





# Searching for Words

*What's at stake  
when languages are lost?*

By Karen McCally '02 (PhD)

It's Friday afternoon, and Farai Mutohori '19 arrives at the crowded entrance to the Wilson Commons Starbucks. Dressed in a white Oxford shirt and tie, the freshman electrical and computer engineering major from Masvingo, Zimbabwe, is taking time out from his usual routine of classes, workshops, and labs to talk about what some of his American friends consider to be one of his more unusual aptitudes. Mutohori speaks five languages in addition to English. "I speak Zulu, Sotho, Ndebele, Shona, and a little bit of Venda," he says.

As exotic as such multilingualism might appear in many contexts, in Mutohori's native region, it's how things are. "It's typical. People in Zimbabwe don't speak one language," he says. "There was a point where—it was very strange for me growing up as a kid. I would use, maybe, three languages in a sentence at some point. But as you go along, you get used to all of them, and you become fluent."

Mutohori's first language is Ndebele. Ndebele is not, according to the definition accepted widely among linguists, an endangered language. But Mutohori feels its position is precarious. When he was growing up, Zimbabwe was embroiled in civil war, and his Ndebele community fell under the domination of Shona-speaking ethnic rivals. His father responded by adopting a Shona surname. Although Ndebele remains, along with Shona and English, an official language of Zimbabwe, Shona domination continues. "Our president is Shona. In politics, Shonas make more of the decisions," Mutohori says.

There are regions of the globe, including much of sub-Saharan Africa, in which so many languages live within such a small geographic area that people there, like Mutohori, may engage in conversations in multiple languages within a single day, depending on the people with whom they're speaking. For linguists, such regions are laboratories for the study of the human capacity to absorb language and to shape and reshape it collectively over time.

But linguists worry there will be fewer such places in the future. The fate Mutohori fears for Ndebele is a present reality for many languages around the globe. There are slightly more than 7,000 distinct languages in use in the world today, according to the *Ethnologue*, a definitive source of information about languages around the world researched and compiled by the Dallas, Texas, based organization SIL International. They come in

**THE MULTILINGUAL NORM:** Farai Mutohori '19 came of age fluent in five languages, and conversant in a sixth. Multilingualism on that scale is still the norm in parts of the world such as his native Zimbabwe. But linguists wonder how long that will be the case.

striking varieties, spanning more than 140 language families, and including well over 100 deaf sign languages.

“We have an enormous world of variability out there,” says Joyce McDonough, a professor of linguistics and brain and cognitive sciences at Rochester. “When I look at a language, I am always astounded at the way they’re put together. They’re remarkable. They’re beautiful. And they’re disappearing.”

The *Ethnologue* groups languages on the basis of 12 categories of language status, from thriving, national languages to languages that no longer have any known speakers. Estimates among linguists range considerably, but the most conservative estimates predict that by the end of this century, half of the languages now in use will have vanished.

The prospect of widespread language loss confronts scholars with an irony. Humanity may have more access to education, more sophisticated tools for learning, and more cross-cultural contact than ever before. But at the same time, greater mobility and cross-cultural contact can place the spectacular range of human thought and invention at risk.

McDonough, who has taught at Rochester since 1997, specializes in North American languages, in particular, Navajo and the related Athabaskan or Dene languages. She’s been working for several years on a study of the remote Dene-speaking communities in the Mackenzie River Basin of northwestern Canada. When a language disappears, she says, a culture disappears with it, as does knowledge that the community has built up over centuries. When a language is lost,

entire oral traditions of storytelling can be lost with it. Highly specialized knowledge of local plant and animal life—knowledge built up over many years and held exclusively, in many cases, within endangered language communities—also vanishes.

There’s an additional problem that weighs heavily on scholars of language. To illustrate, McDonough draws an analogy between language and plants. “Imagine if we only had access to a few different types of plants and we had to describe the phenomena of plant life on the basis of that,” she says. “With the enormous amount of variability out there, we would be missing a lot.” Yet linguistic research has been based disproportionately on

**CRACKING THE LINGUISTIC CODE: Nadine Grimm (top), a specialist in Bantu languages who has conducted field research most recently in Cameroon, oversees a new program in language documentation. Joyce McDonough (bottom), a scholar with a long track record of work with the Navajo and the Dene language communities in northwestern Canada, was instrumental in the program’s formation.**

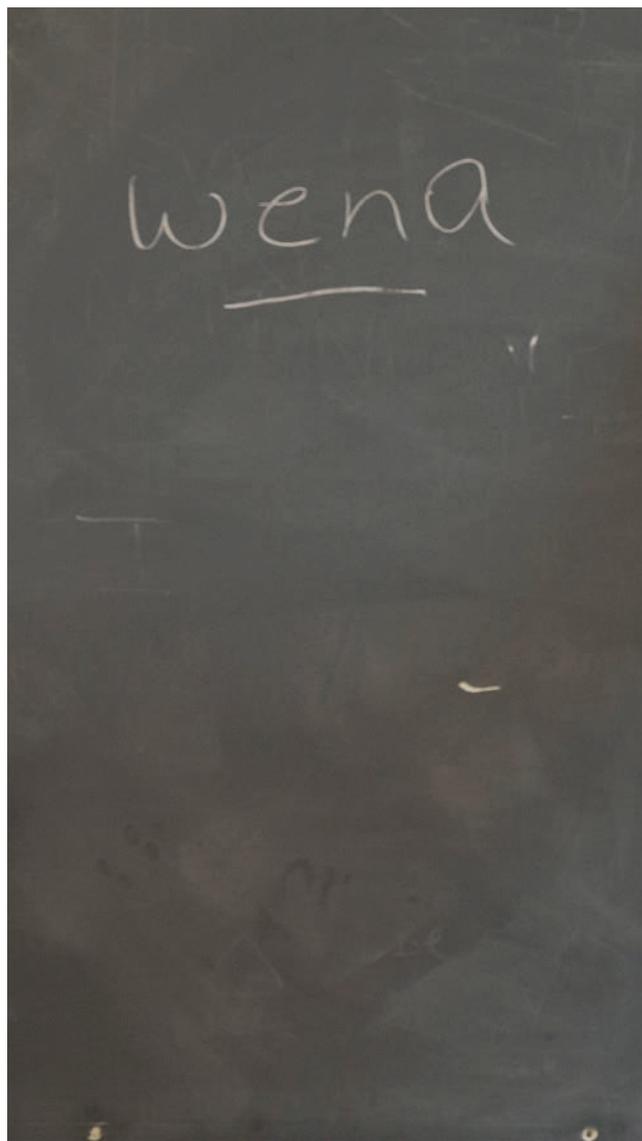
Indo-European languages. “The assumptions we make about what can and cannot be part of language or expressed by language are constantly confounded by new discoveries in under-resourced languages,” she and a coauthor wrote in an article in the *Annual Review of Linguistics* in 2015.

That’s a point echoed by scholars whose work is not directly related to language documentation, but relies on it nonetheless. Florian Jaeger, an associate professor of brain and cognitive sciences, of linguistics, and of computer science at Rochester, develops computational models of human language processing. The integrity of his research, and that of colleagues who develop similar models of language acquisition, rests on the availability of diverse data sets.

“We can think of our brain as a filter,” says Jaeger, whose current research draws on Yucatec Maya, one of approximately 70 Mayan languages spoken in southern Mexico and Central America today. “You could come up with some linguistic innovation right now, but if I can’t make sense out of it, it’s not going to survive. So what is the structure of that filter? We will only figure that out if we have a good idea of the variability that’s actually out there.”



**I**n 1991, recognizing the pace of language loss, the Linguistic Society of America issued a call to action at a symposium titled “Endangered Languages and their Preservation.” The type of language loss occurring in the modern era, participants argued, was not to be confused with the evolution of language that has taken place among virtually all languages throughout human history. Evolution involves the merging of features of multiple





languages, usually within particular geographic areas, and at a relatively slow pace; in modern times, given much greater human mobility, language communities have been overwhelmed and subsumed by their more powerful counterparts with unprecedented speed.

For the past several years, faculty in Rochester's linguistics department have responded, offering seminars to train students to work with endangered and understudied languages. Beginning this fall, the department will ramp up its efforts with a new master's degree program specifically in language documentation.

The program will be overseen by Nadine Grimm, a specialist in African languages, particularly those in the Bantu family, who joined Rochester's linguistics department last year.

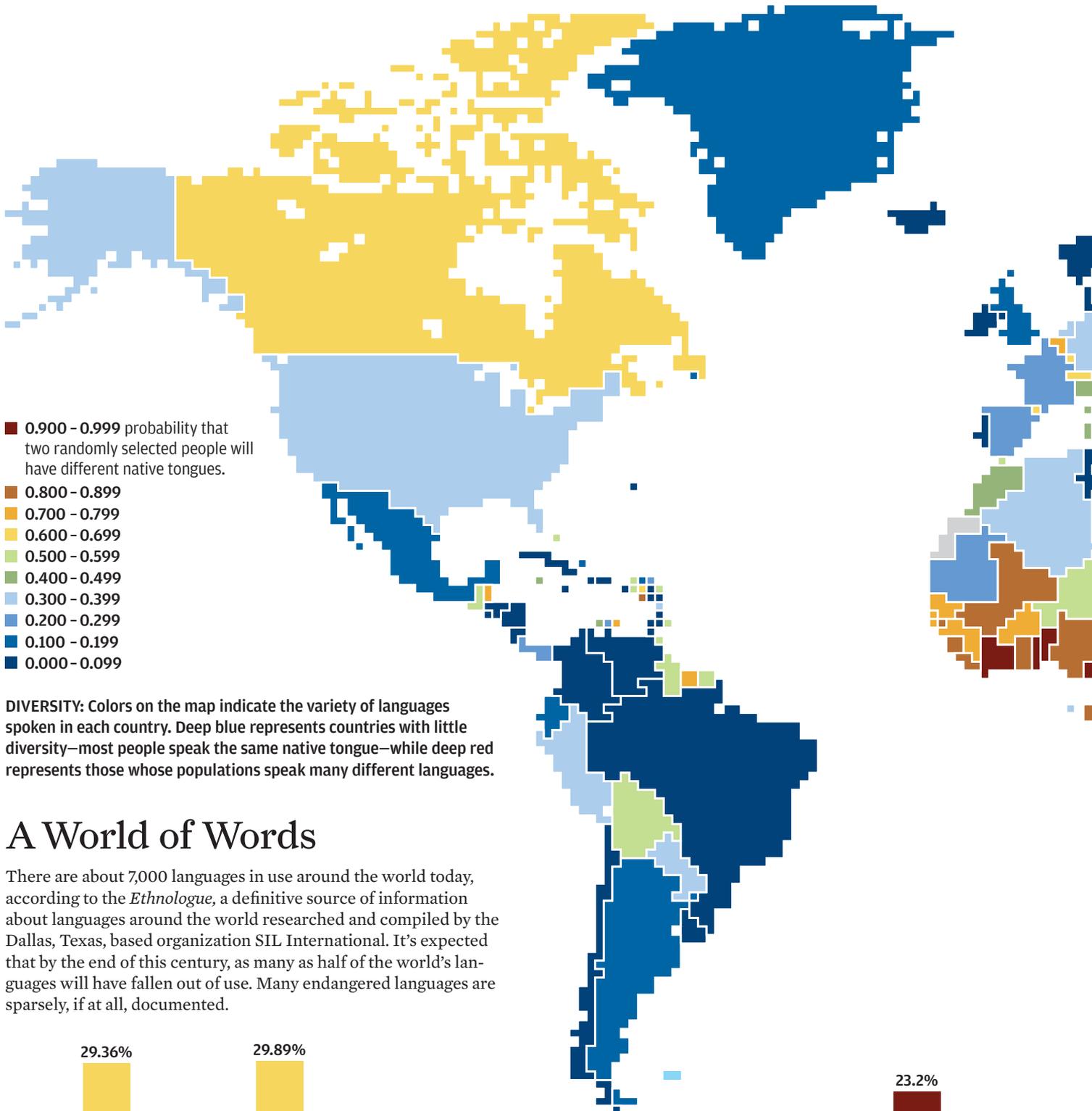
Last fall, Grimm reached out to Mutonhori, looking for native speakers of languages that would be completely unfamiliar to the students registered for her spring seminar, Field Methods in Linguistic Description. She offered Mutonhori a job as linguistic consultant for the course, and he accepted.

The seminar, which has been taught over the years by several members of the department, is structured as a form of role-playing. Students play the role of a team of linguists documenting and analyzing an unfamiliar language. A hired linguistic consultant—in this case, Mutonhori—responds to students' questions in a process called elicitation. Selecting a language has become trickier as Rochester's global student population has increased, but over the years, the course has been taught with speakers of a variety of languages, including Turkish, Albanian, Finnish, Ukrainian, Macedonian, Bengali, and Swahili.

Linguistic description is not the same thing as documentation, Grimm notes. But the two practices are closely related, and she adopts the view that skills in documentation—involving the collection, annotation, and storage of primary audio and video data—require skills in description, which is concerned with the rules and patterns underlying elements such as grammar, structure, sound, and vocabulary. Despite the course's title, by the end, students will have acquired important skills in data collection.

On a typical Thursday afternoon in the seminar, the students arrive with pages of charts. Some of the charts include long columns of English nouns, and adjacent to that column, blank columns for Ndebele translations in singular and plural forms, in forms accompanied by possessive pronouns and adjectives. Another chart lists numbers, beginning with zero to 20, then jumping to 100, 1,000, 10,000, and so on.

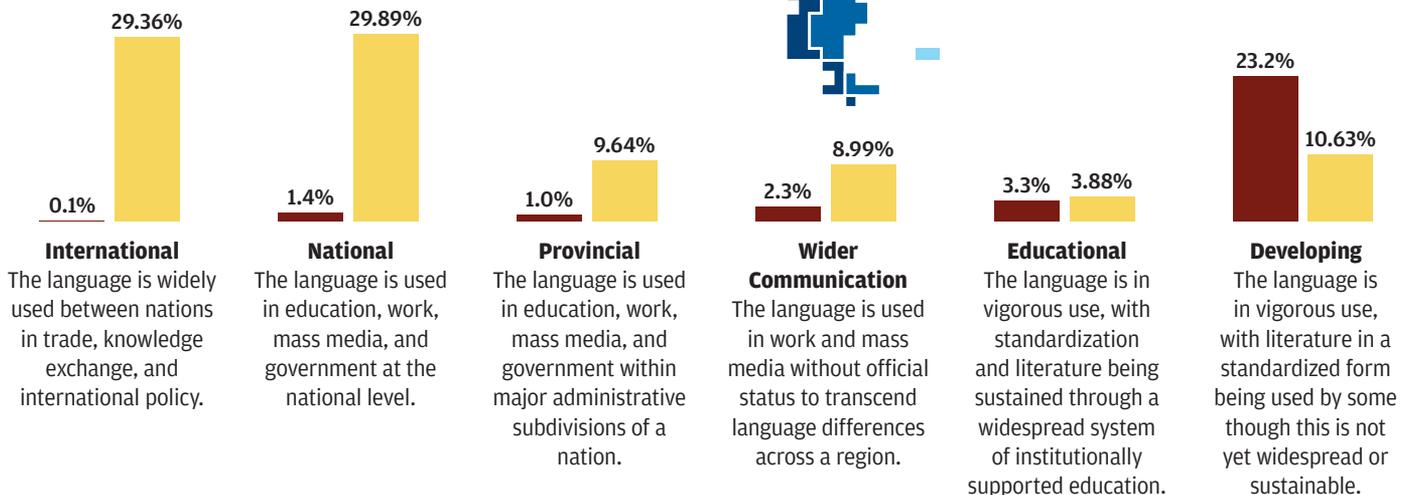
Rebecca Everson '16, a dual major in linguistics and computer science from Rochester, Michigan, says the students work with nouns long before they tackle verbs. "Nouns are the building blocks and verbs are the connectors," she says. "We begin with, (*Continued on Page 34*)

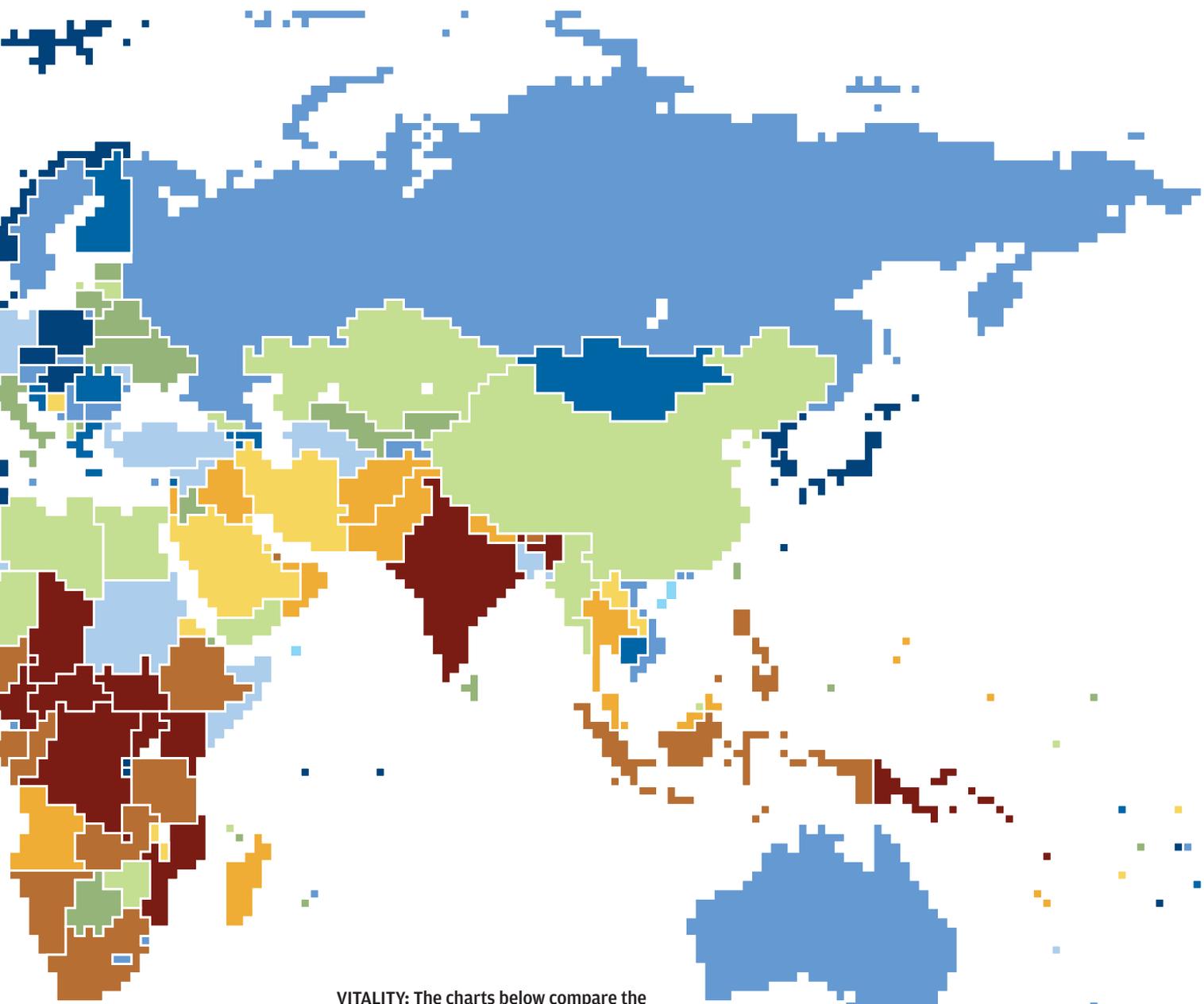


**DIVERSITY:** Colors on the map indicate the variety of languages spoken in each country. Deep blue represents countries with little diversity—most people speak the same native tongue—while deep red represents those whose populations speak many different languages.

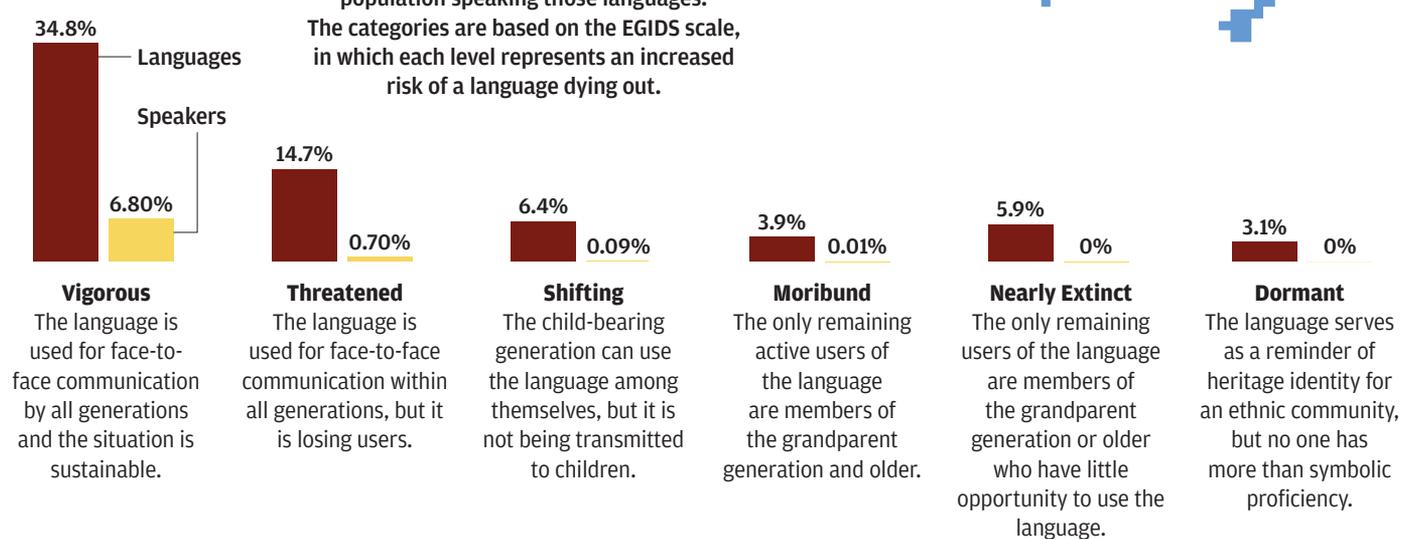
## A World of Words

There are about 7,000 languages in use around the world today, according to the *Ethnologue*, a definitive source of information about languages around the world researched and compiled by the Dallas, Texas, based organization SIL International. It's expected that by the end of this century, as many as half of the world's languages will have fallen out of use. Many endangered languages are sparsely, if at all, documented.





**VITALITY:** The charts below compare the share of the world's languages in each of 12 categories to the share of the world's population speaking those languages. The categories are based on the EGIDS scale, in which each level represents an increased risk of a language dying out.





(Continued from Page 31) for example, ‘apple,’ then ‘apples,’ then ‘two apples,’ then ‘your apples,’ ‘my apples,’ and eventually, ‘I dropped the apple.’”

The students take turns eliciting translations from Mutonhori, recording his responses in oral as well as in written form, using the International Phonetic Alphabet, or IPA. As they listen and write, they take note of patterns that emerge. When they get to numbers, Mutonhori explains the way in which English, which along with Ndebele and Shona, is an official language in Zimbabwe, has seeped into Ndebele.

**INTERDISCIPLINARY:** Rebecca Everson '16 (center), a dual major in linguistics and computer science, sees opportunities to improve the tools of language documentation, such as software often written by computer scientists without a strong linguistics background.

If you continued with the patterns that govern teens, then multiples of ten, and hundreds, he tells the class, “You could spend minutes just trying to say two-thousand.” Grimm interjects to explain that the morphology of numerals is consistent across hundreds of Bantu languages. Although all were originally founded on a base-five system, derived from the human hand, Bantu language speakers have now shifted to a decimal system, “a trend that is observed all over the world,” she says.

Mutonhori says that working with the class has led him to appreciate Ndebele in a new way. “Ndebele is a language I know from the back of my hand,” he says. “When I speak in front of the students, they’ll come up with things that I already know, but by looking at patterns. The way they have learned to come up

with what makes sense—it’s very interesting.”

Mutonhori’s love for the language is rooted in personal history and heritage, but also in some of the features of Ndebele itself. “It’s a very poetic and expressive language,” he says.

His enthusiasm is palpable. “He’s really excited about his language,” Everson says. “That helps us be really excited about asking the questions about it. He’s always giving us new information, and sometimes we forget to press ‘record,’ because we’re so absorbed in what he’s trying to convey.”

Outside of class, the students enter written and oral data into computers using a specialized linguistic software called Praat, which helps them to analyze patterns in sound, the composition of words, sentence structure, and other linguistic elements.

In the linguistics department lounge, shortly before class, Everson pours over data with Emy Lin '16, a senior from Glastonsbury, Connecticut, who, like Everson, is a dual major in linguistics and computer science. They talk about what they’ve found, and some of the challenges that they and the other students have faced. Working on a laptop, Lin calls up a spectrogram that Praat has

produced from her recording data and shows how the spectrogram will help her to elicit phonetic properties of Ndebele. Everson points to a series of blue lines swooping upward and downward across a series of spectrograms. She explains the way in which the lines illustrate tone.

“Tone matters in this language, which is something that took us a little while to figure out,” says Everson. Ndebele has a sing-song quality, and when the students want to clarify tone, Lin says, “we ask Farai to hum the words.”

There are other aspects to Ndebele that pose challenges. It’s a so-called “click” language, in which a series of clicking sounds serve as consonants. And while the students may have encountered languages with two, or maybe three, “genders,” Ndebele, they’ve discovered, has many more than that.

To help solve such puzzles, the students are permitted to consult sources on Bantu languages. But by the end of the semester, they will have constructed an overview of the basic features of Ndebele based almost entirely on the data they’ve collected and analyzed in their work with Mutohori.

“It’s not a lot of time, but they manage,” says Grimm. McDonough, who has taught the language description seminar as well, says the students often begin “in a state of shock. How can we possibly do this?, they wonder. Our job is to help them understand that what they’ve learned here over the past four years actually enables them to do this task.”



**A**mong the more difficult challenges of training students in language documentation is preparing them to work in indigenous communities. As for this aspect of documentation, Grimm comes with a sterling credential. As a doctoral student at Humboldt University in Berlin, she won a grant from the Volkswagen Foundation’s Documentation of Endangered Languages program, or DoBeS, as a participant in a fieldwork project in Cameroon that took place over a five-year period.

“It’s a very difficult grant to get,” says McDonough. According to the senior scholar, the hiring of Grimm is “a real coup. We’re thrilled to have Nadine here because of the training she received under the DoBeS project.”

From 2010 to 2015, Grimm spent months at a time in southern Cameroon as part of a team of three linguists, local French- and Gyeli-speaking interpreters, and members of the Bakola-Bagyeli community of hunter gatherers. Her research, focused on Gyeli grammar, entailed considerable documentation of the language, which is spoken by an estimated 4,200 people, and is considered threatened. The Volkswagen Foundation filmed a documentary of the work of Grimm and her colleagues, as part of their project. In it, Grimm reflected on the dynamics of economic and social change, and language evolution and loss.

In southern Cameroon, much like in Mutohori’s native Zimbabwe, multilingualism is the norm. “It’s a setting where you have eight or nine languages in one area, and where everyone speaks several languages. You meet one person, you speak one language, then you meet someone else, you speak a different language,” she says. There’s also what’s called “code switching,” or shifting from one language to another and back again in a single conversation.

The jostling of so many language communities makes southern Cameroon an especially valuable site for research on the dynamics of language evolution and change. Language loss often conjures up images of Western domination. But in many parts of the world, such as southern Cameroon, language loss proceeds gradually, as one ethnic group is displaced and subsumed by another. Cameroon is undergoing economic changes that have threatened the livelihood of the Bakola-Bagyeli. In the coastal city of Kribi, a large port is under construction. As forests are cleared for the project, the Bakola-Bagyeli have been displaced and compensated with jobs and modern housing in Kribi. “In the course of this massive change,” Grimm says in the documentary, “they also adopt other languages of the area, and that ultimately leads to language shift and language loss.”

There are approximately 500 Bantu languages, of which Gyeli is only one. But it has unique features, underscoring that the significance of a language, from a linguist’s perspective, often bears no relationship to its prominence or vigor.

“The vocabulary of Bagyeli is quite different” from that of many Bantu languages, Grimm says. The Bakola-Bagyeli are hunter gatherers, whereas most Bantu

language speakers are farmers. They’ve invented hunting tools and techniques, and have developed an intimate knowledge of plant life, all unknown in other communities—and, necessarily, an entire vocabulary around such esoteric knowledge and skills. Their knowledge of plant life, some of which has been shown to have medicinal value, has attracted the interest of Western pharmaceutical companies. During her time in Cameroon, it wasn’t unusual for Grimm to be mistaken for someone working for such an enterprise. It happened, she says, “all of the time.”

As part of their documentation, Grimm and her colleagues worked among Gyeli speakers in multiple contexts. They documented conversations among groups, dialogues, storytelling, and explications of daily activities, from hunting to children’s games. That’s an important practice in documentation, she stresses, and part of what distinguishes it from the mere collection of a relatively static group of words and rules. Video and audio recordings are designed to depict “language as it is spoken naturally in the language community,” she says. Spoken language is full of nuances that are derived from context. Those nuances provide linguists with rich insight into some of the core questions motivating language study—questions concerning the relationship between language and social relations, psychodynamics, how language evolves, and the ways in which humans acquire and process language.

Movements toward language preservation and revitalization have taken hold around the world. But they face enormous odds. Few expect that the displacement of the Bakola-Bagyeli, for example, will reverse itself.

“This work, in 100 years, will be what’s left,” says Grimm at the close of the documentary, of the data that she and other linguists collected that will be housed at the Max Plank Institute for Psycholinguistics Language Archive in the Netherlands. “There won’t be a chance to come back and study the nature here. The Bakola will be alive—hopefully—but they’ll be different.”

Stories like this one exert a strong pull on linguistics majors, says Jeffrey Runner, a professor of linguistics and of brain and cognitive sciences, and chair of the linguistics department. Language loss “taps into something.” Students investigating endangered languages “feel like they’re doing something that’s valuable. There’s the social angle of helping save endangered languages, but there’s also the scientific angle of finding new features of language that we don’t know about.”

“Language documentation definitely interests me,” Everson says, reflecting on her long-term plans. She has accepted a job at a software engineering firm in New York City, which she will begin soon after graduation. She’ll stay involved in linguistics through a project she’s working on with Nadine Grimm’s husband, Scott Grimm. He’s also a member of the linguistics department and is producing a dictionary of Dagaare, which is spoken in parts of Ghana and Burkina Faso. Everson’s role, as she describes it, is “to try to make software to make dictionary creation a little bit easier.”

“I’m hoping to go back and get a PhD in documentation and revitalization,” she adds. She’s attracted by the prospect of fieldwork, and convinced of the urgency of language preservation. There’s another, more basic reason, too. “Language is part of a people,” she says. “And letting a language die is not OK.”