



NATIONAL SCIENCE FOUNDATION

NATIONAL ENDOWMENT FOR THE HUMANITIES



Documenting Endangered Languages

SAMPLE APPLICATION NARRATIVE

This sample narrative conforms to a past set of guidelines.

Please consult the guidelines for the current competition on NSF's website at:
<http://www.nsf.gov/pubs/2005/nsf05590/nsf05590.htm>

Project Summary

Intellectual Merit

The proposed project will produce video and audio documentation of two highly endangered indigenous languages, Mohave and Chemehuevi. Both languages are still used on the Colorado River Indian Tribes (CRIT) Reservation where a formal survey, conducted by the tribal library in 2002, revealed just 42 speakers of Mohave and 5-7 speakers of Chemehuevi, all over the age of 60. Work on the syntax of both languages was conducted in the 1970s; some work with phonology and electronic documentation of lexical items was begun in 2003. There has been no documentation of naturally-occurring conversation and earlier audio recordings are technically inadequate for high fidelity archiving.

This project actively involves tribal members in data collection by training them in elicitation procedures using video and audio equipment. The project team will consist of two tribal members from each group (Mohave and Chemehuevi), the PI who has over 35 years of experience in the CRIT community, and two doctoral-level graduate students (one in phonetics and documentation; the other in syntax and pedagogy). The team goals are 1) to carry out video and audio documentation of conversational practices and recorded narrative in both languages, 2) to provide training for tribal members in technology-based documentation of their languages and descriptive linguistics while engaged in fieldwork both on-site and on-line, 3) to use collected materials for constructing language lessons in support of the tribes' goal of language revitalization, 4) to construct text, audio and video databases and, 5) to work with tribal members to construct a standard set of protocols for indigenous communities regarding issues of the public access and use of digital archives and databases.

Broader Impact

The project results will increase the general understanding of the process and value of video documentation for endangered languages, add substantially more linguistic data for Mohave and Chemehuevi, evaluate the use of technology for on-line fieldwork, and generate guidelines for indigenous communities regarding the public use of digital databases and archives. This project builds on previous work which generated a handbook for tribal members on the development of multimedia language lessons. *Technology-Enhanced Language Revitalization* (Penfield, et al. 2004) is available in PDF format online and is now the text for a course at the American Indian Language Development Institute (AILDI) which enrolls students from many different tribes. The current project will result in the development of a similar handbook for tribal members devoted to language documentation. It will also propose a curriculum for an AILDI course on language documentation and will establish a model to be shared at AILDI and on-line for tribal members to employ in determining public access and use of digitally-archived materials.

Project Description

Background

The PI has been working with Mohave and Chemehuevi speakers on the Colorado River Indian Tribes Reservation (CRIT) for since 1969 as an applied linguist and has watched the steady decline of both languages. Mohave, a Yuman language of the Hokan language family, has received the most linguistic attention primarily represented by the published work of Judith Crawford (1976a, 1976b, 1978) and Pamela Munro (1973, 1976a, 1976b, 1976c, 1976d, 1979, 1980a, 1980b, 1992) and unpublished field notes of John P. Harrington (1910-1920). Chemehuevi is a southern Numic language of the Shoshonean branch of Uto-Aztecan. Previous work on Chemehuevi is quite sparse and is also represented in the unpublished field notes of anthropological linguists John P. Harrington and Carobeth Laird from 1910-1920 and in the published works of A. L. Kroeber (1907, 1909), Edward Sapir (1930) and much later by Pamela Munro (1978a, 1978b) and Margaret Press (1979). Notably, these languages have not been the subject of formal linguistic study since the late 1970s and there has never been any work on them which documents naturally-occurring conversation or interaction.

The PI's initial work involved collecting oral history with Mohave speakers (1969-70) in the form of narratives and stories in both English and the native language. This work was not published at the request of the speakers at the time, but is now available for inclusion in this project. The content consists of three Coyote stories in Mohave which have been transcribed in an interlineal framework but not yet committed to a digital format (Penfield and Flores, 1999). Subsequently, the PI worked for the public schools as a teacher and liaison to the tribal community in issues of language and culture. This work led to the PI's unpublished dissertation (Jasper, 1980). The PI is now frequently employed by the tribes as a language preservation consultant (Penfield and Weinberg, 2000). Recently, the PI has been instrumental in developing technology to support both Mohave and Chemehuevi in the CRIT community (Penfield, Cash Cash, and Roberts 2004, Penfield 2003, Penfield et al, 2003, Forger and Penfield, 2003).

In 2002, the PI and the staff at the CRIT library conducted a formal survey of the remaining speakers of Mohave and Chemehuevi at CRIT. The count of fully fluent speakers, all over the age of 60, was Mohave, 42, and Chemehuevi 8-10. Since that time, the numbers have dropped. Mohave speakers are now thought to be about 33; Chemehuevi 4-6. Four fluent speakers living in the CRIT community are known personally by the PI. It should be noted that there are two other communities where Mohave and Chemehuevi are spoken, namely at the Fort Mojave reservation in Needles, California and at the Chemehuevi Valley reservation at Havasu Lake, California. CRIT members stay in touch with these communities and it is known that the numbers of speakers in those communities is far fewer. There is just one known fully fluent speaker of Chemehuevi at Chemehuevi Valley and possibly 10-12 very fluent speakers of Mohave at Ft. Mojave. These are informal estimates based on recent discussions with members from CRIT, Ft. Mojave and Chemehuevi Valley. While work for the present

proposal is centered at CRIT, contact will be made and maintained with members of the other communities of speakers.

Previous work conducted by the PI on two different occasions provides a background for the current proposal:

1) In 1999, with funding from the University of Arizona Faculty Small Grants program, the PI began testing the feasibility of using technology in on-line environments to support linguistic fieldwork (Penfield 2002). This work involved the use of a MOO (Multi-user, Object-Oriented) program which is primarily text-based but can incorporate digitized material through a Web-interface during synchronous on-line discussions. This project led to the development of a new technology, known for now as the OLE board, an asynchronous on-line language environment which incorporates voice, video and text. (Reynaert, Penfield, and Forger 2003) This new technology is currently in the beta form of production and has been used successfully in discussions between the PI and the CRIT tribal library staff. The OLE board is also being implemented in work with two other indigenous language communities, Mutsun (Natasha Warner, personal communication) and Oneida (Williams, 2004).

2) From 2002-2004, in a project funded by the Bill and Melinda Gates Foundation, the PI developed a training manual for tribal members interested in developing multimedia language lessons in support of language revitalization. This book, *Technology-Enhanced Language Revitalization* (Penfield, Cash Cash and Roberts, 2004), is also available in PDF format on line under the project website of the same name and is used as the standard text for a summer course offered at the American Indian Language Development Institute (AILDI) titled, "Computer Applications for Indigenous Communities." The handbook resulted from working with Mohave and Chemehuevi speakers while building a model for training tribal members in basic technology needed to create multimedia language lessons. The funding which supported this project is no longer available in any form. The Gates Foundation funding was part of an initiative offered to tribal libraries and has been discontinued as a grant program. The Gates Foundation has made it clear that their work in support of indigenous communities has been completed; in fact, their funding was never targeted to support language documentation or revitalization, only toward the training of community members in computer technology. At this time, *no further funding is available under any of the Gates Foundation granting programs.*

The PI is well positioned to use technology in the documentation of these two highly endangered languages, Mohave and Chemehuevi. Having worked in this community for a long time, the PI is comfortable establishing fieldwork connections and anxious to involve some graduate students to perpetuate this work. It is imperative to begin this work as soon as possible given the rapid decline in the number of speakers for both of these languages.

Project Goals and Objectives

The intellectual merit of this project is reflected in its goals and objectives. The specific goals of this project are to

- Provide training for tribal members in data collection and language documentation using both digital video and audio equipment following the known 'best practices' as presented for archiving digital material.
- Conduct fieldwork on-site while working with project participants (technical language assistants and tribal elders) and to conduct fieldwork by staying in touch with tribal participants on-line using the new technology represented by the OLE board (discussed in the Work Plan below).
- Use the collected data to
 - Increase the lexical database for electronic dictionaries of both languages and expand on the existing linguistic analysis for these languages.
 - Establish a database for text, audio and video material.
 - Make a corpus of materials available on the web.
 - Further the pedagogical goals set by the tribes for language revitalization by creating more language lessons in both languages.
 - Develop a handbook for indigenous communities focused on language documentation.
 - Work, in concert with tribal members, to establish a model for indigenous communities related to the public use of digital databases and archives.

For the past two years, the PI has been training tribal members in the use of technology to support their languages specifically by using various open source software products in the development of multimedia language lessons. The proposed project would build on and extend this work by moving into the realm of language documentation. Using video and audio recording, concomitant with specific computer software which supports digital documentation, project participants would be carefully trained in data collection. The expectations would clearly be that the participants would, in turn, train others. Training for tribal members in data collection and language documentation will follow the known 'best practices' for archiving digital material.

A review of the 'best practices' for language documentation will be focused on the work done through the E-MELD (Electronic Metastructure for Endangered Languages Data) project. The metadata of the project will be encoded in accordance with the OLAC (Open Languages Archive Community) standards. Additionally, the data will follow 'best practice' recommendations established by the E-MELD project, as in Simons and Bird (2003). In particular, we will use Unicode for character encodings, and archive our texts and analysis using XML. All text encoding will be done using Unicode. Two major goals of the data preparation of this project will be to make the data portable, available for use by other linguists no matter what type of system they may be using and store the data using open source materials to prevent its use from being limited by proprietary software.

Audio and video data will be archived in uncompressed formats, WAV and AVI formats respectively. While the web files will be compressed for ease of access, the original files will be made available. If new information on 'best practice' methods comes to our attention before the onset of the documentation these methods will be applied as they become available.

A fieldwork relationship between the PI and the tribal participants in this project has long been established. Fieldwork for the current project will entail training the technical language assistants to collect data from other tribal members, namely fully fluent tribal elders, using video and audio technology. Working as a team, the project participants will meet on nine regularly occurring visits, in the field, to collect data, discuss the project and continue training of the tribal members involved. This fieldwork collaboration will be continued on-line using the OLE board technology (discussed below) in between site visits. Interested tribal members from both Ft. Mojave and Chemehuevi Valley will also be notified of these training sessions and will be invited to participate in on-site training for all interested CRIT members.

The use of the collected data will be applied in three specific ways:

- To add to the lexical databases for the electronic dictionaries which are currently being developed for both languages. In concert with work provided from a previous project, a beginning Chemehuevi dictionary was established in an on-line context using the 'best practices' presented through the E-MELD project (see references cited). The CRIT tribes have also requested that the Mohave dictionary (Munro et al, 1992) be adapted to an electronic format. As part of this project, lexical items from collected data (both video and audio) will be digitized and used to expand the lexical database for both of these languages. The data will also be immediately available for analysis by the project's technical language assistants who will receive training in descriptive linguistics related to this data. The data will be made available to the linguistic community when the data is posted on the web.
- The collected data will be available for tribal members to use in the development of multimedia language lessons. The CRIT tribes are interested in language revitalization but have struggled to gain support for these efforts. The CRIT library is well equipped with four computers and a wide range of software provided through the Bill and Melinda Gates Foundation grants to tribal libraries (2001). The hope of the CRIT library staff is that the technology available at the library will be used to present language-learning lessons for patrons from both tribes and that electronic versions of the tribal dictionaries will also be readily accessible at this site on CD and on the web.
- With the accessibility to language materials, comes the responsibility for control of these materials. The university team for the current proposal will work collaboratively with project participants to establish a clear model for tribal control over public access and use of these materials. This applies to the use of materials on-site in the reservation community but, perhaps more importantly, to the access

of on-line materials by the general public. As specified for the current grant, data collected for this project will have a public face and contribute to the broader linguistic community's knowledge of these two endangered languages. However, future data collection by tribal members needs to be considered and one possible model would be to assign different levels of access, i.e., some materials designated for tribal use (such as culturally sensitive materials represented in sacred texts or traditional songs) and some for public availability. Part of this proposal is therefore focused on providing a model for 'levels of access' related to the electronic documentation of indigenous language material.

PI's Long Range Goals

The PI expects to continue work with the CRIT community in her capacity as an applied linguist focusing primarily on language revitalization. The PI's current work on Mohave Coyote stories will be added to the database proposed for this project. Permission has been granted by the CRIT tribes to publish two of these stories as a bilingual text. This material will also be digitized and added to the database. However, the PI, who is nearing the end of a long career, recognizes the need to involve graduate students who can assume a long-term commitment to the community and who can assist in perpetuating the work on both Mohave and Chemehuevi. This project offers an opportunity for the PI to work in the field with the graduate students and gradually transfer the role of liaison between CRIT and the university's linguistic community to these highly qualified and committed graduate students so that this relationship can be maintained for many years to come.

Significance of Project

This project is significant because

- It will provide documentation in the form of audio and video recordings for two highly endangered languages.
- It will substantially increase the body of linguistic data currently available for both of these highly endangered languages.
- It will test and evaluate the use of a new technology, the OLE board, for sustaining fieldwork from a distance.
- It will increase the materials devoted to language pedagogy for both languages.
- It will explore the dimensions of training tribal members in audio and video documentation and lead to the development of a handbook for tribal members focused on language documentation.
- It will provide a model for how other communities can effectively document their own languages and manage the resulting archives in terms of public access and use.
- It will also clearly involve and encourage tribal members to exercise self-determination over the use of linguistic materials once electronically archived.

Work Plan

Year 1 will focus primarily on Chemehuevi and will involve hiring two technical language assistants who are tribally enrolled as Chemehuevi and are either speakers or semi-speakers of the language and a database specialist from the university. In Year 2, two Mohave technical language assistants will be added with the same qualifications while continuing the work of the database specialist. In Year 3, the Mohave team and database specialist will be retained. This plan allows for the project team to give clear focus to each language individually for at least one year. In Year 3, additional work will focus on the refining and presentation of digitally collected materials, on the completion of the handbook on language documentation and on the articulation of the model for access and use of electronically archived materials in indigenous communities.

Fieldwork

Initial fieldwork will take place during the summer of the first year. Three weeks will be necessary for the introduction of the graduate students to the community members and for the establishment of contacts and initial data collection. For the summer proceeding Year 2, two weeks of on-site fieldwork will take place. During the year, fieldwork will occur during the eight regular visits and will be supported through on-line contact using the OLE board.

OLE Board

The use of on-line technology to support fieldwork of the type proposed is in need of testing. However, the technology has been developed and used in a number of formats to support various types of on-line language learning and research (Reynaert et al, 2003).

The use of the OLE board (also called the MBS board) will be central to this project on three fronts:

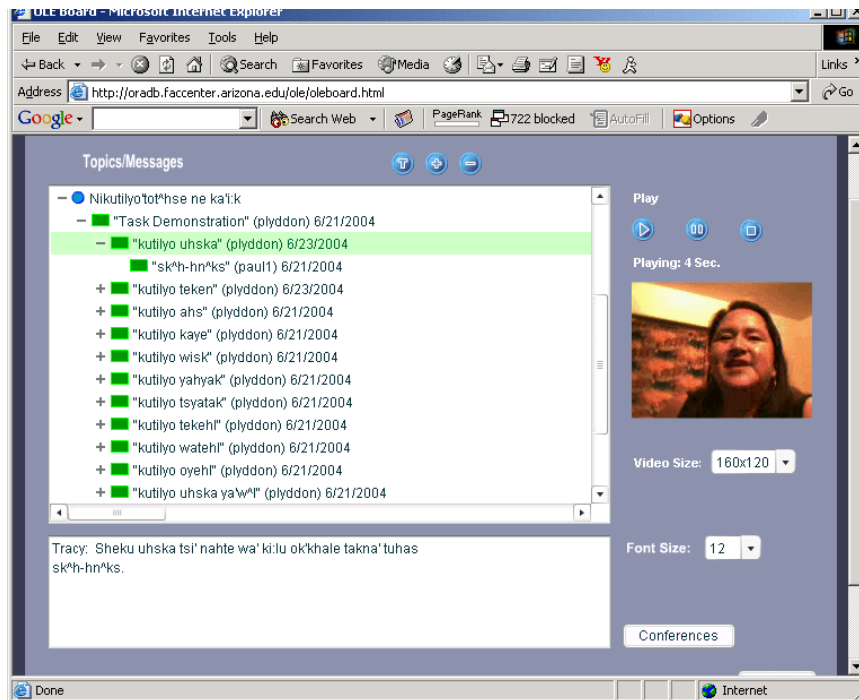
- As a tool for team members to use in order to maintain contact (with individuals or for group discussions).
- As a vehicle for conducting data collection from a distance.
- As a vehicle for the development of on-line language lessons.

This technology has been in existence for the past two years and continues to be refined. It is currently being employed primarily for on-line language instruction in two Native American languages: Mutsun (Warner, personal communication) and Oneida (Williams, 2004). The PI has been part of the development team for this technology since its inception and has already demonstrated that the OLE board is compatible with the CRIT library computers where it has been used in on-line discussions with tribal elders and CRIT library staff.

Unlike previous work by the PI which involved a synchronous system (Penfield, 2001), the OLE board provides an on-line presence marked by voice, video and text in an asynchronous format. This allows users to post messages or lessons using OLE at any time. The additional required equipment includes an individual web-cam and good

quality microphone. Users create threaded conversations as they respond to the instructor or other users. Most attractive for work with indigenous languages, is that it allows for oral language use on-line where literacy is either absent or not desired. Text may be included if needed or appropriate. Video from outside sources may also be included. Therefore, it is expected that videotapes collected for language documentation can be viewed and discussed with project participants on-line.

Figure 1: Display of the OLE board showing a vocabulary lesson in Oneida.



In terms of developing this technology, the focus has been on designing the OLE board in such a way that it delivers an instructional structure along with reliability, performance and scalability. OLE is made up of a four-tiered architecture. Its Internet service delivers a Macromedia Flash Client, which connects to a Cold Fusion MX Service via Flash Remoting. The Cold Fusion MX Service in turn connects to a DBMS, which organizes the instructional structures. The Flash MX Client also connects to a Flash Communication Service in order to retrieve streaming audio and video. From the experience with the initial prototyping of OLE, the PI has determined that this technology is appropriate for the needs that we have identified. On the issue of the portability of this software, Forger has indicated that, "The University of Arizona is committed to creating and using open source software where it best meets the requirements of our institution. The OLE Board has been developed on the Macromedia Flash Communication Server platform. While the Macromedia Server is a commercially-produced software, the OLE Board application is being developed as an open source tool." (Forger, personal communication).

All project participants will be given an account through the University of Arizona which will allow access to the use of the OLE board. Technological support for the OLE board

will be provided by the Learning Technologies Center, at the university, where this new software was developed and continues to be refined.

Part of the current proposal is to test this new software as a vehicle for on-line fieldwork. It is fully expected that on-line conversations using the OLE board will allow the fieldworkers to maintain contact with native speakers, resolve issues in translation and transcription, support work in language documentation and data collection from a distance and thus sustain momentum for the project over time and space.

Training

Training will take place as part of the fieldwork. This project is envisioned as a team effort and training will be available to tribal participants in the project during the three weeks of initial fieldwork and on each on-site visit following that. It will be beneficial for at least one member of the university team to visit the CRIT community for two days, on eight remaining regularly occurring visits, during the course of the academic year. All tribal project participants will be included in training from the beginning. Please note that an invitation to participate in training will also be extended to all members of CRIT and to contacts at both Ft. Mojave and Chemehuevi Valley.

The CRIT technical language assistants central to the project have been carefully selected. All four were awarded laptop computers and have received some basic computer training in the development of language lessons through previous funding. In addition, each has unique qualifications making their inclusion in this project particularly valuable and each has demonstrated a good working relationship with tribal elders. The proposed technical language assistants are

Mohave:

Amelia Flores is the CRIT tribal librarian and archivist. She has attended AILDI twice and has taken a course in basic descriptive linguistics.

Gilford Harper is a computer technician at the CRIT library. He has extensive training in the use of video technology as well.

Chemehuevi:

Johnny Hill, Jr. is the youngest known fully fluent Chemehuevi speaker and amateur videographer. Mr. Hill participated in the creation of the only Chemehuevi lesson to be formatted on MaxAuthor.

Nora Vasquez is a leader of the CRIT Chemehuevi language group and a former member of a Master-apprentice language team. Ms. Vasquez has also attended AILDI.

The graduate students have also been very carefully selected based on their unique qualifications relating to the goals of this project. They are:

Angelina Chtareva is a specialist in morphology and syntax who is beginning her dissertation on verbal morphology and argument structure in Uto-Aztecan languages. She also brings expertise in language pedagogy and a desire to assist in converting documented materials into usable materials for pedagogical purposes.

Benjamin Tucker is a specialist in phonetics and phonology with an emphasis on documentation and phonetic analysis. He also has experience recording and producing audio and video segments in a studio setting and in the field. Most recently, he has been an Editorial Associate for Linguistics Abstracts and has been responsible for the maintenance of an on-line database of over 35,000 abstracts and helped to design and implement the new XML based database for on-line use of databases. Ben also has experience teaching computer instruction at AILDI 2004, which included instruction in the use of MaxAuthor software.

Data Collection

Audio elicitations will be a main part of the data collection. All data will be recorded using a DAT recorder which does not compress the sound files unlike minidisk and other portable systems. Recordings will then be digitized and stored in WAV format on CDs. Care will be taken in the field to follow 'best practice' procedures in making audio recording of speech. For example, consideration will be given to the location of the recording in order to control for background noise. Recordings will then be transcribed and annotated.

Video recording will also be used. To do this we will use a DV video recorder. Additionally we will use a shot-gun microphone with a boom connected to the video camera to record the audio. This will allow us to bypass the microphone on the DV recorder which will not provide the sound quality necessary for linguistic analysis of the data. All video recording will be transcribed and annotated.

Print versions of the lexicon and other text related materials (e.g. linguistic description, pedagogical and transcriptions of discourses) will also be generated and made available in regular hard copy format but also in PDF and other digital formats.

Text Entry and Analysis

As data is collected and analyzed, the project team will focus on searching the archived and analyzed texts for lexical items to be included in dictionary databases, searching analyzed texts for grammatical forms and contexts to include in a proposed grammar database, and creating sound files to accompany dictionary entries. All terms and symbols used will be mapped to a common ontology of linguistic terms.

Materials Development

Beginning in Year 1, the project team will collaborate to create pedagogical materials as the appropriate data is collected. These materials will be formatted for computer-assisted language instruction and also be made available in hard-copy forms for use by the tribal elders and others who are engaged in immersion language classes. The

development of pedagogical materials will be an on-going process extending through all three years. Compilation of materials related to language documentation will also begin immediately but culminate in the third project year in the form of a handbook for tribal members focused on language documentation.

Development of materials

The outcomes of this project will have the following pedagogical impact:

- A training manual will be developed for tribal members who are interested in documenting indigenous languages through the use of video and audio technology. This manual will follow the design created for *Technology-Enhanced Language Revitalization* (Penfield, Cash Cash and Roberts, 2004). This design will foster user-friendly guide to language documentation incorporating examples from indigenous language communities. Conceptually, this volume would begin with a section on the place of technology in language documentation and a discussion of the main topics of concern related to language documentation and indigenous languages. Specifically, this first section would offer a presentation of the model developed during the current project which establishes guidelines for public access and use of electronically-archived materials. The following sections would introduce tribal members to methods of digital documentation and archiving, to the basic use of digital equipment and 'best practice' standards, to available open source software which supports language documentation and offer self-instruction in the basic use of these software products. The final section would be a listing of websites, other software, funding sources and additional references regarding language documentation for indigenous languages. This manual could be useful to indigenous people and to field linguists for use with their field colleagues.

Once completed, this manual would be available on-line in PDF format. Like the earlier handbook, this manual would become a core text for a proposed course on language documentation to be offered through the American Indian Language Development Institute. AILDI students have requested more technology-based courses (Zepeda, personal communication). There is a need and desire to learn the latest 'best practices' in terms of using technology to support indigenous language preservation through both documentation and pedagogy. Therefore, a portion of this course would focus on teaching standards established by OLAC and recommendations from the E-MELD project for language documentation and on how students can use the products of their documentation to create pedagogical materials. There is a broad audience at AILDI since students from many different tribal communities throughout the US and Canada attend this institute every summer.

- Project participants will receive training and assistance in adapting collected audio and video data to technology-based language lessons. The CRIT tribe has encouraged indigenous language learning in two contexts: 1) by tribal elders through immersion with youth either in school or community settings and 2) through computer-assisted language instruction at the tribal library. The present proposal aims to assist the CRIT tribes by offering pedagogical materials for both

of the above contexts. This will be done through the development of language lessons for computer-assisted instruction using presentational software such as PowerPoint or the more interactive software, MaxAuthor. The latter software, MaxAuthor, is an open source product of the Critical Languages Program at the University of Arizona where it is used for less commonly taught languages. There are now over 40 world languages which have CDs available for language instruction based on the MaxAuthor format. MaxAuthor easily adapts digitally-recorded video and audio files into its basic format (Scott Brill, personal communication). For the present project, we will be able to provide a training workshop for tribal members in the use of MaxAuthor each year of the project. The pedagogical need is great for both of these languages and the absence of fluent speakers available to teach makes the development of interactive, multimedia lessons not just attractive but necessary.

Technology is seen as an aid to language teaching in the CRIT community, but not a replacement for immersion classes which are occasionally taught. The adaptation of materials developed for the computer-based language lessons to hard-copy versions might serve as instructional support for elders or others who are teaching in language-immersion classes.

- Data collected and formatted for electronic archives and databases of Mohave and Chemehuevi will have pedagogical importance for students of linguistics as well. Importantly, both of these languages are from different language families. Neither language, however, has been the focus of linguistic work (phonetics, phonology, morphology, or syntax) since the late 1970s. Data from both of these languages needs to be assessed in light of current linguistic theory and practice.

Database and website

A database will be used for storage of the material and in order to export it more efficiently to web applications. Periodically, material in the database will be exported to an XML archive. A web version of the data will be made available to the public in order to allow access to the data. The data will be accompanied by the relevant digital source. The source will be available in compressed format on the website to decrease long download times. The original uncompressed files will also be made available upon request. The website will contain on-line lexicons of each language, discourse recorded with speakers, and other language related material.

This on-line database will be structured so that to access the material, one must first have permission. General access permission will be granted to anyone willing to agree to the terms of use as specified by the tribal community. We would like our native documenters to have the ability to document material that is related to religious ceremonies and other sacred information, and have the ability to limit access to certain materials to those with permission given directly from the tribe. This can be done by having different 'levels of access' so that only tribal members will be able to access this particular information. This decreases the usability for linguistic documentation of parts of the data collected but allows the tribe a method of restricting access to certain types of information. It also allows tribal members a way to preserve the sacredness of certain information but still have the ability to document and store it digitally.

Plans for preservation and archiving

We will collect elicitation from native speakers to produce word lists, lexicon, discourse with audio and video files supporting this data. The data produced will be made available via the web as authorized by the tribal participants in the project. A site that will be housed at the University of Arizona which will help allow for a more stable URL and for the use of most recent technology available through the university. All of the data collected and material produced (including but not limited to audio, video and text materials) as a result of this project will be archived and copies of this archive (including website versions) will be given to the University of Arizona (Arizona State Museum) archive and to the CRIT Tribal Library. Digital archiving will insure the preservation of linguistic material.

Project Products

Several products will result from this project:

- Digitally-mastered dictionaries for Mohave and Chemehuevi augmented with sound recordings and including more linguistic and analytic data.
- Text databases created for each language using the best known practices for archiving, processing and analyzing of linguistic data to provide even richer descriptions and documentation of these languages. Video recordings and audio files will be provided with these texts in order to provide future researchers a method to check content and assure material accountability.
- Language lessons available for computer-assisted language through Power Point or MaxAuthor software, and as hard copies for language immersion classes.
- A handbook designed specifically for tribal members focusing on the construction and use of electronic archiving. In addition to providing tribal members with instruction in data collection equipment and techniques, this handbook will outline the model for communities to use in determining access to digitally–compiled linguistic databases.

Dissemination

There are five projected avenues for dissemination:

1. In printed scholarly publications which will serve the community of linguists and indigenous people as well.
2. In the creation of a website which will allow for the dissemination of linguistic material and pedagogical material via the web.
3. In multimedia formats captured on CD-ROMs prepared using MaxAuthor and made available through the University of Arizona's Critical Language program.

4. The resulting handbook will be used as a text for the proposed AILDI course and also be made available on-line in PDF format.
5. In the form of archived material provided to the CRIT tribal library and to the Arizona State Museum archives.

Broader Impacts

- The most immediate broad impact will be the creation of multimedia databases for two highly endangered and only partially documented indigenous languages, Mohave and Chemehuevi. This work will contribute to the breadth for knowledge of Yuman and Numic languages generally. It will provide a wide-range of new data for comparison with other related languages and offer a searchable lexical database and text databases, supported with multimedia examples incorporating video and audio recordings. The broader linguistic community will benefit from this body of data immensely by providing access to the data and through our linguistic analysis. By involving tribal members in the data collection and construction of these databases, that impact will not only be felt and noticed by the linguistic community. This project culminates with the development of a 'community of practice' which includes linguists and tribal members interested in extending the documentation and analysis of Mohave and Chemehuevi
- A second major impact would be the use of the new technology, the OLE board, in support of fieldwork. This will further promote the use of technology for indigenous communities by involving tribal members in on-line research and language learning. This represents an expansion of previous work at CRIT and builds on the community's effort to use the tribal library as a center for technology and linguistic activity. The use of this technology has already been presented to AILDI as a potential tool for language instruction. The present project would test its application in field situations and as a tool for distance language research. This technology promises to have far-reaching impacts on any field of language study where oral language learning is valued. This technology has been proposed as a vehicle for other on-line language work involving sign language, literacy training, and language interpretation and translation. The present project will test its usability in the narrower context of linguistic fieldwork but in so doing will contribute understanding the value and use of technology in this capacity.
- The creation of a handbook will have a strong impact on at least two distinct audiences. It will be of significant value to indigenous communities interested in documenting their own languages and looking for a place to start. While designed primarily for an indigenous audience, it can also be used as an educational tool for field linguists to use in training research assistants in the field.
- The final important impact will be the development of a model for tribal communities which will set standards and protocols for the use of digitally-archived data by the public and the community. The growing body of electronic materials open to public access is of great concern to many tribal communities

and as yet there are no standard guidelines for the access and use of these materials. A current impediment to the electronic compilation of linguistic materials for endangered tribal languages is often community concern regarding its use and availability to the public. Setting clear, workable guidelines which will delineate which materials can be marked for public use, which materials should be limited to use by community members and which should be archived but restricted perhaps for future use, would offer a sense of security and allow for an increased body of linguistic data to be recorded and appropriately archived. By sharing this model with AILDI participants, it will impact a wide range of tribes in both the United States and Canada. inquisitive