Living archives: A community-based language archive model

MARY S. LINN


Link to this article: http://www.elpublishing.org/PID/137

This electronic version first published: July 2014

This article is published under a Creative Commons License CC-BY-NC (Attribution-NonCommercial). The licence permits users to use, reproduce, disseminate or display the article provided that the author is attributed as the original creator and that the reuse is restricted to non-commercial purposes i.e. research or educational use. See http://creativecommons.org/licenses/by-nc/4.0/

EL Publishing

For more EL Publishing articles and services:

Website: http://www.elpublishing.org
Terms of use: http://www.elpublishing.org/terms
Submissions: http://www.elpublishing.org/submissions
Living archives: A community-based language archive model

Mary S. Linn

Sam Noble Oklahoma Museum of Natural History

1. Introduction

Over the past 15 years, linguistics as a discipline has been considering the roles of documentation and description and consequently the roles of linguists in language communities. Similarly, endangered language archives have been improving access to language materials for community members and discussing how language communities who are stakeholders in such materials (we are calling them ‘heritage language communities’) can play roles in the shaping, processing and delivery of the archival record. Building on the models of Community-Based Language Research (Czaykowska-Higgins 2009) and Participatory Archives (Huvila 2008), I propose a Community-Based Language Archive model (CBLA), and outline its implications for archiving, documentary linguistics and language maintenance, revitalization and renewal.

I use the term ‘language archive’ to refer to any holdings of primary and secondary language documentation materials by memory institutions such as archives, museums, libraries, and special collections within libraries. I use the term ‘archivist’ to include any person actively involved in the collection, arrangement, maintenance, or dissemination of these holdings.

Implementation of a community-based language archive will vary according to the geographic location, local community, and archive, and it will change as new technologies advance. While giving some other, primarily North American, examples, I mainly discuss my endeavors, as the Associate Curator, Native American Languages, at the Sam Noble Oklahoma Museum of Natural History, to create a community-based language archive in the Department of Native American Languages (NAL) at our museum (Linn 2009; Linn 2011). Many other wonderful examples exist,

1 The literature in these areas is growing. A starting place for language documentation should include Himmelmann (2006, 2010), Woodbury (2003, this volume), and for the role of linguists Hale et al. (1992), Ostler (1998), Speas (2009), and Rice (2009).

and it should be noted from the outset that most if not all endangered language archives and archivists that I am aware of engage with communities to some degree (see Wilbur, this volume, for discussion of another case study). Most of us come from backgrounds in language description and revitalization and have brought to archiving our experience of being charged with the care of materials that we understand to be of vital importance to members of the speech communities in which they were created, and to their descendants. Many others involved in endangered languages archiving who have archiving, library or IT backgrounds have come over time to care deeply for the people that continually and so gratefully use these once esoteric collections and thus breathe new life into archived materials. So, this paper does not pretend to present something that is radically new, but to articulate endangered language archives’ roles in language documentation and description, as well as their place in the wider archive community, and in language revitalization. Rather, I argue that community-based language archives are at the forefront of an emerging broader trend to ‘participatory delivery’ of information within which archives can be increasingly relevant.

2. COMMUNITY-BASED AND PARTICIPATORY MODELS

In this section, I review models of linguistic research, focusing in particular on the Community-Based Language Research model (Czaykowska-Higgins 2009) and Participatory Archives (Shilton and Srinivasan 2007; Huvila 2008). I discuss each in terms of their application to endangered language archives.

2.1 Community-Based Language Research

Czaykowska-Higgins (2009) situates linguistic field research within other social science research models that involve community engagement and response, namely Participatory Research (where community members participate in the design of research), Participatory Action Research (empowering communities in the design of the research and its goals) and Community-Based Research (focusing on community and equalizing roles of researcher and community experts). These provide a long overdue articulation of the theoretical, ethical, and social implications of linguistic research in endangered language communities for both researchers and communities.

Building on Cameron et al. (1992), Grinevald (2003) and Rice (2009), among others, Czaykowska-Higgins describes five levels of community engagement, the ethical implications of each level, and the reasons for selecting among them in given field situations. At the first “Linguist-Focused” level, a linguist comes to a community as the expert and remains
detached from them in the name of scientific objectivity. The linguist is primarily responsible to the scientific community and not the language community (Czaykowska-Higgins 2009: 20-22). At the second, or “Linguist-Focused Research”, level, attention is also on the language and its speakers as subjects of research, but the linguist works to minimize any damage or inconvenience arising from his/her presence or research (Czaykowska-Higgins 2009: 22-23), and to share their documentation materials with the community. This second level, in fact, is all that is required of research to follow ethical guidelines for human subject research in many US-based institutions.

The third level, Advocacy Research, begins to break down barriers between researchers and communities. Research is now also for communities. This is often a natural outgrowth of working in a community, and it requires that the researcher be sensitive and empathetic to the community’s needs. Czaykowska-Higgins (1990: 23) cites Labov’s work on Black English being used on behalf of African-American communities, and the use of research on Canadian place names in land-claims.

We can easily extend this third level to archives. For example, we might be asked to be a back-up repository for an Indigenous archive, to give advice, or to provide editing skills on an archive grant application. Archivists may notify communities that relevant collections exist and make sure that finding aids reside in the community (see also Wilbur’s contribution to this volume).

The fourth level, Empowering Research, focuses on language research for and with the community. Here, the activity is more embedded in community needs; for example the linguist may actively teach community members about linguistics (Czaykowska-Higgins 2009: 23-24), or participate more closely with community educators in designing curriculum and lessons, creating learner dictionaries, language teaching, or be involved in other forms of language advocacy (see Gardiner and Thorpe, this volume, for discussion of the Australian context).

For archivists, this fourth level may involve training community members how to use linguistic materials, such as in the Breath of Life/Silent No More workshops initiated at the Survey of California and Other Indian Languages at the University of California, Berkeley, and also run at the Sam Noble Museum and at the Smithsonian Institution. It may involve more active engagement with the community in matters of description of and accessibility to existing collections.
The fifth model involves engagement between researchers and communities from the outset and in all steps along the way. Community needs and wishes drive the whole project. This does not exclude theoretical work by the linguist, however it is just not at the center of the collaborative research. The linguist and community members are considered to be equal experts in the endeavor; in fact, the community trains the linguist. Czaykowska-Higgins (2009: 24) defines the Community-Based Language Research (CBLR) model as:

Research that is on a language, and that is conducted for, with, and by the language-speaking community within which the research takes place and which it affects.

The ultimate goal of CBLR in most endangered language communities is to help bring about community-driven social change through maintaining, revitalizing, or renewing language. Czaykowska-Higgins states that reorienting methodology in this way leads to new research questions, and ultimately to institution-wide changes – a ‘new culture of research’ (Czaykowska-Higgins 2009: 44). Language archives can also take part in this new culture (see also Gardiner and Thorpe, this volume).

2.2 Participatory Archives

Discussions about the representation of information and the participation of communities in the archiving process have also been taking place among memory institutions. The digital revolution has created a world of new users for formerly inaccessible information, by making data accessible from any location for little cost. Even within academia, institutions’ web pages and on-line catalogs have drawn new types of users. The growth of internet-based archives and wiki technology allowing users to interact with information has radically expanded accessibility, driven debates, and spawned the participatory archive. Theimer (2011: 9) defines a participatory archive as an ‘organization, site, or collection in which people other than archive professionals contribute knowledge or resources, resulting in increased understanding about archival materials, usually in an online environment.’

---

2 A large literature exists on presentation and power in museum exhibits. See Karp and Levine (1990) and Karp, Kreamer and Levine (1992) for the beginnings of this discussion.
Shilton and Srinivasan (2007) argue for a Participatory Archiving Model (PAM) for creating and preserving the contextual value of archived materials. At the Southeast Asian Archives at the University of California, Irvine, they worked collaboratively with leaders and members of Vietnamese, Laotian, and Cambodian communities to decide what narratives to collect to best represent the communities and to design outreach for the archives (Shilton and Srinivasan 2007: 93). They describe the value of collaboration in building enriched archives (Shilton and Srinivasan 2007: 87):

We believe that this process can help build culturally relevant records repositories while enabling marginalized communities to share their experiences with a wider public.

Their paper discusses how the traditional archiving steps of appraisal, arrangement, and description can be done collaboratively with an engaged community. Appraisal is the process of choosing what is to be acquired and archived (and what is not). Some archives have very little choice in this as their decisions are governed by what is given to them in accordance with their acquisition or collecting criteria. Others have mandates to build collections or fill in gaps with a particular focus, and some have acquisition budgets to do so. In these cases, the archivist certainly does shape the record. As The Pinky Show (2008) tells us, museum professionals ‘are the ones that tell [us] what is worth remembering and thinking about, and therefore, also what is worth forgetting.’

Linguistics, too, deals with inclusion and exclusion in the record (see also Dobrin et al. (2009) on ‘archivism’ in language documentation). What a linguist and speakers record together often becomes the record of the language (Mithun 2001). This has particularly dramatic consequences when there are few or no first language speakers of a language. CBLR ameliorates this by putting community members in control of what gets recorded. Extending the analogy with appraisal further, the linguistic record has been unalterably shaped, or misshaped, by the omission of language in anthropology divisions in natural history museums, at least in North America. Until 2002 when the Sam Noble Museum began its Native American Languages collection, only the Smithsonian Institution actively archived and curated language materials. Anthropology departments let linguistics atrophy and many departments dispensed with linguistics and the four-field approach altogether. In parallel, linguistics departments ceased to emphasize ethnographic methods, documentation and description, and certainly there was little training in field linguistics (Newman 2009), and no training to be had in archiving and metadata for linguists.
Therefore, our current haste to document endangered languages can be reframed as an attempt to address the gravest gaps in shaping the historical and cultural record of humankind. Shilton and Srinivasan (2007: 89) cannot be more truthful when they write:

Beyond the harm to an archivist of lost contextual knowledge and a consequently distorted historical record, marginalization of the dispossessed within traditional repositories of historical memory robs communities of their cultural identity.

The second and third aspects of classical archiving are arrangement, or organizing a collection, and description, the ethnographic information and other metadata about the items. An archivist traditionally keeps the original order, labeling, and descriptions supplied by the creator, because these have descriptive and historical value. This is part of the provenance of the material, and if strictly applied can be understood as taking a creator-centered approach, parallel to the linguist-centered approach in language research (see also Garrett’s contribution to this volume). Shilton and Srinivasan (2007: 96) state that when a project is collaborative, the order and description should make sense to the community. Asking a community to provide information about the language data, or meta-documentation (Austin 2013), such as background and history of the speakers in the recordings, how and when the recordings were made, additional narratives, and context for the existing information in the recordings starts this process. With digital records, we can easily make different arrangements possible in the presentation of materials without losing the record of the creator’s original arrangement. Shilton and Srinivasan (2007: 95) give an example from the Portland Museum of Art Northwest Coast Indian Collection. Non-Native curators asked the Tlingit community to expand their descriptions of artifacts in the museum’s collection. However, the Tlingit community members used the request as a springboard for telling historical and personal narratives. While perhaps not deemed ‘on task’ by curators, this is a significant addition to the description of Tlingit life from which these items were created and used.

Overall, the approach of Shilton and Srinivasan (2007) is along the same lines as the Empowering Research model. Choice of materials, arrangement, and the necessity of creating rich community-focused context and metadata are all done together with the community. The authors’ reasons go beyond a belief in preserving historical context to empowering disenfranchised groups by connecting presentation (website, exhibit, teaching materials, language renewal) to their cultural identity.

Huvila (2008) builds on the work of Shilton and Srinivasan but outlines an approach more akin to CBLR. He calls his model Participatory Archives (PA), stressing that the archives as a whole are participatory in nature, rather than
embracing just the processes outlined in PAM. Huvila worked on digital archiving projects centered on the archaeology and history of a Finnish manor house and castle. The needs that drove the development of the participatory archives concept are shared by language archives everywhere: geographical dispersion of creators and users, variety of types of archive objects (in his case ranging from manuscripts to physical objects and archeological measurement data), multiple and parallel interpretations of these objects, and expertise held by different individuals (Huvila 2008: 20). Fulfilling these needs, within the realities and possibilities of digital archives, led Huvila (2008: 25) to define three characteristics of participatory archives: (1) decentralized curation, (2) radical user orientation, and (3) contextualization of both the records and the archiving process.

In decentralized curation, curatorial responsibilities are shared between archivists and the participants in an archive. For Huvila, it is not the individuals but the collective that has the most in-depth knowledge about the records or items, their contexts, and their uses. This position is the most radical departure from Shilton and Srinivasan and from archiving practices in general. Whereas in PAM, the community acts as (continual) advisors to the archivists, for Huvila, there is no need for an archivist at all. In fact, in participatory archives there would be ‘information managers’ to maintain technical platforms and provide tools for working with the archive, but these people are not archivists, and have no greater role or claim to expertise than anyone else (Huvila 2008: 26; see also Garrett, this volume). This has a parallel in CBLR, where the linguist bears but one of many types of expertise and has no greater influence than community experts. While perhaps harder to embrace fully for large established archives, decentralized curation can easily be envisioned in newly created CLBR projects where archiving can and should be included in all aspects of documentation. And decentralized curation is, in fact, what several of the large digital language repositories practise.

Radical user orientation dictates that findability and usability are the highest priorities. They guide even appraisal and preservation choices. In short, the archive is ‘oriented and reoriented to its users all the time’ (Huvila 2008: 26). Whereas a given archive’s methods and extent of findability and usability may not always match with community goals (for example, in the case of sensitive materials), it is good for a collaborative team to have these goals in mind. Accessibility has been one of the major constraints in community use of archives. From making on-site use easier and more open to non-academics, to training community members to use linguistic materials, to accessibility via the web, we are only beginning to think about accessibility (see also Nathan 2013). Accessibility for community users ultimately depends on community decisions about what is collected, what is presented, and the
nature of presentation. It is not an impossible burden to design projects from the beginning with community-wide usage in mind.

Radical user orientation is compatible with archivists’ codes of ethics. The Society of American Archivists Code of Ethics, part VI Access, states that it is the responsibility of archivists to ‘promote the use of records as a fundamental purpose of the keeping of archives’ (Society of American Archivists 2011). The International Council of Archives (1996: 2) in their Code of Ethics describes access ethics (emphasis added):

[Archivists] should observe faithfully and apply impartially all the agreements made at the time of the acquisition, but, in the interest of liberalisation of access, should renegotiate conditions in accordance with changes of circumstance.

Although most historical language-material depositor agreements were limited by older notions of accessibility, in the broadest sense, these ethical codes always encouraged archivists to consider accessibility and how to respond to new advances and new orientations. Thus, language archivists are continuously faced with reinterpreting accessibility, hopefully now in collaboration with the families and communities in which the materials were made.

Contextualization of records and the archival process puts the voice of the community into all aspects: concept, design, organization of the records, including aspects ‘beyond provenance’. Context, Huvila (2008: 95) argues, should be that of the originators, creators, and users. Here again, he departs from Shilton and Srinivasan’s PAM, which assumes the existence of community-wide points of view and the appointment of community members with the authority to represent or make choices about these views. In PA, there is no ‘predetermined consensual community’ (Huvila 2008: 26). The community, like the archive, emerges out of all the structures, descriptions, viewpoints, creators, users, and usages to form an ever-evolving corpus.

Such an approach may not be applicable to every situation. In some North American Indigenous communities, there is a clearly defined group of people who can articulate traditional community-wide viewpoints and provide a historical record, but in others there is not. However, a participatory archive can be useful in situations where there is no internal consensus about who can speak for the community. In many communities (and especially, it could be argued, communities where the languages are endangered), individual beliefs and rights regarding language and language use can collide with beliefs about communal rights in the language, and having all viewpoints stated can help resolve such conflicts. Participatory archives can also provide a place for younger language users to have a voice; one that is not unappreciated by elders but often fails to get heard in more formal or traditional settings.
3. COMMUNITY-BASED LANGUAGE ARCHIVES

3.1 Defining Community-Based Language Archives

Having situated language archives in the broader landscape of language research and memory institutions, I propose a Community-Based Language Archive model (CBLA).\(^3\) In this model, as in CBLR, the archive actively engages with the relevant community in conducting all levels of documentation, describing and contextualizing, maintenance, and dissemination of information. A CBLA archive can be defined as an archive or collection that is focused on a language, and that cares for and disseminates documentation that is conducted for, with, and by the language-speaking community within which the documentation takes place and which it affects. The notion of accessibility is expanded and achieved through engaging community members as primary shapers and users of the archives with the goal of sustaining and renewing their languages.

I have retained the term ‘community-based’ as used in current linguistics and social sciences, instead of using the term ‘participatory’ favored in the archiving literature discussed above. This is mainly because most language archivists have come from a linguistics background. We typically archive materials that were collected using some linguistic methodology; presumably in the future more and more materials will result from a CBLR method. But more importantly, in this model, archivists and archiving are involved in the language work from the beginning; all language research should include documentation and dissemination, and these entail archives. Ideally, community-based archives are embedded in CBLR, ensuring community-wide accessibility of language resources, extending language environments to the internet, and enabling community presentation of language and culture. Like Huvila, I emphasize that archives should be continuously expanding and evolving with new modes and notions of interpretation, accessibility and dissemination.

\(^3\) This model should perhaps be named Community-Based Endangered Language Archives. However, as all languages have communities of users and issues of identity and contextualization, this model will to some degree serve all language archives, including archives of English (and other prestige language) corpora and databases of cross-language phenomena.
3.2 Transforming Existing Collections

The Carolyn Quintero Collection of Osage at the Sam Noble Museum illustrates the potential of applying CLBA to existing archive collections, whether those collections are newly deposited or were deposited years ago. Carolyn Quintero worked on the Osage language most of her adult life. She produced *First Course in Osage* (1997) for community classes, *Osage Grammar* (2005), and *Osage Dictionary* (2009). After her death, her family donated the Osage materials, which included 42 manuscript boxes of unpublished materials and approximately 140 hours of audio recordings, to the Sam Noble Museum. These materials are invaluable as there are no longer any fluent first language speakers of Osage. I contacted the Osage Nation Language Program (ONLP) to let them know that we were expecting the materials and to start collaborating with them from the beginning. At the core of the very active ONLP is a small group of young, functionally fluent second language learners, but they desperately need more primary language documentation materials.

A graduate student created an inventory and preliminary finding aid based on Quintero’s organization, and then ONLP and Department of Native American Languages (NAL) staff met to discuss collaborating on the materials. ONLP decided which materials were most valuable to them. They made a list of their high priority items, and thus established the order for digitization. Highest priority was given to the notebooks of Robert Bristow, an amateur linguist and second language Osage speaker. Bristow’s work is accurate, easy to read, and contains rich language documentation material from well-known speakers; his materials were immediately useful for Osage participants in the 2010 Oklahoma *Breath of Life/Silent No More* workshop. One Osage participant used NAL’s studio and staff to record a traditional story collected by Bristow. Nearly five minutes long, her final project improved her spoken abilities in Osage, created audio accompaniment for her classes, and provided new oral documentation of the Osage language in the Quintero collection.

After the meeting, ONLP asked NAL to send a memo outlining the nature of this collaboration. This helped situate their work with us and with the Osage tribal government. It helped detail the nature of collaborative work for the Museum. Most importantly, it recognized our new relationship. I recommend doing this in all collaborative endeavors.

Processing and digitizing collections is costly and time consuming. NAL staff and volunteers, ONLP staff, Oklahoma University graduate students, and Osage Nation interns (Osage undergraduate students at Oklahoma University, paid by the tribe) have all shared the work and cost. A former ONLP staff member and graduate student in library science is currently working with the NAL collection manager to catalog the collection. She and ONLP members
are instrumental in providing narrative and context for the materials, and tribal and family histories related to the content. After this work is completed, ONLP and NAL staff and project workers will begin to collaborate on an online, interactive archive for the collection. A wiki will allow a wider Osage audience to provide new narratives about information in the collection and create new usages for the materials (Holton 2009).

The *Wick R. Miller Collection: Returning to the Community* is a similar example of transforming an existing language collection. The Shoshoni Language Project, formerly at the Center for American Indian Languages at the University of Utah and now part of the Western American Languages Research Group at the University of Utah, disseminates materials and engages community youth in the collection. Through the Shoshoni-Goshute Youth Language Apprenticeship Program (developed by Katherine Matsumoto-Gray), high school students take language classes and internships on language projects, including claymation videos, storybooks, and children’s pattern books. The community is creating new uses for Wick Miller’s and others’ documentation, and in the process learning the language and creating new language materials and curriculum; these will, in turn, generate new documentation from new speakers (Shoshoni Language Project 2013).

### 3.2 Building New Archives and Collections

We want language archives to be accessible to the source communities in order to support maximal usage by those communities for revitalization and other purposes. The core of accessibility does not lie simply in the ability to download a digital file or watch a video on YouTube, but in community involvement in the creation and care of collections, and decisions about their distribution (see also Nathan’s contribution to this volume). These decisions may, and often do, include access through innovative uses of the language in videos posted on YouTube, Facebook pages and posts, Twitter, and other new forms of social media, but it is the interaction with and continued use of the language resources enabled by these channels that truly constitutes community access.

CBLA can be applied to all small documentation projects and collections. NAL houses many family-made collections, from families who came to NAL to digitize older recordings. We digitize the families’ recordings of loved ones and special events, and provide them with digital copies which they can then also give to other language projects or archives. We re-house the physical originals in archival storage and teach family members how to care for them in the Oklahoma climate. They name the collection, help us with metadata by giving any and all information that they want, and dictate access conditions.
Archive creation can also be done together with youth, such as in the Shoshoni Project internship program mentioned above. In the Native American Youth Video Documentation Workshops at NAL, we train students to use video cameras, how to create story lines for their projects, how to interview (including getting permission), and how to use Final Cut Pro video editing software. Students have several weeks to finish their final projects and present them at a public screening at the museum. Their projects are added to the collection. They can borrow the video/computer kits any time they want after they have completed the course, and a copy of any subsequent work goes into the collection as well. In addition, students who participate in the Oklahoma Native American Youth Language Fair sponsored by the museum get to have their videos and presentations deposited in the museum’s collection, adding to its vast collection of books, video documentation of Oklahoma languages, and documentation of songs, skits, narratives and other forms of public speaking.

Community and individual decision-making about what is created and archived has several positive outcomes. First, it leads to pride in language, both on a community level and on the family or individual level. It elevates the status of the language, especially among youth. It is clear that when youth and families know what is in archives and how they work, the collections get used more. This in turn leads to an increased sense of ownership and support by heritage language communities.

Archives can provide long-term connections between language documentation, communities, linguists and other professionals. CBLA encourages archiving from the very first stages of documentation, description, or revitalization. On-line participatory archives have the potential to draw in members of a language community to use materials for linguistic purposes and to define personal and cultural identity. CBLA greatly expands usage of materials: at this time when we need as many people as possible to document and revitalize languages, CBLA can tap into a pool of youth and community experts as language documenters, learners, and advocates. Breath of Life and other archive-situated ways of teaching community members how to read and use linguistic materials give new life to old and under-used sources, just as they give new voice to silent languages. Thus, CBLA promotes and tests the assertion that access to language documentation leads to the creation of new materials and new users, ultimately leading to new language documentation and new speakers.

CBLA is also in keeping with the Protocols for Native American Archival Materials (First Archivists Circle 2007). Although aimed at Canadian and US archives, this wonderful document outlines issues of sovereignty, ownership, handling, access, use and presentation that are applicable to all Indigenous collections. The authors stress the importance of building collaborative
relationships with Indigenous communities. The protocols ask archives to respect diverse forms of knowledge and perspectives and to incorporate alternative approaches to knowledge management (see also Gardiner and Thorpe, this volume, for related work in Australia).

4. ADVOCATING FOR LANGUAGE ARCHIVES

Michael Kurtz’s 2011 keynote address to the US National Archives and Records Administration calls on leaders in the archives and preservation community to be able to articulate why archives play a critical role in society. This is crucial in the current economic and political climate, where in the US the survival of memory institutions is threatened by severe funding cutbacks on the local, state and national levels. We cannot take for granted that financial support will continue for memory institutions, language documentation and description, or indeed any social science research.

Archives have continuously evolved to remain relevant even as media environments, access concepts, and user demographics have changed. Endangered language archives are well-placed to participate in and articulate these shifts. CBLA with heritage language communities can advocate for the languages whose records are in our keeping. In supporting expanded community use of archival materials, archives help link communities to many kinds of professionals, and in turn link these professionals to lesser-known knowledge and new perspectives. We are at the forefront of a grand experiment to explore what can be accomplished through collaboration in access to digital information.

References


