Preserving endangered languages as 3D shapes

A nylon prototype made via the latest industrial type of 3D printing (Multi Jet Fusion),
based on a language sample from an Amazonian language (Tariana). Image courtesy: University College London.

LONDON.- Half of the world's languages are endangered and more than a thousand are expected to be lost in coming decades. A team at UCL is using animation software to preserve these languages in an entirely new way.

It's estimated that there are over 7,000 documented languages spoken across the world. Yet around half of these languages are endangered. Between 1950 and 2021, around 230 languages were wiped from human memory. About 1,500 endangered and rare languages are at a high risk of being lost in the next century.

The consequences of language loss—also known as linguicide—are felt around the world, affecting identity and well-being as well as culture and linguistic diversity.

In response to this growing trend, an anthropologist and architect at UCL have come together to present and preserve some of these languages in a truly unique way, using inspiration and technology from animation in the gaming industry.

Reimagining language

Language is often only imagined in the shape dictated by our writing system—as words on a page or as sound. But the "shape" of humanity's natural languages and their high-dimensional form mostly remain to be fully explored and modeled as visual material—similar to how we envision the double helix of DNA.

Rather than only listening to a recording, scientists have now made it possible to also handle a language excerpt physically, by building 3D printed models based on language patterns and grammar.

As part of a study published in Humanities and Social Sciences Communications, scientists explored how the structure of language can be represented in 3D using software commonly used for game, film and video animation.

To create their 3D designs, they drew on the sounds of a language—the number of syllables in a line—as well as its grammar, focusing on a particular grammatical system called evidentials. They gave a numerical value to "evidential weight"—that is, the nature of the evidence being conveyed. In one of the languages they focused on—the Amazonian language Tariana—speakers are grammatically obliged to indicate the nature of the evidence they're conveying at all times, unlike in English, where this isn't required. Tariana contains a hierarchy of preferred evidentials, ranging from information obtained through direct visual observation to the repetition of information related by someone else.

They plotted points in three dimensions, with the numerical values derived from evidentials along the Z axis, the number of syllables along the Y axis and timeline on the X axis. The design software then turned these points into a 3D shape, virtually filling in the digital weave of warp and weft to appear as a smooth, woven undulating surface.

The researchers were then able to create four prototypes and digital 3D images from this, which they printed as 3D objects.

The first of its kind in the world, the project and its outputs not only enable the visual demonstration of the architecture of language, but also enables its preservation in a permanent, solid form.

Crossing boundaries

The multi-disciplinary study brought together Dr. Alex Pillen (UCL Anthropology) and Ph.D. candidate Emma-Kate Matthews

(Bartlett School of Architecture at UCL). Dr. Pillen explains: "The seed of the idea for this work came from our initial conversations—with each of us coming at this from our own different disciplines.

"At the time, I was just finishing a book project, and I felt that I couldn't quite come up with a single image that would go on the cover.

"And I thought, well, what if I could create an image that roughly would be the architecture of the Kurdish language incurred through an abstract design that could be used as the cover?

"Then more people were saying, "Well, we should put in some comparative materials. It shouldn't just be about Kurdish."

"We then contacted colleagues who are looking at Amazonian languages—and then I had a Ph.D. student who worked on Akkadian, an ancient language, who agreed to code it. Then a colleague suggested we could add English as well, which we did."

Dr. Pillen explains: "We've focused on one single grammatical system that is most prominent in Native American, Amazonian and Aboriginal languages—evidentials. Grammar is complex in these languages, and these are precisely the languages that are under threat.

"Many endangered languages have only a couple of hundred speakers left. We are now faced with a major loss for both marginalized communities and the history of humanity."

The researchers say that visualizing words and grammar in 3D offers a palpable new way to preserve languages that will not leave behind a written record, or whose complex patterns are often lost in translation—documenting endangered languages and enhancing grammatical understanding—and bringing to life ancient languages no longer spoken for future generations.

"We do, of course, have recordings of these languages to preserve the way they sound, while the vocabulary can be preserved in a dictionary. But what is more difficult to preserve is the grammar, because this is often presented in a very dry manner by linguists.

"You get transcripts, annotated with very technical terminology. But by producing the geometry of grammar in 3D, we allow people to have an immediate intuitive relationship to these languages that are under threat—or that might disappear."

As well as its implications for recording language in an era of rapid loss of linguistic diversity, the scientists say that 3D printing words and phrases also enables them and their students to use spatial intuition to get to know the patterns of languages that we do or do not speak.

"In terms of spatial intuition, you can immediately see in terms of—for example, the Tariana example—if you hold it in your hand, there is an immediate sensation of its complexity, which I would say the kind of books written by linguists, with all the technical detail about grammar, don't achieve," Dr. Pillen adds.

Dr. Pillen also notes that it isn't just languages under threat of extinction that are experiencing linguicide.

"This is not only about preservation of languages that are endangered—but also about linguistic loss," she says.

"Some Amazonian language, Aboriginal languages, North American languages are facing extinction—but there are several other languages spoken by sizeable populations who have been faced with language loss, such as the Kurdish language.

"In Turkey and Syria, where language policies obstruct the standardization of language, and schooling in the Kurdish mother tongue, communities suffer a loss of culture, and loss of linguistic heritage.

"It's for people like that that these models are also important as a solid object to reflect that predicament of loss. Our method allows for aspects of an intangible heritage to become tangible—a rare form of solidity for people who have often lost so much."

The researchers say that there is more they'd like to do with the technology they've developed to further enhance learning and language preservation.

"At the moment, these are language excerpts that are frozen in time representing very short speech moments," Dr. Pillen adds. "Now, what would be amazing is to use larger datasets to create moving images of evolving shapes in natural language."