Utilization of language archives in endangered language research, revitalization, and documentation

Pavilion East
2:00 – 5:00 PM

Organizers: Andrea L. Berez (University of Hawai‘i)
Gary Holton (University of Alaska Fairbanks)
Bradley McDonnell (University of California, Santa Barbara)
Ruth Rouvier (Education Development Center)

Participants: Daryl Baldwin (Miami University)
Uldis Balodis (University of Helsinki)
Andrea Berez (University of Hawai‘i)
Lynnika Butler (Wiyot Tribe)
Natalia Fernandez (Oregon State University)
Susan Gehr (Humboldt State University)
Ryan Henke (University of Hawai‘i at Mānoa)
Gary Holton (University of Alaska Fairbanks)
Carmen Jany (California State University, San Bernardino)
Susan Smythe Kung (Archive of the Indigenous Languages of Latin America)
Lewis Lawyer (University of California, Davis)
Mary S. Linn (Smithsonian Institution)
Megan Lukaniec (University of California, Santa Barbara)
Bradley McDonnell (University of California, Santa Barbara)
Jennifer O’Neal (University of Oregon)
Gina Rappaport (Smithsonian Institution)
Miranda Rectenwald (Washington University in St. Louis)
Ruth Rouvier (Education Development Center)
Clare Sandy (University of California, Berkeley)
James Sarmento (University of California, Davis)
Michael Shepard (University of British Columbia)
Justin Spence (University of California, Davis)
Lajos Szoboszlai (University of California, Davis)

A renewed emphasis on language documentation (Hale et al. 1992, Himmelmann 1998) has led to the availability of an unprecedented wealth of linguistic data. Over the past decade the field has made significant advances in designing and adopting protocols for ensuring long-term preservation of these data, most of which are now created in digital form, through archiving materials in dedicated repositories. The value of archiving as a platform for meeting the needs of endangered language documentation and preservation is now understood as an integral part of the research workflow (e.g., Conathan 2012), and much has been written to encourage and instruct linguists to archive their materials (e.g., Nathan 2012, Woodbury 2011).

Until recently, language archives have been primarily valued for storage of endangered language data, with access and utility primarily relegated to the responsibility of those with the time, inclination, and resources to harvest materials in the service of a narrow set of research questions. Now, however, we can begin to consider new ways to interact with archived data, making possible a data-driven science of linguistics through which hypotheses can be tested against a wide range of data from the world’s languages. At the same time these data can form a foundation for the maintenance and revitalization of endangered languages. By providing new access to large quantities of data, language archives have the potential to transform the way we do linguistics (Seifart 2012). Archived data can enrich research and revitalization programs, and research and revitalization efforts can in turn enrich the archival record. In addition, self-reflexive examination of archive use can influence how we create new collections. Several important questions emerge. Are archives meeting the needs of current users? How can archive usage inform our approaches to language documentation? How can archival materials best be utilized for revitalization projects? What kinds of linguistic questions can we ask now that we couldn’t before? What ethical dilemmas does “open access” raise for endangered language communities?
This panel and poster session examines this transformation by exploring the various ways in which archives are being utilized by linguists, scholars from other disciplines, language learners, and language activists. The presentations and posters are clustered around six topics: integration of archived materials into linguistic research; use of archives for sleeping languages; use of archives by language communities; accessibility; user-centered design of language archives; and enhancement of archival resources based on linguistic research. Members of the panel and poster session include members of language communities, students, archivists, linguists, and other scholars.

References


Abstracts: Papers

Andrea Berez (University of Hawai‘i)

Gary Holton (University of Alaska Fairbanks)

Bradley McDonnell (University of California, Santa Barbara)

Ruth Rouvier (Education Development Center)

Ryan Henke (University of Hawai‘i at Mānoa)

*The use of archives in endangered language preservation: The state of the art*

This presentation is intended as an introduction to the panel and poster session. We cover the rise of discussion about linguistic archiving as it relates to language documentation, description and revitalization. We talk about the change in the past two decades in (i) the nature of archives (from analog to digital), (ii) users of archives (from academic researchers to language communities and beyond), and (iii) conceptualization of archives (from storage facilities to living centers of access). We introduce the questions listed in the session abstract and preview the speakers and poster presentations on those topics.

Justin Spence (University of California, Davis)

*Managing uncertainty in archival linguistic research*

Archival linguistic documentation can be fraught with uncertainties, especially where materials were haphazardly collected over many decades by researchers with varying interests, aptitudes, and training in linguistics. While interpretive research can yield valuable insights about particular languages, it is sometimes unclear what such archival collections can contribute to more broadly generalizable knowledge – “research” as a theory-oriented enterprise as it is traditionally understood in academia. Drawing on archival work exploring variation and recent innovations in Hupa and other Pacific Coast Athabaskan languages, I argue that in some cases at least they can contribute a great deal, provided that the research questions one seeks to answer are framed in an appropriate way. Often enough reasonable assumptions guided by insights in one theoretical domain or sub-domain (grammatical, typological, or historical) can produce interesting results in others; in principle this is no different from many other kinds of linguistic research drawing on non-archival data.
Daryl Baldwin (Miami University)
aapisaataweyankwi mahsinaakanenkonci: Reviving our language from documentation

The Myaamia (Miami) language has been undergoing a community revival since the mid 1990s. This language ceased to be spoken in the early 1960s forcing community educators to turn towards the vast language documentation as a source for its revitalization effort. From documentation, to community programing, and emerging as a living language again, is a process requiring a great deal of tools, support, training, and caution from specially trained individuals. This talk demonstrates how the Miami Tribe of Oklahoma has supported the development of this 20 year effort and how the use of extant language archives has lead to the reclamation of their indigenous knowledge system and is empowering a younger generation of myaamia youth.

Michael Shepard (University of British Columbia)
The significance of language ideology and culture in endangered language archiving practices

Endangered language documentation is occurring globally to address processes of language shift. Increasingly these documentation efforts are occurring in the context of small community based archives. My presentation describes research on community archive initiatives in Alaska and Washington, and draws comparisons to selected international archive platforms. I focus on the identification of existing language ideologies and cultural values in the community to increase the efficacy of language preservation efforts. I find that language archive strategies closely reflect local ideologies and values, especially in relation to sovereignty and self-determination. As we work toward mobilizing the knowledge found in archives to support revitalization efforts, cultural factors that make archival collections relevant, accessible and participatory are increasingly important. Additionally, initiatives that facilitate application of archival collections for language education can benefit from processes that are culturally grounded and pedagogically sound.

Jennifer O’Neal (University of Oregon)
Archives, access, and collaborative solutions

One of the main tenets of the archival profession is to provide open and equitable access to archival records and to ensure that accurate descriptions and inventories accompany those collections. This access is especially important for language records that can often be hidden or buried in other collections, such as fieldnotes, research files, and reports. However, researchers can find it difficult to navigate these often large collections that are frequently not inventoried at an item level. In addition, some cultural heritage communities may request that access limitations be placed on language archives in non-tribal repositories based on culturally sensitive content. This talk will highlight access parameters for cultural heritage language archives and suggest best practices for successful collaboration between repositories and communities.

Abstracts: Posters

Carmen Jany (California State University, San Bernardino)
Uldis Balodis (University of Helsinki)
Grammar writing for Yuki and Chimariko based solely on archival material

One of the main disadvantages for grammar writing solely based on archival materials is the inability to collect additional data for verification purposes and to fill gaps. For some languages, however, archives represent the only source of data available to communities and linguists today. California has historically been characterized by a great deal of linguistic diversity (Golla 2011). During the decades prior and following the turn of the twentieth century, John P. Harrington and Alfred L. Kroeber extensively documented many of these native California languages, collecting a considerable amount of elicited material and connected speech (Golla 2011). The material is so vast in scope that even these linguists themselves did not have the opportunity to fully analyze it during their lifetimes. This paper examines the utilization of this archival data for the purpose of grammar writing for two languages: Yuki and Chimariko.
Bringing archival materials together to expand the Karuk text corpus

Many archival materials on Karuk exist, but their accessibility for research is often limited, e.g., audio without transcription or translation. I demonstrate how connecting related materials housed at different archives has been essential for rendering texts useful to researchers. An audio recording of Karuk, made by Franz Boas in 1929 and archived at the Phoebe Hearst Museum, is matched with a corresponding transcription in John P. Harrington's field notes archived at the Smithsonian. Recordings made by William Bright and archived at the California Language Archive are reunited with his transcriptions, archived at the American Philosophical Society. Audio recordings of texts in Bright's field notes archived at the Survey for California and Other Indian Languages have also been located at the American Philosophical Society. These archival materials are used to expand and improve the Karuk text corpus, which is vital to research on the phonology and morphology of this highly endangered language.

Relying on the past: Quests for knowledge using language archives

Archived language materials can provide stop-gaps to language disappearance in cases where intergenerational transmission has ceased. Yet these materials are merely records of past knowledge; they do not constitute knowledge in themselves. The task of the archival linguist is to illuminate the knowledge encapsulated in the archival record, and to make sure that this knowledge is accessible to those who need it. This poster outlines four aspects of the process of archival linguistic research. 1. Finding and assembling sources for specific languages from scattered locations into a central accessible repository. 2. Confronting the task of digitizing materials written in a variety of incompatible orthographies (and representing sounds for which current orthographic conventions are inadequate). 3. Transforming archives into community-friendly pedagogical materials for use in reviving ancestral languages. 4. Addressing issues of ownership and access of culturally sensitive texts in light of conflicting interests of tribal communities and academic institutions.

Using archival data to revitalize Wiyot

Until recently, archival records of the Wiyot language were largely inaccessible to the Wiyot tribal community, comprising a range of materials written in a variety of orthographies, together with untranscribed audio materials. This has posed a major obstacle to revitalization efforts, since the last known Wiyot speaker died over fifty years ago, making archival records the only source of information about the language. Since 2008, a major effort of the Wiyot Tribe's Language Program has been to digitize all known source materials in the form of a language database. Thanks to the comparative analysis and rapid search functions made possible by this database, we now use archival language data as the basis for beginning Wiyot language classes, language learning materials, mini-dictionaries, answers to language inquiries, and interpretive signage and other projects. This poster will illustrate selected examples of recent Wiyot language revitalization and outreach projects made possible by archival data.

From archives to adult and child language learning: Reconstructing and revitalizing Wendat (Iroquoian)

Recent discussions about language documentation reiterate the need for appropriate, accessible and responsible archival practices (Austin 2006; Bird & Simons 2003; Bowern 2008; Johnson 2004; Nathan 2011; Thieberger & Berez 2012), focusing on the trajectory from fieldwork to archives. This paper, however, discusses a different archival perspective, one in which materials are 'taken off the shelf' and subsequently transformed into pedagogical materials and taught in language courses. The different stages of the revitalization process, from archival access and manuscript processing to linguistic reconstruction, teacher training, and language teaching, are described for the dormant language Wendat (Iroquoian). Moreover, this presentation illustrates the complexities of using such material for linguistic reconstruction as well as the implications for developing a methodology to proceed from manuscripts to language courses. Furthermore, examples drawn from the Wendat context provide valuable insights into how archival materials inform, shape, and constrain the reconstruction and revitalization process.
**Natalia Fernandez** (Oregon State University)  
*Oregon Tribal Archives Institute: Assisting tribal communities through archival education*

In August of 2012, the Oregon Multicultural Archives at Oregon State University, with assistance from various partners, hosted the Oregon Tribal Archives Institute, the culmination of a two year grant project made possible by an Oregon State Library LSTA (Library Services & Technology Act) grant. The weeklong Institute was designed to address the need for in-depth archives training for Oregon’s nine federally recognized tribes to support the development, preservation, and use of their community archives. Over the course of two years, the project included: site visits during the summer of 2011 to each of the tribes to conduct an archives needs assessment, the development of the curriculum and planning of the Institute, and hosting the Institute. This project is a potential model for other institutions to assist their local tribal communities in their professional development through archival education. For more information visit: http://scarc.library.oregonstate.edu/oma/tai/index.html.

**Susan Gehr** (Humboldt State University)  
*Breath of Life: Revitalizing California's Native languages through archives*

The Advocates for Indigenous California Language Survival held its first Breath of Life Workshop in 1996. The Breath of Life Workshop connects tribal participants with Native American language archival collections at the University of California at Berkeley for the purpose of bringing back into use languages with no or very few living fluent speakers. During Breath of Life, participants use the materials to complete a research project that can be completed in a week. Participants finish the workshop prepared to conduct archival research independently. Participants use their archival research to develop curriculum for language learning, to study language around cultural practices such as basket weaving or regalia making, and restore ceremonies that were interrupted by genocide and assimilation.

**Mary S. Linn** (Smithsonian Institution)  
*Ethnography and metadata: Simple ways to capture the most*

This poster will seek to address ways in which language researchers and documenters can improve the quality of their metadata during their research. I will concentrate on ethnographic and language genre information that we capture in the catalog (database) the Native American Languages collections at the Sam Noble Oklahoma Museum of Natural History and are often overlooked by linguists in the field. Thus, the poster will encourage language documentation situated in ethnography. In addition, I will give two examples of metadata collection from language researchers -- a linguist and a language educator -- who have donated their materials to the museum. These examples emphasize how researchers can capture and easily organize metadata without it becoming an overwhelming task.

**Susan Smythe Kung** (Archive of the Indigenous Languages of Latin America)  
*How language researchers can enhance and improve archival materials: Suggestions from the Archive of the Indigenous Languages of Latin America*

In this poster, staff of the Archive of the Indigenous Languages of Latin America offer suggestions for how archival language collections can be enhanced and improved beyond the obvious addition of transcriptions, translations, and interlinearizations, and they offer ideas on how language researchers can both use and enhance archival materials--either their own or someone else's--in their own fieldwork.

**Miranda Rectenwald** (Washington University in St. Louis)  
*Thinking outside the (archival) box: Innovative uses of Jules Henry’s Field Notes*

Archived documents can provide multi-faceted, dynamic opportunities for teaching and learning in academia and indigenous communities. Anthropologist Jules Henry compiled extensive linguistic and cultural notes in the 1930s while living among the Xokleng Laklânõ (Brazil) and Pilaga (Argentina) communities. Until recently, these documents archived at Washington University in St. Louis were seldom used. However, by starting a collaborative digital project with Unicamp State University (São Paulo, Brazil) a number of innovative uses have emerged, including: The Unicamp Linguistics Department turning digitized documents into teaching materials for Xokleng community schools. A Washington University professor incorporated the project into a course on indigenous communities and human rights. Anthropology undergraduates are helping digitize documents, and gaining hands-on experience with 21st century skills, such as metadata. By expanding the classic archival view of how collections can be used, a wealth of possibilities emerge from what, at first glance, are only archival boxes of paper.
Gina Rappaport (Smithsonian Institution)

*Endangered language resources at the National Anthropological Archives: Scope, use and collaborations*

The Smithsonian Institution’s National Anthropological Archives (NAA) historical linguistic collections relating to endangered Native North American languages is one of the largest of its kind in the world. Because of this, the NAA is committed to making these resources available to their source communities and traditional knowledge keepers, and is in continual collaboration with individuals, tribes, organizations, universities, cultural centers, and others to develop innovative ways to share the information in the collections. This poster will present an overview of the scope of the NAA’s endangered language resources, provide examples of their use in language revitalization projects, and highlight important collaborative projects and programs that aim to enhance access to and use of these collections.