Technology
PUBLISHED BY MIT Review

January/February 2008

The Genetics of Language

Researchers are beginning to crack the code that gives humans our way with words.

By Jon Cohen

Daniel Geschwind reaches up to his office bookshelf, takes down a three-dimensional puzzle of the human brain, and begins trying to snap the plastic pieces together. A neurogeneticist at the University of California, Los Angeles, Geschwind hopes the puzzle will help him describe the parts of the brain that control speech and language. But for the life of him, he can't figure out how the left and right hemispheres attach. "I'm really bad spatially, so don't make fun of me," he pleads. "It's like I'm having a little stroke or something. I'll get it together, and then I'll figure it out."

The plastic model may have momentarily flummoxed Geschwind, but when it comes to the genes that govern the brain's development and functions, he excels at putting the pieces together. Over the past few years, he has emerged as one of the leading geneticists in a nascent field that aims to spell out which genes are related to speech and language development--and how our intelligence and communication skills evolved beyond those of our ape relatives, giving us the unique ability to speak.

Research like Geschwind's sits at the intersection of two fields: behavioral genetics and evolutionary biology. Each field depends on the other to make sense of the flood of studies on the genetics of language now pouring out of labs around the world. To peer into the human brain and see how it typically stores, uses, and comprehends words, Geschwind investigates not only normal human brains but also those where the process goes awry, studying the genes of families afflicted by autism, dyslexia, schizophrenia, and other conditions that can involve speech and language disorders. This research may help make diagnosis and treatment of language-related disorders more precise, but it also has a more basic purpose. "Studying disease is really a fundamental way to understand normal function," says Geschwind. "Disease has given us extraordinary insight to understand how the brain works or might not work."

While behavioral genetics compares the genes of people with different abilities, evolutionary biology compares the genes of different species. Researchers use this data to determine what limits other species' communication skills and what expanded ours so dramatically that language became one of our defining characteristics. Geschwind's own forays into evolutionary biology have led him to look at DNA in the brains of chimpanzees, monkeys, and even songbirds. "A lot of people think our lab is all over the place," he says. "It's actually pretty integrated. Language is complex, and the only way we're going to have a hit is when two or three findings point to the same place."

With the help of improved techniques for detecting DNA, as well as cutting-edge analytical tools and the genome sequences of species from humans to mice, Geschwind and other researchers have begun to tease out how we evolved the capacity for sophisticated speech. But though neuroscientists working in the postgenomic era have made a lot of progress, they have only begun to scratch the surface of how the relevant genes are collectively put into action.

FOXP2 Hunting

Despite more than a decade of effort and many tantalizing leads, neurogeneticists have so far definitively linked only a single gene to speech and language. The story of its discovery begins in 1990, when clinical geneticists at the Institute of Child Health in London first reported a speech disorder that appeared in three generations of Britons known as the KE family. The doctors took note of 15 affected members who seemed to have inherited problems with grammar, syntax, and vocabulary that were tied to poor control of facial muscles and difficulty pronouncing words. Although it seemed clear that there had to be a genetic link, researchers hunted for more than a decade before they found the gene responsible.

The big break came in 1998, when University of Oxford geneticists led by Anthony Monaco and Simon Fisher identified a distinct chunk of chromosome 7 linked to the speech and language problems found in the KE family. Yet the region held dozens of genes, and they couldn't pinpoint the one bad actor. Enter Jane Hurst, a clinical geneticist who worked at a hospital on Oxford's grounds and, coincidentally, had coauthored the first report on the KE family.

The chromosome 7 paper led Hurst to reëxamine the results of an amniocentesis, for a pregnant woman unrelated to the KE family, that she had reviewed four years earlier. Hurst had found that the fetus had a chromosomal hiccup called a translocation, and she later learned that the child developed speech and language problems strikingly similar to those seen in the KE family. Looking at the results again, she saw that the translocation

had occurred in the very same region of chromosome 7 that Fisher had identified. "I phoned up Simon and said, 'I found you the patient who's going to get you the gene,' " recalls Hurst, adding that she wasn't serious. But that's precisely what happened: the translocation in the boy disrupted a gene called *FOXP2*, which it turned out had been mutated in the 15 members of the KE family who exhibited severe problems.

When Monaco, Fisher, Hurst, and coworkers reported the convergent *FOXP2* findings in the October 4, 2001, issue of *Nature*, it made international headlines--and, more important, announced the start of a new era in speech and language research.

Even then, the scientists knew that *FOXP2* does not single-handedly wire the brain for language. In the grand theater of the genome, it is cast as a transcription factor, turning other genes on or off by telling them whether to transcribe their DNA into messenger RNA, which leads to the production of proteins. And *FOXP2* has a broad repertoire in embryonic development, playing critical roles in the formation of the lungs, heart, and intestines.

Yet *FOXP2* is clearly involved, too, in the molecular pathways behind speech and language. Clinicians in several countries have now reported patients with aberrant *FOXP2* genes and KE-like speech and language problems. Geschwind has taken some of the first steps in uncovering the connection between *FOXP2* and language. He and Fisher recently studied human fetal brains and neural-cell cultures to identify which genes the *FOXP2* protein turns on or off in the brain. They connected *FOXP2* to more than 200 genes that control the development of neurons, the release of neurotransmitters that send messages between nerves, and the changes in synapses that underlie learning and memory. Some of these genes will very likely turn out to be involved in speech and language. To sift this genetic river for the gems, Geschwind is zooming in on about 15 genes that also have ties to schizophrenia, as well as 34 genes to which *FOXP2* binds in two areas of the brain that other studies have shown are involved with language and speech.

To date, the discovery of *FOXP2*'s link to speech and language has yielded more questions than answers. But it has kicked open a door that neuroscientists had been knocking on for over a century.

The Knotty Mind

In 1861, Pierre Paul Broca came to a meeting of the Anthropological Society of Paris with another man's brain. Broca, a surgeon and neurologist who was the society's founder, had retrieved the brain from an unusual patient who had been hospitalized for

30 years. The patient was known as Tan because he would answer "Tan, tan" to any question put to him. He eventually lost the ability to speak altogether, although he understood almost everything he heard. Broca first met Tan only five days before his death, when he arrived in the surgery unit because of a massive, gangrenous infection. On autopsy, Broca found that Tan's brain contained a number of lesions, the most extensive and oldest of which was in the middle of the left frontal lobe. Broca asserted that this damage caused Tan's loss of speech.

Thirteen years later, the German physician Carl Wernicke described the brain of a stroke patient who could speak but had immense difficulty understanding what was said to him. Again, a lesion in the left hemisphere stood out, although it was farther back, near the intersection of the temporal and parietal lobes.

As Geschwind explains the importance of what are now known as Broca's and Wernicke's areas, he points out the cerebral real estate they occupy on the plastic brain he has finally assembled. Subsequent research has shown that both areas do play critical roles in speech and language. Though damage to either does not always cause problems, the neural circuitry for speech typically runs along the left Sylvian fissure--a sort of neural Grand Canyon that stretches from Broca's area to Wernicke's.

Geschwind has been captivated by this asymmetry, and by its relationship to handedness. Roughly 90 percent of us are right-handed, and nearly all righties depend on that left "perisylvian" region for speech and language. (About 40 percent of lefties instead rely on the right perisylvian region or use both hemispheres.) "There's some kind of benefit to the kind of processing that's going on in language--which is extremely rapid processing--to keep everything in one circuit in one hemisphere," he concludes.

The process that creates asymmetry often goes amiss in people with dyslexia, schizophrenia, or autism--all disorders with links to language problems. So Geschwind and others have set about hunting for genetic aberrations implicated in language disorders and for genes linked to differences in brain asymmetry, such as those related to handedness.

While the discovery of the mutation in *FOXP2* required great effort (and a dollop of luck), all told it involved analyzing the DNA of no more than 50 people. In contrast, no simple mutation of a single gene is likely to disrupt brain asymmetry or cause dyslexia, schizophrenia, or autism. Rather, these problems are caused by subtle aberrations in genes and networks of genes working in concert. That subtlety forces researchers to collect and sort through DNA from hundreds if not thousands of people. For example,

the Autism Genome Project, a large international collaboration in which Geschwind participates, performed an analysis of more than 1,400 families that have at least two members affected by autism-spectrum disorders. This massive study didn't isolate a single mutant gene, but it did find intriguing links between the disorders and missing or extra copies of a region of chromosome 11. Such variations can increase or decrease the amount of protein produced by genes, with unpredictable effects.

Geschwind also contributed to a study, led by Oxford's Clyde Francks, that revealed some of the intricate connections among language-related disorders, brain asymmetry, and handedness. The study began as a hunt for a gene that controls handedness in dyslexics. Previous reports had suggested that dyslexics are more likely to be left-handed and that left-handed people are more likely to have reduced asymmetry. Francks and his colleagues could not corroborate that suggestion, but they did find a region of chromosome 2 that seemed linked to left-handedness. They then examined the DNA of pairs of healthy left-handed brothers: the same linkage to chromosome 2 surfaced, evidence that a gene or genes in that region might influence handedness. Adding still more bizarre connections, the team performed a study of siblings with schizophrenia, which implicated the same region.

To find the gene or genes at the heart of this knot of links, the researchers compared the same region of chromosome 2 in healthy right-handed people, healthy left-handed people, and people with schizophrenia. They found four DNA differences that distinguished the schizophrenics from the mentally healthy lefties; the location of these variations led them to a gene called *LRRTM1*. Geschwind collaborated in the work that helped identify where in the human brain *LRRTM1* was turned on, or expressed: it probably helps shape forebrain structures and influences how neurons connect. He suspects that in early gestation, it also contributes to brain asymmetry.

Francks and his colleagues think that certain variants of *LRRTM1* somehow decrease production of the *LRRTM1* protein during fetal brain development. Presumably, reduced levels of *LRRTM1* could have contributed to reduced brain asymmetry, tilting the developmental scales toward left-handedness and schizophrenia--and potentially toward a variety of speech and language problems.

All this adds up to little more than a list of genes that may or may not be involved in creating speech and language: *FOXP2*; genes that *FOXP2* interacts with; genes with copy number anomalies implicated in autism; and an aberrant gene connected to schizophrenia and left-handedness. Moving from correlations between genes and disorders to knowledge of the neural circuitry that allows a human but not a chimp to ask, "To be, or

not to be?" requires researchers to find connections between seemingly disparate findings. To that end, Geschwind and others are turning to evolutionary studies that analyze these genes in other species and compare them with the human versions. Such studies may also provide clues to how humans evolved the capacity for language.

The Origin of Speech

Like songbirds, dolphins, whales, bats, elephants, and--of course--humans, monkeys and apes can learn sounds and use them to communicate. For many decades, researchers have attempted to decode such animal messages. They have also tried to teach chimpanzees, bonobos, gorillas, and orangutans to use symbols, lexigrams, and sign language, and a few poster apes like Koko, Washoe, and Kanzi have no small measure of fame thanks to PBS documentaries, magazine cover stories, and books about their communication skills. Some have even shown what appears to be a remarkable ability to understand spoken words.

Nevertheless, an impassable border separates our speech and language abilities from theirs. The best-trained apes can learn only a few hundred words. Most any human three-year-old has a larger vocabulary, and the average high-school graduate has a mental lexicon of about 60,000 words. Linguists and psychologists who have studied "talking apes," including researchers who have taught them to communicate, stress that the animals rarely combine even two words into a semantic whole and never utter the type of complex "recursive" sentence--like this one--that embeds one thought in another.

In the hope of beginning to explain this discrepancy, Geschwind investigated which genes are turned on in the brains of humans and in those of chimpanzees, our closest genetic relatives. He found hundreds of differences but had no way to determine which ones mattered--which were most significant in driving evolution and determining brain function. Overwhelmed, he turned to a mathematician friend at UCLA, Steve Horvath.

With Horvath's guidance, Geschwind and his grad student Michael Oldham arrived at a new way to approach the problem. Rather than looking at differences between individual genes, they analyzed differences between networks of genes expressed at the same time. Specifically, they looked at autopsied slices of human and chimp brains and compared these "coexpressed" genes in specific "modules," including the cerebral cortex, the cerebellum, and the primary visual cortex.

They found that within each module's networks, some genes served as hubs, connecting to many other genes. Diagrams of the networks look much like maps of airline routes, and both the human and chimp maps have a ridiculous number of hubs and spokes. But

the diagrams make it easy to see the most important genes--those at the hubs. And when the team took the human map of a module and removed all the chimp connections for the same module, only a few genes were left. It became startlingly clear not only which genes are uniquely human, but also which of those are most important.

This approach yielded insights that weren't possible with older techniques; simply comparing human and chimp expression of individual genes misses the vast majority of variation that takes place between groups of genes. Though no new connections between genes and language have emerged yet, Geschwind and his colleagues did find that most of the differences occurred in the cerebral cortex--the very part of the brain that expanded the most in humans, and in which Broca's and Wernicke's areas reside. Geschwind is hopeful that taking a broader view of not only the genome but also the transcriptome--the set of genes that are turned on at any given time--will lead to more insights into the genetics of language. "We need to understand the transcriptome in the same way we understand the genome," he says.

So far, however, the most intriguing and concrete genetic clues to the evolution of speech and language have emerged from simple, direct comparisons of animal and human versions of *FOXP2*. "*FOXP2* is paradigmatic," says Geschwind. "It's this beacon, and the first proof that this area of research might lead to great insights about human beings and evolution."

Soon after Fisher, Monaco, and their colleagues linked *FOXP2* to human speech and language, they teamed up with a leading evolutionary-biology group headed by Svante Pääbo at the Max Planck Institute in Leipzig, Germany. They found that the protein made by the *FOXP2* gene in chimps is virtually identical to that made in mice: just one amino acid differs between the two. Biologists believe that if proteins undergo little alteration over an evolutionary span of tens of millions of years, they must perform such essential functions that they simply cannot tolerate change. But two amino acids in human *FOXP2* differ from those in the chimp protein--a total of three changes from the mouse version. That the gene withstood such dramatic change in such a short time span (evolutionarily speaking) suggests that the change helped us survive--as the development of language surely did.

Then, in October 2007, Pääbo and coworkers published a jaw-dropping paper about *FOXP2* in Neanderthals, evolutionary relatives of modern humans that died out 30,000 years ago. The researchers isolated parts of the *FOXP2* gene from the bones of two Neanderthals. Although they have yet to sequence the entire gene, they found that Neanderthals and modern humans matched at the two critical spots that separate humans

and chimpanzees. Though often depicted as knuckleheads, our closest hominid relatives may have shared at least some of our capacity for speech and language. "There is no reason to think that Neanderthals did not have language as we do," says Pääbo. But he adds that the many unknown genes involved in language will eventually have to be found and looked at in Neanderthals.

Geschwind is continuing his hunt for those unknown genes, applying to his behavioral-genetics work the technique he developed to compare human and chimp gene expression. His lab is now doing the same sort of coexpression studies on brains from healthy humans and schizophrenics, which he hopes will uncover connections that are broken in schizophrenia and perhaps lead to still more genetic pathways related to speech and language. He hopes eventually to do similar analyses with autopsied brains from people who had autism-spectrum disorders.

So far, Geschwind and his colleagues have found what amount to some interesting genetic words that they've been able to string into a few sentences to explain the roots of speech and language. They can't yet tell a coherent story. Still, confidence is building that in the not-too-distant future, scientists will be able to write a lengthy book about how we evolved our phenomenal gift of gab, highlighting the critical suites of genes that make it possible. If they do, they could also find ways to correct disruptions to this network--disruptions that can leave people at a serious loss for words.

Upcoming Events

O'Reilly Emerging Technology Conference (http://conferences.oreilly.com/etech)
San Jose, CA
Monday, March 09, 2009 - Thursday, March 12, 2009
http://conferences.oreilly.com/etech (http://conferences.oreilly.com/etech)

South By Southwest (http://www.sxsw.com/interactive)
Austin, Texas
Friday, March 13, 2009 - Tuesday, March 17, 2009
http://www.sxsw.com/interactive (http://www.sxsw.com/interactive)