

Designing & Sustaining Language Archives

This presentation will outline the design and development of the Pacific and Regional Archive for Digital Sources in Endangered Cultures (PARADISEC) as an example of a small digital archive, built by linguists and musicologists. Recognising that many analog recordings were not being looked after we found tapes that were essentially orphaned and got permission from their creators to preserve them. We built systems to automate as much of the work as possible to anticipate periods with no funding. This allowed users to enter metadata, controlled as much as possible by dropdown menus to avoid errors.

Archiving is not enough in itself to ensure that records are properly created, providing training to researchers is now an essential aspect of preparing material for accession

We initially built a catalog in a simple database system (FMPro) that allowed us to build our metadata system and to export a static repository to the Open Language Archives Community. Over time this was converted to an online PHP-SQL system, with a separate data management system which was combined a few years ago into our current system.

Each collection has conditions set by the depositor. There has to be clear consent from depositors to allow re-use of the material. We have developed various ways of presenting records, including large manuscript collections as page images and text and time-aligned media.

What has worked:

- locating, digitising and accessioning 4,000 hours of analog audio
- simple data structure
- building an online system that depositors can fill in themselves
- detailed online instructions for file naming, file formats
- collaborative project (3 universities) with distributed ingestion and a shared online operations manual
- establishing a collection of national and international significance (part of UNESCO's Memory of the World programme; European Data Seal of Approval)

Sites:

Page images: <http://paradisec.org.au/fieldnotes/AC2.htm>

Time-aligned media: <http://www.eopas.org/transcripts/21>

Thieberger, Nick. 2014. Curation of oral tradition from legacy recordings: an Australian example. In *Oral Tradition* Volume 28.2: 253-260, October 2013.

Thieberger, Nick and Linda Barwick. 2012. Keeping records of language diversity in Melanesia, the Pacific and Regional Archive for Digital Sources in Endangered Cultures (PARADISEC), pp. 239-253 in Nicholas Evans and Marian Klamer (eds). *Melanesian languages on the edge of Asia: Challenges for the 21st Century*. [LD&C Special Publication No. 5]. Honolulu: University of Hawai'i Press.