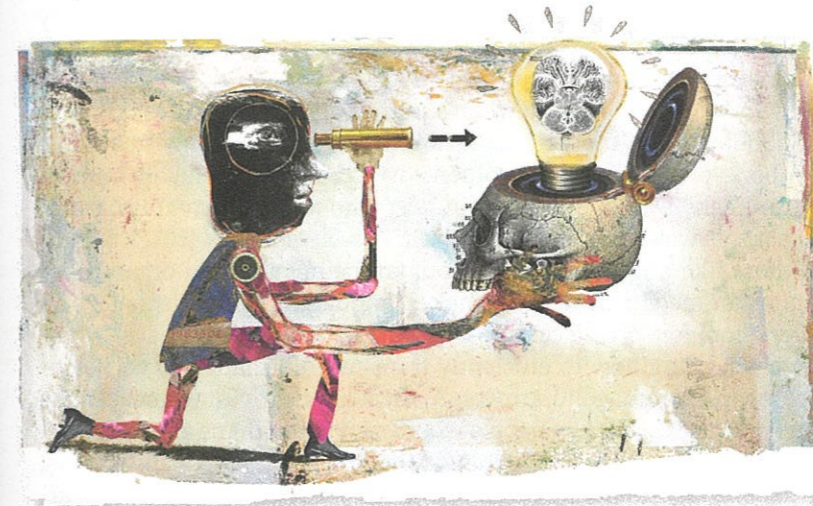# 2 Origins of Human Language

As far as we know, no other species on Earth has language; only humans talk. Sure, many animals communicate with each other in subtle and intricate ways. But we're the only ones who gossip, take seminars, interview celebrities, convene board meetings, recite poems, negotiate treaties, conduct marriage ceremonies, hold criminal trials—all activities where just about the only thing going on is talking.

Fine, we also do many other things that our fellow Earth-creatures don't. We play chess and soccer, sing the blues, go paragliding, design bridges, paint portraits, drive cars, and plant gardens, to name just a few. What makes language so special? Here's the thing: language is deeply distinct from these other activities for the simple reason that *all humans do it*. There is no known society of *Homo sapiens*, past or present, in which people don't talk to each other, though there are many societies where no one plays chess or designs bridges. And all individuals *within* any given human society talk, though again, many people don't play chess or design bridges, for reasons of choice or aptitude.

So, language is one of the few things about us that appears to be a true defining trait of what it means to be human—so much so that it seems it *must* be part of our very DNA. In fact, language has often been described as an innate instinct, something that we are inherently programmed to do. This **nativist view** is presented in Steven Pinker's book *The Language Instinct* (1994). In its strongest version, the nativist position says that not only do our genes program us to have a capacity for language, we're genetically programmed for the thing itself—its general structures, the building blocks that go into it, the mental process of acquiring it, and so on. One way to express this way of thinking is to say that as children, we don't so much *learn* language (the way we learn chess or

nativist view The view that not only are humans genetically programmed to have a general capacity for language, particular aspects of language ability are also genetically specified.

anti-nativist view The view that the ability of humans to learn language is not the result of a genetically programmed "language template," but is an aspect (or by-product) of our extensive cognitive abilities, including general abilities of learning and memory.

piano-playing) as *grow* language based on a genetic blueprint, much as birds grow wings, elephants grow trunks, and female humans grow breasts. This view of language as a genetically specified mental "organ," or as a preprogrammed instinct, captures why it is that language is not only common to all humans but also is unique to humans—no "language genes," no talking.

But many language researchers see it differently. The **anti-nativist view** is that language is not a specialized "organ," but a magnificent by-product of our impressive cognitive abilities. Humans alone learn language, not because we inherit a preprogrammed language template, but because we are the superlearners of the animal kingdom. What separates us from other animals is that our brains have evolved to become the equivalent of swift, powerful supercomputers in comparison with our fellow creatures, who are stuck with more rudimentary technology. Current computers can do qualitatively different things that older models could never aspire to accomplish. This supercomputer theory is one explanation for why we have language while squirrels and chimpanzees don't.

But what about the fact that language is universal among humans, unlike chess or trombone-playing (accomplishments which, though uniquely human, are hardly universal)? Daniel Everett, a linguist who takes a firm anti-nativist position, puts it this way in his book *Language: The Cultural Tool* (2012): Maybe language is more like a tool invented by human beings than an organ that's genetically programmed to grow. What makes language universal is that it's an incredibly *useful* tool for solving certain problems that all humans have—foremost among them being how to efficiently transmit information to each other. Everett compares language to arrows. Arrows are nearly universal among hunter-gatherer societies, but few people would say that humans are genetically programmed to make arrows *specifically*. More likely, making arrows is just part of our general tool-making, problem-solving competence. Bows and arrows can be found in so many different societies because at some point, people who didn't grow their own protein had to figure out a way to catch protein that ran faster than they did. Since it was well within the bounds of human intelligence to solve this problem, humans inevitably did—just as, Everett argues, humans inevitably came to speak with each other as a way of achieving certain pressing goals.

The question of how we came to have language is a huge and fascinating one. If you're hoping that the mystery will be solved by the end of this chapter, you'll be sorely disappointed. It's a question that has no agreed-upon answer among language scientists and, as you'll see, there's a range of subtle and complex views among scientists beyond the two extreme positions I've just presented.

In truth, the various fields that make up the language sciences are not yet even in a position to be able to resolve the debate. To get there, we first need to answer questions like: What *is* language? What do all human languages have in common? What's involved in learning it? What physical and mental machinery is needed to successfully speak, be understood, and understand someone else who's speaking? What's the role of genes in shaping any of the above behaviors? Without doing a lot of detailed legwork to get a handle on all of these smaller pieces of the puzzle, any attempts to answer the larger question about the origins of language can only amount to something like a happy hour discussion—heated and entertaining, but ultimately not that convincing one way or the other. In fact, in 1866, the Linguistic Society of Paris decreed that no papers about the origins of language were allowed to be presented at its conferences. It might seem ludicrous that an academic society would banish an entire topic from discussion. But the decision was essentially a way of saying, "We'll get nowhere talking about language origins until we learn more about language itself, so go learn something about language."

A hundred and fifty years later, we now know quite a bit more about language, and by the end of this book, you'll have a sense of the broad outlines of this body of knowledge. For now, we're in a position to lay out at least a bit of what might be involved in answering the question of why people speak.

2.1 Why Us?

The "language" of bees

Let's start by asking what it is about our language use that's different from what animals do when they communicate. Is it different down to its fundamental core, or is it just a more sophisticated version of what animals are capable of? An interesting starting point might be the "dance language" of honeybees, as identified by Karl von Frisch (1967).

When a worker bee finds a good source of flower nectar at some distance from her hive, she returns home to communicate its whereabouts to her fellow workers by performing a patterned waggle dance (see **Figure 2.1**). During this dance, she repetitively traces a specific path while shaking her body. The elements of this dance communicate at least three things:
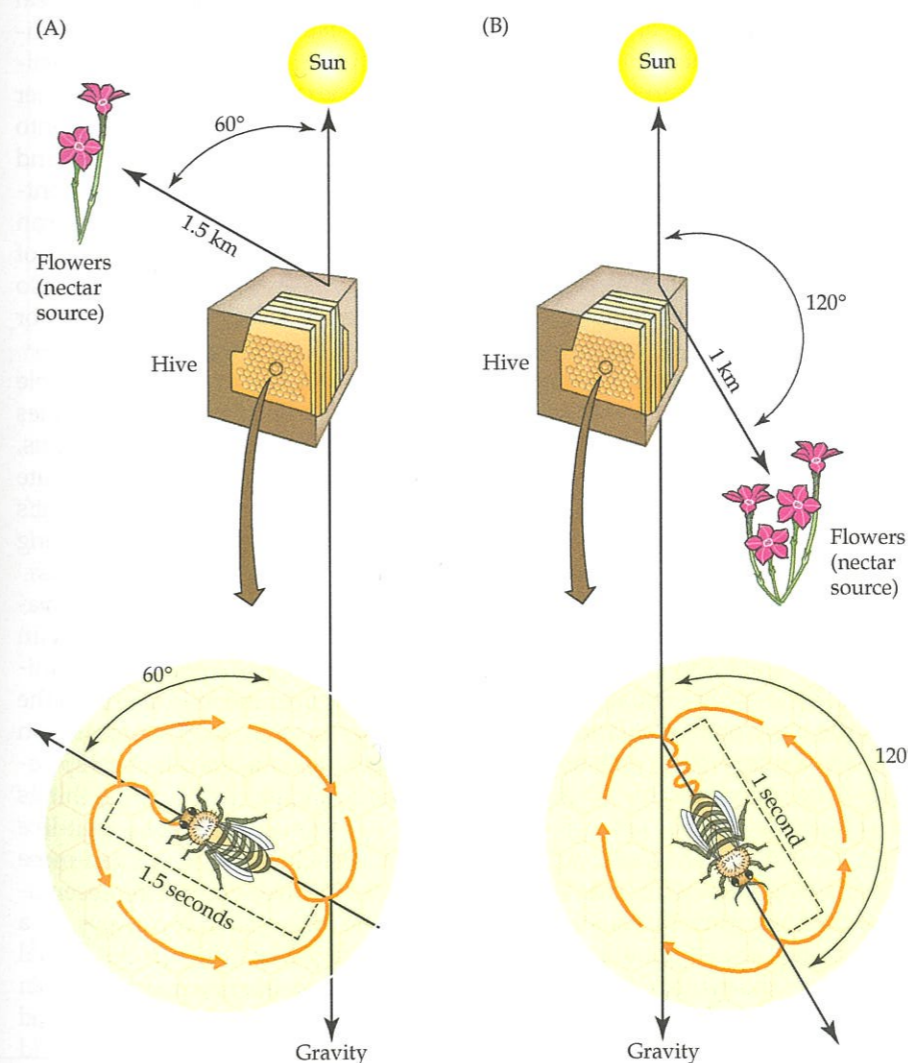

Figure 2.1 The waggle dance of honeybees is used by a returning worker bee to communicate the location and quality of a food source. The worker dances on the surface of the comb to convey information about the direction and distance of the food source, as shown in the examples here. (A) The nectar source is approximately 1.5 km from the hive flying at the indicated angle to the sun. (B) The nectar source is closer and the dance is shorter; in this case the flowers will be found by flying away from the sun. The energy in the bee's waggles (orange curves along the line of the dance) is in proportion to the perceived quality of the find.

1. *The direction in which the nectar source is located*. If the bee moves up toward the top of the hive, this indicates that the nectar source can be found by heading straight toward the sun. The angle of deviation away from a straight vertical path shows the direction relative to the sun.

2. *The distance to the source*. The longer the bee dances along the path from an initial starting point before returning to re-trace the path again, the farther away the source is.

3. *The quality of the source*. If the bee has hit the nectar jackpot, she shakes with great vigor, whereas a lesser source of nectar elicits a more lethargic body wiggle.

Different bee species have different variations on this dance (for example, they might vary in how long they dance along a directional path in order to indicate a distance of 200 meters). It seems that bees have innate knowledge of their own particular dance "dialect," and bees introduced into a hive populated by another species will dance in the manner of their genetic ancestors, not in the style of the adopted hive (though there's some intriguing evidence that bees can learn to *interpret* foreign dialects of other bees; see Fu et al., 2008).

In some striking ways, the honeybee dance is similar to what we do in human language, which is presumably why von Frisch used the term *language* to describe it. The dance uses body movements to represent something in the real world, just as a map or a set of directions does. Human language also critically relies on symbolic representation to get off the ground—for us, it's usually sequences of sounds made in the mouth (for example, "eat fruit"), rather than waggling body movements, that serve as the symbolic units that map onto things, actions, and events in the world. And, in both human languages and bee dances, a smaller number of communicative elements can be independently varied and combined to create a large number of messages—where bees can combine different intensities of wiggling with different angles and durations of the dance path, we can piece together different phrases to similar effect: "Go three miles northwest and you'll find a pretty good Chinese restaurant"; or "There are some amazing raspberry bushes about thirty feet to your left."

Honeybee communicative behavior shows that a complex behavior capable of transmitting information about the real world can be encoded in the genes and innately specified, presumably through an evolutionary process. Like us, honeybees are highly cooperative and benefit from being able to communicate with each other. But bees are hardly among our closest genetic relatives, so it's worth asking just how similar their communicative behavior is to ours. Along with the parallels I've just mentioned, there are also major differences.

Most importantly, bee communication operates within much more rigid parameters than human language. The elements in the dance, while symbolic in some sense, are still closely bound to the information that's being communicated. The angle of the dance path describes the angle of the food source to the sun; the duration of the dance describes the distance to the food source. But in human language, there's usually a purely arbitrary or accidental relationship between the communicative elements (that is, words and phrases) and the things they describe; the word *fruit*, for example, is not any more inherently fruit-like than the word *leg*. In this sense, what bees do is less like using words and more like drawing maps with their bodies. A map does involve symbolic representation, but the forms it uses are constrained by the information it conveys. In a map, there's always some transparent, non-arbitrary way in which the spatial relations in the symbolic image relate to the real world. No one makes maps in which, for example, all objects colored red—regardless of where they're placed in the image—are actually found in the northeast quadrant of the real-world
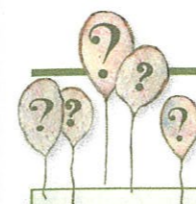
BOX 2.1
Hockett's design features of human language

1. **Vocal-auditory channel** Language is produced in the vocal tract and transmitted as sound; it's perceived through the auditory channel.

2. **Broadcast transmission and directional reception** Language can be heard from many directions, but it is perceived as coming from one particular location.

3. **Rapid fading** The sound produced by speech fades quickly.

4. **Interchangeability** A user of a language can send and receive the same message.

5. **Total feedback** Senders of a message can hear and internalize the message they've sent.

6. **Specialization** The production of the sounds of language serves no purpose other than to communicate.

7. **Semanticity** There are fixed associations between units of language and aspects of the world.

8. **Arbitrariness** The meaningful associations between language and the world are arbitrary.

9. **Discreteness** The units of language are separate and distinct from one another rather than being part of a continuous whole.

10. **Displacement** Language can be used to communicate about things that are not present in time and/or space.

11. **Productivity** Language can be used to say things that have never been said before, and yet are understandable to the receiver.

12. **Traditional transmission** The specific language that's adopted by the user has to be learned by exposure to other users of that language; its precise details are not available through genetic transmission.

13. **Duality of patterning** Many meaningful units (words) are made by the combination of a small number of elements (sounds) into various sequences. For example, pat, tap, and apt use the same sound elements combined in different ways to make different word units. In this way, tens of thousands of words can be created from several dozen sounds.

14. **Prevarication** Language can deliberately be used to make false statements.

15. **Reflexiveness** Language can be used to refer or describe itself.

16. **Learnability** Users of one language can learn to use another, different language.

Adapted from Hockett, 1960, and Hockett & Altmann, 1968.

space being described, while the color yellow is used to signal objects in the southwest quadrant, regardless of where they appear in the image.

Another severe limitation of bee dances is that bees only "talk" about one thing: where to find food (or water) sources. Human language, on the other hand, can be recruited to talk about an almost infinite variety of topics for a wide range of purposes, from giving directions, to making requests, to expressing sympathy, to issuing a promise, and so on. Finally, human language involves a complexity of structure that's just not there in the bees' dance language.

To help frame the discussion about how much overlap there is between animal communication systems and human language, the well-known linguist Charles Hockett listed a set of "design features" that he argued are common to all human languages. The full list of **Hockett's design features** is given in **Box 2.1**; you may find it useful to refer back to this list as the course progresses. Even though some of the features are open to challenge, they provide a useful starting point for fleshing out what human language looks like.

Primate vocalizations

If we look at primates—much closer to us genetically than bees—a survey of their vocal communication shows a pretty limited repertoire. Monkeys and apes do make meaningful vocal sounds, but they don't make very many, and

Hockett's design features A set of characteristics proposed by linguist Charles Hockett to be universally shared by all human languages. Some (but not all) of the features are also found in various animal communication systems.

the ones they use seem to be limited to very specific purposes. Strikingly absent are many of the features described by Hockett that allow for inventiveness, or the capacity to re-use elements in an open-ended way to communicate a varied assortment of messages.

For example, vervet monkeys produce a set of alarm calls to warn each other of nearby predators, with three distinct calls used to signal whether the predator is a leopard, an eagle, or a snake (as found by Seyfarth, Cheney, and Marler, 1980). Vervets within earshot of these calls behave differently depending on the specific call: they run into trees if they hear the leopard call, look up if they hear the eagle call, and peer around in the grass when they hear the snake alarm. These calls do exhibit Hockett's feature of *semanticity*, as well as an *arbitrariness* in the relationship between the signals and the meaning they transmit. But they clearly lack Hockett's feature of *displacement*, since the calls are only used to warn about a clear and present danger and not, for example, to suggest to a fellow vervet that an eagle *might* be hidden in that tree branch up there, or to remind a fellow vervet that this was the place where we saw a snake the other day. There's also no evidence of *duality of patterning*, in which each call would be made by combining similar units together in different ways. And vervets certainly don't show any signs of *productivity* in their language, in which the calls are adapted to communicate new messages that have never been heard before but that can be easily understood by the hearer vervets. In fact, vervets don't even seem to have the capacity to *learn* to make the various alarm calls; the sounds of the alarm calls are fixed from birth and are instinctively linked to certain categories of predators, though baby vervets do have to learn, for example, that the eagle alarm shouldn't be made in response to a pigeon overhead. So, they come by these calls not through the process of *cultural transmission*, which is how humans learn words (no French child is born knowing that *chien* is the sound you make when you see a dog), but by being genetically wired to make specific sounds that are associated with specific meanings.

This last point has some very interesting implications. Throughout the animal world, it seems that the exact shape of a communicative message often has a strong genetic component. If we want to say that humans are genetically wired for language, then that genetic programming is going to have to be much more fluid and adaptable than that of other animals, allowing humans to learn a variety of different languages through exposure. Instead of being programmed for a specific language, we're born with the capacity to learn any language. This very fact might look like overwhelming support for the anti-nativist view, which says that language is simply an outgrowth of our general ability to learn complex things. But not necessarily. The position of nativists is more subtle than simply arguing that we're born with knowledge of a specific language. Rather, the claim is that there are common structural ingredients to all human languages, and that it's these basic building blocks of language that we're all born with, whether we use them to learn French or Sanskrit. More on this later.

One striking aspect of primate vocalizations is the fact that monkeys and apes show much greater flexibility and capacity for learning when it comes to *interpreting* signals than in producing them. (A thorough discussion of this asymmetry can be found in a paper by primatologists Robert Seyfarth and Dorothy Cheney, 2010.) Oddly enough, even though vervets are born knowing which sounds to make in the presence of various predators, they don't seem to be born with a solid understanding of the meanings of these alarms, at least

as far as we can tell from their responses to the calls. It takes young vervets several months before they start showing the adult-like responses of looking up, searching in the grass, and so on. Early on, they respond to the alarm calls simply by running to their mothers, or reacting in some other way that doesn't show that they know that an eagle call, for example, is used to warn specifically about bird-like predators. Over time, though, their ability to extend their understanding of new calls to new situations exceeds their adaptability in producing calls. For instance, vervets can learn to understand the meanings of alarm calls of other species, as well as the calls of their predators—again, even though they never learn to *produce* the calls of other species.

Seyfarth and Cheney suggest that the information that primates can pull out from the communicative signals they hear can be very subtle. An especially intriguing example comes from an experiment involving the call behavior of baboons. Baboons, as it happens, have a very strict status hierarchy within their groups, and it's not unusual for a higher-status baboon to try to intimidate a lower-status baboon by issuing a threat-grunt, to which the lower-ranking animal usually responds with a scream. The vocalizations of individual baboons are distinctive enough that they're easily recognized by all members of the group. For the purpose of the study, the researchers created a set of auditory stimuli in which they cut and spliced together prerecorded threat-grunts and screams from various baboons within the group. The sounds were reassembled so that sometimes the threat-call of a baboon was followed by a scream from another baboon higher up in the status hierarchy. This is a situation that overhearing baboons would be very unlikely to ever hear in nature. Yet they reacted to this unusual pairing of sounds with surprise, which seems to show that the baboons had inferred from the sequence of sounds that a lower-status animal was trying to intimidate a higher-status animal—and understood that this was a bizarre state of affairs.

It may seem strange that animals' ability to understand something about the world based on a communicative sound is so much more impressive than their ability to *convey* something about the world by creating a sound. But this asymmetry seems rampant within the animal kingdom. Many dog owners are intimately familiar with this fact. It's not hard to get your dog to recognize and respond to dozens of verbal commands. It's getting your dog to talk back to you that's more difficult. Any account of the evolution of language will have to grapple with the fact that speaking and understanding are not necessarily just the mirror image of each other.

Can language be taught to apes?

As you've seen, when left to themselves in the wild, non-human primates don't indulge in much language-like vocalization. This would suggest that the linguistic capabilities of humans and other primates are markedly different. Still, a non-nativist might object and argue that looking at what monkeys and apes do among themselves, without the benefit of any exposure to real language, doesn't really provide a realistic picture of what they *can* learn about language. After all, when we evaluate human infants' capacity for language, we don't normally separate them from competent language users—in other words, adults— and see what they come up with on their own. Suppose language really is more like a tool than a biological organ, with each generation of humans benefiting from the knowledge of the previous generation. In that case, in order to see whether primates are truly capable of attaining language, we need to see what they can learn when they're allowed to have many rich interactions with individuals who have already solved the problem of language.

This line of thinking has led to a number of studies that have looked at how apes communicate, not with other non-linguistic apes, but with their more verbose human relatives. In these studies, research scientists and their assistants have raised young apes (i.e., chimpanzees, bonobos, orangutans, and gorillas) among humans in a language-rich environment. Some of the studies have included intensive formal teaching sessions, with a heavy emphasis on rewarding and shaping communicative behavior, while other researchers have raised the apes much as one would a human child, letting them learn language through observation and interaction. Such studies often raise tricky methodological challenges, as discussed in Method 2.1. For example, what kind of evidence is needed in order to conclude that apes know the meaning of a word in the sense that humans understand that word? Nevertheless, a number of interesting findings have come from this body of work (a brief summary can be found in a review article by Kathleen Gibson, 2012).

First, environment matters: there's no doubt that the communicative behavior of apes raised in human environments starts to look a lot more human-like than that of apes in the wild. For example, a number of apes of several different species have mastered hundreds of words or arbitrary symbols. They spontaneously use these symbols to communicate a variety of different functions, not just to request objects or food that they want, but also to comment on the world around them. They also refer to objects that are not physically present at the time, showing evidence of Hockett's feature of *displacement*, which was conspicuously absent from the wild vervets' alarm calls. They can even use their symbolic skills to lie—for instance, one chimp was found to regularly blame the messes she made on others. Perhaps even more impressively, all of the species studied have shown at least some suggestion of another of Hockett's features, **productivity**—that is, of using the symbols they know in new combinations to communicate ideas for which they don't already have symbols. For example, Koko, a gorilla, created the combination "finger bracelet" to refer to a ring; Washoe, a chimpanzee, called a Brazil nut a "rock berry." Sequences of verbs and nouns often come to be used by apes in somewhat systematic sequences, suggesting that the order of combination isn't random.

As in the wild, trained apes show that they can master comprehension skills much more readily than they achieve the production of language-like units. In particular, it quickly became obvious that trying to teach apes to use vocal sounds to represent meanings wasn't getting anywhere. Apes, it turns out, have extremely limited control over their vocalizations and simply can't articulate different-sounding words. But the trained apes were able to build up a sizable vocabulary when signed language was substituted for spoken language, or when researchers adopted custom-made artificial "languages" using visual symbols arranged in systematic structures. This raises the very interesting question of how closely the evolution of language is specifically tied to the evolution of speech, an issue we will probe in more detail in Section 2.4.

But even with non-vocal languages, the apes were able to handle much more complexity in their understanding of language than in their production of it. They rarely produced more than two or three symbols strung together, but several apes were able to understand commands like "make the doggie bite the snake," and they could distinguish that from "make the snake bite the doggie." They could also follow commands that involved moving objects to or from specific locations. Sarah, a chimpanzee, could reportedly even understand "if/then" statements.

Looking at this collection of results, it becomes apparent that with the benefit of human teachers, ape communication takes a great leap toward human language—human-reared apes don't just acquire more words or symbols than they do in the wild, they also show that they can master a number of Hock-

productivity The ability to use known symbols or linguistic units in new combinations to communicate ideas.

METHOD 2.1

Minding the gap between behavior and knowledge

If a chimpanzee produces the sign for *banana* and appears satisfied when you retrieve one from the kitchen and give it to her, does this mean the chimp knows that the sign is a symbol for the *idea* of banana and is using it as you and I would use the word? It's certainly tempting to think so, but the careful researcher will make sure not to overinterpret the data and jump to conclusions about sophisticated cognitive abilities when the same behavior could potentially be explained by much less impressive abilities.

Since chimpanzees look and act so much like us in so many ways, it's tempting to conclude that when they behave like us, it's because they *think* like us. But suppose instead of interacting with a chimp you were observing a pigeon that had learned to peck on keys of different colors in order to get food rewards. How willing would you be to conclude that the pigeon is treating the colored keys as symbols that are equivalent to words? My guess is, not very. Instead, it seems easy enough to explain the pigeon's behavior by saying that it's learned to associate the action of pecking a particular key with getting a certain reward. This is a far cry from saying that the bird is using its action as a symbol that imparts a certain thought into the mind of the human seeing it, intending that this implanted thought might encourage the human to hand over food. In linking behavior to cognition, researchers need to be able to suspend the presumption of a certain kind of intelligence and to treat chimps and pigeons in exactly the same way. In both cases, they need to ask: What evidence do we need to have in order to be convinced that the animal is using a word in the same way that a human does? And how do we rule out other explanations of the behavior that are based on much simpler cognitive machinery than that of humans?

Sue Savage-Rumbaugh is one of the leading scientists in the study of primate language capabilities. In 1980, she and her colleagues wrote a paper cautioning researchers against making overly enthusiastic claims when studying the linguistic capabilities of apes. They worried about several possible methodological flaws. One of these was

a tendency to overattribute human-like cognition to simple behaviors, as discussed in the paragraph above. They argued that in order to have evidence that an ape is using a sign or symbol referentially, you need to be able to show that the animal not only produces the sign in order to achieve a specific result, but also shows evidence of understanding it—for example, that it can pick out the right object for a word in a complex situation that involves choosing from among many possibilities. You also need to be able to show that the ape can produce the word in a wide variety of situations, not just to bring about a specific result. Moreover, you need to look at *all* the instances in which the ape uses a particular sign. It's not enough to see that *sometimes* the chimp uses the sign in a sensible way; if the same chimp also uses the sign in situations where it seems inappropriate or not meaningful, it lessens the confidence that the chimp truly knows the meaning of that sign.

Savage-Rumbaugh and her colleagues also worried about the possibility that researchers might unknowingly provide cues that nudge the ape to produce a sign that seems sensible given the context. Imagine a possible scenario in which a chimpanzee can produce a set of signs that might be appropriate in a certain context—for example, signs that result in getting food. If the chimp doesn't really know the meanings of any of these individual signs, it might start sloppily producing some approximate hand movements while watching the researcher's face. The researcher might inadvertently communicate approval or disapproval, thereby steering the chimp's signing behavior in the right direction. In order to be able to fairly evaluate what a chimp knows, the researcher has to set up rigorous testing situations in which the possibility of such cues has been eliminated.

Savage-Rumbaugh and her colleagues ultimately concluded that apes *are* able to learn to use symbols in a language-like way. But their remarks about good methodological practices apply to any aspect of language research with primates—including the young human variety, as we'll see in upcoming chapters.

ett's design features that are completely absent from their naturalistic behavior. This is very revealing, because it helps to answer the question of when some of these features of human language—or rather, the *capability* for these features—might have evolved.

Figure 2.2 The evolutionary history of hominids. The term *hominids* refers to the group consisting of all modern and extinct great apes (including humans and their more immediate ancestors). This evolutionary tree illustrates the common ancestral history and approximate times of divergence of *hominins* (including modern humans and the now-extinct Neanderthals) from the other great apes. Note that a number of extinct hominin species are not represented here.
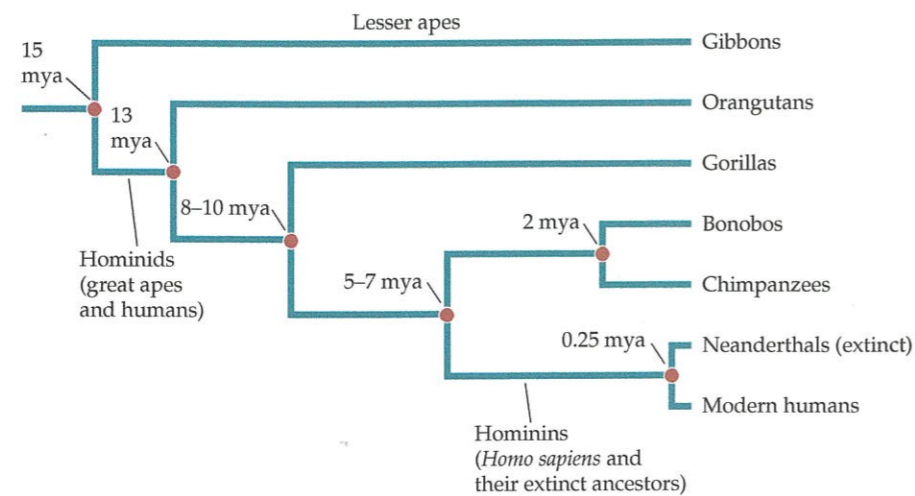
Biologists estimate that humans, chimpanzees, and bonobos shared a common ancestor between 5 and 7 million years ago. The last common ancestor with gorillas probably occurred between 8 and 10 million years ago, and the shared ancestor with orangutans even earlier than that (see **Figure 2.2**). Evidence about *when* the features of human language evolved helps to answer questions about whether they evolved specifically because these features support language.

Among nativists, the most common view is that humans have some innate capabilities for language that evolved as adaptations. **Evolutionary adaptations** are genetically transmitted traits that give their bearers an advantage—specifically, an adaptive trait helps individuals with that trait to stay alive long enough to reproduce and/or to have many offspring. The gene for the advantageous trait spreads throughout a population, as over time members of the species with that trait will out-survive and out-reproduce the members without that trait. But not all adaptations that help us to use language necessarily came about *because* they gave our ancestors a communicative edge over their peers. Think about it like this: humans have hands that are capable of playing the piano, given instruction and practice. But that doesn't mean that our hands evolved as they did because playing the piano gave our ancestors an advantage over non-piano-playing humans. Presumably, our nimble fingers came about as a result of various adaptations, but the advantages these adaptations provided had nothing to do with playing the piano. Rather, they were the result of the general benefits of having dexterous hands that could easily manipulate a variety of objects. Once in possession of enhanced manual agility, however, humans discovered that hands can be put to many wonderful uses that don't necessarily help us survive into adulthood or breed successfully.

The piano-playing analogy may help to make sense of the question, "If language-related capabilities evolved long before humans diverged from other apes, then why do only humans make use of them in their natural environments?" That is, if apes are capable of amassing bulky vocabularies and using them creatively, why are they such linguistic underachievers in the wild? The contrast between their communicative potential and their lack of spontaneous language in the wild suggests that certain cognitive skills that are required to master language—at least, those skills that are within the mental grasp of apes—didn't necessarily evolve *for* language. Left to their own devices, apes don't appear to use these skills for the purpose of communicating with each other. But when the cultural environment calls for it, these skills can be recruit-

ed in the service of language—much as in the right cultural context, humans can use their hands to play the piano. This state of affairs poses a challenge to the "language-as-organ" view.

Nevertheless, it's entirely possible that the skills that support language fall into two categories: (1) those that are necessary to get language off the ground but aren't really specific to language; and (2) traits that evolved particularly because they make language more powerful and efficient. It may be that we share the skills in the first category with our primate relatives, but that only humans began to use those skills for the purpose of communication. Once this happened, there may have been selective pressure on other traits that provided an additional boost to the expressive capacity of language—and it's these later skills that are both language-specific and uniquely human.

It seems, then, that when we talk about language evolution, it doesn't make sense to treat language as an all-or-nothing phenomenon. Language may well involve a number of very different cognitive skills, with different evolutionary trajectories and different relationships to other, non-linguistic abilities.

Throughout this book, you'll get a much more intimate sense of the different cognitive skills that go into human language knowledge and use. As a first step, this chapter will start by breaking things down into three very general categories of language-related abilities: the ability to understand communicative intent, a grasp of linguistic structure, and the ability to control voice and/or gesture.

2.2 The Social Underpinnings of Language

Understanding the communicative urge

Imagine this scene from long ago: an early hominid is sitting at the mouth of his cave with his female companion when a loud roar tears through the night air. He nods soberly and says, "Leopard." This is a word that he's just invented to refer to that animal. In fact, it's the first word that's passed between them, as our male character is one of language's very earliest adopters. It's a breakthrough: from here on, the couple can use the word to report leopard sightings to each other, or to warn their children about the dangerous predator. But none of this can happen unless the female can clue in to the fact that the sounds in *leopard* were intentionally formed to communicate an idea—and were not due to a sneeze or a cough, or some random set of sounds. What's more, she has to be able to connect these intentional and communicative sounds with what's going on around her, and make a reasonable guess about what her companion is most likely to be trying to convey.

From your perspective as a modern human, all of this may seem pretty obvious, requiring no special abilities. But it's far from straightforward, as revealed by some very surprising tests that chimpanzees fail at miserably, despite their substantial intellectual gifts. For example, try this next time you meet a chimp: Show the animal a piece of food, and then put it in one of two opaque containers. Shuffle the two containers around so as to make it hard to tell where it's hidden. Now stop, and point to the container with the food. The chimpanzee will likely choose randomly between the two containers, totally oblivious to the very helpful clue you've been kind enough to provide. This is exactly what Michael Tomasello (2006) and his colleagues found when they used a similar test with chimpanzees. Their primate subjects ignored the conspicuous hint even though the experimenters went out of their way to establish that the "helper" who pointed had indeed proven herself to be helpful on earlier occasions by tilting the containers so that the chimp could see which container had the food (information that the chimps had no trouble seizing upon).

evolutionary adaptation A genetically transmitted trait that gives its bearers an advantage—specifically, it helps those with the trait to stay alive long enough to reproduce and/or to have many offspring.

Chimpanzees' failure to follow the pointing cue is startling because chimps are very smart, perfectly capable of making subtle inferences in similar situations. For example, if an experimenter puts food in one of two containers and then shakes one of them but the shaken container produces no rattling sound, the chimpanzee knows to choose the other one (Call, 2004). Or, consider this variation: Brian Hare and Michael Tomasello (2004) set up a competitive situation between chimpanzees and a human experimenter, with both human and chimp trying to retrieve food from buckets. If the human extended her arm toward a bucket but couldn't touch it because she had to stick her hand through a hole that didn't allow her to reach far enough, the chimpanzees were able to infer that this was the bucket that must contain the food, and reached for it. Why can chimpanzees make *this* inference, which involves figuring out the human's intended—but thwarted—goal, but not be able to understand pointing? Tomasello and his colleagues argued that, whereas chimpanzees can often understand the intentions and goals—and even the knowledge states—of other primates, what they can't do is understand that pointing involves an *intention to communicate*. In other words, they don't get that the pointing behavior is something that's done not just for the purpose of satisfying the pointer's goal, but to help the chimpanzee satisfy *its* goal (see **Box 2.2**).
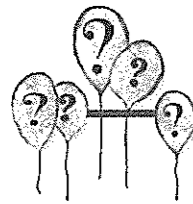
BOX 2.2
Dogs versus chimps: A pointed difference

Perhaps you were surprised to learn that chimpanzees don't understand pointing. Maybe you even thought, "That's impossible—even my dog understands that." Now, everyone knows that dog owners are prone to inflate their pets' intelligence in their own minds, but in this case, you'd be right. Dogs *do* understand pointing, as confirmed by studies using the same hidden food test that chimpanzees fail. Apparently, dogs respond to pointing even when the experimenter walks toward the wrong food container but points in the opposite direction at the right container.

In an interesting article, Brian Hare and Michael Tomasello (2005) describe a number of situations in which dogs outperform chimps on tests of social cognition—for example, dogs but not chimps are more likely to stay away from forbidden food when a nearby human's eyes are open than when they're closed. But it's not that dogs are smarter than chimps overall—dogs fail certain tests that chimps pass with flying colors. For example, chimps are more likely to pull on a string that's connected to food than one that's not, showing that they can understand simple cause-effect relationships; but dogs perform dismally on this test. So, the two species show an interesting inversion: dogs are better than chimps on tests having to do with social cognition, whereas chimps are better on tests that probe their understanding of the physical world.

Is the social intelligence of dogs due to good dog-rearing practices by humans? Apparently not. Puppies do well on the pointing test at a very young age, whether or not they've had much contact with humans. This suggests that the behavior has a genetic basis and isn't just the result of socialization. A second reasonable hypothesis is that complex social skills have evolved in dogs as a result of their pack structure in the wild. But the interesting thing is that domestic dogs are *better* than human-raised wolves at using social cues to find food, even though wolves perform at least as well as dogs on non-social problem-solving tests. So, the social skills of dogs do seem to have a hereditary basis, but *these skills appear to have evolved specifically in canines that have contact with humans*, during the process of domestication.

Added evidence for this conclusion comes from intriguing experiments with foxes. In 1959, Siberian scientists began an experimental breeding program in which they selected individual foxes that showed an inclination to approach humans without fear or aggression. Decades later, Brian Hare collaborated with these scientists to test foxes from this experimental population on their ability to use social cues (Hare et al., 2005). The "tame" foxes performed better on these tests than foxes of the same species that had not been part of this special breeding program—in fact, they did as well as puppies born to domestic dogs. In other words, selection for the general trait of seeking out human company seems to have resulted in human-like social skills.

To some researchers, it's exactly this ability to understand communicative intentions that represents the "magic moment" in the evolution of language, when our ancestors' evolutionary paths veered off from those of other great apes, and their cognitive skills and motivational drives came to be refined, either specifically for the purpose of communication, or more generally to support complex social coordination.

Some language scientists have argued that a rich communication system is built on a foundation of advanced skills in social cognition, and that among humans these skills evolved in a super-accelerated way, far outpacing other gains we made in overall intelligence and working memory capacity. To test this claim, Esther Hermann and her colleagues (2007) compared the cognitive abilities of adult chimpanzees, adult orangutans, and human toddlers aged two and a half. All of these primates were given a battery of tests evaluating two kinds of cognitive skills: those needed for understanding the physical world, and those for understanding the social world. For example, a test item in the physical world category might involve discriminating between a smaller and a larger quantity of some desirable reward, or locating the reward after it had been moved, or using a stick to retrieve an out-of-reach reward. The socially oriented test items looked for accomplishments like solving a problem by imitating someone else's solution, following a person's eye gaze to find a reward, or using or interpreting communicative gestures to locate a reward. The researchers found that in demonstrating their mastery over the physical world, the human toddlers and adult chimpanzees were about even with each other, and slightly ahead of the adult orangutans. But when it came to the social test items, the young humans left their fellow primates in the dust (with chimps and orangutans showing similar performance).

There's quite a bit of additional evidence showing that even very young humans behave in ways that are quite different from how other primates act in similar situations. For example, apes don't seem to be inclined to communicate with other apes for the purpose of helping the *others* achieve a goal, when there's nothing obvious in it for themselves. But little humans almost feel compelled to. In one study by Ulf Lizskowski and colleagues (2008), 12-month-olds who hadn't yet begun to talk watched while an adult sat at a table stapling papers without involving the child in any way. At one point, the adult left the room, then another person came in, moved the stapler from the table to a nearby shelf, and left. A little later, the first adult came back, looked around, and made quizzical gestures to the child. In response, most of the children pointed to the stapler in its new location. Apparently, this is very un-ape-like behavior; according to Michael Tomasello (2006), apes never point with each other, and when they do "point" to communicate with humans (usually without extending the index finger), it's because they want the human to fetch or hand them something that's out of their reach.

Skills for a complex social world

So, humans are inclined to share information with one another whereas other primates seem not to have discovered the vast benefits of doing so. What's preventing our evolutionary cousins from cooperating in this way? One possibility is that they're simply less motivated to engage in deeply social behavior than we humans are. Among mammals, we as a species are very unusual in the amount of importance we place on social behavior. For example, chimpanzees are considerably less altruistic than humans when it comes to sharing food, and they don't seem to care that much about norms of reciprocity or fairness. When children are in a situation where one child is dividing up treats to share and extends an offer that is much smaller than the share he's claimed for himself,

the other child is apt to reject the offer, preferring to give it up in order to make the point that the meager amount is an affront to fairness. A chimp will take what it can get (Tomasello, 2009).

In fact, when you think about the daily life of most humans in comparison to a day in the life of a chimpanzee, it becomes apparent that our human experiences are shaped very profoundly by a layer of social reality, while a chimpanzee may be more grounded in the physical realities of its environment. In his book *Why We Cooperate* (2009), Michael Tomasello points out how different the human experience of shopping is from the chimpanzee's experience of foraging for food:

> Let us suppose a scenario as follows. We enter a store, pick up a few items, stand in line at the checkout, hand the clerk a credit card to pay, take our items, and leave. This could be described in chimpanzee terms fairly simply as going somewhere, fetching objects, and returning from whence one came. But humans understand shopping, more or less explicitly, on a whole other level, on the level of institutional reality. First, entering the store subjects me to a whole set of rights and obligations: I have the right to purchase items for the posted price, and the obligation not to steal or destroy items, because they are the property of the store owner. Second, I can expect the items to be safe to eat because the government has a department that ensures this; if a good proves unsafe, I can sue someone. Third, money has a whole institutional structure behind it that everyone trusts so much that they hand goods over for this special paper, or even for electronic marks somewhere from my credit card. Fourth, I stand in line in deference to widely held norms, and if I try to jump the line people will rebuke me, I will feel guilty, and my reputation as a nice person will suffer. I could go on listing, practically indefinitely, all of the institutional realities inhibiting the public sphere, realities that foraging chimpanzees presumably do not experience at all.

Put in these terms, it becomes obvious that in order to successfully navigate through the human world, we need to have a level of social aptitude that chimpanzees manage without.

At some level, the same socially oriented leanings that drive humans to "invent" things like laws and money also make it possible for them to communicate through language. Language, law, and currency all require people to buy into an artificial system that exists only because everyone agrees to abide by it. Think about it: unlike vervets with their alarm calls, we're not genetically programmed to produce specific sounds triggered by specific aspects of our environment. Nor do our words have any natural connection to the world, in the way that honeybee dance language does. Our words are quite literally figments of human imagination, and they have meaning only because we all agree to use the same word for the same thing.

But it may not just be an issue of general social motivation that's keeping our primate relatives from creating languages or laws of their own. It's possible that they also lack a specific cognitive ingredient that would allow them to engage in complex social coordination. In order to do something as basic as make a smart guess about what another person's voluntary mouth noises might be intended to mean, humans needed to have the capacity for **joint attention**: the awareness between two (or more) individuals that they are both paying attention to the same thing. Again, this doesn't seem especially difficult, but

joint attention The awareness between two or more individuals that they are paying attention to the same thing.

Tomasello and his colleagues have argued that, to any reliable extent, this capacity is found *only* in humans. Chimps can easily track the gaze of a human or another ape to check out what's holding the interest of the other; they can also keep track of what the others know or have seen. In other words, chimps can know what others know. But there's no clear evidence that they participate in situations where Chimp A knows that Chimp B knows that Chimp A is staring at the same thing. Presumably, this is exactly the kind of attunement that our ancestors sitting by their caves would have needed to have in order to agree on a word for the concept of leopard.

It turns out that joint attention skills are very much in evidence in extremely young humans. Toward the end of their first year (on average), babies unambiguously direct other people's attention to objects by pointing, often with elaborate vocalization. Months prior to this, they often respond appropriately when others try to direct their attention by pointing or getting them to look at something specific. It's especially relevant that success with language seems to be closely linked to the degree to which children get a handle on joint attention. For example, Michael Morales and his colleagues (2000) tracked a group of children from 6 to 30 months of age. The researchers tested how often individual babies responded to their parents' attempts to engage them in joint attention, beginning at 6 months of age; they then evaluated the size of the children's vocabularies at 30 months and found that the more responsive the babies were at 6 months, the larger their vocabularies were later on. Another study by Cristina Colonnesi and her colleagues (2010) documented evidence of a connection between children's pointing behaviors and the emergence of language skills. Interestingly, the connection was apparent for *declarative* pointing—that is, pointing for the purpose of "commenting" on an object—but not for *imperative* pointing to direct someone to do something. This is intriguing because when apes do communicate with humans, they seem to do much less commenting and much more directing of actions than human children do, even early in their communication. Both are very clearly *communicative* acts, and yet they may have different implications for linguistic sophistication.

It's increasingly apparent, then, that being able to take part in complex social activities that rely on mutual coordination is closely tied to the emergence of language. Researchers like Michael Tomasello have argued that there is a sharp distinction between humans and other apes when it comes to these abilities. But there's controversy over just how sharp this distinction is. It's also not obvious whether these abilities are all genetically determined or whether skills such as joint attention also result from socialization. What should we make, for example, of the fact that apes raised by humans are able to engage in much more sophisticated communication than apes raised by non-humans? Furthermore, we don't have a tremendous amount of detailed data about apes' capabilities for social cognition. Much research has focused on the limitations of non-human primates, but newer studies often show surprisingly good social abilities. For instance, many scientists used to think that chimpanzees weren't able to represent the mental states of others, but it now appears that they're better at it than had been thought (see **Method 2.2**; Call & Tomasello, 2010).

But whether or not the rich social skills of humans reveal a uniquely human adaptation, and whether or not these adaptations occurred largely to support communication (or more generally to support complex social activity), there are other skills that we need in order to be able to command the full expressive power of language. These other skills, in turn, may or may not be rooted in uniquely human adaptations for language, a theme we'll take up in the coming section.

METHOD 2.2

Exploring what primates can't (or won't) do

Discussions about the evolution of language rely heavily on comparative studies, in which the abilities of different species are evaluated against each other in targeted experiments. These comparisons usually involve designing a test that seems appropriate for all of the species being studied, then applying the same test to members of all the species. The performance of the test subjects is then presumed to reveal something about their cognitive abilities.

But, as any schoolteacher knows, test results don't always tell the whole story when it comes to assessing a student's abilities or knowledge. Some students are highly intelligent but simply don't care, and so their test results are mediocre. Other students become anxious in test situations and freeze up, unable to fully access their knowledge. Test results for all of these students will underestimate what they know or are capable of doing. It's worth keeping this in mind when it comes to using performance on experimental tests as a way to gauge the abilities of various species.

Suppose we're trying to compare a human 2-year-old to an adult chimpanzee. Can we assume that the same test given to both will be an equally good measure of the targeted skills of each? Not necessarily. For instance, the very fact that a *human* experimenter is in charge might have very different effects on the two subjects. The child might be more motivated than the chimp to perform well for the human. The chimp might be more anxious than the child in the presence of a human.

Eliminating these potential problems isn't always possible—after all, we can't readily train a chimp to administer an experiment. But some potential problems can be avoided by paying close attention to the design

of the experiments. Specifically, in testing to see whether primates have a particular skill, it makes sense to ask in what situations that skill would be useful to them in the wild. These are the situations in which we'd expect that they'd be motivated to use the relevant skill, and be less disoriented by the experimental setup. This strategy has led researchers to conclude that some past experiments have underestimated the social cognition of chimpanzees.

For example, in one study (Povinelli & Eddy, 1996), chimpanzees appeared to be unable to imagine the visual experience of another individual. The test involved putting a chimpanzee subject in the presence of two humans with food placed between them. One of the humans had a bucket over his head, and hence couldn't see the chimp. The experimenters were interested in learning whether the chimps would choose to beg for food from the human who could see them. Interestingly, their study subjects begged randomly from either human, as if oblivious to the fact that the person wearing the bucket was temporarily blind.

But a number of researchers have argued that, in the wild, a chimpanzee's social understanding is most useful in competitive situations than cooperative ones. A later study (Bräuer et al., 2007) built its experimental design around this premise. In the competitive setup, a lower-ranking chimpanzee competed with a higher-ranking chimpanzee for food. Two pieces of food were present, one of which could be seen by both chimps but the other of which was hidden from the higher-ranking chimp by a barrier. More often than not, the subordinate chimp decided to go for the food that the dominant chimp couldn't see.

The moral of the story seems to be, if you want to find out what a chimp is thinking, you have to think like a chimp as you set up your study.

2.3 The Structure of Language

Combining units

Being able to settle on arbitrary symbols as stand-ins for meaning is just one part of the language puzzle. There's another very important aspect to language, and it's all about the art of combination.

Combining smaller elements to make larger linguistic units takes place at two levels. The first level deals with making words from sounds. In principle, we *could* choose to communicate with each other by creating completely different sounds as symbols for our intended meanings—a high-pitched yowl might mean "arm," a guttural purr might mean "broccoli," a yodel might mean

"smile," and so on. In fact, this is very much how vervets use sound for their alarm calls. But at some point, we might find we'd exhausted our ability to invent new sounds but still had meanings we wanted to express. To get around this limitation, we can take a different approach to making words: simply recruit a relatively small number of sounds, and combine them in new and interesting ways. For example, if we take just ten different sounds to create words made up of five sounds each without repeating any of the sounds within a word, we can end up with a collection of more than 30,000 words. This nifty trick illustrates Hockett's notion of *duality of patterning*, in which a small number of units that don't convey meanings on their own can be used to create a very large number of meaningful symbols. In spoken language, we can take a few meaningless units like sounds (notice, for example, that there are no specific meanings associated with the sounds of the letters *p*, *a*, and *t*) and combine them into a number of meaningful words. Needless to say, it's a sensible approach if you're trying to build a beefy vocabulary.

But the combinatorial tricks don't end there. There may be times when we'd like to communicate something more complex than just the concepts of leopard or broccoli. We may want to convey, for instance, that Johnny kicked Freddy in the shin really hard, or that Simon promised to bake Jennifer a cake for her birthday. What are the options open to us? Well, we could invent a different word for every complex idea, so a sequence of sounds like *beflo* would communicate the first of these complex ideas, and another—say, *gromi*—would communicate the second. But that means that we'd need a truly enormous vocabulary, essentially containing a separate word for each idea we might ever want to communicate. Learning such a vocabulary would be difficult or impossible—at some point in the learning process, everyone using the language would need to have the opportunity to figure out what *beflo* and *gromi* meant. This means that as a language learner, you'd have to find yourself in situations (probably more than once) in which it was clear that the speaker wanted to communicate the specific complex idea that *gromi* was supposed to encode. If such a situation happened to *never* arise, you'd be out of luck as far as learning that particular word.

A more efficient solution would be to combine *meaningful* elements (such as separate words) to make other, larger meaningful elements. Even better would be to combine them in such a way that the meaning of the complex idea could be easily understood from the way in which the words are combined—so that rather than simply tossing together words like *Jennifer, birthday, promised, bake, Simon,* and *cake,* and leaving the hearer to figure out how they relate to each other, it would be good to have some structured way of assembling sentences out of their component parts that would make their meanings clear from their structure. This added element of predictability of meaning requires a **syntax**—or a set of "rules"—about how to combine meaningful units together in systematic ways so that their meanings can be transparent. (For example, once we have a syntax in place, we can easily differentiate between the meanings of *Simon promised Jennifer to bake a birthday cake* and *Jennifer promised Simon to bake a birthday cake.*) Once we've added this second level into our communication system, not only have we removed the need to learn and memorize separate words for complex ideas, but we've also introduced the possibility of combining the existing elements of our language to talk about ideas that have never before been expressed by anyone.

Structured patterns

It should be obvious that the possibility of combining elements in these two ways gives language an enormous amount of expressive power. But it also has

syntax In a given language, the set of "rules" that specify how linguistic elements are put together so that their meaning can be clearly understood.

some interesting consequences. Now, anyone learning a language has to be able to learn its underlying structural patterns. And since, for the most part, human children don't seem to learn their native language by having their parents or teachers explicitly teach them the rules of language (the way, for example, they learn the rules of arithmetic), they have to somehow intuit the structures on their own, simply by hearing many examples of different sentences. You might think of language learning as being a lot like the process of reverse engineering a computer program: Suppose you wanted to replicate some software, but you didn't have access to the code. You could try to deduce what the underlying code looked like by analyzing how the program behaved under different conditions. Needless to say, the more complicated the program, the more time you'd need to spend testing what it did.

In case you're tempted to think that language is a fairly simple program, I invite you to spend a few hours trying to characterize the structure of your own native tongue (and see Box 2.3). The syntactic structures of human languages involve a lot more than just basic word orders. Once you start looking up close, the rules of language require some extremely subtle and detailed knowledge. For example, how come you can say:

Who did the hired assassin kill the mayor for?

meaning "Who wanted the mayor dead?" But you can't say:

Who did the hired assassin kill the mayor and?

intending to mean "Aside from the mayor, who else did the hired assassin kill?" Or, consider these two sentences:

Naheed is eager to please.
Naheed is easy to please.

These two sentences look almost identical, so why does the first involve Naheed pleasing someone else, while the second involves someone else pleasing Naheed? Or, how is it that sometimes the same sentence can have two very different meanings? As in:

Smoking is more dangerous for women than men.

meaning either that smoking is more dangerous for women than it is for men, or that smoking is even more hazardous than men are for women.

It's not only in the area of syntax that kids have to acquire specific knowledge about how units can be put together. This is the case for sound combinations as well. Languages don't allow sounds to be combined in just any sequence whatsoever. There are constraints. For instance, take the sounds that we normally associate with the letters *r, p, m, s, t,* and *o.* If there were no restrictions on sequencing, these could be combined in ways such as *mprots, stromp, spormt, tromps, rpmsto, tormps, torpsm, ospmtr,* and many others. But not all of these "sound" equally good as words, and if you and your classmates were to rank them from best- to worst-sounding, the list would be far from random.

Here's another bit of knowledge that English speakers have somehow picked up: even though they're represented by the same letter, the last sound in *fats* is different from the last sound in *fads* (the latter actually sounds like the way we usually say the letter *z*). This is part of a general pattern in English, a pattern that's clearly been internalized by its learners: if I were to ask any adult native speaker of English exactly how to pronounce the newly invented word *gebs,* it's almost certain that I'd get the "z" sound rather than the "s" sound.

BOX 2.3
The recursive power of syntax

Among the most impressive linguistic accomplishments of human-reared apes is their ability to combine signs in new ways (often producing them in some regular order) and their ability to distinguish the meanings of sentences based solely on how the words are combined. We might conclude from this that apes "get" the idea of syntactic structure, and that the only difference between apes and humans when it comes to combinatorial structure is that we can produce and understand longer, more complicated sentences than they can. But we'd be leaping to unwarranted conclusions. In Method 2.1, you were cautioned against the temptation to overinterpret the data, and to consider the possibility that two species behaving in similar ways at times doesn't mean that the behaviors are driven by the same underlying cognitive apparatus. This applies to evaluating apes' syntactic abilities just as well as it does to their symbolic powers.

Let's suppose that a chimp regularly produces neatly ordered sequences of signs that combine action signs with person/chimp signs. For example:

Washoe tickle. Lana eat. Sue give. Dog growl.

These look a lot like perfectly well-behaved (if minimalist) sentences of English that even a linguistics professor might sometimes utter:

John stinks. Bianca jogs. Students complain.

But this doesn't mean that the system of rules that has generated these sentences looks the same for the chimp and the linguistics professor. All of these sentences *could* be generated by a rule that says:

Symbol for person/chimp/animal doing the action + action

But let's suppose that the professor, in addition to making simple sentences like the ones above, is also capable of producing sentences like these:

John's brother stinks.

Bianca and Leila jog.

Students who are too lazy to do the readings complain.

In fact, our professor can probably even say things like:

John's brother's dog stinks.

Bianca and the girl who lives next door to John's brother jog.

Students who are too lazy to do the readings but want to get an A anyway complain.

Clearly, these sentences can't be produced by the simple rule that is perfectly sufficient to capture the chimp's combinations. It's not just a matter of stringing more words together, or even of having a larger number of rules that allow words to be strung together in a greater variety of ways. It's that the rules have to be different *in kind* from the one I proposed as sufficient for describing the chimp's combinations. For the professor, the linguistic units that get combined with the action words can't just be symbols corresponding to people/chimps/animals, etc. They have to be much more abstract objects—like, whatever it is that you might call the grouping *Bianca and the girl who lives next door to John's brother* or *students who are too lazy to do the readings but want to get an A anyway.* These abstract objects have to in turn be composed of other abstract objects such as *the girl who lives next door to John's brother*—which, by the way, also has to be an abstract object, composed of combinations such as *John's brother.* And so on. Rules that operate at this level of abstraction are called **recursive** because they allow linguistic objects of the same kind (for example, *John's brother*) to be nested within other linguistic objects of the same kind (*the girl who lives next door to John's brother*)—notice, for example, that either of these phrases by itself can be combined with an action word, which is a hint that they are the same kind of phrase.

Rules that work to combine not just words, but higher-order groupings of words, take the language to a whole new level. Whereas word-combining rules allow language users to put together existing words in ways they haven't seen before, rules that allow recursion literally allow users to produce an *infinite* array of meanings because small, abstract units can always be combined with others to make larger abstract units, which can always be combined with others to make larger abstract units, which can always be combined... you get the picture.

Although apes can manage a small number of word-combining rules, recursive rules appear to be beyond their cognitive reach, leading a number of researchers (e.g., Hauser, Chomsky & Fitch, 2002) to suggest that the property of recursion is part of a uniquely human set of skills for language.

recursion "Nesting" of related clauses or other linguistic units within each other.

So there's structure inherent at the level of sound as well as syntax, and all of this has to somehow be learned by new speakers of a language. In many cases, it's hard to imagine exactly *how* a child might learn it without being taught—and with such efficiency to boot. For many scholars of the nativist persuasion, a child's almost miraculous acquisition of language is one of the reasons to suspect that the whole learning process must be guided by some innate knowledge.

One of the leading nativists, Noam Chomsky (1986), has suggested that the problem of learning language structure is similar to a demonstration found in one of Plato's classic dialogues. (The ancient Greek philosopher Plato used "dialogues" written as scenarios in which his teacher Socrates verbally dueled with others as a way of expounding on various philosophical ideas.) In the relevant dialogue, Socrates is arguing that knowledge can't be taught, but only *recollected*, reflecting his belief that a person's soul has existed prior to the current lifetime and arrives in the current life with all of its preexisting knowledge. To demonstrate this, Socrates asks an uneducated slave boy a series of questions that reveals the boy's knowledge of the Pythagorean theorem, despite the fact that the boy could not possibly have been taught it.

Chomsky applied the term "Plato's problem" to any situation in which there's an apparent gap between experience and knowledge, and suggested that language was such a case. Children seem to know many things about language that they've never been taught—for instance, while a parent might utter a sentence like "What did you hit Billy for?" she's unlikely to continue by pointing out, "Oh, by the way, notice that you can't say *Who did you hit Billy and?*" Yet have you ever heard a child make a mistake like this?

Moreover, Chomsky argued that children have an uncanny ability to home in on exactly the right generalizations and patterns about their language, correctly chosen from among the vast array of logical possibilities. In fact, he's argued that children arrive at the right structures even in situations where it's extremely unlikely that they've even *heard* enough linguistic input to be able to choose from among the various possible ways to structure that input. In reverse-engineering terms, they seem to know something about the underlying language program that they couldn't have had an opportunity to test. It's as if they are ruling out some types of structures as impossible right from the beginning of the learning process. Therefore, they must have some prior innate knowledge of linguistic structure.

Are we programmed for language structure?

If we do have some innate knowledge of linguistic structure, what does this knowledge look like? It's obvious that, unlike vervets with their alarm calls, or bees with their dances, humans aren't born programmed for specific languages, since all humans easily learn the language that's spoken around them, regardless of their genetic heritage. Instead of being born with a preconception of a specific human language, Chomsky has argued, humans are prepackaged with knowledge of the kinds of structures that make up human languages. As it turns out, when you look at all the ways in which languages could possibly combine elements, there are some kinds of combinations that don't ever seem to occur. Some patterns are more inherently "natural" than others. The argument is that, though different languages vary quite a bit, the shape of any given human language is constrained by certain universal principles or tendencies. So, what the child is born with is not a specific grammar that corresponds to any one particular language, but rather a **universal grammar**, which specifies the bounds of human language in general. This universal grammar manifests itself as a predisposition to learn certain kinds of structure and not others.

universal grammar An innately understood system of combining linguistic units that constrains the structural patterns of all human languages.

If the idea that we could be genetically programmed to learn certain kinds of language patterns more easily than others strikes you as weird, it might help to consider some analogous examples from the animal kingdom. James Gould and Peter Marler (1987) have pointed out that there's plenty of evidence from a broad variety of species where animals show interesting learning biases. For example, rats apparently have trouble associating visual and auditory cues with foods that make them sick, even though they easily link smell-related cues with bad food. They also have no trouble learning to press a bar to get food but don't easily learn to press a bar to avoid an electric shock; they can learn to jump to avoid a shock but can't seem to get it through their heads to jump to get food. Pigeons also show evidence of learning biases: they easily learn to associate sounds but not color with danger, whereas the reverse is true for food, in which case they'll ignore sound but pay attention to color. So, among other animals, there seems to be evidence that not all information is equal for all purposes, and that creatures sometimes harbor useful prejudices, favoring certain kinds of information over others (for instance, for rats, which are nocturnal animals, smell is more useful as a cue about food than color, so it would be adaptive to favor scent cues over color).

Proposing a universal grammar as a general, overarching idea is one thing, but making systematic progress in listing what's in it is quite another. There's no general agreement among language scientists about what an innate set of biases for structure might look like. Perhaps this is not really surprising, because to make a convincing case that a specific piece of knowledge is innate, that it's unique to humans, and furthermore, that it evolved as an adaptation *for language*, quite a few empirical hurdles would need to be jumped. Over the last few decades, claims about an innate universal grammar have met with resistance on several fronts.

First of all, many researchers have argued that nativists have underestimated the amount and quality of the linguistic input that kids are exposed to and, especially, that they've lowballed children's ability to learn about structure on the basis of that input. As you'll see in upcoming chapters, extremely young children are able to grasp quite a lot of information about structure by relying on very robust learning machinery. This reduces the need to propose some preexisting knowledge or learning biases.

Second, some of the knowledge that at first seemed to be very language specific has been found to have a basis in more general perception or cognition, applying to non-linguistic information as well.

Third, some of the knowledge that was thought to be language-specific has been found to be available to other animals, not just humans. This and the more general applicability of the knowledge make it less likely that the knowledge has become hardwired specifically because it is needed for language.

Fourth, earlier claims about universal patterns have been tested against more data from a wider set of human languages, and some researchers now argue that human languages are not as similar to each other as may have been believed. In many cases, apparent universals still show up as very strong tendencies, but the existence of even one or a couple of outliers—languages that seem to be learned just fine by children who are confronted with them—raises questions about how hardwired such "universals" can be. In light of these exceptions, it becomes harder to make the case that language similarities arise from a genetically constrained universal grammar. Maybe they come instead from very strong constraints on how human cognition works, constraints that tend to mold language in particular ways, but that can be overridden.

Finally, researchers have become more and more sophisticated at explaining how certain common patterns across languages might arise from the fact
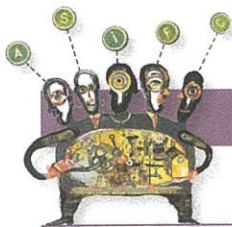
LANGUAGE AT LARGE 2.1

Engineering the perfect language

In all likelihood, the earliest users of languages never sat down to deliberately invent a system of communication in the way that humans invented Morse code or even the system of writing that we use to visually express our spoken languages. More likely, people intuitively fumbled around for the most natural-seeming way to express themselves, and language as we now know it was the eventual result. And the result *is* impressive. The combinatorial properties of human languages make them enormously powerful as communicative systems. By combining sounds into words and then words into sentences, we can create tens of thousands of meaningful symbols that can be combined in potentially infinite ways, all within the bounds of human ability to learn, use, and understand. An elegant solution indeed. But can it be improved upon?

It's interesting that when modern humans have turned their deliberate attention to language, they've often concluded that naturally occurring language is messy and poorly constructed. Many have pointed out its maddening irregularities and lapses of logic. For example, why in the world do we have the nicely behaved singular/plural forms *dog/dogs, book/books, lamp/lamps,* and *toe/toes* on the one hand—but then have *foot/feet, child/children, bacterium/bacteria, fungus/fungi,* and *sheep/sheep*? Why is language shot through with ambiguity and unpredictability, allowing us to talk about noses running and feet smelling? And why do we use the same word (*head*) for such different concepts in phrases like *head of hair, head of the class, head of the nail, head table, head him off at the pass*?

In a fascinating survey of invented languages throughout history, Arika Okrent (2010) described the various ways in which humans have sought to improve on the unruly languages they were made to learn. Many of these languages, such as Esperanto, were designed with the intention of creating tidier, more predictable, and less ambiguous systems. But unlike Esperanto, which was based heavily on European languages, some invented languages reject even the most basic properties of natural languages in their quest for linguistic perfection.

For example, in the 1600s, John Wilkins, an English philosopher and ambitious scholar, famously proposed a universal language because he was displeased with the fact that words arbitrarily stand in for concepts. In a more enlightened language, he felt, the words themselves should *illuminate* the meanings of the concepts. He set about creating an elaborate categorization of thousands of concepts, taking large categories such as "beasts" and subdividing them down into smaller categories so that each concept fit into an enormous hierarchically organized system. He assigned specific sounds to the various categories and subcategories, which were then systematically combined to form words. The end result is that the sounds of the words themselves don't just arbitrarily pick out a concept; instead, they provide very specific information about exactly where in the hierarchical structure the concept happens to fall. For example, the sounds for the concept of dog were transcribed by Wilkins as *Zita,* where *Zi* corresponds to the category of "beasts," *t* corresponds to the "oblong-headed" subcategory, and *a* corresponds to a sub-subcategory meaning "bigger kind."

Wilkins's project was a sincere effort to create a new universal language that would communicate meaning with admirable transparency, and he held high hopes that it might eventually be used for the international dissemination of scientific results. And many very educated people praised his system as a gorgeous piece of linguistic engineering. But notice that the Wilkins creation dispenses with Hockett's feature of *duality of patterning,* which requires that meaningful units (words) are formed by combining together a number of inherently meaningless units (sounds). Wilkins used intrinsically mean*ingful* sounds as the building blocks for his words, seeing this as an enormously desirable improvement.

But the fact that languages around the world *don't* look like this raises some interesting questions: Why not? And do languages that are based on duality of patterning somehow fit together better with human brains than languages that don't, no matter how logically the latter might be constructed? As far as I know, nobody ever tried to teach Wilkins's language to children as their native tongue, so we have no way of knowing whether it was learnable by young human minds. But surely, to anyone tempted to build the ideal linguistic system, learnability would have to be a serious design consideration.

that all languages are trying to solve certain communicative problems. We can come back to our much simpler analogy of the seeming universality of arrows. Arrows, presumably invented independently by a great many human groups, tend to have certain striking similarities—they have a sharp point at the front end and something to stabilize the back end; they tend to be similar lengths, etc. But these properties simply reflect the optimal solutions for the problem at hand. Language is far more complex than arrows, and it's hard to see *intuitively* how the specific shape of languages might have arisen as a result of the nature of the communicative problems that they solve—namely, how to express a great many ideas in ways that don't overly tax the human cognitive system. But an increasing amount of thinking and hypothesis testing is being done to develop ideas on this front.

2.4 The Evolution of Speech

The ability to speak: Humans versus the other primates

In the previous sections, we explored two separate skills that contribute to human language: (1) the ability to use and understand intentional symbols to communicate meanings, perhaps made possible by complex social coordination skills; and (2) the ability to combine linguistic units to express a great variety of complex meanings. In this section, we consider a third attribute: a finely tuned delivery system through which the linguistic signal is transmitted.

To many, it seems intuitively obvious that speech is central to human language. Hockett believed human language to be inherently dependent on the vocal-auditory tract, and listed this as the very first of his universal design features. And, just as humans seem to differ markedly from the great apes when it comes to symbols and structure, we also seem to be unique among primates in controlling the capacity for speech—or, more generally, for making and controlling a large variety of subtly distinct vocal noises. In an early and revealing experiment, Keith and Cathy Hayes (1951) raised a young female chimpanzee named Viki in their home, socializing her as they would a young child. Despite heroic efforts to get her to speak, Viki was eventually able to utter only four indistinct words: *mama, papa, up,* and *cup.* To understand why humans can easily make a range of speechlike sounds while great apes can't, it makes sense to start with an overview of how these sounds are made.

Most human speech sounds are produced by pushing air out of our lungs and through the vocal folds in our larynx. The vocal folds are commonly called the "vocal *cords,*" but this is a misnomer. Vocal sounds are definitely not made by "plucking" cord-like tissue to make it vibrate, but by passing air through the vocal folds, which act like flaps and vibrate as the air is pushed up. (The concept is a bit like that of making vibrating noises through the mouth of a balloon where air is let out of it.) The vibrations of the vocal folds create vocal sound—you can do this even without opening your mouth, when you making a humming sound. But to make different speech sounds, you need to control the shape of your mouth, lips, and tongue as the air passes through the vocal tract. To see this, try resting a lollipop on your tongue while uttering the vowels in the words *bad, bed,* and *bead*—the lollipop stick moves progressively higher with each vowel, reflecting how high in your mouth the tongue is. In addition to tongue height, you can also change the shape of a vowel by varying how much you round your lips (for instance, try saying *bead,* but round your lips like you do when you make the sound "w"), or by varying whether the tongue is extended forward in the mouth or pulled back. To make the full range of consonants and