A new study suggests regions where Indigenous languages are under threat face the greatest consequences from language loss. Cherokee, an endangered language, is pictured above in an undated document from Yale's Kilpatrick Collection of Cherokee Manuscripts. Image courtesy:
Yale University.

NEW HAVEN, CT.- Languages, like animal species, can go extinct. More than half of the world's approximately 7,000 signed and spoken languages are currently endangered. And without intervention they are likely to become extinct, meaning nobody will speak or sign them any longer.

While language loss is happening across the world, the costs vary strikingly in different places, according to a new study co-authored by Yale linguist Claire Bowern. Regions where all Indigenous language are endangered — including parts of South America and the United States —

face the greatest consequences.

The study, recently published in the journal Science Advances, is the first to use Grambank — the world's largest and most comprehensive database of language structure — to better understand global linguistic diversity and the threat that language loss poses to humanity's collective knowledge of history, culture, and cognition.

"Grambank allows us to find patterns across many languages and language families in a way that hasn't been possible at scale until now," said Bowern, a professor of linguistics in Yale's Faculty of Arts and Sciences, a member of the international team that built the database, and a senior author of the study.

"It demonstrates the beauty and complexity of language and how languages can be a window into the past. It also shows us how that window is very much under threat."

The novel database currently covers 2,467 language varieties spanning 215 different language families and 101 isolated languages from all inhabited continents and geographic areas. It captures 195 language properties — including word order, verbal tense, and whether a language features gendered pronouns — allowing researchers to draw comparisons between and across the languages.

"Grambank is like a DNA code of languages," she said. "We can use it to make comparisons to build language trees or examine how languages that split from a common ancestor differ from each other. We can identify features that are very rare in languages across the globe and figure out which of those features are particularly associated with endangered languages."

To calculate the potential effects of language loss globally and regionally, the researchers applied a metric called "functional richness," similar to one used in ecology that measures the number of species occupying niche space in an area. Substituting languages for species, they found that while functional richness would decline only moderately worldwide with the loss of all languages currently under threat, the consequences would vary greatly across regions. Northeast South America, Oregon and Alaska in the United States, and northern Australia — regions where all Indigenous languages are under threat — face the harshest consequences.

The loss of those languages would reduce functional richness to zero in these places, according to the analysis.

"Once linguistic diversity is lost, it's not easily recovered," said Bowern, a historical linguist whose work focuses on language change and language documentation in Indigenous Australia. "It will take sustained efforts to document and revitalize endangered languages to avoid a devastating loss to the communities where these languages are part of the cultural fabric and to further prevent the erosion of humanity's great diversity of languages,

which represents an incredible wealth of global human cultural heritage."

The analysis also revealed that there is a lot more variation across languages than was widely believed and provides important insights into how languages evolve and diversify. For example, the researchers show that genealogy — the gradual changing and splitting of languages over time — plays a larger role in shaping linguistic diversity than does geography, through which languages borrow words and grammatical constructs via contact between people speaking different languages.

Grambank's developers hope that other researchers will began to use the database to discover new patterns in linguistic diversity, Bowern said.

The project was a collaboration among the Max Planck institutes in Leipzig and Nijmegen, the Australian National University, the University of Auckland, Harvard University, Yale University, the University of Turku, Kiel University, Uppsala University, SOAS at the University of London, and the Endangered Languages Documentation Programme, and more than 100 scholars from around the world.