

# 1

## *In the beginning*

Few topics about which scholars have puzzled can be quite so intriguing and so tantalizing, but at the same time so frustrating, as the evolution of the human capacity for language. Nothing so decisively sets us apart from our primate cousins as our constant chatter. It is no exaggeration to credit language for the very humanity that distinguishes us from the beasts from which we sprang. If we are even a tiny bit curious about our own origins, we have to be curious about the origins of language.

This is why it so frustrating to have no direct evidence for the language of our early ancestors. Fading as soon as it is uttered, spoken language leaves no trace. A few of our remote forebears left their bones in places where we could find them, and as more and more of these bones have been moved to museums, we have gained a clearer understanding of the several million years of evolution during which our bodies diverged from those of chimpanzees. The tools that early humans knocked from stones have survived in their thousands, and they tell us a good deal about early technology. But it was only after writing was invented, a mere five or six thousand years ago, that earlier languages could leave any trace. By then, the human capacity for spoken language had already had a very long history. Not even the earliest writing can tell us anything about the far more ancient periods when people first began to talk.

The lack of direct evidence for such a crucial part of our heritage has left the topic open to speculation, some of it reasonable, some that might be called “imaginative,” and some downright crazy, and this has

*In the beginning*

brought the subject a certain disrepute. Every student of linguistics is told about the famous prohibition of the Linguistic Society of Paris that, in 1866, banned the topic of language origins as a subject unfit for the Society's meetings. The ban is often cited as a sorry example of intellectual censorship, but anyone who has read widely in the literature on language origins cannot escape a sneaking sympathy with the Paris linguists. Reams of nonsense have been written on the subject.

Nevertheless, we cannot forever taboo a topic of such great interest, and even if we have no direct evidence, enough indirect evidence has accumulated since the prohibition of Paris to invite us to think carefully about what the early forms of language might have been like. In the last two decades, scholars of apparently sound mind from such serious disciplines as paleontology, primatology, cognitive science, archeology, and linguistics have once more turned their curiosity to the origins of language.

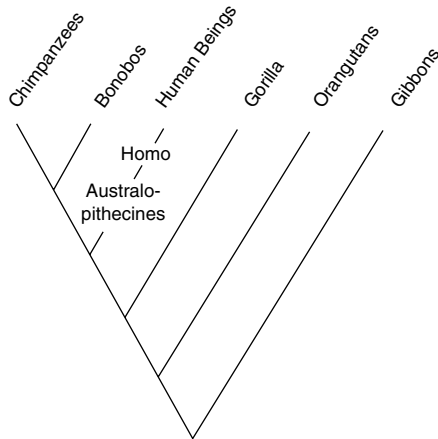
Any attempt to figure out how language evolved does have two reasonably solid anchor points. To get an idea about where it all began, we can observe the behavior of our closest primate cousins, the chimpanzees and bonobos. The latter, once rather misleadingly called "pygmy chimpanzees", are now recognized as a separate species. For the ending point, we can look at our own languages. We want to know how animals with something similar to the capacities of chimps and bonobos could have evolved into animals that could talk. Right away, we need to qualify. That "something similar" is important, for we did not evolve from chimps or bonobos, but only from "something similar" to them. These two closely related, but strangely different, species split from each other about three million years ago. Their common ancestor, in turn, split from the human line two or three million years before that. Since the time of that earlier separation, chimps and bonobos have had just as long to evolve as we have, so we certainly did not evolve from the apes we know today, and the differences between chimpanzees and bonobos show us that they have also evolved. Nevertheless, in the time since we separated, our ape cousins have changed their living circumstances much less than we have, and they have probably changed their bodies and behavior much less as well. Gorillas are considerably further from us genealogically than chimpanzees and bonobos. Orangutans are even more distant,

*In the beginning*

and they probably split off from the rest of us at least thirteen million years ago, perhaps more. (See Illustration 1.) In spite of this great phylogenetic distance, gorillas and orangs share more of their behavior with chimps and bonobos than they do with us.

All the great apes have their own special characteristics, of course, the legacy of their own millions of years of separate evolution. All of us have evolved, in our diverging directions, away from our distant common ancestor, but where the apes show common features in their anatomy or behavior, it is a good bet that their, and our, common ancestor had that trait too. Since all the apes are so much hairier than we are, for example, we can be reasonably confident that it is we who have lost our hair rather than the apes who have acquired theirs. Since we can talk and they cannot, we can be equally confident that it is we who have acquired language, not they who have lost it. The apes, especially the chimpanzees and bonobos, give us the best idea we have about the behavior of our common ancestor. We can use the behavior of these apes, cautiously, as a plausible starting point.

Our other anchor is modern human language. We can listen to people talk and try to figure out how language works. Even here we know less than we would like. We have only a primitive idea of how the brain processes language, for example, and we know much more about



**Illustration 1.** Phylogenetic Relationships among Humans and Apes

*In the beginning*

how the vocal tract produces the sounds of speech than about how our ears and brains manage to interpret them. Still, we do know a good deal about the sounds, words, and grammar of our languages, so the question that has to be asked by anyone interested in the origin of language is “How did we get from an ordinary primate that could not talk to the strange human primate that can’t shut up?”

Between our two anchors we have stones and bones. The fossil record of human bones has grown wonderfully richer in the past half century. We now have a reasonable understanding of how our bodies evolved from an ape that was much better than we are at clambering around trees, but not as good as us at walking on two legs. We also know more about how, some millions of years after leaving the trees, the serious expansion of the brain got underway. We can make some plausible inferences about what these bodily changes implied for behavior. We can also inspect a couple of million years of glacially slow development of stone tool technology and then the sudden speeding up of technological change in the last hundred thousand years. From the evidence of stones and bones, we can infer less about communication than about technology and subsistence, but the technology gives us some idea of the cultural context within which language grew.

The central argument of this book is that comprehension, rather than production, was the driving force for the evolution of the human ability to use language. To put comprehension first bumps up against a widespread, but barely recognized, bias that usually consigns comprehension to second place. We always act as if speaking is what really matters. We are more likely to ask a friend “Can you speak French?” than “Can you understand French?” Statistics tell us how many “speakers” each language has. They never count up the “understanders.” We have no less than four common words that refer specifically to the production of spoken language: “speak,” “say,” “talk,” and “tell,” but not a single word for what happens when language reaches our ears. We make do with “listen,” “hear,” “receive,” “understand,” or “comprehend” but none of these is specific to language. We can hear or listen to music or to the passing traffic as well to language.

*In the beginning*

“Understand” and “comprehend” are no better since, unlike “speak,” “say,” and “talk,” they are used for more than language. We can “understand” something as nonlinguistic as the workings of a mouse-trap. Our language makes it seem that when we “speak” or “talk” we do something special that can only apply to language. When we “understand” or “listen” we seem to use ordinary skills that serve many other purposes than just language.

Grammars that describe the world’s languages are packed with rules that explain how words are built up from prefixes, bases, and suffixes. Other rules show how words are joined to form sentences. You must search out very specialized literature to find suggestions about how a sentence might be decomposed into its words, or how the words can be taken apart into their smaller bits. Speaking, admittedly, is much easier to study than comprehension. We just listen to what people say, or even peer into our own minds to find out what we can say ourselves, and then try to figure out how in the world we do it. That is how linguists spend most of their time. It is much more difficult to know how, or even whether, people understand.

Nevertheless, speaking is only one half of the communicative process. Language needs a listener as much as it needs a speaker, and whenever we pay close attention to understanding, we find that, everyone—children, adults, and even animals—can understand more than they can say. Comprehension always surpasses the ability to produce. Sometimes, we can even interpret another’s actions when he would much rather we understood nothing at all. Consider some examples.

*Children.* Children learn their first language with what seems, to adoring parents, like magical ease. Parents eagerly follow the steps by which their children learn to use words, first alone and then in increasingly complex combinations. Linguists can scribble the child’s words and sentences on paper or record them on tape, and then use the records of paper and tape to build up a picture of their step-by-step progress toward a mature ability to speak. It is much harder to know just how much children understand, but parents quickly discover that those little ears can pick up far more than the child could possibly say. Astonished by their baby’s precocious understanding, parents quickly learn to guard their tongues.

*In the beginning*

Linguists have not always shared the parental confidence in their children's ability to understand. Linguistic skepticism derives, in part, from a vaguely behaviorist bias that makes the "behavior" of speaking seem more important than mere "passive" comprehension, but they have a better reason than this for doubting the reports of doting parents. Comprehension is very difficult to study. How do we know what a child understands? How much does she grasp from the language and how much from the wider context? Hold out a cookie to a child and ask "Do you want a cookie?" When she eagerly responds with gurgles and outstretched hands, it would be foolish to leap to the conclusion that she understands the inquiring words. No one can doubt that she understands the situation.

Distinguishing what a child learns by means of the words from what she picks up with the help of the context is so difficult that linguists, to say nothing of hard-nosed experimental psychologists, are always tempted to turn to the active part of communication, to the "behavior" that can be listened to and recorded. The linguist can even point out that children, now and then, use a word that they do not seem able to understand. Sometimes they even produce a grammatically impeccable construction where its meaning seems out of place. If they can learn words or bits of grammar before they grasp what they mean, it can seem as if production can actually come before comprehension. That, however, is a misinterpretation of what is going on.

To master a word, a child needs to know two things, both its meaning and its pronunciation. Children occasionally learn to pronounce a word without having learned its meaning. This can give the illusion that production is possible without comprehension, but it is really pronunciation that has been learned without meaning. A child who can pronounce a word has certainly heard that pronunciation often enough to recognize it even if he has not yet learned what the word means. By the time a child uses a word with its correct meaning, he has certainly learned to understand its meaning. Pronunciation can be learned before meaning, but children cannot use a word with its correct meaning before being able to recognize the meaning when they hear the word. What goes for words also goes for grammar. Children may use a grammatical form in an inappropriate place because they have not yet learned to understand its meaning. They

*In the beginning*

need to be able to understand its meaning before they can use it in the correct context.

Parents never doubt that their children understand far more than they can say, and I place my trust in the judgement of parents rather than in the judgement of skeptical linguists. One example is especially vivid to me. My own grandson, then called “Jamie,” was very late to talk. At the age of two years and two months, his total productive vocabulary consisted of exactly three words, only one of which was at all frequent: a loud insistent *da-da-da-da* meaning “Give it to me,” “I want it.” At the same age, however, he was able to point to the right place, not only when asked to show his eye, nose, or mouth, but also when asked about his elbow, knee, or shoulder. He could point not only to a window or door, but also to the wall, ceiling, or floor. He understood the names of dozens of people. He could follow quite elaborate instructions, and his total receptive vocabulary was certainly in the hundreds, if not thousands. A month later he apparently decided that it was time to talk, and he very quickly advanced to full sentences.

Jamie’s delay in production was extreme, but not unique. All children so consistently understand more than they can say that we have to conclude that an essential part of language learning takes place silently as children absorb the sounds, the words, and the grammatical patterns of the language that swirls around them. Language starts to be absorbed even before a child is born, as the baby inside becomes familiar with the rhythm and intonation of the language that penetrates from the outside. This, at least, seems to be the only explanation for the ability of four-day-old infants to react differently to their mother’s native language (as spoken by someone other than their own mothers) than to other languages. Four-and-a-half-month-old babies react differently when they hear their own names than when they hear another name with the same number of syllables and the same stress pattern, so by the age of four and a half months they have already learned to distinguish at least one word. Halfway through their second year, when most children start to produce words, they have had more than a year and a half of intensive exposure to their language. Starting to speak is a magical moment, but it is not the day when language learning begins.

We ought to regard speaking as only the final step in the long process of acquisition. It is the point at which language that is already

*In the beginning*

understood is finally made active. We may not really know how children learn their language until we know how they learn to understand it. We might do better with second-language instruction if we would more often allow learners to be quiet, at first, while they listen to large amounts of the language.

*Adults.* Even as adults, we always understand more than we can say. English speakers from the opposite shores of the Atlantic and from the opposite sides of the equator can generally understand each other with no more than an occasional minor hitch, but few of us ever try to speak another dialect. We all understand words that we would not risk using ourselves, not only words from distant dialects, but the slang of other ethnic groups or other generations. We understand technical terms from specialties with which we are only partially familiar. We understand and admire rhetorical styles that we cannot duplicate. In New Guinea, people have a nice way of distinguishing receptive and productive skill. They say “I can hear that language but I cannot talk it,” acknowledging that it is possible to have a skilled ability to understand a language without the ability to speak.

*Chimpanzees and bonobos.* As people, over the years, have labored to teach language to apes, their efforts have most often been directed toward production. Comprehension has generally been an afterthought. One of the most ambitious early experiments was with a chimpanzee named Viki who was born in 1971 and raised by Keith and Catherine Hayes. Viki grew up in the Hayes home and in everything from bottles to diapers, the Hayes did their best to treat Viki like their own child. They subjected her to all sorts of experiments, but the experiment that caught the most attention was an intense effort to teach her to articulate a few words of English. She was coaxed to make word-like noises, but the best she ever managed was four breathy vocalizations that, with charity, could be taken as attempts to say “Mama,” “Papa,” “cut,” and “up.”

How much could she understand? Unfortunately, the study of chimpanzee comprehension is even more difficult than the study of comprehension by human beings. Like human children, chimps are skillful at drawing inferences from the context, so it is every bit as hard with apes, as with children, to know how much they grasp from the context and how much they are helped by the language. Responding

*In the beginning*

correctly to “Close the drawer,” when a drawer had just been opened and when a gesture accompanies the request, may demonstrate a firm grasp of social routines, but it is hardly proof of a high level of language comprehension. The Hayeses did report that Viki could understand a considerable amount of spoken English but they were so eager to teach her to articulate words that they failed to study her comprehension systematically. Their reports of comprehension are anecdotal, and Viki is always remembered for her failure to speak rather than for her success at understanding. Our behaviorist biases have prevented us from taking “mere” comprehension seriously enough.

Now, however, we can put aside our skepticism about comprehension, for in a report that should have attracted more attention than it has, Sue Savage-Rumbaugh and her colleagues have described how the famous bonobo named Kanzi dramatically confirmed the ability of at least one ape to comprehend a significant amount of spoken English. Kanzi was born in captivity at the Yerkes Primate Center in Georgia in October 1980. Between the ages of six months and two and a half years, he was cared for by his foster mother, Matata, while she was being trained to use lexigrams. These are arbitrary visual symbols that were attached to the keys of a keyboard and that lit up when pressed. Both humans and apes could punch these keys and use them, like words, to communicate with each other. Several chimpanzees and bonobos had learned to recognize and use lexigrams, and while Matata was being trained, her baby Kanzi was allowed to wander about the laboratory. In this way he was regularly exposed both to the lexigrams and to natural spoken human language, but he received no deliberate instruction. He did take a certain delight in punching the buttons on the keyboard with their lexigram labels, but he seemed to punch them quite randomly, and he gave no sign that he had learned their meaning.

When Kanzi was about two and a half years old, Matata was separated from him so as to give her a chance to get pregnant, and Kanzi was left behind. He still visited the laboratory and the familiar people who worked there, and it was then, quite suddenly, that he showed more interest in the lexigrams. To the general astonishment of all who knew him, Kanzi, who had never been given any deliberate training, turned out to have learned a number of the lexigrams. Even more surprisingly,

*In the beginning*

he had also absorbed far more spoken English than anyone had imagined.

Unlike Viki, Kanzi was not home-raised. He was housed with the other bonobos but he spent many of his waking hours with humans. In spite of, or perhaps because of, having received much less deliberate instruction than most language-trained apes, he continued to learn even after his mother had gone. By the time he was eight years old, his apparent ability to understand spoken English called for a careful study, and to give a standard for comparison, a parallel study was made of a human child who appeared to understand approximately as much as Kanzi. When the study began, Kanzi was eight and Alia, the child, was between one and a half and two. Alia's mother was one of Kanzi's caretakers, and Alia had watched the apes through a window, but she had never interacted with them directly. Like Kanzi, she had some experience with lexigrams but, like all children, and like Kanzi, she was learning spoken language by living among talkative people rather than by deliberate instruction. Though differing markedly in age, the study confirmed that they had reached about the same level in their ability to understand spoken English.

Kanzi and Alia were tested in similar home-like laboratories. Each laboratory had a bathroom, TV, refrigerator, and places to sit, and each was provided with food and with other small objects to manipulate. The ape and the child were tested separately but in closely parallel ways. After a period of training that gave Kanzi a chance to become familiar with the procedures, Sue Savage-Rumbaugh gave him a series of instructions from behind a one-way mirror so that she could not inadvertently give visual clues to the meaning. After a similar period of training, Alia's mother, Jeannine Murphy, did the same for Alia. At first, both Kanzi and Alia were uncomfortable about responding to a disembodied voice from an unseen person, but with time and practice they both grew willing to cooperate. Other people were sometimes present so that Kanzi and Alia could interact with them. Sometimes these people covered their eyes so that they could not inadvertently give clues to the subjects, and this, too, was disconcerting to both Kanzi and Alia. Now and then, they both simply refused to cooperate.

Manipulable objects, such as food items and toy animals, could be taken from one place to another in the laboratories or in the yards

*In the beginning*

outside. Once Kanzi and Alia had become comfortable with the situation, the experimenter issued an instruction from behind the one-way mirror: “Put the ball on the pine needles.” “Give the lighter to Rose.” “Give Rose a hug.” “Get Rose with the snake.” “Knife the sweet potato.” “The surprise is hiding in the dishwasher.” “Take the [toy] snake outdoors.” “Go to the refrigerator and get a banana.” “Go get the carrot that’s in the microwave.” “Make the doggie bite the snake.”

Responses to these instructions ranged from prompt, immediate, and correct action to hesitation (giving a correct response only after the instruction had been repeated or clarified), to doing what was requested but doing something else in addition, doing only part of what was requested; doing the right thing with the wrong object or using the right object for the wrong thing; performing the parts of a task in the wrong order; refusing to cooperate; and finally, doing something that was completely wrong.

Over the course of the experiment Kanzi was given 415 trials. He performed 246 of these, 59%, promptly and correctly. He carried out another 61 instructions correctly but only after some hesitation or after the instruction had been repeated or clarified. 100 were carried out partially but not fully. On 4 trials Kanzi did not respond at all, and 4 responses were totally incorrect. Alia was given 407 blind trials. As Table 1 shows, she made a few more flat-out mistakes than Kanzi did, but overall, her results were not much different.

For both Kanzi and Alia, the results are overwhelmingly better than could have been expected by chance. Dozens of responses might have been made to any of the instructions. Many objects might have been

**Table 1.** Kanzi’s and Alia’s Comprehension

	Kanzi		Alia	
	N	%	N	%
Prompt and correct	246	59%	220	54%
Correct after hesitation or a repeated instruction	61	15%	47	11%
Partially correct	100	24%	108	27%
No response	4	1%	8	2%
Completely wrong	4	1%	24	6%

*In the beginning*

manipulated, many locations were available for the objects, and many different actions could have been performed with each of them. No one would hesitate to attribute Alia's success to her ability to understand English. Can there be any further reason to doubt that an eight-year-old bonobo can understand as much?

To be sure, Alia and Kanzi did not behave in identical ways. During the period of her testing, Alia was rapidly learning not only to understand, but to speak, and unlike Kanzi, she frequently commented verbally on what she was doing. By speaking, she confirmed her ability with English, and since Kanzi did not speak, we cannot look for the same kind of confirmation from him. Nevertheless, it would seem perverse to deny Kanzi's ability to understand a large number of English words representing several parts of speech, and even to use the order of words to infer something about the way they are related. Kanzi was able to go out of the room to retrieve an object from a particular place, even though, when he first heard the request, he could see neither the object nor the place. The circumstances in which natural animal calls are used are narrowly restricted. Alarm calls, for example, are produced only in a situation of danger. Because animals use their calls in such limited ways, many of us have been unwilling to accept them as comparable to words. Kanzi, like Alia, understood words in much more varied circumstances than animals understand alarm calls, and the kind of skepticism that has been directed toward animal calls is no more justified for Kanzi's understanding, than for Alia's.

We do not have to credit either Kanzi or Alia with the ability to understand every word in a sentence or every detail of its syntax. A request such as "Take the sparklers outdoors" might be obeyed correctly with little more than an understanding of "sparklers" and "outdoors." That, itself, is no mean achievement, but we do not have to assume that they understood "the," and even "take" might have been guessed from the context, although other guesses would have been plausible. If only "sparklers" and "outdoors" were understood, the sentence might mean "Look for the sparklers outdoors," or "Bring the sparklers from outdoors."

It is clear that Kanzi, like Alia, could do more than simply understand words. He was able to respond correctly to three types of

*In the beginning*

sentences where word order was significant. He responded correctly to 33 out of 42 examples (79 per cent) with forms like “Put the ball on the rock” or “Put the rock on your ball” where word order was crucial. Even his mistakes were not usually simple reversals, but various other kinds of errors. With instructions such as “Take the umbrella outdoors” or “Go outdoors and get the umbrella,” Kanzi responded correctly on 38 out of 46 trials. When he heard a sentence of the first kind, he would survey the objects around him, apparently searching, but with sentences of the second kind he would more often move directly to the place instructed without bothering to look at the things nearby. In such sentences, to be sure, the verbs are also different, so that word order is not the only clue to the meaning, but neither Kanzi nor Alia had trouble understanding a considerable number of different words and a variety of sentences types.

Kanzi’s receptive skills give better evidence of linguistic ability than has ever been demonstrated by any other nonhuman primate who has been trained to use symbols, whether these were spoken words, deaf signs, plastic chips, or buttons that needed pressing. Indeed, Kanzi’s comprehension demonstrates a degree of linguistic competence that linguists have often presumed to be exclusively human. No one need fear that a bonobo, or any other ape, is about to give serious competition to human children. Kanzi, after all, was eight years old when he was tested, while Alia was not yet two. Nevertheless, it is hard to deny that Kanzi had learned a good deal of English.

The pattern is consistent. It is not only humans who understand more than they can say. So does Kanzi. If Kanzi could learn to understand so much, it is reasonable to suppose that when our ancestors first separated from the line that led to the chimps and bonobos they would already have been able to understand a good deal. To make use of their abilities, however, they had to wait until there was something to be understood.

*Animal signals.* It is not only among language-using humans and apes that comprehension runs ahead of production. Comprehension also came first as animal signals became established under the pressures of natural selection. The gestures and vocalizations by which animals communicate probably began as purely instrumental acts—movements and vocalizations that are part of the ordinary business of

*In the beginning*

living: moving around, eating, scratching, yawning. Instrumental acts are performed to meet the needs of the animal, but with no intention to communicate. Even without intending to communicate, animals may profit by being able to interpret one another's movements and sounds, but only after some behavior has come to convey a meaning to another animal, can it develop into a specifically communicative signal. The classic example is the retracted lip of a dog's snarl.

At first, the lip must have been drawn back as a simple instrumental gesture. It would have been nothing more than one part of getting ready to bite. A dog that did not want to bite his own lip needed to get the lip out of the way of his teeth. A few million years could have passed before potential victims began to recognize a retracted lip as a sign of an imminent bite, but any victim that was clever enough to read the lip movement as a warning might have had a chance to flee and to avoid the bite. By escaping, an animal would improve his chance of staying alive long enough to reproduce, so natural selection would have insured the spread of skillful comprehension.

Comprehension had to be the first step, but once the instrumental act was understood, a new opportunity was opened to an aggressor. By retracting his lip as if to bite, the aggressor might frighten off his enemy even while avoiding the much riskier activity of really biting. Now it was the turn of natural selection to favor those individuals who were clever enough to pull back their lip in order to scare away their enemies without a fight. From that point on, production and comprehension of the signal could evolve together. To make the sign less ambiguous, aggressors might even develop exaggerated or stereotyped lip movements. The sign would then have evolved from a purely instrumental act into a stereotypic communicative signal.

By evolving into a communicative signal, the retracted lip became useful for both the aggressor and his potential victim. The victim might escape, and the aggressor might avoid a fight. All this happened, of course, under the slow but relentless pressure of natural selection. The end result was to build in a signal, but learning could have nudged the process along. At first, some victims may have been able to learn by experience that lip retraction is likely to be followed by pain. If they also learned that running away was the best way to avoid pain, they would boost their own chance for survival and give their genes a better

*In the beginning*

chance to reach the next generation. Natural selection would favor those who could learn most quickly and easily. Eventually, no more than a slight triggering experience would be needed to persuade a potential victim that bared teeth were a serious matter, and finally, animals might react to the gesture with flight even without needing any experience at all. Potential victims could then be said to have the fear of a curled lip firmly built in.

Like their victims, the earliest lip curlers also probably needed to learn by experience. They would learn that some animals flee at the sight of a retracted lip, and discover how useful the gesture could be even when they had no intention of biting. Again, natural selection would have favored those who could most easily learn how the gesture affected others, but after some thousands of generations the behavior became almost, or fully, automatic. Behavior is not neatly divided into some bits that are learned and other bits that are built in. Rather, some bits are learned more easily and some less easily. The history of the curled lip shows us how hopeless it is to ask whether heredity or environment is to be credited for some behavior. Both are essential.

The process by which natural selection builds in signals is known as “ritualization.” The ritualization of the lip twitch turned an instrumental act into a communicative signal, but ritualization could not even begin until the twitch was understood. Other animal signals began much as did the retracted lip. Only after meaning is discovered in instrumental gestures or vocalizations can they be ritualized into stereotypic signals.

I will be returning to comprehension repeatedly in the later chapters of this book, but any discussion of language evolution must also consider many other issues, and I want to clear the ground now by staking out positions on four of these issues: first, whether language evolved from animal calls or as a part of an evolving mind; second, whether technology or social relationships gained the most from language; third, the speed with which language developed, whether gradually or abruptly; and fourth, the relative significance of vocabulary and syntax in the development of language. All four of these issues have been debated

*In the beginning*

more vigorously by students of language origins than has comprehension, but on all of them, opinion has been divided.

*Animal calls and the mind.* The first kind of evidence to which people turn, when looking for hints about the antecedents of language is, more often than not, the communication of other animals, particularly the communication of our nearest animal kin, the primates. Like most other mammals, primates communicate with their voices as well as by movements and gestures. With their cries, whoops, and chatters they coordinate their activities, call for help, show their anger, make threats, and even warn one another of danger. Since most human languages are also produced with the voice, and since language is also used to coordinate activities, call for help, and show emotions, hardly anyone can resist searching among primate calls for the forerunners of human language. It seems only reasonable to ask how natural selection might have transformed a set of primate calls into the kind of language that humans speak.

Language, however, is organized in such utterly different ways than primate or mammalian calls and it conveys such utterly different kinds of meanings, that I find it impossible to imagine a realistic sequence by which natural or sexual selection could have converted a call system into a language. Human beings, moreover still have a fine set of primate calls that remains quite separate from language. Primate calls have much less in common with human language than with human screams, sighs, sobs, and laughter. Our own audible cries, howls, giggles, and snorts, along with our visible scowls, smiles, and stares, all belong securely to our primate heritage. They form the primate communication of the human primate. We produce our gestures and noises with the same parts of our bodies that other primates use for their signals, and we use them to convey the same kinds of messages. Primate communication and some aspects of human nonverbal communication resemble each other very closely, but human language is different from either of them. We will understand more about the origins of language by considering the ways in which language differs from the cries and gestures of human and nonhuman primates than by looking for ways in which they are alike.

Even if language owes little to primate calls, however, we still have plenty to learn from other primates. Apes, in particular, have

*In the beginning*

minds that resemble our own, sometimes more closely than we find comfortable. Chimps and bonobos show us the kind of mind where the first glimmerings of language must have appeared. In spite of ingenious and energetic efforts by field primatologists, little of what we have learned about wild chimps, bonobos, gorillas, or orangutans suggests that their natural communication is more like human language than is the communication of other, more distantly related mammals. Primate calls and communicative gestures may be a bit more complex than the calls and gestures used by other mammals, but they do not seem to be organized in a different way. Primate minds, on the other hand, share a great deal with ours. We will learn much more about the beginning of language by examining how primates use their minds than by learning how they use their voices. Language, after all, is not only a way to communicate. It has also become a tool that helps us to think clearly. As language has developed, the human mind has been transformed. Human beings also have another kind of communication that remains more like the communication of other animals. I will look on language as one product of the evolving human mind.

*Technology and social relations.* A second easy assumption about the origin of language is that it developed primarily as an instrument to assist with the business of making a living. Today, language is so essential for technical and practical affairs that we easily presume that, from its earliest days, its main use was to coordinate the hunt, agree on where to meet, tell others where to find ripe fruit, warn about a marauding lion, or instruct the young on how to make a hand-axe. I join others such as the evolutionary psychologists Nicholas Humphrey, Robin Dunbar, and Geoffrey Miller, who argue, instead, that we need language most urgently for dealing with one other. Language has always been used, first of all, for the fine-tuning of social relationships.

Language is more than simply a means of exchanging information. It is also a medium for art, humor, poetry, storytelling, and oratory. It gave our ancestors, as it gives us, a means to display themselves. Language may have helped our ancestors to cooperate, but at least as important, it gave them new ways to compete. With language we have acquired a powerful new tool with which we can try to outmaneuver

*In the beginning*

and outmanipulate one another. If you want to engage in social climbing, you had better hone your language skills. I want to understand the emergence of language as one aspect of the evolution of our minds, and as a means by which our minds can build increasingly complex social relationships. This should mean that the better we understand the evolution of our minds and our social relationships, the better we will understand the emergence of language.

*Sudden or gradual emergence.* The question of whether language emerged gradually or suddenly has generated more debate than it should have among students of language evolution. Some, particularly linguists, are so impressed by the uniqueness of language and by the interdependence of its parts, that they have argued that it must have arisen quite suddenly. Derek Bickerton once even insisted that the most important aspects of language, in particular its syntax, could only have come about as the result of a single crucial mutation. Others, particularly primatologists who have searched for continuities between ape and human communication, have argued that language must have developed slowly and incrementally. I side with the gradualists, not because I want to derive language from primate calls, but because I believe that this is the way evolution works. Wings distinguish birds from reptiles as sharply as language distinguishes human beings from other primates, and the parts of a wing are as interdependent as the parts of language, but nobody would argue that wings were achieved by a single mutation. Of course we now have fossils that show us some of the steps by which front legs evolved into wings, but even without the fossils we would be certain that there were, once upon a time, animals that were intermediate between walking reptiles and flying birds. We will never have fossil evidence for language, but it makes no more sense to imagine that language arose abruptly than to suppose that wings arose abruptly. There had to be intermediate stages.

Readers familiar with the debates about language evolution will have noticed that I take opposing positions on two issues that are often presumed to be linked. Many of those, especially linguists who have been most impressed by the unique features of language, have been dismissive of the relevance of primate calls, and it is they who have most often argued in favor of the abrupt beginnings of language. Many

*In the beginning*

of those, especially primatologists, who have searched most diligently for language-like features in primate communication, have argued for a slower, more gradual development of language. On the issue of the relevance of primate calls, I side with the linguists, but on the issue of gradualism I side with the primatologists. If wings could evolve through many intermediate steps, so could language, but this does not mean that language had to evolve from calls.

*Words and syntax.* Linguists, or at least some linguists, love syntax. Their devotion can seem odd to others who have less-than-fond memories of diagramming sentences or of the grammar that came with high-school French. For a certain kind of mind, however, syntax has an intricate beauty. It is extraordinarily complex, but close investigation reveals startling regularities. It is so utterly unlike anything used by other animals that it is easy to draw the conclusion that syntax is what makes language unique.

It is well to remember, however, just what syntax is used for. It is needed to serve the lexicon. We use syntax to arrange our words in efficient and unambiguous ways, and it would have no purpose at all without the words. For most ordinary speakers, if not for linguists, meaning is what language is all about, and while syntax certainly contributes to meaning, words do more. Because syntax needs words, while even single lonely words that have no syntax can easily convey meanings, words must have come first in the course of human evolution.

This is not to dismiss the evolution of the specific capacity for syntax as unimportant. Indeed, its very complexity gives syntactic evolution a special fascination and presents us with special puzzles. Far more scholarly ingenuity has gone into trying to imagine how syntax could have evolved than into puzzling about the evolution of the capacity to use words. We can imagine that vocabulary just needs a lot of cranial storage capacity but that syntax demands some very special cortical circuitry. Nevertheless, words had to come first and I will put a bit more weight on the lexicon than have some others who have shared my interest in language origins.

In the chapters that follow, these four themes will recur, interwoven with the insistence that comprehension always comes before production. Together, these suggest a plausible step-by-step picture of the

*In the beginning*

emergence of language, and of the central role that language played in the evolution of our species.

A puzzle has always hovered over the first appearance of language: If no one else was around with the skills to understand, what could the first speaker have hoped to accomplish with her first words? The puzzle dissolves as soon as we recognize that communication does not begin when someone makes a meaningful vocalization or gesture, but when someone interprets another's behavior as meaningful. The original behavior that was interpreted in a language-like way could not have been intended to communicate anything at all. A lonely producer who experimented with a new kind of signal would surely fail to communicate, but she would be unlikely ever to try. A lonely interpreter, on the other hand, might gain considerable advantage by being able to understand another's actions, even when no communication had been intended. At every stage of evolution, the selective pressures favoring skill at comprehension must have been more insistent than the selective pressures favoring skill at production. Producers often benefit by not giving themselves away. Only profit can result from successful interpretation.

The precocity of comprehension implies that at every point along the evolutionary path toward language, interpretation needed to be at least one step ahead of production. Only when behavior was being interpreted correctly, could an animal deliberately use that behavior as a communicative signal. More precisely: *The only innovations in signal production that can be successful, and so consolidated by natural selection, are those that conform to the pre-existing receptive competence of other individuals.* At every point, production would have been limited by and directed by the interpretive ability that already existed in the population. Only after others had grasped the meaning of some bit of instrumental behavior, would anyone be able to expand his productive repertoire with a new linguistic trick. This disposes of any mystery about the communicative usefulness of the first word-like signs. They would not have been produced with communicative intent. Their communicative value came from the skill of the receiver, not from the intent of the producer.

*In the beginning*

The question that we should ask, therefore, is not “Why did the first speaker try to communicate if no one was around who shared her talents?” The answer to this is easy: “She didn’t. It would have been useless.” A much better question is “Why would anyone engage in language-like behavior when she had no intention of communicating?” Here, a plausible answer is that the first acts that could be interpreted in a language-like way were instrumental ones. Everyone has to act instrumentally and others can benefit by being able to interpret their acts. Only as skill at interpretation grew, did the time finally come when actors could conventionalize the actions that others were already understanding, and turn these actions into deliberate communicative signals.

This implies that word-like signs began in the same way as animal signals such as a retracted lip, but from the beginning, proto-words differed from animal signals in one crucial respect. Most animal signals have been ritualized by the long process of natural selection. Early word-like signs, as much as modern words, needed to be learned by each individual within a single lifetime. Our ancestors must have been selected for their ability to give conventional meanings to visible or audible signals, but the conventionalizations of even the earliest words were never passed down by the genes. Rather, they had to be learned anew by the members of each generation. In Chapter 6, I will return to a more detailed consideration of the process by which conventionalized and language-like signals could have begun.

If comprehension must always run ahead of production, not only in the history of each animal sign and in the language abilities of our children, but also at the earliest stages of language, we ought to ask rather different questions than we have usually asked about language origins. We should ask what selective pressures could have driven our prehuman and early human ancestors toward ever greater skill at interpreting the instrumental acts of other individuals. We should ask how, at each later stage, people could have understood the gestures or vocalizations of their fellows in increasingly word-like, and then sentence-like, ways. The origins of comprehension are, after all, considerably less mysterious than the origins of production. Animals,

*In the beginning*

including human animals, have little to lose and a great deal to gain by learning as much as possible from the behavior of others: What is that fellow likely to do? What does she want? Why is she moving off in that direction? What does that grunt mean? The more that an animal can infer from the behavior of others, the more adroitly it can plan its own behavior.

A focus on improved comprehension gives us a different picture of the sequence by which new features enter language than does a focus on production. If, for example, we assume that our forebears used single words before joining them together into orderly sequences, too much attention to production might lead us to puzzle about why anyone would bother to join them. If, instead, we ask how comprehension evolved, we might wonder how understanders could start to make inferences from words that they heard in vague proximity, even when the producer had made no effort to arrange them in any sort of orderly way.

These are questions that arise naturally as soon as we recognize the crucial role of comprehension, and I will return repeatedly to such questions in chapters that follow. Before that, however, I need to establish the beginning and end points, first by describing the various forms of human communication and the place of language among the other forms, and second by showing just how different language is from animal communication. These will be the topics of the next two chapters. Once these topics have been addressed, it will be possible to focus on the steps by which a species without language could have evolved into a species whose members could talk.