Transcript Summary 1
Introduction and Background

Workshop on User-Centered Design of Language Archives
20-21 February 2016
Co-Organizers: Christina Wasson and Gary Holton

Transcription by Heather Roth
15 July 2016

This material is based upon work supported by the National Science Foundation under Grants No. BCS-1543763 and BCS-1543828.

The material is available for use under a Creative Commons Attribution-NonCommercial-ShareAlike 4.0 International License.
<table>
<thead>
<tr>
<th>Document Information</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Title</strong></td>
</tr>
<tr>
<td><strong>Author(s)</strong></td>
</tr>
<tr>
<td><strong>Original Creation Date</strong></td>
</tr>
<tr>
<td><strong>Version</strong></td>
</tr>
<tr>
<td><strong>Date of Current Version</strong></td>
</tr>
<tr>
<td><strong>Revised By</strong></td>
</tr>
<tr>
<td><strong>Rights Information</strong></td>
</tr>
<tr>
<td><strong>Licensing Information</strong></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Revision History</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Previous Version</strong></td>
</tr>
<tr>
<td><strong>Changes</strong></td>
</tr>
</tbody>
</table>
Table of Contents

Introduction and Background .................................................................................................................. 1
  Overview of Language Archives ........................................................................................................ 1
  Overview of User-Centered Design .................................................................................................... 4

Explanation of Links Between Transcript Summaries and Video/Audio Files

The transcript summaries of the workshop have been linked to the relevant video files. Hyperlinks are provided in blue text, and time code for the video is noted periodically.

At some moments during the workshop, the videocamera failed to record. For those spots only (which are in transcript summaries 2 and 3), we have provided hyperlinks to and time code for the relevant audio files.

See the Table of Contents: Transcripts and Recordings for a full mapping of the relationships among transcript summaries, video, and audio files.
Introduction and Background

This portion of the workshop took place on February 20, 2016.

Start Video 1

[00:00 – 4:30]
[Everyone goes around the room to provide brief introductions of their name/title/affiliations]

[4:30-7:00]
Gary: thanks everyone for coming to the workshop. “This was actually inspired by Shobhana a couple of years ago, the idea that we had an opportunity to think about designing language archives by bringing together people who knew something about archives, people who knew something about language, and people that know something about how to design things.” Gary admits User-Centered Design is something he didn’t know about and that the UCD people may not have known that language archives existed, so this meeting is a way to “collaborate in an interesting way.” There have been many meetings over the past few years about how to approach language archives. For instance, a meeting with the Linguistic Society of America. Even so, the question remains on how to better engage users and “how to make language archives more useful to people.”

[7:05]
Christina: explains that she hopes the workshop will kick off a longer research trajectory. Would like to identify where to go, gaps, etc. to then apply for another grant “culminating in the design or redesign of a couple of language archives.” A sort of test to figure out how UCD applies to language archives. Ultimately she hopes to come up with a set of guidelines as best practices while also recognizing diversity in types of language archives. The workshop is organized to move forward with identifying “what is known and what is not known.”

Gary will provide an overview presentation about language archives, Christina will provide an overview presentation about user-centered design, and each workshop participant will give a 7-minute presentation accompanied by their own personal PowerPoint slides to answer questions about language archiving from the perspective of their main stakeholder group. There were five stakeholder groups identified in total:

1. Linguists
2. Archivists and Curators
3. Language Communities
4. User-Centered Design Experts
5. Representatives of Funding Agencies

Overview of Language Archiving

[11:43]
Gary first defines what an archive is. Think about archives as more of a process to make things more available and accessible. Historically, archives have been around for a long time. Here are some examples:

- 1743 American Philosophical Society
• 1879 National Anthropological Archives
• Collections of Melville Jacobs (a student of Boas) – housed at the University of Washington Libraries

Archiving was seen for a long time as something taken for granted, why go out and document languages without preserving them? There was also a focus on access/utilization. Therefore, archives of the past are viewed as “brick and mortar” archives.

The emergence of digital language archiving came about due to three independent developments:
  1. Attention to endangered languages following the 1991 meeting of the Linguistic Society of America (LSA)
  2. Digital data transformation during the 90s
  3. Urgent need to access language resource for language maintenance

Students often assume that we have always been doing language documentation, but “there was a dark period out there” when linguists didn’t care about language archives. “Some could argue that still happens today…but I really think we’ve changed a lot in 25 years.” Gary questions whether what’s going on now is something new or a return to the past. “It may not seem obvious [to the outsider] but the field of linguistics has undergone an amazing transformation in the past 25 years. Largely due to linguists and language community members.”

Susan: says that although she would agree, she went to a cognitive linguistics conference recently and she was the only person there who did anything related to language archiving.

Daryl: points out that there was “an earlier period of documentation” with highly trained folks from the Jesuit era. There are manuscripts that are “quite extensive.” Then there was a dark period before formalized anthropology and linguistics. This is a rich area that has not been published on and the documentation occurred at a time when communities, such as his own, were still living closer to their traditional lifestyle prior to assimilation.

Gary believes the digital transformation was sort of an “accident” that put documentation and description to the forefront again, raising questions about how to preserve language. Other issues to grapple with in the past 20 years were real-time archiving and balancing preservation and access. He points out how the creation of apps and mobile phones is exploding, but there are few standards for preservation to ensure such materials stick around.

Linguists then went out on their own to create digital language archives. He points out some big names in archiving such as DoBeS, AILLA, and Kaipuleohone, all of which contain representatives in the room. “These were efforts that were built up over the last 15 years to do digital language archives.” However, Gary points out that often these standards were produced by linguists but not archivists. Some big changes were:

• Shift in focus to a language rather than on a particular person
• Documentation being done for the purposes of archiving

There was also an evolution going on with traditional archives. There are standards that do exist for “trusted repositories.” There is a greater reliance on content specialists and a greater move
toward participatory archiving – looking at depositors as collaborators. There is also an increased focus on users. “Rethinking how people are approaching the materials.”

[29:00]
Gary provides a set of diagrams to represent different conceptions of archives. The first shows a singular scholar who has access to an archive that was meant just for them. Then he shows a diagram of the “participatory model” which includes multiple arrows showing how involved this model is. He also points out there has been a sort of parallel evolution among linguistics and archives.

So some challenges that are encountered with digital language archives are issues of access, “it should be easy.” Here the group questions, “what is access?” [32:20] because the term contains myriad meanings.

Mandana: points out other issues of ways information is organized and presented. “It’s very Western, North American…our entire interface language is in English. So if you come from Mexico and are a Chatino speaker…you cannot access your own materials if you don’t speak English.” She points out the difficulties inherent on other terms used in Western archiving such as “browsing” and “community.”

The ways in which archives are used is also highly varied. Gary notes that “much of the interest is not just language…it can be content, history, knowledge, biographies.” He also cites issues of interpretation and lack of context, with recordings for example. He cites a quote from Michael Shepard’s work, where the quote actually comes from Edward Alexander, “an archive is an expression of sovereignty. Access is also about control of materials.” In another quote from Edward, Gary cites, “in participatory archiving, it’s more than just making things accessible, it’s engaging conversation, it’s facilitating that conversation.” He says these all inspired him in putting together the meeting. When thinking about users, we also need to think about users through time, the past, present, and future. For instance, depositors who are no longer alive or unborn users.

[37:50]
Gary says it’s important to separate” access” and “accessibility.” Because access means something is there, but that doesn’t mean the item is accessible. He cites a new digital archive using the Mukurtu CMS platform.

He comes back to the idea of participatory archiving – “what have language archives done to engage with that conversation?” Again he cites the LSA meeting from the previous year as well as Mary Linn’s work in community-based language archiving (Linn 2014). In his discussions with Christina, he says “we have a chance to actually rethink what language archives are, the whole notion of archives, in some ways, is not very user-centered. It’s kind of top-down.” He asks a variety of questions, concluding with “is an archive even the right model?” Although this may be too much to consider at the workshop, it’s something to think about. Lastly he reinforces that there is a “multidimensional space” of users.

Shobhana: points out that we are using the term “language archive” specifically for endangered languages, but we should be focusing on all languages and naturalistic data. Particularly in the ways in which access and analysis can occur in different languages. “That would really revolutionize archiving.”
Mandana: when looking at more common languages they have various corpora, “they have worked out procedures and tools.” She cites child language data, and how other types of research have contributed to such languages. “If you want to get a typologist to start working empirically in making the data available, I don’t think that is related to the accessibility of language archives. I think that’s related to the conceptual framework within their reasoning.”

Gary admits he has been assuming endangered languages although there are other languages out there as well. He says “archives will have to serve multiple communities by design.”

Christina: points out there is also a difference between an archive that serves only one language and one that serves a community that brings together multiple languages.

**Overview of User-Centered Design**

[45:53]

Christina first brings up the parallel tracks between linguistics and archives from Gary’s presentation. “So often there are these almost social movements or changes that happen in so many places at once, and people start working parallel. And so one of those is this idea of participatory something.” She notes that the idea of “participatory” is nice but “the enactment of it is often challenging.”

As a starting point, Christina provides a definition from Wikipedia, “the chief difference from other product design philosophies is that user-centered design tries to optimize the product around how users can, want, or need to use the product, rather than forcing users to change their behavior to accommodate the product.”

UCD emerged in the 1980s and 90s in the business world in various contexts, but she will focus only on technology. She uses the example of changes in computers to demonstrate UCD principles, but also as a way to provide context around the emergence of UCD. For example, the shift from giant, mainframe computers to the desktop was fraught with difficulties, even though the desktop was made for everyone, perhaps due to the difficulty of DOS. Therefore, Apple’s introduction of the graphical user interface (GUI) “was a huge innovation. Because they thought ‘what is easy for people to use? What is intuitive for people to use?’ and so they came up with a metaphor of the desktop.” Ideas surrounding users emerged, then, because typical users “were not engineers…and their minds didn’t work the way the mind of an engineer might work.”

She lists various individuals who contributed to the field during their careers:

- Gould and Lewis
- Lucy Suchman – an anthropologist at Xerox PARC

Here Christina points out that user-centered and human-centered are interchangeable terms, however she prefers the term “user” because it is more specific and “evokes work practices, leisure practices.”

Going back to the emergent paradigm shift on “designing products and technologies that were easy and accessible” she points out that in the business world there was a “profit motive” if products could fit in the lives of users. Interestingly, in language archives, there is no profit.
motive. She posits the following question, “are there parallel motivations? Or what are the motivations? To think about developing language archives that are a better fit for their users.”

Designers themselves used to follow a sort of “cult of the creative genius” where they would create cool designs which had a high failure rate because they were often created without users involved. To solve these problems designers “started to connect with social scientists” but this began with cognitive psychologists and the field of human factors. She defines human factors as “broader than cognitive psychology, because it encompasses both considerations of psychology but even the physical body.” Christina shares a classic methodology where a test subject is in front of a computer, and a guide asks the subject to complete a task and there are observers on the other side of a two-way window. She points out this can be useful, “but it is very much a one person using one machine, and it is this experimental situation.”

Then designers realized that “anthropologists bring a complementary set of expertise about the user” such as:

- Providing information about the “actual context of use.” For instance, watching a person use a product in their own home.
- Discovering how users interact in social groups with a product or service.
- Uncovering cultural practices and meanings.

She shares a real-world example from her own work where she studied how visitors interacted with museum exhibits at the Field Museum in Chicago. A photo shows a family interacting with an exhibit, and she points out how studying just one person, out of context, may not uncover “social phenomena”, for instance, how parents explain to children parts of the exhibit. She also posits a variety of questions that came up when considering cultural practices and meanings. For instance, “what do museums mean in American culture? Why do people go to museums? What is appropriate museum visitor behavior?” etc.

Christina shares four basic components of user-centered design:

1. Identify user groups – who is currently using the technology? Who might benefit? Are there target audiences we are missing?

Here Christina points out that, in her limited experience with language archives, “there is, to me, surprisingly little holistic consideration of who are all the user groups of an archive.” She points out a simplified list of two to start: linguists and members of language communities by citing an article by Peter Austin (2011). Regional archives seemed to be used more by language communities and their descendants, whereas larger archives are used mostly by researchers. She points out that a majority of scholarly articles focused either just on linguists or just on language communities. Therefore, if one takes a second pass at identifying user groups there is much more complexity. For instance, user groups not yet considered or divisions within user groups. Another example are users who maybe don’t need to use the archive, but just want to know that it’s there.

Another concern she points out is “archives are really only user-centered for some users,” mainly, linguists and archivists. She claims that the design of language archives, then, is a result of Western science principles of linguistics. “There’s almost a colonial kind of feeling to archiving.” She admits this aspect makes her a bit “nervous” but that she is “hopeful there is a path” that will be explored in the workshop.

Christina points out that metadata, although appealing to linguists and archivists, may not be as useful to other people. So maybe it could be expanded to be useful to others. On the screen she
shows a videogame, *Never Alone*, as an example of one way an Alaskan language is stored and used to reach “users from any background through compelling narrative.” She concludes, “maybe there could be other ways of thinking of what is an archive and what it could be used for.”

2. Must consider diverse needs of multiple user groups.

Christina explains the bigger umbrella of user experience research. An example of some research questions might be, “what are some way users use the technology? What would they like it to do that it doesn’t do right now? What kinds of problems do they encounter? What’s the broader cultural context?” She also notes that levels of participation vary due to the types of user groups involved. For instance, with a small community, users could be engaged in the design process. “It can be fully participatory,” but there is often a “huge population” in the business context, and there isn’t a well-defined community, leading to the use of ethnographic research to understand the users and be their voice.

Susan: asks “how do they identify who to study for these large, amorphous populations?

Christina responds by saying marketing people will often have a good idea of who their target demographic is. Therefore a representative sample is used based off of that idea.

Crysta: her rule of thumb is to get a small group of people who are incredibly diverse. If there are patterns in the data then you’re on the right track.

3. UCD is interdisciplinary and collaborative

Christina notes that UCD is “a group process” where no one person contains all of the expertise needed to successfully complete a project. In any one project, there might be technology specialists, researchers, representatives of user groups, etc.

4. UCD is an iterative process

Christina shows a diagram that illustrates the process of UCD research, which contains various circles, characterized by trying and testing. There is never a true endpoint because products and services must continuously respond to a changing social and technological environment.

*End Video 1 [1:15:32]*