

# 'Talking dictionaries' document vanishing languages

February 17 2012

---

Digital technology is coming to the rescue of some of the world's most endangered languages. Linguists from National Geographic's Enduring Voices project who are racing to document and revitalize struggling languages are unveiling an effective new tool: talking dictionaries.

Of the nearly 7,000 tongues spoken today on Earth, more than half may be gone by century's end, victims of cultural changes, ethnic shame, government repression and other factors. National Geographic Fellows K. David Harrison and Gregory Anderson, the [linguists](#) who are creating these dictionaries, say that some of them represent the first time that the language has been recorded or written down anywhere.

Harrison, associate professor of linguistics at Swarthmore College, and Anderson, president of the Living Tongues Institute for [Endangered Languages](#), have traveled to some of Earth's most remote corners, visiting language hotspots and seeking out the last speakers of vanishing languages. The last speakers and their threatened [cultural heritage](#) are photographed by National Geographic Fellow Chris Rainier.

Occasionally the team surfaces tongues not known to science. In 2010 they announced with National Geographic the first documentation of a highly endangered language known as Koro, spoken by only a few hundred people in northeastern India.

Harrison unveiled eight new talking dictionaries Feb. 17 at the annual meeting of the [American Association for the Advancement of Science](#)

(AAAS) in Vancouver, British Columbia. The dictionaries contain more than 32,000 word entries in eight endangered languages, more than 24,000 audio recordings of [native speakers](#) pronouncing words and sentences, and photographs of cultural objects.

"Endangered language communities are adopting digital technology to aid their survival and to make their voices heard around the world," Harrison said. "This is a positive effect of globalization."

The AAAS meeting featured a panel on using digital tools to save languages that included Alfred "Bud" Lane, among the last known fluent speakers of the Native American language known as Siletz Dee-ni, spoken in Oregon. "The talking dictionary is and will be one of the best resources we have in our struggle to keep Siletz alive," Lane has written. "We are teaching the language in the Siletz Valley School two full days a week now, and our young people are learning faster than I had ever imagined."

The talking dictionaries are produced by National Geographic's Enduring Voices project and the Living Tongues Institute for Endangered Languages. Other support for the efforts has come from Swarthmore College, the National Science Foundation, the Confederated Tribes of Siletz Indians and National Geographic's Genographic Legacy Fund. Besides Siletz (<http://siletz.swarthmore.edu>) the new talking dictionaries include:

Matukar Panau, an Oceanic language of Papua New Guinea. Only 600 speakers remain, living in just two small villages. Until the Enduring Voices team began documenting it three years ago, the language had not been recorded or written. The community requested that the language be placed on the Internet, even though they had not seen the Internet. They finally saw and heard their language in a digital medium after electricity arrived in their village in 2010, followed by computers the next year.

Matukar Panau dictionary: 3,045 entries; 3,035 audio files; 67 images.  
<http://matukar.swarthmore.edu>

Chamacoco, a language of Paraguay's remote northern desert, still spoken by about 1,200 people but highly endangered. The Chamacoco people still practice hunting, fishing and gathering, while also adopting modern technologies like mobile phones and text messaging. The Enduring Voices project is providing equipment and training to local language activists who are writing and recording the dictionary.

Chamacoco dictionary: 912 entries; 912 audio files.  
<http://chamacoco.swarthmore.edu>

Remo, a highly endangered and poorly documented language of India. 4,008 entries; 1,157 audio files; one image. <http://remo.swarthmore.edu>

Sora, a tribal language of India under pressure to assimilate. With National Geographic's help, the Enduring Voices team employed a native speaker to lead the field recording efforts. Currently 453 entries; 453 audio files, and rapidly expanding. <http://sora.swarthmore.edu>

Ho, a tribal [language](#) of India with about a million speakers but under pressure from larger tongues. The traditional Ho script can't be typed on computers yet, so the project is petitioning the Unicode consortium to add it. 3,020 entries; 3,012 audio files; four images.  
<http://ho.swarthmore.edu>

Tuvan, an indigenous tongue spoken by nomadic peoples in Siberia and Mongolia. 7,459 entries; 2,972 audio files; 49 images.  
<http://tuvan.swarthmore.edu>

An eighth dictionary is dedicated to Celtic tongues, and more are in production, Harrison and Anderson report.

Provided by National Geographic Society

Citation: 'Talking dictionaries' document vanishing languages (2012, February 17) retrieved 20 March 2024 from <https://phys.org/news/2012-02-dictionaries-document-languages.html>

This document is subject to copyright. Apart from any fair dealing for the purpose of private study or research, no part may be reproduced without the written permission. The content is provided for information purposes only.