

The Open Language Archives Community

San Francisco Hyatt Regency, Grand Ballroom B

2:00-4:30pm, 4 January 2002

2:10 **Introduction**

Steven Bird (Penn) and Gary Simons (SIL International)

1. WHAT IS THE OPEN LANGUAGE ARCHIVES COMMUNITY?

2:15 **The Seven Pillars of Open Language Archiving**

Gary Simons, SIL International

Digital archiving of language documentation and description on the World-Wide Web holds the promise of unparalleled access to language information. But if it is not done well, it also offers the specter of frustration and chaos on an unparalleled scale. This talk presents an executive summary of our vision for the kind of infrastructure that would unlock the promise. Special focus is given to the seven pillars on which such an infrastructure would be erected: DATA, TOOLS, ADVICE, GATEWAY, METADATA, REVIEW, and STANDARDS.

2:35 **OLAC and Linguist List**

Helen Dry and Anthony Aristar, LINGUIST List

Over the past 11 years, the Linguist List has become the primary source of information for the linguistics community, reaching out to over 15,000 subscribers worldwide, and having four complete mirror sites. The Linguist List will be augmenting its service by providing the primary entry point for OLAC, and permitting linguists to browse distributed language resources at a single place. This talk will include a demonstration of a new Linguist List “service provider”, and also report progress on a new NSF-sponsored project to create a “Showroom of Best Practice” for language documentation.

2. WHY IS LANGUAGE ARCHIVING IMPORTANT?

2:55 **Web based archiving as a tool for language preservation and maintenance**

Megan Crowhurst, UT Austin and LSA Committee on Endangered Languages and their Preservation

Problems facing linguists and others working on endangered language varieties are the challenges of first documenting them, and second, working to preserve, maintain, or revitalise them, where desired. Even when adequate documentary materials exist, limitations on the distribution of traditionally published print media restrict the availability of these materials to those who need them for preparing the materials appropriate for language maintenance projects. The talk discusses some of the ways in which web based archives for language documentation materials can be used as a tool for language maintenance efforts.

3:10 **Language Archives and Linguistic Anchoring of Digital Archives**

Chu-Ren Huang, Academia Sinica, Taiwan

Language archiving is not purely for linguistic purposes. We believe that language archives are also crucial in the context of a broader digital archives project, since linguistic description and interpretation, like temporal and geographical description and interpretation, underlies all digital archive items. Thus, ideally, when archiving digital knowledge, each knowledge item should be temporally, geographically, *and linguistically* anchored. And any linguistic anchoring is impossible without comprehensive language archives. A second topic to be covered by this talk is a regional perspective on Chinese and Austronesian language archives.

3:25 **Legal, Ethical, and Policy Issues Concerning the Recording and Publication of Primary Language Materials**

Mark Liberman, University of Pennsylvania

Language documentation involves recording, analyzing, archiving and publishing a wide variety of materials, including audio and video recordings, transcripts, linguistic and cultural annotations and commentaries, dictionaries, grammars, and instructional works. Because of advances in digital media, in networking, and in computing technology, language documentation is becoming both easier and more useful. In addition, there has been much recent interest in documentation of endangered languages. This talk will review the legal, ethnical and regulatory context of language documentation projects in the United States.

3. WHAT DOES IT TAKE TO PARTICIPATE IN OLAC?

3:40 **Creating an OLAC Data Provider at ANLC**

Gary Holton, Alaska Native Language Center, UA Fairbanks

The Alaska Native Language Center is a major center for the study of Eskimo and Northern Athabaskan languages. The ANLC maintains an archival collection of more than 10,000 items documenting the twenty Alaska native languages, including virtually everything written in or about these languages. ANLC would like to improve the accessibility of its collection, and is making catalog data available in the OLAC metadata format. This talk will report experience with OLAC from the perspective of the ANLC archive. It will also describe the issues that a linguist/archivist faces in participating in OLAC, concerning computational issues and archive metadata issues.

3:55 **Getting involved in OLAC**

Steven Bird, University of Pennsylvania

This talk will describe a variety of ways for the wider community to participate in OLAC. For individuals and institutions wanting to disseminate their resources, three methods for contributing metadata will be presented. One of these, the OLAC Repository Editor (ORE), will be demonstrated live, to show how easy it is for linguists to document the resources they create. The talk will also describe the OLAC Process, a method for identifying best practices for the digital storage of linguistic documentation.

4:10 **Open Discussion and Questions from the Floor**

4:25 **Closing Statements from the Panel**

Contact Details for Discussants

Steven Bird, LDC, University of Pennsylvania, 3615 Market Street, Suite 200, Philadelphia, PA 19104-2608.
sb@ldc.upenn.edu

Megan Crowhurst, Department of Linguistics, University of Texas at Austin, 402 Calhoun Hall, Austin TX 78712-1196.
mcrowhurst@mail.utexas.edu

Helen Dry, Department of Linguistics, Eastern Michigan University, Ypsilanti, MI 48197.
hdry@linguistlist.org

Gary Holton, Alaska Native Language Center, U. Alaska Fairbanks, 218 Elelson Building, P.O. Box 757680, Fairbanks, AK 99775-7680.
gary.holton@uaf.edu

Chu-Ren Huang, Institute of Linguistics, Academia Sinica, 115 Nankang, Taipei, Taiwan.
churen@sinica.edu.tw

Mark Liberman, LDC, University of Pennsylvania, 3615 Market Street, Suite 200, Philadelphia, PA 19104-2608.
myl@unagi.cis.upenn.edu

Gary Simons, SIL International, 7500 W. Camp Wisdom Road, Dallas, TX 75236.
Gary_Simons@sil.org