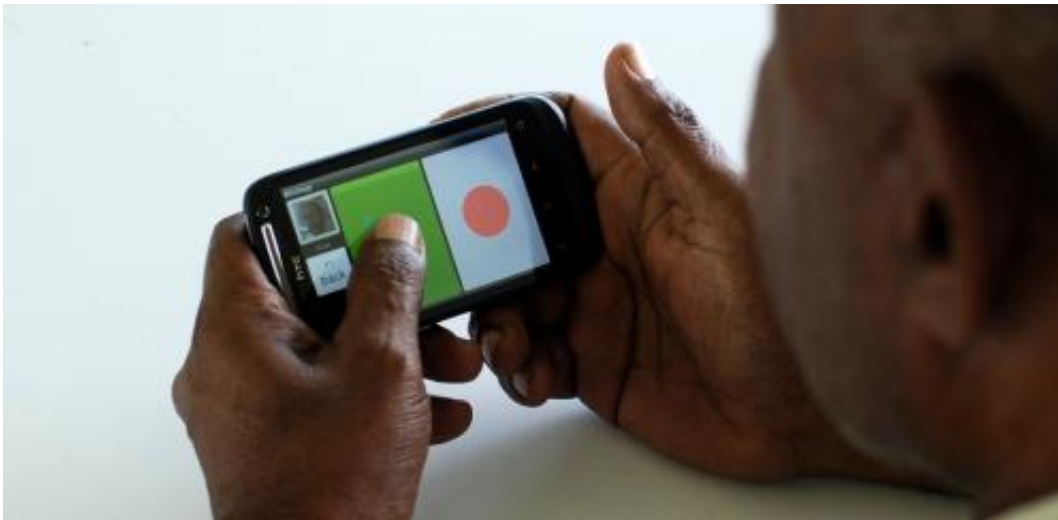# Cyberlinguistics: Recording the world's vanishing voices

March 12 2013, by Steven Bird

Endangered languages are being preserved using cheap mobile phones. Credit: Steven Bird

Of the 7,000 languages spoken on the planet, Tembé is at the small end with just 150 speakers left. In a few days, I will head into the Brazilian Amazon to record Tembé – via specially-designed technology – for posterity. Welcome to the world of cyberlinguistics.

Our new Android app Aikuma is still in the prototype stage. But it will dramatically speed up the process of collecting and preserving oral literature from endangered languages, if last year's field trip to Papua New Guinea is anything to go by.

Going off-grid in Papua New Guinea

In the steep hill country of the New Guinea Highlands, we journeyed to meet a man called Waks, who ran a small family farm and – in his spare time – taught other adults how to read and write his ancestral language. He is among the last generation of Usarufa speakers, who number only about 1,000.

No one learns Usarufa any more: the children are raised to speak Tok Pisin, a trade language (one used by merchants and the like without a common tongue) spoken throughout Papua New Guinea. Before our arrival in Usarufa country, no linguist had set foot there since the 1970s, around the time Waks was born. Usarufa is on the way out, yet no one is making any record of its culture and literature.

Travelling with me were my PhD student Florian Hanke, linguist Matt Taylor, and my 20 year-old son Andrew who had funded his own travel and served as our cook.

Waks and his family live in a circular hut divided into two rooms, with a mud floor, roofed with dry grass. His window opens to a million-dollar view across the deep valley to the west, and a breathtaking series of ridges and peaks.

As we looked down from that view we saw Waks' desk, covered with books and notes. Despite his lack of formal education beyond the early years of high school, the desk was that of a serious scholar, and I felt suddenly humbled.

After working our way deeper into the valley on a sharply descending ridge, we arrived at a large bamboo building with an iron roof: the literacy classroom. It was clean, airy and dry, and would be perfect for our needs. At one end we pitched our tents, glorified mosquito nets that

would lower the risk of catching malaria.

Soon our equipment was set up: solar panel feeding car battery feeding inverter feeding Wi-Fi router feeding network storage and mobile phones. Relieved that everything was working, we powered it all down to save electricity, and sat in near dark around the cooking fire.

Traditional culture and technical curiousities

The next day we hosted a village feast, and brought together those people who were well-known as storytellers. We contributed a large box of frozen lamb meat, and this was cooked with green vegetables in a traditional ground oven, or "mumu".

Speakers of Usarufa, a language with only 1,000 speakers remaining (Waks is in the centre of the back row). Credit: Steven Bird

After eating, we invited people to tell stories. Waks asked if we would like the children to be sent away since they couldn't understand Usarufa, but I insisted they should stay. A dozen people got up in turn, all highly animated, all drawing enthusiastic responses from the 50-strong audience.

In the afternoon, eight Usarufa men and women stayed back with us in the literacy classroom. All had volunteered to give us three days of their time to work on their language. We gave a short demonstration of our Aikuma mobile app and handed out the phones.

They went off on in ones and twos to find shady places outside. They recorded stories, songs, legends, personal narratives and dialogues. As they spoke into the phone their voices were transmitted to the storage attached to the Wi-Fi router.

When the group reassembled we showed them the refresh button, and all the stored recordings appeared on each of the devices. They sat on the floor and listened, replaying favourites and marking them using the "like" button.

Amid the excitement, recordings quickly accumulated. The most frequently liked recordings rose to the top of the list, thanks to the wisdom of the crowd.

On another occasion, I played recordings of languages from other places – languages that have now fallen silent – and the penny dropped for the Usarufa speakers. As one elder said: "We must record our stories now so that they can be preserved for our children, since they are not learning our language!"

The penny dropped for me at that same moment. I suddenly realised we now had our elusive informed consent! For months I had been wrestling

with the problem of how to obtain permission to store someone's recordings in a digital archive on the internet when they have not even experienced a computer. What did they really understand they were agreeing to when signing a permission form?

But there was no need to have that esoteric discussion. Everyone already understood that speaking into this device was tantamount to public speaking, only directed at an audience of people inhabiting other times and places.

They hoped the audience would include their descendents, who might one day want to study their cultural heritage and possibly even breathe new life into the Usarufa language.

Storytelling after the village feast. Credit: Steven Bird

Going further: A new task, a fresh location

A year on, and I am embarking on a trip to the Tembé village in the Amazonian rainforest, where the community has asked for technical help with documenting their language.

Working with local linguists who have an ongoing relationship with the Tembé people, I will teach people to use our app to record and translate their stories, and get ideas for improvements.

Participants will try out the latest version which includes voice-activated translation: while listening to a recording, the user can interrupt to give a simultaneous interpretation of the recording in another language (in this case, Portuguese).

This interpretation is captured by the phone and linked back to the original recording, phrase by phrase. In this way, the collected recordings are guaranteed to be interpretable even once the language is no longer spoken. This interpretability is what gives the recordings their archival value.

All materials we collect in this way will be left for the community and also lodged with the Museu Goeldi, a local research centre where they will be permanently available to the community.

Upscaling: The future

If enough people use Aikuma we will accumulate a large number of

recordings from the world's small languages, including Usarufa and Tembé. The result promises to be a digital-audio Rosetta Stone.

With permission, we will store the recordings and translations in the Internet Archive, a digital repository that has been preserving snapshots of the web since its inception in the early 1990s, and which is the most credible place to store digital content in perpetuity.

Cyberlinguists of the future may be able to discover the words and structures of dead languages from this data, and even construct dictionaries and grammars.

But back to the present, the living speakers of today's disappearing languages are equipped to preserve their voices, their unique perspective on the world, and how they have managed to live sustainably in their homeland for centuries.

Source: The Conversation

Citation: Cyberlinguistics: Recording the world's vanishing voices (2013, March 12) retrieved 20 May 2024 from https://phys.org/news/2013-03-cyberlinguistics-world-voices.html