
[Databases selected:](#) ProQuest Newspapers

Science File; Project Preserves Diversity of Language in Miniature; [HOME EDITION]

EMILY SINGER. *Los Angeles Times*. Los Angeles, Calif.: Aug 12, 2002. pg. A.9

Abstract (Summary)

The design itself has a functional significance: "It implies the instructions of the disk," said Jim Mason, director of the Rosetta Project. The eight major languages around the outside of the circle are English, Spanish, Russian, Hindi, Chinese, Hebrew, Arabic and Swahili.

Researchers at the Long Now Foundation have data on 1,500 languages. Before production of the first-edition disk this fall, the top 1,000 languages will be selected for inclusion. Candidates will be chosen based on which have the most complete data, the most diverse elements and the widest geographic distribution.

Language collection won't stop there. Mason hopes to have data on 3,000 to 4,000 of the world's 5,000 to 7,000 languages in the next few years. And the project will include as many as possible in upcoming editions of the disk. The project has been funded through a \$350,000 grant from the nonprofit Lazy Eight Foundation.

Full Text (740 words)

(Copyright (c) 2002 Los Angeles Times)

The Yupik language, once common among Inuit communities on the Alaskan coast, is now spoken solely by adults. Children learn only English, making it probable that Yupik will be lost in the next generation.

Yupik is one of thousands of languages that are expected to die over the next century. While linguists struggle to keep the languages in these small communities alive, a group of researchers in San Francisco is determined to save some record of current linguistic diversity for thousands of years into the future--long after its speakers are gone.

Their brainchild is the Rosetta Disk--a 3-inch disk containing micro-etched translations of the Book of Genesis in 1,000 languages.

The original Rosetta Stone, discovered in 1799, contains the same text in Egyptian hieroglyphics, Greek and Demotic, an everyday form of ancient Egyptian. The parallel organization of the basalt slab, inscribed in 196 BC, became the key in deciphering hieroglyphics.

The contemporary Rosetta Disk packs even more of a linguistic punch than the ancient original. Squeezed onto the small disk are many linguistic elements--pages of words and pronunciations, as well as examples of native texts.

The project is sponsored by the San Francisco-based Long Now Foundation, a nonprofit organization founded in 1996 by computer pioneers Danny Hillis and Stewart Brand to support long-term thinking in a fast-paced world.

The organization's other main project is the creation of a clock designed to tick once a century for 10,000 years.

The lasting perspective is especially crucial in linguistics, where languages are becoming extinct faster than they can be documented. James McCloskey, a professor of linguistics at UC Santa Cruz, said: "There is a real sense of urgency ... the majority of languages have only a few hundred speakers."

The disk, which should last for more than 2,000 years, is made of nickel and requires only a microscope to be read. Each page of text is miniaturized and etched onto the disk as an image.

This simplicity guards against the threat of changing technologies, which could make a digital disk unreadable by computers in the future.

The disk is protected against the elements in a stainless steel and glass orb, which also acts as a magnifying lens.

Stowed in the bottom of the orb is a metal ribbon, where owners can inscribe their names as a pedigree of the object.

The design itself has a functional significance: "It implies the instructions of the disk," said Jim Mason, director of the Rosetta Project. The eight major languages around the outside of the circle are English, Spanish, Russian, Hindi, Chinese, Hebrew, Arabic and Swahili.

The text spirals down into rings of text that taper to microscopic scale. The final level requires a 1,000-power microscope to read. An image of the Earth at the center of the disk radiates spokes that indicate the geographical origin of each language.

Researchers at the Long Now Foundation have data on 1,500 languages. Before production of the first-edition disk this fall, the top 1,000 languages will be selected for inclusion. Candidates will be chosen based on which have the most complete data, the most diverse elements and the widest geographic distribution.

But language collection won't stop there. Mason hopes to have data on 3,000 to 4,000 of the world's 5,000 to 7,000 languages in the next few years. And the project will include as many as possible in upcoming editions of the disk. The project has been funded through a \$350,000 grant from the nonprofit Lazy Eight Foundation.

What was meant to be a piece of long-term preservation has become a much larger project, said Alexander Rose, director of the Long Now Foundation.

When collecting information for the disk, the Rosetta team came up against a problem: No one knew all the languages. So they created an online database to house the information.

Linguists and the public can now see the data and make comments and additions.

Users have pored over the Web site, correcting annotations of the wrong language, the wrong dialect and even the wrong orientation of a piece of text, Rose said.

People can help get their favorite language included by making a \$100 donation to the project.

And true language aficionados can order a first-edition Rosetta Disk--at \$25,000 a copy.

Mason and Rose hope to distribute the disks globally--to archives, museums and representatives of each language--as soon the production process becomes economically feasible.

[Illustration]

Caption: GRAPHIC: Preserving Language; CREDIT: RAOUL RANOA / Los Angeles Times

Credit: TIMES STAFF WRITER

Indexing (document details)

Subjects: Translations, Nonprofit organizations, Language, Data collection
Companies: Long Now Foundation
Author(s): EMILY SINGER
Document types: Feature
Section: *Main News; Main News; National Desk*
Publication title: Los Angeles Times. Los Angeles, Calif.: Aug 12, 2002. pg. A.9
Source type: Newspaper
ISSN: 04583035
ProQuest document ID: 148413561
Text Word Count 740
Document URL: <http://0-proquest.umi.com.opac.sfsu.edu/pqdweb?did=148413561&sid=1&Fmt=3&clientId=17866&RQT=309&VName=PQD>

