Linguists rush to study endangered languages, which may hold clues to whether grammar is innate or learned, and how speech influences thought

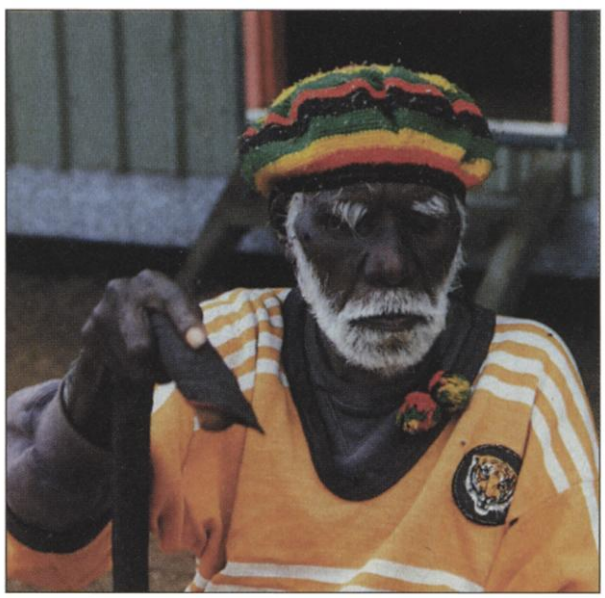
Learning the World's Languages—Before They Vanish

Pat Gabori loves to talk, and when he does, linguists listen. Gabori, who estimates his age at about 80, is the last native-born male speaker of Kayardild, an aboriginal language spoken only on Bentinck Island off north Australia. Once a skilled dugong hunter, now blind, Gabori has relayed a wealth of stories—and the grammar of his language—to linguist Nick Evans of the University of Melbourne in Victoria. Evans found Kayardild full of grammatical rarities. For example, whereas most languages change only the verb to indicate past or future tense, Kayardild marks tense on other words too, including nouns. Thus Gabori might say "The boy spear-ed the fish-ed." Or rather, he would probably say "Speared fish-ed the boy," as Kayardild also allows speakers to explode the traditional structure of phrases.

Whatever the word order, only one other known language marks nouns with tense—a sister language to Kayardild called Lardil that has only one fully competent speaker. "If you didn't know about these two languages, you would say this phenomenon is impossible," Evans says. And if every known language followed the rule that only verbs express tense, you might conclude, following the famous theory proposed by linguist Noam Chomsky almost 50 years ago, that the rule is an innate part of language, genetically programmed into the brain of every child. But to Evans and a vocal minority of other linguists, the possibilities seen in languages like Kayardild challenge the "universal" rules of grammar, and suggest that far more of grammar is learned and culturally variable.

Even the handful of linguists who study Kayardild and Lardil disagree on the significance of their idiosyncrasies, however. To Ken Hale of the Massachusetts Institute of Technology (MIT), who has studied Lardil for 30 years, the languages, far from undermining Chomsky's theory, actually subtly reinforce many of the "universals" of grammar. "Linguists rarely find things that challenge universals," he says.

Although linguists spar about the deeper meaning of Kayardild grammar, they agree on one thing: The data they need to settle linguistic arguments are fast disappearing. In Australia alone, aboriginal people spoke about 260 languages at the time of European contact in the late 1700s; today about 160 of those tongues are extinct, and only about 20 have a reasonable number of speakers. The world's 6 billion people speak approximately 6000 to 7000 languages, and most experts expect that at least half—and perhaps up to 90%—will disappear in the 21st century. War or scattering can demolish a linguistic community in a generation or two, so that even a language with thousands of speakers today may be at risk tomorrow; most linguists consider a language "endangered" when fewer and fewer children learn it.



The power of speech. Pat Gabori's words, in Kayardild, give linguists examples of rare grammatical structures.

CREDIT: P. JOHNSON

The loss of languages is not only a crisis for many communities, but also presents a major challenge for researchers intent on analyzing the structure of languages and how they convey meaning. Just as biologists study species to understand evolution, so linguists scrutinize grammars and vocabularies to understand what aspects of language are innate and what are learned. "We're losing our natural laboratory of variation, our Galápagos," says Steve Levinson of the Max Planck Institute for Psycholinguistics in Nijmegen, the Netherlands.

The causes of this global loss are many: wars, diasporas, education that emphasizes national languages, and assimilation into dominant cultures. To date, with less than 1000 fully described languages, "every language lost is a loss of clues as to how the human language faculty works," says Jerry Sadock, a linguist at the University of Chicago. This realization is prompting a small community of linguists to increase their fieldwork, compiling grammars and lexicons and putting spoken languages into writing, helping to preserve some languages as they learn more about them. But the work is slow going. "Describing a language is a 10- to 15-year job," says Evans, who wrote a 600-page grammar of Kayardild.

Linguists argue passionately that each language is precious. The particulars of a language "encapsulate a long history of people in an ecology, a way of living and a way of thinking," Levinson says. Different languages reflect peoples' perception of the world, capture their prehistory (see sidebar on p. 1158), and may subtly shape thought itself. "When these languages die, we'll lose these glimpses into the capacities of the human mind," says linguist Marianne Mithun of the University of California, Santa Barbara.

Universal questions

Whether grammar—the structure of language—is innate or learned is the longest running battle in modern linguistics. It was fired up by Chomsky back in the 1950s and hasn't calmed down since. He proposed that universal grammar is a manifestation of linguistic ability that is hard-wired into the human brain, a genetic endowment that allows every child to master language with ease and also dramatically restricts the types of language that are possible. Ever since, lin-

guists have sought to test his ideas in the world's languages. Today, most researchers agree that at least some grammar, such as rules for how questions may be framed, is universal and therefore probably innate. But they clash fiercely over just how extensive this "universal grammar" is.

For example, one of the most basic aspects of grammar concerns sentence structure—the order in which the subject (S), verb (V), and object (O) appear. Most languages are either SVO (like English), SOV (like Japanese), or VSO (like Irish Gaelic). Linguists had thought that other orders were prohibited. But OVS does occur—in fewer than 1% of the world's languages, all of them endangered, notes Norvin Richards of MIT. One such language is Hixkaryana, spoken by some 300 people on a tributary of the Amazon River in Brazil. "If linguists had waited another couple of decades, languages with this construction would all be dead, and we would say it is not possible," says Richards.

Another recently appreciated realm of variability concerns affixes such as prefixes and suffixes. Affixes added to the beginning or the end of a word typically change its meaning in particular ways, such as *pre*natal or sturdi*ness*. Some linguists have thought that all languages share a set of affix meanings that are part of universal grammar, and that there is always a clear distinction between stem words and affixes. But the Yup'ik language, spoken by about 10,000 people in Alaska, breaks with that presumption, according to work published by Mithun last year.

Mithun points out that Yup'ik has suffixes with meanings like "eat" and "hunt" that in English would seem like stems rather than add-ons. When speakers add the suffix "hunt" to the word for seal, it means "seal hunt," added to "egg," it means "egg hunt." But in addition to a suffix for hunt, Yup'ik also has a stem for hunt, so that there are different ways to say "seal hunt," each with its own shade of meaning. "These kinds of differences offer speakers tremendous rhetorical alternatives," she says.

Whereas some Yup'ik suffixes relate to subsistence or the environment, others refer to ways of being or behavior. There are suffixes that mean "willfully" and "secretively." One suffix means "to be inept at": Added to the verb "sleep," it yields "to have insomnia." Another suffix means, "finally, after desiring to do so, but being prevented by circumstance," and can be added to verbs such as "go." Not only does this complex suffix system challenge the idea that affixes play a restricted role in carrying meaning, it preserves a record of the cultural transmission that creates grammars, says Mithun.

Another linchpin of universal grammar is the distinction between nouns, verbs, and adjectives. Some languages have thousands of verbs and some less than a dozen, but few wholly dispense with that distinction. But David Gil, a linguist at the Max Planck Institute for Evolutionary Anthropology in Leipzig, Germany, says that Riau, a dialect of Indonesian, does just that, allowing nearly any word to function as a verb, a noun, or an adjective, depending on the context. The same word can be used to mean "eat" or "food," and any word can appear anywhere in the sentence. "The key to meaning lies in the larger context of the sentence," explains Bernard Comrie, director of the institute's linguistics department.

Other languages retain nouns and verbs, but bend the grammatical structures thought to govern their use, says Sadock. For example, in most languages the verb agrees with the subject. But in Aleut—now spoken by only about 100 people in the Aleutian Islands off the coast of Alaska and heard in daily conversation in just a single village, Atka—the verb can agree not only with the subject or the object, but also with the possessor of the subject or object. Thus instead of saying, "Their house is big," an Aleut speaker says, "Their house are big"—something Sadock calls "virtually unprecedented." Yet, he adds, "the fact that Aleut is so seriously endangered means that we might never know its true genius."

Sadock, who chairs the Committee on Endangered Languages and Their Preservation for the Linguistic Society of America, says that such "wild differences" in grammar as seen in Aleut and Yup'ik "challenge our notions about what kind of language the human mind can construct." He thinks that the realm of universal grammar has been exaggerated and that many grammatical rules are culturally variable. "Most grammar arises out of human cognition and interaction with people and objects, rather than being an innate hard-wired system that has nothing to do with daily life," agrees Dan Slobin, a psycholinguist at the University of California, Berkeley.

But others argue that although various languages may be unusual, the differences are superficial. For instance, Hale and MIT colleague Richards think that Lardil and Kayardild actually conform to universal grammar rather than breaking the rules. Although these languages allow speakers to jumble the words of a sentence, the words inside a phrase are first modified by tense markers, notes Richards. Thus, he argues, traces of the phrase structure remain.

To take an even more basic example, linguist Mark Baker of Rutgers University in New Brunswick, New Jersey, makes the case that many languages reported to lack distinctions between parts of speech such as nouns and adjectives do indeed make such distinctions—if you know where to look. "I can show that even those languages have adjectives," he says. For example, Nunggubuyu, an endangered aboriginal language in Australia, has been described as using nouns and adjectives interchangeably. But Baker reports in a monograph to be published by Cambridge University Press that the elusive distinction between these parts of speech becomes apparent in a particular way of forming compound words called noun incorporation.

Nunggubuyu speakers can say, "I bought meat" or "I meatbought," incorporating the noun into the verb. They can also say "I

A SAMPLING OF ENDANGERED LANGUAGES

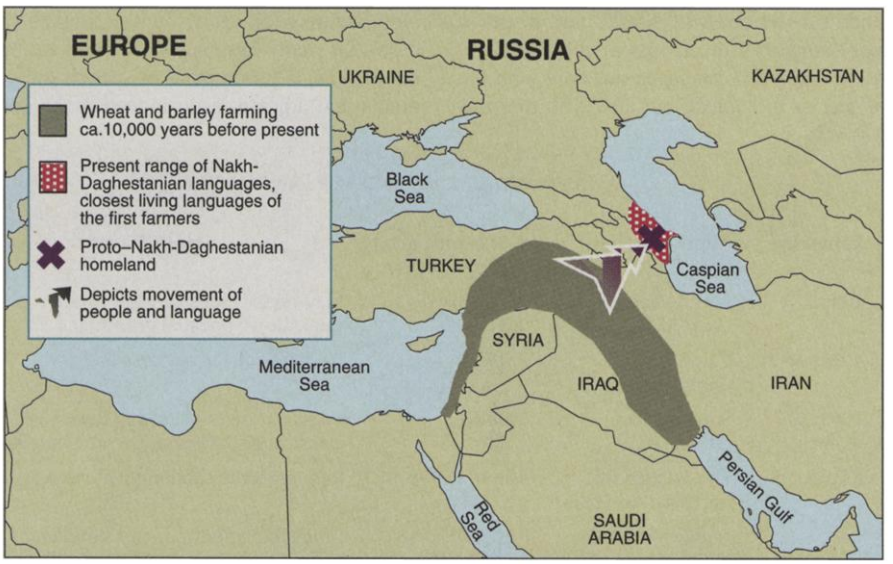
Language	Where spoken	No. of ethnic group/ No. of speakers	Scientific interest
Ahlon	Southwest Togo, West Africa	2000–3000/ unknown	May reveal clues to West African migrations; unusual grammar
Cambap	Cameroon, Central Africa	250/30	Evidence of the region's history
Harsusi	South-central Oman	600/600	Clues to the relationship between classical and modern Semitic languages
Kayardild	Two islands off northern Australia	About 150/4–6	Rare features challenge universal grammar
Leco	Andes Mountains in Bolivia	80/20	Clues to ancient contact between Amazonian and mountain-dwelling people
Mohawk	New York, Quebec, and Toronto	Tens of thousands/ 1000–2000	Complex word constructions add insights into universal grammar
Yup'ik	Southwestern Alaska	20,000/10,000	Unusual grammar challenges universal rules
Yélî dnye	Rossel Island, Papua New Guinea	4000/4000	Unique sounds; vocabulary that upsets universal color terminology

Peering Into the Past, With Words

Prehistorians typically rely on stones, bones, and DNA to piece together the past, but linguists argue that words preserve history too. Two new studies, both based on endangered languages, offer new insights into the identity of mysterious ancient peoples, from the first farmers to early inhabitants of the British Isles.

Archaeologists have long known that some 10,000 years ago, ancient people in Mesopotamia discovered farming, raising sheep, cattle, wheat, and barley. And researchers knew that by 8000 years ago agriculture had spread north to the Caucasus Mountains. But they had little inkling of whether traces of this first farming culture lived on in any particular culture today. People have migrated extensively through the region over the millennia, and there's no continuous archaeological record of any single culture. Linguistically, most languages in the region and in the Fertile Crescent itself are relatively recent arrivals from elsewhere.

Now, however, linguist Johanna Nichols of the University of California, Berkeley, has used language to connect modern people of the Caucasus region to the ancient farmers of the Fertile Crescent. She analyzed the Nakh-Daghestanian linguistic family, which today includes Chechen, Ingush, and Batsbi on the Nakh side and some 24 languages on the Daghestanian side; all are spoken in parts of Russia (such as Chechnya), Georgia, and Azerbaijan.



Farm talk. Languages heard near the Caspian Sea today trace their ancestry back to the first farmers of the Fertile Crescent.

Nichols had previously established the family tree of Nakh-Daghestanian by analyzing similarities in the related languages much the way biologists create a phylogeny of species. She found that three languages converge at the very base of the tree. Today, speakers of all three live side by side in the southeastern foothills of the Caucasus Mountains, suggesting that this was the homeland of the ancestral language—on the very fringes of the Fertile Crescent. To get a rough estimate of when the language arose, Nichols used a linguistic method that assumes a semiregular rate of vocabulary loss per 1000 years, and she dated the ancestral language to about 8000 years ago.

Nichols also found that the ancestral language contains a host of words for farming. The Chechen words *muq* (barley), *stu* (bull), and *tkha* (wool), for example, all have closely related forms in the earliest branches of Daghestanian, as do words for pear, apple, dairy product, and oxen yoke—all elements of the farming package developed in the Fertile Crescent. Thus location, time, and vocabulary all suggest that the farmers of the region were proto–Nakh-Daghestanians. "The Nakh-Daghestanian languages are the closest thing we have to a direct continuation of the cultural and linguistic community that gave rise to Western civilization," Nichols says.

Population geneticist Henry Harpending of the University of Utah, Salt Lake City, has just begun the job of unraveling the genetic ancestry of Daghestanian speakers and is impressed with Nichols's work. "For years I wished linguists would get in the game. Nichols sure is."

Nichols is now reconstructing the ancestral language, hoping for more clues to the culture of these early farmers. But she has to work fast, for the three Nakh languages are vanishing. Although there are still about 900,000 Chechen speakers left, the other two tongues have fewer speakers, and all three are being eroded by war, economic chaos, and Russian educational practices, Nichols says.

More than 3200 kilometers away, another linguist is mining Celtic languages—which are also all considered endangered—for clues to the inhabitants of the early British Isles. Artifacts show that the islands were occupied long before Celts from the European continent made landfall about 700 B.C. But mysteries remain as to their identity.

So Orin Gensler of the Max Planck Institute for Evolutionary Anthropology in Leipzig, Germany, analyzed Celtic languages, including Irish Gaelic, Scottish Gaelic, Welsh, and Breton. Once prevalent throughout Europe, these languages are now spoken only in the British Isles and Brittany in France. Linguists have noted surprising grammatical differences between Celtic languages and related languages such as French, while at the same time seeing striking resemblances between Celtic and Afro-Asiatic languages spoken for millennia across a swath of coastal Northern Africa and the Near East.

In a forthcoming monograph, Gensler studied 20 grammatical features found in both Celtic and Afro-Asiatic languages. He sought these linguistic traits in 85 unrelated languages from around the world, reasoning that if the features were widespread, their appearance in both Celtic and Afro-Asiatic languages might be mere coincidence. But if the shared features are rare, coincidence is unlikely. Overall, Gensler found that about half the shared features are rare elsewhere. "I think the case against coincidence is about as good as it could be," he says.

And a closer look at a number of features, including word order, offers a provocative theory for just how the Celtic islanders acquired these linguistic traits. In Gaelic and Welsh—and many Afro-Asiatic languages—the standard sentence structure is verb-subject-object. But Celtic languages spoken in Continental Europe in antiquity have the verb in the final or middle position. The best explanation for the shift to verb-initial order, says Gensler, is that when Celtic speakers made landfall on the British Isles, Afro-Asiatic speakers were already there. As these people learned Celtic, they perpetuated aspects of their own grammar into the new language.

Although others are interested in Gensler's idea, so far "there is no significant northwest African genetic signature ... in Celtic populations," says Peter Underhill, a molecular geneticist at Stanford University in California. But in this instance, he adds, the linguists may be ahead of the geneticists, for researchers need more genetic markers before they can confirm or refute Gensler's idea.

–B.W.

bought big" (akin to "I bought the big one"), but they cannot say "I bigbought," incorporating the adjective into the verb. "If big had the same status as meat, it should be able to incorporate the same way," Baker says, arguing that this is indeed a distinction between the two word categories.

Baker recently applied this kind of analysis to a range of extremely different languages all thought to lack distinctions in parts of speech, including Mohawk and Salish, two endangered Native American languages; Edo, spoken in Nigeria; and Chichewa, spoken in Malawi. His results show that all these languages have noun, verb, and adjective categories, although they may be subtly expressed. "In any language where we can frame these questions precisely enough to measure, we can find a lot that's universal," Baker says.

As the argument continues, both sides realize that their evidence is vanishing. "In a few years everyone will speak English, Mandarin, and Spanish, and the similarities among languages will exist by accident, not because they reflect limits of human cognition," says Baker. Adds Mithun: "Without the variability, proposals about universals have been and will be hopelessly naïve."

From words to thoughts

In addition to adding much-needed data to the question of universal grammar, the diversity of languages also offers insight into the relationship between language and thought—how speakers perceive and understand the world. "Put bluntly, does the language you speak affect the way you think?" Levinson asks.

He has approached this by studying how unwritten, often endangered languages express spatial concepts. Many cognitive scientists have presumed that all languages express spatial concepts in similar ways, because thinking spatially is a necessity for higher animals and therefore is likely to be hard-wired into the human brain. "This is the very last area in which we would expect to find significant cultural variation," Levinson says. But find it he did.

Many languages, including English, express spatial concepts using relative coordinates established through the planes of the body, such as left-right and front-back. But Levinson found languages that use very different systems. Guugu Yimithirr, an endangered Australian language spoken by fewer than 800 people, uses a fixed environmental system of four named directions that resemble north, south, east, and west. Speakers modify the four words to yield some 50 terms that indicate such things as motion toward or away from a direction. They use the same terminology to describe both landscape and small-scale space, for example: "The school is to the west of the river" and "There's an ant on your eastern leg."

The Guugu Yimithirr terminology reflects an entirely different way of conceptualizing a scene, says Levinson. It requires laying out all memories in terms of the four directions and continually running a mental compass and a positioning system.

Levinson has now investigated a similar phenomenon in the Mayan language Tzeltal, in work in press and co-authored with Penelope Brown, also of the Max Planck Institute for Psycholinguistics. Brown gathered 600 hours of videotape of 15 Tzeltal-speaking children performing various tasks and found that children as young as 4 years old have mastered the positioning system. Children were asked to describe the arrangement of toys on a table, then turn 180 degrees and describe an identical arrangement arrayed in front of them. English-speaking children rotate their coordinate system as they turn—left becomes right. But for Tzeltal-speaking children, north was always north and south remained south.

By word of mouth. Children learn an unusual directional system along with Tzeltal.

Levinson concludes that even spatial thinking is learned, not innate. Rather than starting from a biologically set concept of space, children quickly learn the system used in their culture. Concurs psycholinguist Slobin: "The results show flexibility for how we can organize spatial concepts for talking and probably thinking."

The art of language

As linguists uncover such cultural differences, many find themselves fascinated with the diverse linguistic systems that humans are capable of creating. Each language preserves a society's history, culture, and knowledge and is itself something akin to art, they say. "Language is one of the most intimate parts of culture," says Mithun.

Languages can gracefully encapsulate a society's values, for example. Mayali has a special set of words that simultaneously indicates the relationship of both the speaker and the hearer to the person being discussed. A speaker uses a different term for mother when speaking to her grandmother, her brother, or her husband. The terms, which are so complicated that Mayali speakers don't learn them until adulthood, "represent the ability to simultaneously take two points of view. Using them forces you to pay attention to where everyone fits into the kinship system and to make incredibly complex calculations," Evans says.

Mayali culture values proper use of the kinship terms, which indicate consideration for others by subtly factoring in their point of view. But this feature is being lost as younger speakers either adopt English or speak a simplified form of Mayali.

Although linguists delight in such examples, all this richness means that the task of understanding and preserving languages is huge. "Most of the currently receding languages will disappear without even being recorded," says Matthias Brenzinger of the University of Cologne in Germany. Last February, he convened linguistics experts from all regions of the world to begin the difficult task of rating the degree of danger faced by the world's languages, to help linguists focus their efforts. Dozens of researchers are already working with native speakers to preserve languages. Hale and Richards compiled a dictionary with Lardil speakers, for example, and Richards plans to put together readers for teaching.

Such efforts can be successful: Back in the early 1970s, Mithun worked with the Mohawk community in Quebec, Canada, to codify Mohawk spelling conventions. That text and others later became part of a public school curriculum, and now a Quebec elementary school teaches Mohawk to hundreds of students; older students can even study Mohawk linguistics at McGill University.

Success stories are hard won, however. And as researchers contemplate the impending loss of the world's linguistic diversity, they see their opportunities for gaining insights into human linguistic development slipping away. "If you understand what the constraints of a possible language are," says Levinson, "you would understand a fundamental part of what it is to be human."

–BERNICE WUETHRICH

Bernice Wuethrich is a science writer in Washington, D.C.

CREDIT: S. LEVINSON