NEWS

The Origin of Speech

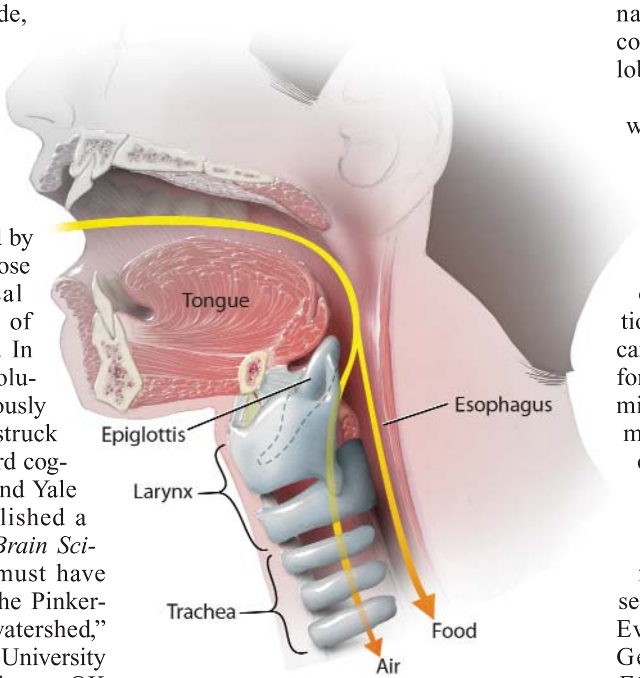
How did the remarkable ability to communicate in words first evolve? Researchers probing the neurological basis of language are focusing on seemingly unrelated abilities such as mimicry and movement

In the 1860s, both the British Academy and the Société de Linguistique de Paris warned their members not to discuss the origins of language, because the topic was so seductive—and so speculative—that it spawned endless, futile theorizing. More than a century later, Noam Chomsky, the most influential linguist of the last 50 years, wrote that language evolution and the brain mechanisms underlying it "appear to be beyond serious inquiry at the moment."

But the time now appears ripe for this endeavor. In the past decade, an unprecedented number of researchers from many disciplines have begun to tackle the origin of speech, spurred by new techniques as well as new ways of thinking. Among linguists, the question of language origins was long obscured by the dominance of Chomsky, whose theory of an innate "universal grammar" ignored the problem of how this language ability arose. In 1990, however, the wave of evolutionary thinking that had previously swept through biology finally struck linguistics too: That year, Harvard cognitive scientist Steven Pinker and Yale psychologist Paul Bloom published a long article in *Behavioral and Brain Sciences* arguing that language must have evolved by natural selection. The Pinker-Bloom paper was "a kind of watershed," says linguist James Hurford the University of Edinburgh, U.K. "Suddenly it was OK to talk about evolution of language in Chomskyian circles."

Meanwhile, advances in brain imaging, neuroscience, and genetics have enabled a new contingent of researchers to go ever deeper into our brains and our biological past. For a long time, researchers treated language ability as some sort of "miracle," says neuroscientist Michael Arbib of the University of Southern California (USC) in Los Angeles. Now, he says, researchers are breaking that miracle down into a series of smaller, more manageable "miracles," involving disparate capacities such as the ability to imitate facial expressions or to string movements together. They're not fantasizing that the human brain at some point suddenly found that it could speak with the tongues of angels, he says; rather, it achieved a more modest state some researchers call "language readiness," which opened the door to further advances in linguistic ability.

Language origins are "certainly worth talking about now," says Hurford, who in 1996 launched the first of a series of biennial conferences on language evolution* that have grown steadily. Hurford's Edinburgh colleague Simon Kirby has documented the leap in interest with a citation search: The number of papers dealing with both "language" and "evolution" more than doubled from the 1980s to 1990s. (See also Book Review, p. 1299.)



Dangerous talk. Side view of human vocal tract shows that because of our lowered larynx, food and drink must pass over the trachea, risking a fall into the lungs if the epiglottis is open.

Yet despite all the activity, the new lines of evidence remain indirect, leaving plenty of room for interpretation—and conflict. "If you want a consensus, you won't get it," says cognitive scientist Philip Lieberman of Brown University. With no fossils of speech, the origin of language remains "a mystery with all the fingerprints wiped off," says brain scientist Terrence Deacon of the University of California, Berkeley.

The long view

Archaeologists have identified various milestones in human behavior in the 5-million-year evolutionary void between animal communication and human speech, but there is no consensus on which achievements imply the capacity for language. For example, the first stone tools date to 2.4 million years ago; some researchers think this may indicate linguistic facility, but others argue that toolmaking is far removed from speech. Another possible starting point is 2 million years ago, when the hominid brain began a period of rapid expansion, including in the primary brain areas associated with producing or processing language—namely Broca's area in the left frontal cortex and Wernicke's in the left temporal lobe (see brain model, p. 1318).

As for actually producing the sounds of words, or phonemes, skeletal studies reveal that by about 300,000 years ago, our ancestors had become more or less "modern" anatomically, and they possessed a larynx located at the top of the trachea, lower than in other primates (see diagram). This position increases the range of sounds humans can make, although it also makes it easier for food going down the esophagus to be misdirected into the windpipe, leaving us more vulnerable than other mammals to choking. Such anatomy could have developed for no other purpose than speech, says Deacon.

Other possible milestones come from genetic studies. For example, researchers at the Max Planck Institute for Evolutionary Anthropology in Leipzig, Germany, reported last year that the *FOXP2* "speech gene," which affects both language and the ability to articulate (*Science*, 16 August 2002, p. 1105), was apparently a target of natural selection. This gene may have undergone its final mutation fewer than 100,000 years ago—and no more than 200,000 years ago—perhaps laying the groundwork for a new level of linguistic fluency.

Most researchers are inclined to the view that language gradually emerged over perhaps a couple of hundred thousand years (*Science*, 20 November 1998, p. 1455). But all we know for certain, says Pinker, is that fully developed language was in place by at least 50,000 years ago, when humans in Europe were creating art and burying their dead, symbolic behaviors that point unequivocally to fluent language.

ILLUSTRATION: C. SLAYDEN/SCIENCE

* www.ling.ed.ac.uk/evolang

The motor route

Understanding when language emerged will probably have to await better understanding of how it emerged. In recent years many researchers have become increasingly attracted to the notion that changes in the brain's motor areas were crucial for language capability.

Although we tend to associate language first with sound rather than movement, speech may be better understood as a motor activity, says Deacon. Like other fine motor activities such as threading a needle or playing the violin, speech demands extraordinarily fine and rapid motor control. Elaborate movements of the larynx, mouth, face, tongue, and breath must be synchronized with cognitive activity.

Thus researchers are probing the links between language and areas in the brain that control gestures, either hand movements or the articulatory gestures of mouth and tongue. Linguist Robert Kluender of the University of California, San Diego, says explorations of gestures, including sign language, offer glimpses into what might have been the "intermediate behavioral manifestations" between animal communication and speech.

Many researchers think hand and face gestures offer behavior that is more analogous to speech than are animal vocalizations. In all other mammals, both breathing and articulation are directed by brain areas quite separate from those associated with human speech, notes Pinker. Lieberman argues that nonhuman primates engage in "a limited number of stereotyped calls" such as alarm calls and that they don't have the interactive or combinatorial quality of language. Apes' anatomy is such that they "could produce a [phonetically] reduced form of human speech," adds Lieberman. "But they don't." They're much better at signing, because apes' motor behaviors have more flexibility and are more involved in social interaction—through gaze, mouth and facial movements, and limb gestures—than their calls, Lieberman says.

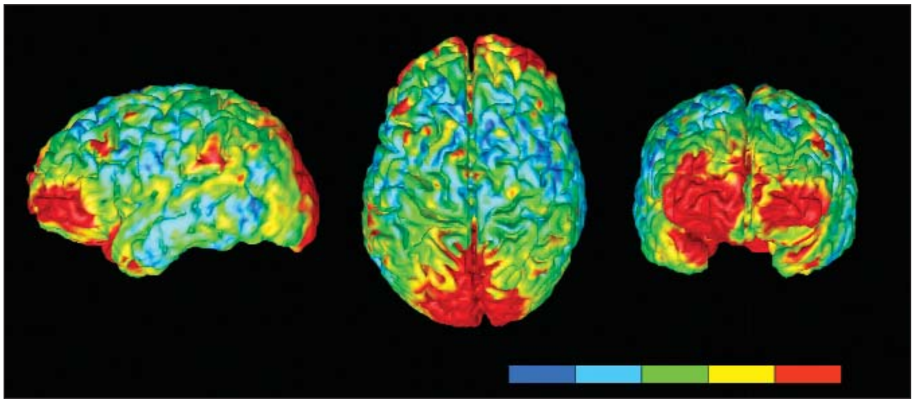
Lieberman argues that the crucial changes that laid the groundwork for language ability occurred in brain circuits connected with the basal ganglia, subcortical structures involved in movement. In his view the basal ganglia is the "sequencing engine" that makes combinations—both verbal and gestural—possible. As evidence he points to the fact that patients with Parkinson's disease, which disrupts the basal ganglia, suffer erosion of syntactical abilities, as well as problems with balance and movement.

Pinker's research, with cognitive scientist Michael Ullman of Georgetown University in Washington, D.C., lends weight to this view. They have shown that Parkinson's patients with basal ganglia damage have more trouble with regular verbs than with irregular ones. Conjugating a regular verb such as "walk," Pinker explains, is a combinatorial, sequential task that calls for adding the "ed" for past tense. But retrieving the past tense for an irregular verb such as "come" simply calls on long-term memory. Such tasks require other brain areas as well, but Lieberman argues that the basal ganglia are a common element in both movement and language disorders.

Indeed, although many other brain areas, including those responsible for articulation, hearing, planning, and memory, had to develop to support language, there is abundant behavioral evidence for an intimate connection between language and motor abilities, says Pinker. For example, psychologist David McNeill of the University of Chicago cites the case of a man who lost all sense of touch below the neck due to a strange virus. Although the man had to relearn the simplest movements, using cognitive and visual feedback to substitute for lost senses, he continued to gesture automatically when he spoke, even when researchers hid his hands from his own and listeners' view. "The hands are really precisely linked to speech articulation," says McNeill. "Gesture is not a behavioral fossil that was superseded by language but an indispensable part of language."

But not everyone is ready to dismiss the meaningfulness of animal calls, with differing views often dependent on a scientist's specialty. Primatologist Marc Hauser of Harvard, for example, believes that primate calls are better candidates for speech precursors than any gestures are. With chimp gestures, "nothing gives a suggestion of anything referential"—that is, having an explicit association with a concept or thing—he says. Primate alarm calls, in his view, "kind of look like words." For example, he cites work by psychologist Klaus Zuberbühler of the University of St. Andrews, U.K., who has reported that African Diana monkeys can modulate their alarm calls to indicate what type of animal (leopard or eagle) is threatening. Such sounds, says Hauser, "have a far greater … connection to language than any discovery on nonvocal signals."

Many linguists, too, are unmoved by motor arguments, which they do not believe can explain how the brain developed syntax. "Motor organs are for muscular movements," says Derek Bickerton of the University of Hawaii, Manoa, and that puts them at the "end of the pipeline" of language production. "Whatever organizes motor movements is on a par with what organizes throwing movements," says Bickerton. "The purpose is to put things in a regular invariant sequence." That, he says, is very different from making sentences, which requires "putting



From ape to human. Magnetic resonance images of a bonobo brain are warped onto the shape of a human cortex, viewed from (left to right) the side, top, and front. Red and yellow areas in the temporal region (linked to language) and in the prefrontal and occipital regions had to be stretched the most to reach the human configuration, whereas blue areas are similar in apes and humans.

Hand and mouth. Chimps gesture with both face and hands to help express themselves.

CREDITS: (TOP TO BOTTOM) KARL ZILLES, PETER PIEPERHOFF, AND HARTMUT MOHLBERG/RESEARCH CENTER JÜLICH, INSTITUTE OF MEDICINE, GERMANY; NIGEL J. DENNIS/PHOTO RESEARCHERS INC.

things into an extremely plastic order determined by your conceptual structure."

Mirror, mirror

Despite such caveats, the motor-language connection continues to draw attention, in part because of a 1996 discovery that many see as the first hard data in years to bolster the theory. This is the so-called mirror neuron system found in monkeys' brains.

Mirror neurons' link to language depends on imitation, a skill largely unique to humans and considered vital to language. Although parrots and dolphins can do vocal mimicry, imitation is not as a rule a mammalian attribute: Even nonhuman primates do it poorly (contrary to the implication of the term "to ape"). But imitation is the way babies learn their first words. And it's the only way a common meaning can emerge for an abstract symbol, a phenomenon that linguists call "parity." "Imitation is the common thread for people writing about language origins," says neuroscientist Marco Iacoboni of the University of California, Los Angeles.

So researchers were excited when a team led by Giacomo Rizzolatti of the University of Parma, Italy, found what they considered a plausible antecedent for the human ability to imitate in the brains of monkeys. The team recorded electrical activity in macaques from 532 neurons in an area called F5, which is homologous to Broca's area in humans. Neurons in F5 are known to fire during monkeys' "goal-directed" hand and mouth movements—for example, when reaching for food.

What intrigued the researchers is that a subset of these neurons, which they dubbed mirror neurons, also became active when a monkey merely watched another monkey (or a human) perform the action. This finding "opened a whole new approach to the language evolution story," says Arbib of USC. "What would a mirror system for grasping be doing in the speech area of the brain?" The researchers concluded that these mirror cells form a system for matching the observation and execution of mouth and hand actions—the first steps toward imitation.

So far, mirror neurons have been found in only two brain areas in macaques, and the single-cell brain recording technique that revealed the macaque neurons isn't done on humans. But Iacoboni believes he has identified a similar circuit—"a core neural architecture for imitation"—in people. He combined the results of single-cell brain recordings in monkeys with functional magnetic resonance imaging in humans while they watched or imitated finger movements or facial expressions. Iacoboni says that in addition to Broca's, the circuit comprises an area in the superior temporal cortex (which overlaps with Wernicke's and has neurons that respond to face and body movements) and one in the parietal cortex, the homolog to the macaque area called PF, which combines visual and bodily information. "The neural architecture for imitation … overlaps very well with well-known language areas in the human brain," says Iacoboni, who concludes that the dual-use nature of Broca's area in particular "suggests an evolutionary continuity between action recognition, imitation, and language."

Wired for imitation? Classic language areas—Broca's and Wernicke's (yellow)—overlap (orange) with areas critical for imitation (red).

Mirror neurons provide the "neural missing link" between movement and speech control, says Arbib. They also fit well with an old theory, the "motor theory of speech perception," developed in the 1950s by the late Alvin Liberman of Yale University's Haskins Laboratories. Psychologist Michael Studdert-Kennedy of Haskins Labs explains that when children imitate their first words, experiments have shown that they (unlike another imitator, the parrot) are guided by the "gestural" features of the sound—that is, by the actions of the mouth rather than by a sound's acoustic features. A well-known trick to demonstrate this is known as the McGurk effect: If you watch someone pronounce the syllable "ga" while listening to a recording of someone saying "ba," you will likely hear "da," a sound anatomically between the other two.

This means "you perceive speech by referring the sounds you hear to your own production mechanism," says Studdert-Kennedy. Humans, unlike other animals, are equipped with an intuitive sense of how their body parts correspond with those of others. Thus a small child knows how to raise its hand in response to a parental wave. "There's obviously a direct representation of your body in its body," says Studdert-Kennedy.

The theory developed new life when Studdert-Kennedy brought it to bear on questions of language evolution. Mirror neurons, he says, "for the first time provide an example of a direct physiological hookup between input and output": the observation of an action and its imitation. Indeed, Rizzolatti's group recently reported that the macaque has "audiovisual" mirror neurons: Some of the cells in F5 fire not only when a macaque watches a meaningful grasping action, but when it hears the sound of one, such as the sound of breaking peanuts (*Science*, 2 August 2002, p. 846). Arbib believes that mirror systems probably exist in other parts of the brain for many other behaviors.

He and others feel that mirror neurons offer the first concrete neurological evidence of abilities crucial to language, but it's a long way from a few firing neurons to speech. Some scientists think the potential significance of mirror neurons may be exaggerated. Macaques, after all, can't speak and they can't imitate either, notes Pinker. In his view, mirror neurons' "relevance to language is still pretty fuzzy."

The first syntax: words or waves?

Despite such drawbacks, mirror neuron theory is being invoked by both sides in the schism over whether the earliest language—that is, symbolic sounds or gestures connected by some sort of rules of syntax—used the voice or the hands.

Those who favor gestural origins, such as psychologist Michael Corballis of the University of Auckland, New Zealand, point out that mirror neurons are found in brain areas responsible for grasping. "I think it's extremely likely that language evolved in our early ancestors as a manual system, not as a vocal one," as far back as a million years ago, says Corballis. He notes that when robbed of speech, people quickly develop sign language, as has been shown by the case of a community of deaf Nicaraguans who created their own language.

Given the strong role of manual and facial gesture in speech and the relatively recent final mutation of the *FOXP2* gene, Corballis argues that "autonomous" speech may not have become fully developed until the cultural explosion beginning 50,000 years ago. The mirror system, he believes, reinforces his theory, because it apparently evolved first for manual control. It "probably picked up vocal and facial control quite late in hominid evolu-

CREDIT: MARCO JACOBONI, JOHN BACHELLER, AND ARTHUR TOGA/UCLA

tion," he says, as speech became the preferred modality for communication for various reasons, such as the need to free the hands for work or to talk in the dark.

But others believe equally strongly that even if movement and language are inseparable, language is primarily an oral, not manual, behavior. Psychologist Peter MacNeilage of the University of Texas, Austin, has developed a theory that monkey oral behaviors (not vocalizations) are precursors of human syllables, and he argues that the mirror neuron system—especially the recent discovery of neurons that respond to lip smacking and nut cracking—bolsters his ideas.

MacNeilage suggests that the brain's supplementary motor area (an area adjacent to the primary motor cortex that is important for motor memory and sequential movements) controls the physical constraints on vocal expression. The actions of chewing, sucking, and licking took on communicative content—a job for Broca's predecessor—in the form of lip smacks, tongue smacks, and teeth chatters. The next stage, says MacNeilage, was to give voice to these behaviors by bringing the larynx into play. This theory fits well with the fact that the unique sounds of click languages, which some speculate may have been the original mother tongue (see next story), do not use the larynx. Once the larynx was involved, a phonology—a set of sounds that could be combined in endless ways to form a large vocabulary—developed, and this in turn paved the way for the emergence of syntax.

"I don't believe manual gestural communication got to the point of the combinatorial phonology that I'm talking about, because if it did we'd still have it," says MacNeilage. In his view, if sign language had become that complex, there would have been no reason strong enough—the desire to talk in the dark notwithstanding—to cause a transition to vocal speech. "Nobody who argues that we went from sign to speech has given us an adequate translation theory," he says.

Others say the "which came first" debate is beside the point. "Evolution selected the ability to combine speech and gesture under a meaning," says McNeill. "The combination was the essential property"; neither gesture nor speech could have evolved without the other, he says. It doesn't matter which came first, agrees Zuberbühler: "Once an individual reaches a certain threshold in its cognitive sophistication, it will inevitably express itself in a sophisticated way," through any means at its command, he says.

The deepest questions—such as how humans became symbolic thinkers and developed "theory of mind," or awareness of others' thought processes—remain far from resolved. Researchers say one way to tackle them will be through ever-finer brain imaging technology so they can, as Bickerton puts it, "find out the flow chart for a sentence in the brain." Harvard's Hauser and colleagues believe that research in animals may identify behavioral analogs for "recursion": the ability to string words together in infinite hierarchical combinations. Arbib predicts that the discovery of other types of mirror systems, in both humans and animals, will help yield a better "taxonomy" of the language conundrum, especially if bolstered by computational modeling. But answers won't come all at once. "I see this as a process of gradual convergence. The problem space is shrinking" at long last, says Bickerton. "It will be solved when that space goes to zero, not when someone comes up with the killer solution."

–Constance Holden

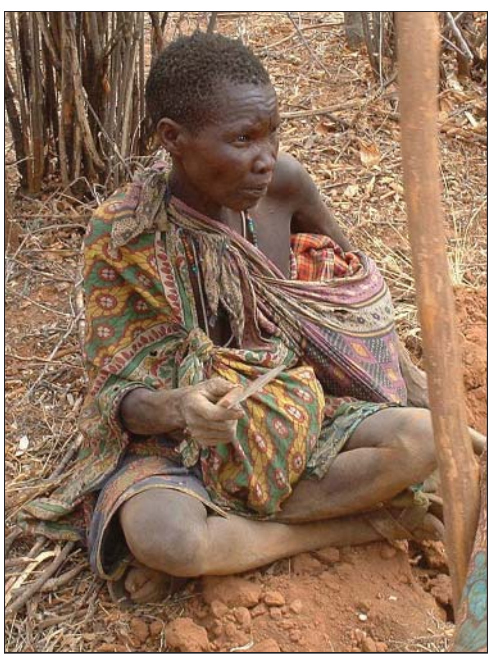
NEWS

The First Language?

Genetic and linguistic data indicate—but can't quite prove—that our ancient ancestors spoke with strange clicking noises

In the 1980 movie *The Gods Must Be Crazy*, a soda bottle falls out of the sky and lands among some strange-sounding Africans. Their excited chatter, punctuated by rapid-fire sucking and clicking noises, sounded intriguing but alien to audiences around the world. But a handful of studies of this seemingly esoteric language suggest that our early ancestors depended on such clicks to communicate. The latest linguistic work points to clicks as having deep roots, originating at the limits of linguistic analysis sometime earlier than 10,000 years ago, and genetic data suggest that click-speaking populations go back to a common ancestor perhaps 50,000 or more years ago.

Although the idea is far from proven, "it seems plausible that the population that was ancestral to all living humans lived in the savanna and used clicks," says vertebrate systematist Alec Knight of Stanford University. Knight estimates that today only about 120,000 people rely on these odd sounds. Even so, they are providing new insights into how humans evolved the gift of gab, particularly when researchers add up the results of different kinds of data. "There's a lot of mileage to be gained by cross-referencing linguistic, genetic, and archaeological data and theories," says Nigel Crawhall, a graduate student studying click languages at the University of Cape Town, South Africa.



CREDIT: NIGEL CRAWHALL

All alone. Researchers ponder why the Hadzabe live so far from other click speakers.

Clicks in context

Today clicks are part of typical conversation for about 30 groups of people, most from Botswana, Namibia, South Africa, and nearby. The only recognized non-African click language is Damin, an extinct Australian aboriginal language used only during manhood initiation ceremonies. Among African click speakers, daily conversations can be dominated by clicks, and sometimes verbal sounds drop out completely.

Adept tongue and inward air movements distinguish clicks from other nonverbal utterances. They are really just very strongly pronounced consonants, says Amanda Miller-Ockhuizen, a linguist at Cornell University in Ithaca, New York. Click speakers have click sounds in common, but they have different words and therefore very different languages.* Some researchers argue that click languages are far more different from each other than English is from Japanese.

But that diversity is only now being

* To hear click sounds, go to hctv.humnet.ucla.edu/departments/linguistics/VowelsandConsonants/index.html