

December 1, 2005

Culture

Online Project Seeks To Preserve Disappearing Languages

by Danielle Belopotosky

Araona, Kamasa and Potawatomi are names that do not slide off the tongue easily to English-speakers, but to a small team of linguists in California, they are part of an ambitious project to preserve disappearing languages.

More than 50 percent of the world's languages are endangered, according to the United Nations Educational, Scientific and Cultural Organization. Furthermore, 96 percent of the world's languages are only spoken by 4 percent of the population.

With the rate of one language disappearing every two weeks, according to UNESCO, linguists are tediously working at the Rosetta Project to build a digital archive of the world's 7,000 known languages. To date, the archive consists of nearly 2,500 languages and more than 400,000 distinct words have been documented.

A team of linguists has been collecting information on a few thousand languages by culling special libraries for documentation on language family relationships, dictionaries and writing systems, said Laura Buszard-Welcher, the project's curator. The Web site includes word lists, language rules and the phonology, or function, of sounds.

"With a project of this scope, you really have to have a worldwide community to monitor and build it," Buszard-Welcher said. A new Web site is under development to improve navigation and functionality, including language trees and a community aspect to the site, she said. Users will be able to comment on the site, much like the "wikipedia," which is an online encyclopedia collaboratively written by its users.

It also will include audio recordings of languages. The project has inherited a treasure-chest from the family of Alan Lomax, the folklorist and musicologist. The family has donated about 200 reel-to-reel language recordings to the project.

The first instantiation of the language project was a 3-inch nickel disk capable of holding up to 350,000 images, she said. The Los Alamos Laboratories and Norsam Technology developed micro-etches of analog images and text on the disk, which are only readable via an electron microscope and has been made to

withstand elevated temperatures and salt and tap water for long durations.

The project comes at a time of heightened awareness of cultural and language preservation. During the U.N. World Summit on the Information Society last month, some leaders warned world languages could be lost. Certain languages dominate the Internet as it increasingly becomes a virtual means of communications across cultures and peoples. UNESCO estimates that 90 percent of the world's languages are not represented on the Internet. The United Nations has established the Register of Good Practices in Language Preservation, a project of UNESCO's Endangered Language Program to identify practices to preserve languages.

The Rosetta Project has its own roots in the legendary Rosetta Stone, which was discovered in Egypt in 1799 by French soldiers and carved in 196 B.C. On it are two languages, Egyptian and Greek, and three scripts, hieroglyphic, demotic and Greek.

The project was initiated by the Long Now Foundation and received \$1 million in funding from the National Science Foundation .

For the ethno-linguistically curious, Araona, Kamasa and Potawatomi are spoken in Bolivia, Papua New Guinea and the Midwest, respectively.