
Plan Overview

A Data Management Plan created using DMPTool

Title: Language Endangerment and Revitalization at the Dialectal Level: A current analysis of Salasaca Kichwa (ISO 639-3 = qxl)

Creator: Robin Aronow

Affiliation: Temple University (temple.edu)

Funder: National Science Foundation (nsf.gov)

Funding opportunity number: EAGER

Template: NSF-SBE: Social, Behavioral, Economic Sciences

Last modified: 06-07-2015

Copyright information:

The above plan creator(s) have agreed that others may use as much of the text of this plan as they would like in their own plans, and customize it as necessary. You do not need to credit the creator(s) as the source of the language used, but using any of the plan's text does not imply that the creator(s) endorse, or have any relationship to, your project or proposal

Language Endangerment and Revitalization at the Dialectal Level: A current analysis of Salasaca Kichwa (ISO 639-3 = qxl)

Roles and responsibilities

The Data Management Plan should outline the rights and obligations of all parties as to their roles and responsibilities in the management and retention of research data. It must also consider changes to roles and responsibilities that will occur should a principal investigator or co-PI leave the institution.

The PI, R. Aronow, is responsible for the oversight and success of the DMP. There are three types of responsibilities to be carried out in the field: data gathering, data storage, and backing up electronic files. Our research team consists of four researchers. All researchers are required to store any field notes taken in a lockbox owned by the PI. Surveys are permitted to be administered by any researcher, but must be delegated by the PI for field storage on a daily basis. TASCAM and camcorder interviews and recordings may also be administered by any researcher. C. Rodgers will be responsible for ensuring that A) all equipment is functional and batteries are charged prior to daily field work, B) there is ample room on memory devices prior to daily field work, and C) electronic data is backed up to an external hard drive on a daily basis.

After the period of data retention, the PI will be responsible for data sharing and managing accounts with the institutions listed under 'Long Term Archiving and Data Sharing and Access'. B. Okum is responsible for alerting the participant community that data and results are available to access, as well as ensuring that up-to-date contact information is provided to participants and archiving and sharing institutions. Ensuring that the data is properly archived in private and institutional domains, responsibilities related to sharing/publishing, and any other decisions regarding the handling of data post-archiving and post-sharing will rest with the PI. Should the PI become unavailable for any reason, these responsibilities will be passed on to B. Okum.

Expected data

The Data Management Plan should describe the types of data, samples, physical collections, software, curriculum materials, and other materials to be produced in the course of the project. It should then describe the expected types of data to be retained.

Data will be generated in four forms: handwritten surveys, handwritten field notes, recorded audio, and recorded video/audio. Our surveys will be indexed and numbered prior to administration in the field for organizational purposes. The surveys are 5 pages each, and we expect to collect 40. They will be distributed and collected on site, and all 200 pages will be stored in a single lockbox on a daily basis until we return to the United States. Field notes taken in notebooks will likewise be stored in the lockbox daily until our return. The TASCAM audio recorder will be used to interview approximately 40 participants for 10 to 15 minutes each – an estimated total of 10 hours of audio that will be stored on a microSD and backed up daily on an external hard drive. We estimate the size of the audio to be less than 20GB. The Canon HD camcorder will be used in conjunction with the TASCAM for three reasons, 1) provide a secondary source of audio, 2) provide a visual component for the phonetic analysis, and 3) provide video footage that can be tailored and used for educational purposes as well as any academic presentations where our research is presented. The video footage will be stored on a second microSD card and will be backed up daily on the same external hard drive. We expect to record less than 100GB of video footage.

Period of data retention

SBE is committed to timely and rapid data distribution. However, it recognizes that types of data can vary widely and that acceptable norms also vary by scientific discipline. It is strongly committed, however, to the underlying principle of timely access, and applicants should address how this will be met in their DMP statement.

All data will be archived with restricted access for a period of 6 months or until results have been published.

Data format and dissemination

The Data Management Plan should describe data formats, media, and dissemination approaches that will be used to make data and metadata available to others. Policies for public access and sharing should be described, including provisions for appropriate protection of privacy, confidentiality, security, intellectual property, or other rights or requirements. Research centers and major partnerships with industry or other user communities must also address how data are to be shared and managed with partners, center members, and other major stakeholders.

Data Formats:

Surveys and field notes will be generated to .docx files and will be stored on an external hard drive. They will then be converted to a single PDF file each. PDF files are easily readable, are not easily retroactively edited, and consolidating all files will streamline the archiving process. TASCAM audio is recorded in uncompressed b-WAV format. WAV is the highest quality, least accessible and largest format, thus the audio will be analyzed

in WAV, and all publicly accessible recordings will be archived in 356kbs MP3 format. The video/audio camcorder records in AVI, which is compatible FinalCut-Pro.

Metadata Tools:

We will use EAD to generate metadata for domain-local archiving. After the results of our research are published, we will use a TEI tool to generate metadata for a public access archiving. This tool is optimal for creating machine-readable text for stories, oral history and other narratives by interview participants.

Sensitive Information:

A large amount of the data collected, including surveys and interviews, may contain personal information about research participants. In accordance to the obligations laid out by IRB protocol, we will obtain written and/or verbal consent prior to administering surveys or conducting interviews, which grants us permission to use and publish any information provided for educational and academic purposes. Our IRB protocol number is 22892; minimal risk research. Its status is under review with anticipated approval within 30 days.

Data storage and preservation of access

The Data Management Plan should describe physical and cyber resources and facilities that will be used for the effective preservation and storage of research data. These can include third party facilities and repositories.

Short Term Archiving:

During the retention period all data will be archived privately through Temple University Online OwlBox, with access to the PI and team researchers, and access to public members only with special permissions. Backups of data will be maintained by the PI on an external hard drive, which will be stored with completed surveys and field notes at the PI's office location at Temple University.

Data that will be submitted to scientific literature will be open access and available at no or minimal cost, including: A comparative study on the phonology of the target language, a synchronic study of the status of the target language, and an article proposing new language revitalization methodology.

Long Term Archiving and Data Sharing and Access:

After the period of retention our data will either be restricted or open access depending upon its content. Restricted access data, including raw audio and video recordings and physical and digitalized field notes and surveys will remain archived with Temple University Online OwlBox. This content will be openly accessible only to the research team and those who obtain special permission from the PI. A second complete set of all data will be archived and partially shared from the Temple University Library Digital Archives.

Content stored as 'open access' (not as restricted access) will be available to anyone at no cost upon request by contacting either the PI or B. Okum. This publicly available content includes details regarding research methodology and any additional files or media that should be publicly available but have not been archived or shared at another institution. All content shared with the foregoing institutions will also be available to anyone at no cost through Temple University Library Digital Archives.

Select data will be archived through The Archive of Indigenous Languages of Latin America (AILLA). This content will be freely and publicly accessible and includes: edited video and audio files and transcripts containing content that is relevant to the language, lore, history and culture of the participant community.

Select data will be archived through Ethnologue and SIL International. This content will be freely and publicly accessible and includes: data that pertains to the vitality of the target language and areas where the language is spoken.

Additional possible data management requirements

More stringent data management requirements may be specified in particular NSF solicitations or result from local policies and best practices at the PI's home institution. Additional requirements will be specified in the program solicitation and award conditions. Principal Investigators to be supported by such programs must discuss how they will meet these additional requirements in their Data Management Plans.

Question not answered.