

Talk to the Hand: New Insights into the Evolution of Language and Gesture

May 01, 2008

In his book *Me Talk Pretty One Day*, humorist David Sedaris chronicled his pain at trying to learn French, in France, at age 41. His commiseration with a fellow language student sounds like it could be a dialogue between, say, two australopithecines, dimly anticipating the communicative achievements of their hominid descendants:

“Sometimes me cry alone at night.”

“That be common for I, also, but be more strong, you. Much work and someday you talk pretty. People start love you soon. Maybe tomorrow, okay.” (Sedaris, 2000)

Philosophers have always esteemed language among our most defining attributes, and the storytellers of every culture have tried to explain how humans acquired the gift. In Judeo-Christian myth, God granted Man the right to name things as he pleased, and later confused the world’s tongues in retribution for human pride — leading to David Sedaris’s predicament. Darwin supposed that language’s origins were a more gradual and less deliberate outgrowth of animal communication: “Man not only uses inarticulate cries, gestures, and expressions, but has invented articulate language; if, indeed, the word *invented* can be applied to a process, completed by innumerable steps, half-consciously made” (Darwin, 1872/1998). Accounting for these innumerable steps has been a challenge in the evolutionary psychology of language. What led our ancestors to become articulate? How did we finally learn to, you know, talk pretty?

It is intuitive to look to the vocal calls of primates for clues, and some primatologists still see this form of communication as the likeliest precursor for human language abilities. Yet evidence is accumulating that the “inarticulate cries” of monkeys appear to be controlled by different brain systems than those governing human language ability (Rizzolatti & Arbib, 1998), and psychologists interested in human and ape communication are turning with new interest to the properties of gesture. The story of language’s “invention” may turn out to be more complicated than even Darwin could have imagined.

What is Language?

A language is a system that can express an infinite range of ideas using a finite set of sounds or word elements — a *discrete combinatorial system*, as APS Fellow and Charter Member Steven Pinker, Harvard University, calls it (Pinker, 1994). Simple sound elements like phonemes (for example, the sounds *ba*, *da*, and *pa*; some researchers even focus on smaller units called *articulatory primitives* — see Poeppel & Monahan, 2008), are combined into words standing for things or actions, which are combined into larger groupings like sentences that express ideas of varying levels of complexity — theoretically, infinite complexity (a language feature known as *recursion*).

The system that enables this infinite recombination from finite raw materials is *grammar*, and it is the

element most conspicuously absent from all forms of animal communication. Vervet monkeys, for instance, have what could be called a vocabulary, a handful of distinct warning calls that are tied to specific threats in their environment like leopards, snakes, and eagles (see Cheney & Seyfarth, 2005); but there is no vervet grammar — the monkeys cannot mix and match their calls or use them to express new ideas (but see Zuberbühler, 2005). And as complex as some bird and whale songs are, grammatical rules enabling sentence-like recombination of ideas don't appear to exist in such animals either.

The linguist Noam Chomsky argued that humans uniquely are born with a *universal grammar*, an underlying set of rules that serves as the basis for language acquisition. His classic example was the made-up sentence “Colorless green ideas sleep furiously.” It makes no sense — indeed, it consists of self-contradictory ideas — but the brain accepts it because it is grammatical; it obeys the rules of syntax. More recently, Pinker has upheld the Chomskyan view of the innateness of language ability in humans, calling language an “instinct” that humans are born with (Pinker, 1994).

According to Pinker, the instinct for language evolved as an adaptation for social coordination in our hunter-gatherer ancestors (Pinker, 1994), and its deep structure still bears evidence of the fundamental human priorities of manipulating the social and physical environment (Pinker, 2007). Somewhat controversially, Pinker also argues that language is a modular system that evolved independently from other human cognitive abilities — that it is its own unique tool in the toolbox that is the human brain. His view of the modularity of mental adaptations has been compared to that of University of California, Santa Barbara evolutionary psychologists (and APS Fellows) Leda Cosmides and John Tooby, who have likened the mind to a “Swiss Army knife” comprising numerous special-purpose adaptations for solving particular challenges.

Those who argue for a unique language-processing module in the brain make their case in opposition to connectionists, who emphasize that language arises from multiple distributed cognitive abilities and is inseparable from all the other intelligent feats humans can perform. Among those who are passionate about such subtleties, it is a hot debate. Neuroscientists are generally converging on a connectionist view of most cognitive abilities such as object recognition, categorization, and memory (see the April 2008 special issue of *Current Directions in Psychological Science*: “The Interface Between Neuroscience and Psychological Science”). It appears that different aspects of language also are handled by widely distributed, functionally interconnected brain areas. Speech perception and language comprehension, for example, are now known to involve a complex network of brain areas operating in parallel, including a “dorsal pathway” that maps auditory sound representations onto motor representations for producing speech sounds, and a “ventral pathway” that maps speech sound representations onto representations of word concepts (Poeppel & Monahan, 2008; see also Holt & Lotto, 2008).

Different systems also appear to handle semantics (meaning) and syntax (grammar). Recordings of event-related potentials — brain waves recorded with electrodes placed on the scalp — reveal that violations of semantics such as the sentence “He spread his warm bread with socks” causes something in the brain to balk, with a negative potential peaking at 400 milliseconds after the sense-violating word (in this case, the word “socks”; see Hagoort, 2008). However, an entirely different brain wave response betrays the brain's complaint at a violation of syntax. Like Chomsky's observation about colorless green ideas, something about the sentence “The boiled watering can smokes the telephone in the cat” is perfectly acceptable to brain's syntax enforcer even though it makes absolutely no sense. But a violation like “The boiled watering can *smoke* the telephone in the cat” causes a negative-amplitude spike at 600

milliseconds after the offending (grammatically incorrect) word “smoke” — evidence that syntax and sense-making are distinct cognitive functions (see Hagoort, 2008).

Particularity

Language has been called an instinct because it is so readily learned. Infants quickly begin to acquire language, without being actively taught: At 10 months of age, they know around 50 words (even though they do not say much), but by 30 months, they are already “social sophisticates,” speaking in complete sentences with a production vocabulary of 550 words (Golinkoff & Hirsh-Pasek, 2006). Yet while it is instinctively hungry to acquire language, the newborn brain is also completely unbiased to respond to the particular subset of possible sounds that constitute the spoken language of its parents — that is, when it comes to *phonetics*, it is a blank slate. That changes as the plastic brain quickly rewires (or prunes itself) to recognize only those sounds used in the language being spoken around it; older infants can only discriminate sounds from their own language (see Kraus & Banai, 2007), and adults learning a new language may have difficulty mastering its foreign sound distinctions (e.g., Japanese-speakers often have trouble distinguishing English “l” from “r”).

This is one of the most interesting facts about language: you can’t learn language without learning *a* language. Thus, while language requires an underlying mechanism (or mechanisms) common to everyone, and while language is used for the same purposes and in the same ways everywhere (i.e., it is a psychological universal), it is also a cultural system whose hallmark is particularity. There is no universal language any more than there is a “typical human.” Languages, like people, are unique.

This is not merely an accident of history or evolution. One of the defining features of language, setting it apart from mere communication, is the feature known as *arbitrariness*. The Swiss linguist Ferdinand de Saussure noted that *signifiers* (e.g., words) by and large bear no necessary or logical connection to *signifieds*, or the things they stand for. There is no more reason to designate something you put on your head a “hat” than there is to call it a “chapeau.” As such, the connection between the thing you put on your head and the word for it used in your community can only be a learned social convention. This is even true of onomatopoeia — words with a resemblance to natural sounds, the seeming exception to Saussure’s rule. A Russian speaker will not recognize the onomatopoeic “bang” as the sound a gun makes, for example; where she grew up, she would have learned this sound as “batz.”

The benefit of having to learn your lexicon instead of being born with it already hard-wired is that you are free to use words in novel and creative ways. It is possible to come up with other words for hats, or to lie about hats, or imagine a hat that doesn’t exist yet, or wax nostalgic about hats in the past. It would be hard to imagine such human behaviors as tool-making, art, humor, long-term planning, or consciousness of self without the ability to represent abstractions, objects, and states of mind by words and other symbols that can be manipulated independently of what they stand for. The roots not only of abstract thought but also of culture lie in this radical disconnect between words and things.

Missing Links

On January 21st of this year, Alaskan octogenarian Marie Smith Jones died at her Anchorage home, at age 89. As a result, language conservationists moved her native tongue, Eyak, from the lists of “Endangered” to “Extinct.” Smith had been the last living speaker of a language that, in prehistoric times, may have been spoken over much of Alaska’s southeastern coast. The minor flurry of news reports of her passing briefly helped publicize the issue of language diversity and its rapid worldwide

decline.

Despite the disappearance of languages like Eyak, there are still 6,000 languages spoken in the world today, give or take. They vary widely in the size of their lexicons, but in fundamental respects all these languages are pretty much alike. They are all fully modern and capable of expressing ideas of whatever complexity they are called upon to express. Linguists have observed the emergence of new languages, creoles, out of simplified pidgins that arise in trading communities and other situations when people who don't share a common language live with each other. But even creoles exhibit the complexity and syntactical capabilities of languages having long histories. No anthropologist has ever found, in some remote tribe, an evolutionary "missing link" between modern languages and the more rigid and stereotyped modes of communication used by animals.

Researchers attempting to explain how language could have evolved in humans inevitably return to the linguistic capabilities of other living primates for clues. The obvious social intelligence of apes, in particular, made them appealing candidates in some of the early experimental attempts to assess the language abilities of animals. Apes' vocal tracts cannot produce the sounds needed for spoken language, so in 1967, Beatrice and Allen Gardner (University of Nevada, Reno) tried raising a young chimp, Washoe, to communicate using American Sign Language (ASL), training her using operant conditioning techniques. Washoe's trainers reported that, by the time she died last year at age 42, she had learned around 250 signs and could even apply some of them in novel situations. Koko, a 37-year-old lowland gorilla, has been claimed by Stanford psychologist Francine Patterson to know over 1,000 ASL signs and to recognize over twice that many words of spoken English. Koko has gained a degree of fame for her sign-language abilities, being the subject of television documentaries and even taking part in an online "chat."

But many scientists have rejected the claims of the Gardners, Patterson, and other proponents of animal language, saying that researchers (and an eager public) have projected human-like mentality onto these animals in the absence of compelling evidence that they are doing much more than parroting their trainers or using linguistic signs in relatively rigid, nonlinguistic ways. The only ASL-fluent member of the Gardners' research team, for example, disputed Washoe's use of true ASL signs (see Pinker, 1994). These animals may display remarkable communication abilities, but communication — the ability to affect others' behavior — is not the same thing as language.

The most scientifically compelling case for rudimentary language abilities in apes comes from Kanzi, a bonobo who learned a system of communicating by pressing lexigrams (arbitrary symbols) on a keyboard. Unlike other ape language subjects, Kanzi was not raised by human parents, nor was his "language acquisition" a product of active training — he learned his first keyboard signs passively, as an infant, while his mother was being taught them by researcher Sue Savage-Rumbaugh at Georgia State University's Language Research Center, in the early 1980s. Kanzi, now 27, has shown remarkable abilities to understand spoken language, to link spoken words and things to corresponding lexigrams, and possibly to construct novel messages from combinations of signs. Savage-Rumbaugh claims that Kanzi is even able to understand the grammatical structure of some sentences (Savage-Rumbaugh, 1989). Whether the finite number of lexigrams available to Kanzi limits his ability to produce sentences as complex as the ones he understands (as Savage-Rumbaugh suggests), or whether there is some more basic linguistic threshold separating his abilities from full-blown language, remains an open question. But the case of Kanzi does appear to refute the notion that apes are only capable of parroting their

human companions.

Much has been gained from observing the way primates communicate with each other in their natural environments. Many African monkeys, as well as chimpanzees, have been found to have repertoires of acoustically different alarm calls for different threats. APS Fellow Robert Seyfarth and his University of Pennsylvania colleague and wife Dorothy L. Cheney have found that such calls are highly dependent on social context; vervet monkeys, for example, seldom give an alarm if they are alone, and they are more likely to call in the presence of their own kin than in the presence of unrelated individuals (Cheney & Seyfarth, 2005).

Do monkeys understand the meaning of calls in the same way that humans understand the meaning of words? Is a vervet “leopard” alarm a word for a type of jungle cat, a recommendation (“run!”), or simply a symptomatic expression of a particular flavor of anxiety?

Primate calls are not simply reflexive; a monkey can decide whether or not to make a call based on who else is around. Such “audience effects,” displayed by a number of species, are evidence for cognitive control and complexity in communication. But the evidence also suggests that despite the dependence on social context, monkeys lack theory-of-mind ability — the ability to conceptualize what other individuals may be thinking or how their knowledge may be changed by making (or not making) a vocalization. For Cheney and Seyfarth, this inability may be the key thing that distinguishes nonhuman primate communication from human language, and is probably at the root of their inability to generate new signals in creative ways or to utilize signals syntactically (Cheney & Seyfarth, 2005).

An interesting theme emerging in research on primate communication (as well as communication in other vocal animals such as parrots and dolphins) is the extreme asymmetry between vocal production and auditory comprehension. Animals are relatively inflexible and limited in the calls they can produce, yet they are often capable of much greater subtlety when it comes to grasping syntactical (i.e., causal) relationships, understanding the semantic meaning of calls, responding to the pragmatics (intentions and consequences) of calls, and even recognizing calls of other species. For example, Klaus Zuberbühler (University of St. Andrews) has found that Diana monkeys living among chimpanzees often made leopard alarm calls of their own when hearing chimpanzee leopard alarm screams, whereas Diana monkeys with less chimpanzee experience were more likely to hide silently (i.e., from the chimpanzees, who sometimes prey on the monkeys; Zuberbühler, 2005). Ape language experiments (and even everyday experience with pets) reflect the receptiveness and responsiveness to more sophisticated communication than animals are generally able to produce themselves. Zuberbühler suggests that the evolution of language in our species built on a basic competence in comprehension already existing in our primate ancestors.

How Necessary Was Speech?

The transition from hearing and understanding to actually talking required a revolution not only cognitively (e.g., theory-of-mind ability) but also in controlling the face and mouth. Humans uniquely are able to produce and combine a huge array of subtly distinct sounds (over 100 acoustically unique *phones* are listed in the International Phonetic Alphabet). The difference is partly due to the shape and position of the larynx (see below) and to finer motor control of the articulators — lips, tongue, jaw, and other structures that modify sounds. This fine motor control cannot be mastered by monkeys or apes (as the early ape language experiments showed), and it is now known to have a genetic component. In

humans, the gene known as FOXP2 controls the facial and mouth motor abilities necessary for speech; damage to this gene causes inability to speak but few or no other cognitive handicaps. The normal human form of this gene dates to a mutation that was established about 200,000 years ago; this may have been a watershed event in the history of human speech (Zuberbühler, 2005).

But speech is not synonymous with language, and may not even be a prerequisite for it.

Most primates have a repertoire of vocal calls, but only we and our closest relatives, the apes, regularly communicate with our hands as well, suggesting that gesture may be a newer evolutionary development than the ability to vocalize. A counterintuitive theory that is gaining ground among researchers in a range of fields — from primatology, neuroscience, and even paleontology — is the notion that the driving force in language evolution may not have been the inarticulate cries of our primate ancestors, but their gestures (Corballis, 2003).

At Emory University, then-PhD-student Amy Pollick and her mentor Frans de Waal coded over 600 hours of videotaped interactions by chimpanzees and their relatively less-studied relatives, bonobos, in different captive groups. The aim was to compare the animals' gestural and vocal/facial communication. They found that the overwhelming majority of signals used to initiate social interactions in both species were either solely gestural or involved a combination of gestures and facial/vocal signals. According to Pollick, this finding was a surprise: Apes scream and hoot at each other a lot, and it would be easy for a casual observer to assume vocalization is these animals' dominant mode of initiating communication.

Ape vocalizations have been relatively less studied than those of monkeys (Zuberbühler, 2005), but recently Zuberbühler and his colleagues have found evidence for cognitive complexity and audience effects in chimpanzee screams. For example, during aggressive encounters, individuals varied their screams depending on the severity of an encounter, their own role in it, and who else was present to hear them; they even exaggerated calls for support (intensifying the severity of a call compared to the real severity of the encounter) if a higher-ranking male was present (Slocombe & Zuberbühler, 2007). But most research so far shows that chimps' vocal signals are not much more complex than those of monkeys. Sounds are fairly highly stereotyped and are closely tied to particular emotions and situations (Pollick & de Waal, 2007). Social contexts eliciting particular facial/vocal displays in chimps reliably elicit the same displays in bonobos, and vice versa; and most vocalizations don't appear to have a targeted recipient.

By contrast, Pollick and de Waal found a highly nuanced hand-gesture vocabulary in chimps and bonobos, with great situational variation in use of gestures and combination with vocalizations, and a tendency to use gestures dyadically (i.e., more like conversational exchanges). The Emory researchers found that chimp and bonobo gestures were much less tied to particular emotions and situations than their vocalizations were. And hand gestures, even if they clearly evolved from basic object-related manual movements, were much more conventionalized (i.e., less stereotyped) and appeared to be deployed more deliberately — revealing greater cortical control over this mode of communication. Often the meaning of a particular chimp or bonobo gesture could only be extracted from its context.

The Emory researchers also found that, particularly in bonobos but to a lesser extent in chimps, gestures differed between different groups of the same species — evidence that, in these animals, gesture has truly begun to break from biology, becoming cultural. “Far more than facial expressions and vocalizations,”

they write, “gestures seem subject to modification, conventionalization, and social transmission” (Pollick & de Waal, 2007, p. 8188). Pollick and de Waal speculate that the flexible use of gestures and responsiveness to combined signals that they observed “may have characterized our early ancestors, which in turn may have served as a stepping stone for the evolution of symbolic communication” (p. 8188).

Talking With Our Hands

Pollick, who now works in Washington, DC as APS’s Director of Government Relations, admits that her interest in ape gestures and the evolution of language isn’t accidental. She is deaf and from a deaf family, so American Sign Language is, so to speak, her native tongue. “Having grown up with ASL, I was just naturally attuned to issues of communication,” Pollick says. “I was also naturally attuned to gesture. All humans gesture, wherever they are, in all cultures. People gesture when they are not visible to the receiver, such as when they talk on the phone. Blind people gesture when talking to other blind people. This led me to think that gesture is deeply ingrained in human communication, and I began to wonder where this came from.”

Pollick explains that in order to theorize about the relationship between ape gestures and human language, she drew a stricter distinction between hand gestures and other body movements than previous ape communication researchers had drawn. She also made finer-grained distinctions among different gestures — for example, determining that the meaning of an outstretched hand depended on the angle the hand was rotated at. Chimps, for instance, used an extended, upraised palm (i.e., “gimme”) in a variety of situations: to request food, to request sex, to request to be groomed, or to implore the aid of another chimp. Sometimes the gesture was combined with a vocalization such as a scream. Bonobos mainly used the gesture to solicit play.

Humans use the “gimme” gesture too — as well as countless others. Linguists used to ignore the way humans use their hands when communicating, or relegate it to the subordinate category of “body language.” But psychological research on human gesture is revealing that, as with chimps and bonobos, when humans talk with their hands it is far more than just an exception or a sideshow to the main attraction.

Important insights into the nature of language have come from studies of signing in the deaf. Linguists agree that human sign languages such as ASL are every bit as “linguistic” as spoken languages are — that is, they possess all the syntactical complexity and are just as flexible and open-ended as their spoken analogues. They are also just as readily learned. University of Toronto psychologist Laura-Ann A. Petitto found that deaf children exposed to ASL or the Quebec sign language, *Langue des Signes Quebecoise*, learned to sign at the same rate that hearing children acquire spoken language (Petitto, 2000). Other researchers have even suggested that deaf children acquire sign language on a faster developmental schedule than non-deaf learners of spoken language (Meier & Newport, 1990).

If a sign language doesn’t happen to be available in a deaf child’s environment, she will go ahead and invent one. APS Fellow and Charter Member Susan Goldin-Meadow (University of Chicago) studied profoundly deaf children in the United States and Taiwan who were raised by hearing parents and were unexposed to sign language. Such children spontaneously used gesture to communicate, and their gestures displayed the same structural properties (such as recursion and displaced communication — referring to things not present) that characterize natural spoken languages and sign languages (Goldin-

Meadow, 2006).

The sign language instinct appears to exist also in hearing adults who have never learned a sign language. As Pollick notes, most people talk with their hands — that is, gesticulate to provide counterpoint, emphasis, or visual illustration of what they are saying with speech. Goldin-Meadow found that, when gesture accompanies speech, it lacks the fully linguistic properties observed in deaf people's spontaneous signing. In another study by Goldin-Meadow, non-deaf participants were asked to describe an event orally and also to attempt to describe it using only gestures. When they used speech, their hand gestures supported what they were saying imagistically, supporting the main spoken channel of communication by providing a kind of visual aid, but were not by themselves linguistic; however, when they had to describe an event solely through hand gestures, their gestures assumed the linguistic properties the researcher found in her studies with deaf children (Goldin-Meadow, 2006).

Goldin-Meadow's University of Chicago colleague, APS Fellow David McNeill, considers hand gestures to be intrinsic to language, driving thought and speech. Language, he argues, is a dialectic in which images (conveyed by the hands) work with and against speech, the tension between these two modes of thought propelling thought and communication forward. His studies of speakers' hand gestures revealed a temporal structure, distinct from ordinary syntax (sentence structure), in which gestural imagery and spoken content periodically resolved in what he dubs "growth points" — temporal boundaries of unfolding thought sequences that can be detected when a word or phrase synchronizes with gesture in a certain way (McNeill, 2006).

Speakers vary in how they use their hands when they are speaking. In another series of experiments, Goldin-Meadow compared learning and problem-solving ability in speaking children whose gestures matched (i.e., conveyed the same information as) their own or their teachers' verbal explanations of problems with the abilities of children whose gestures conveyed different information than what was being spoken. Children who used mismatching gestures or who were taught by teachers who used mismatching gestures learned faster and were more successful at solving problems than were those whose gestures merely supported the spoken communication. It suggests that this second, silent channel of nonverbal information may be an important helping hand (so to speak) to thought. According to Goldin-Meadow, "A conversation in gesture ... appears to be taking place alongside the conversation in speech whenever speakers use their hands" (Goldin-Meadow, 2006, p. 37).

How Did We Get Here?

APS Fellow and Charter Member Michael Corballis (University of Auckland) is a proponent of the gesture-first scenario of language origins, arguing that it makes sense of a wide range of findings in various fields (Corballis, 2003). There is the paleontological evidence, for one thing: Sometime after the human lineage split from that of chimps and bonobos about six million years ago, our australopithecine ancestors became bipedal; whether or not gestural communication was a factor driving this shift, it did free the hands for greater manipulation of the physical environment, and likely facilitated communicative manipulation of the social environment. Australopithecine brains remained ape-like in most respects, as did their vocal tracts and breathing apparatus — meaning they couldn't talk — but expansion of the cortex, including specific brain areas governing language and gesture, is found in their tool-manufacturing descendent *Homo habilis*. Endocasts (casts of brain cases) of 2-million-year-old *H. habilis* skulls reveal an asymmetry that could correspond to expansion of areas on the left side of the brain, such as Broca's area, that have long been associated with language in humans (see Corballis,

2003).

In the 1990s, a group of neuroscientists at the University of Parma, Italy, made a discovery suggesting that the brain area in monkeys corresponding to Broca's area could have served as the platform for the emergence of language out of gesture in our deep primate past (Fogassi & Ferrari, 2007; Rizzolatti & Arbib, 1998). The homologue of Broca's area in the monkey brain, known as area F5, is involved in controlling manual gestures, not vocalizations; it also possesses a *mirror-neuron system*. Mirror neurons fire both when an animal (or person) initiates an action and when the animal perceives another individual make the same action (see "Mirror Neurons: How We Reflect on Behavior" in the May, 2007, issue of the *Observer*). Mirror neuron systems have been proposed as the basis for various forms of learning, social coordination, and even theory-of-mind abilities in animals and humans.

One category of mirror neurons in the monkey F5 fires both when the monkey makes a motor act with its forelimbs and when it hears the sound produced by the same action (e.g., by another monkey, or on a recording); another type of motor neuron in the same area specifically activates during observation of another monkey's mouth-communicative gestures like lip-smacking or tongue protrusion (Fogassi & Ferrari, 2007). Research in humans has revealed similar properties for Broca's area. It activates when people observe goal-related hand or mouth motor movements by other people, for example. And there is evidence for a matching mechanism whereby heard phonemes activate corresponding tongue motor representations in the cortex (Fogassi & Ferrari, 2007).

The neuroscience findings support a longstanding and influential theory of speech perception called the motor theory (Liberman & Mattingly, 1985). In this theory, linguistic primitives (basic elements) are not represented in the cortex as abstract sounds but as the motor signals that one would use to make those sounds. The picture is turning out to be more complicated — speech perception involves many more parts of the brain than just Broca's area. But the common neural basis of manual dexterity and important aspects of language means that the notion that language is a "tool" could be more than just a metaphor. Could tool use and language be flip sides of the same cognitive coin?

Hand to Mouth

Corballis, like Zuberbühler, sees speech per se as a late development possibly occurring only with the rise of *Homo sapiens* around 170,000 years ago. Yet language, in some form, could well have been around for a long time before that. The beginnings of stone tool manufacture occurred around 2.5 million years ago, followed by a shift from forest dwelling to living and carrying materials on the open savannah around 2 million years ago. These developments hint at a level of abstract thinking and social coordination abilities that could have gone hand in hand (so to speak) with language skills.

One plausible scenario for the transition from gestural to vocal communication is that increased use of the hands for tool making and carrying drove greater use of the face for communication, and this ultimately led to speech. In modern sign languages, manual gestures convey semantic content and facial and body movements act as modifiers. Corballis suggests that facial movements could have become integrated into the manual sign system as carriers of syntax (Corballis, 2003) and this integration could have been an outgrowth of the mechanics of eating — an idea supported by mirror-neuron findings (Fogassi & Ferrari, 2007).

From this gesture–face integration, it would have been a small evolutionary step to add voicing to facial

gesture to provide more range of meaning — perhaps made possible by the FOXP2 mutation mentioned earlier, as well as the descent of the larynx and changes to the muscles controlling breathing.² Corballis suggests that we should think of speaking not as the production of abstract phonemes but as a kind of noisy gesturing with our mouths (see Corballis, 2003).

Whatever its evolutionary origins, speech has acquired a great deal of autonomy from hand gestures. People can normally communicate on the phone, for example, with little loss of meaning, even if gesturing at the same time helps them think. The autonomy of speech has left many people with the natural impression that our linguistic abilities are more closely akin to animal vocalization than to other forms of communication. Since words and gestures don't fossilize, it may never be possible to answer the question "Where did language come from?" definitively. Pollick acknowledges that the theory that our language abilities evolved from gesture remains a "just-so story," even though the scientific evidence for pieces of the theory is compelling. Only further research, across a range of disciplines, can help settle the question of whether talking pretty piggybacked on gesture or the other way around — or whether language evolved in some completely different way.

References

- Cheney, D.L., & Seyfarth, R.M. (2005). Constraints and preadaptations in the earliest stages of language evolution. *The Linguistic Review*, 22, 135–159.
- Corballis, M.C. (2003). From mouth to hand: Gesture, speech, and the evolution of right-handedness. *Behavioral and Brain Sciences*, 26, 199–260.
- Darwin, C. (1998). *The expression of the emotions in man and animals*. New York: Oxford University Press. (Original work published 1872)
- Fogassi, L., & Ferrari, P.F. (2007). Mirror neurons and the evolution of embodied language. *Current Directions in Psychological Science*, 17, 136–141.
- Goldin-Meadow, S. (2006). Talking and thinking with our hands. *Current Directions in Psychological Science*, 15, 34–39.
- Golinkoff, R.M., & Hirsh-Pasek, K. (2006). Baby wordsmith: From associationist to social sophisticate. *Current Directions in Psychological Science*, 15, 30–33.
- Hagoort, P. (2008). Should psychology ignore the language of the brain? *Current Directions in Psychological Science*, 17, 96–101.
- Holt, L.L., & Lotto, A.J. (2008). Speech perception within an auditory cognitive science framework. *Current Directions in Psychological Science*, 17, 42–46.
- Hopkins, W.D., & Cantalupo, C. (in press). Theoretical speculations on the evolutionary origins of hemispheric specialization. *Current Directions in Psychological Science*.
- Kraus, N., & Banai, K. (2007). Auditory-processing malleability: Focus on language and music. *Current Directions in Psychological Science*, 16, 105–110.
- Liberman, A.M., & Mattingly, I.G. (1985). The motor theory of speech perception revisited. *Cognition*, 21, 1–36.
- McNeill, D. (2006, September). *Gesture and thought*. Paper presented at the Summer Institute on Non-verbal Communication and the Biometrical Principle, Vietri sul Mare, Italy. Downloaded April 2, 2008, from http://mcneilllab.uchicago.edu/pdfs/dmcn_vietri_sul_mare.pdf
- Meier, R.P., & Newport, E.L. (1990). Out of the hands of babes: On a possible sign language advantage in language acquisition. *Language*, 66, 1–23.
- Petitto, L.A. (2000). On the biological foundations of human language. In H. Lane & K. Emmorey (Eds.), *The signs of language revisited* (pp. 447–471). Mahwah, NJ: Erlbaum.

- Pinker, S. (1994). *The language instinct*. New York: HarperCollins.
- Pinker, S. (1997). *The stuff of thought*. New York: Viking.
- Poeppel, D., & Monahan, P.J. (2008). Speech perception: Cognitive foundations and cortical implementation. *Current Directions in Psychological Science*, 17, 80–85.
- Pollick, A.S., & de Waal, F.B.M. (2007). Ape gestures and language evolution. *Proceedings of the National Academy of Sciences, USA*, 104, 8184–8189.
- Rizzolatti, G., & Arbib, M.A. (1998). Language within our grasp. *Trends in Neurosciences*, 21, 188–194.
- Savage-Rumbaugh, S., Shanker, S.G., & Taylor, T.J. (1998). *Apes, language, and the human mind*. New York: Oxford University Press.
- Sedaris, D. (2000). Me talk pretty one day. In *Me talk pretty one day* (pp. 166–173). New York: Little, Brown.
- Slocombe, K.E., & Zuberbühler, K. (2007). Chimpanzees modify recruitment screams as a function of audience composition. *Proceedings of the National Academy of Sciences, USA*, 104, 17228–17233.
- Zuberbühler, K. (2005). The phylogenetic roots of language: Evidence from primate communication. *Current Directions in Psychological Science*, 14, 126–130.

1 The common neural origins of language and manual dexterity has been used to explain why nine out of ten humans are right-handed (Corballis, 2003). A bias toward right-handedness makes sense if manual dexterity developed in tandem with language ability, both sharing an underlying cortical substrate in the left hemisphere (which controls the right side of the body). However, recent discovery of handedness and hemispherical asymmetries in many other animal species, including many primates, may complicate this picture (see Hopkins & Cantalupo, in press).

2 The descent of the larynx, incidentally, made modern humans uniquely vulnerable to choking. Unlike other animals, we are prevented from breathing and swallowing at the same time, and, thus, are imperiled whenever we eat.