

<b>Institution: SOAS</b>
<b>Unit of Assessment: 28 Modern Languages and Linguistics</b>
<b>Title of case study: Documenting, Preserving and Sharing Global Linguistic Heritage (ELAR)</b>
<p><b>1. Summary of the impact</b> (indicative maximum 100 words)</p> <p>There is a growing, global crisis of language endangerment: At least half of the world's 7,000 languages are under threat. The Endangered Languages Project at SOAS supports the multimedia documentation of as many endangered languages as possible, drawing on research in the new field of documentary linguistics. A component part of the project, the Endangered Languages Archive (ELAR) preserves and makes available through managed access 10 terabytes of material from 160 endangered languages projects to date. It has benefitted a broad, international user base including endangered language speakers and community members, language activists, poets and others.</p>
<p><b>2. Underpinning research</b> (indicative maximum 500 words)</p> <p>In response to the urgent need to document endangered languages, the new discipline of Documentary Linguistics has emerged in the last fifteen years. It takes advantage of developments in information, media and communication technologies and is concerned with the theoretical, methodological, technological and ethical frameworks for documentation, preservation, support and dissemination of digital samples of endangered languages and cultures. Documentary Linguistics also directly supports endangered language speakers and communities by ensuring that outcomes are useful for their language and heritage goals, through encouraging collaboration, through respect for their rights in recorded materials, and by providing managed access to materials that they consider private, sacred or sensitive.</p> <p>The development of the online Endangered Languages Archive (ELAR) at SOAS occurred in parallel with the growth of Documentary Linguistics to which Professor Peter Austin and David Nathan have contributed substantially. The six publications listed below, dating from 2005, when ELAR was created by them and their team, are representative of their many, often foundational research outputs in the field, which treat both theoretical and practical topics that have guided the development of ELAR.</p> <p>Output a, for example, explores how language documentation materials can be better described, discovered, and interpreted. Metadata records (which categorise and describe electronic materials) when implemented in traditional fixed, limited schemas are insufficient to provide the level of contextualisation, understanding and access control needed for complex and diverse materials. Austin and Nathan propose a broader approach to metadata, enabling enhanced articulation of knowledge about the creation, context and content of resources, to reflect the perspectives of various participant and user groups, especially endangered language speakers. As a result, ELAR supports a flexible, nuanced approach to metadata. Output f extends this argumentation to apply to "metadocumentation" of the language documentation process itself.</p> <p>With the advent of Web 2.0 and social networking, Nathan investigated how an archive can provide a platform for information-sharing relationships between depositors, users, language speakers and others. ELAR implemented such a platform in 2009. Output e of 2010 motivates and describes this platform, proposing that social networking channels incorporated into ELAR provide ways for users to access sensitive materials through negotiation with depositors in conjunction with endangered language speakers.</p> <p>Threaded through the underpinning research is the issue of access protocol - the collection and implementation of information about private/sensitive materials, necessary because language documentation typically consists of recordings of spontaneous private conversations. As one of the</p>

## Impact case study (REF3b)

emerging archives for language documentation, ELAR had to deal with the sensitivities associated with the private, personal, sacred or other sensitive content of many archived recordings. ELAR was the first archive to research the needs for access management and to implement them using contemporary social networking architecture. ELAR's innovative platform allows depositors and users to negotiate directly about access as well as to share information about the archived languages, thus redefining the archive as a platform for information-based relationships between information providers and users. ELAR's unique system of graduated and negotiated access safeguards the rights of language speakers and also opens new exchange channels between those providing, and those wishing to access, materials (see ELAR's access protocol <http://www.elar-archive.org/using-elar/access-protocol.php>)

Other themes directly relevant to the design and development of ELAR in the underpinning research include:

- sustainability (ensuring long-term preservation and usability of electronic resources);
- maximising the quality of language documentation through informed methodologies, enhanced skills, and understanding of equipment and digital document technologies;
- progress towards enhanced metadata and "meta-documentation", especially related to the goals, history, processes, methods, dynamics and structures of language documentation projects.

### 3. References to the research (indicative maximum of six references)

- Nathan, David, and Peter K. Austin. 2005. "Reconceiving Metadata: Language Documentation through Thick and Thin." In *Language Documentation and Description Vol 2*, edited by Peter. K. Austin, 179-87. London: SOAS, 2005.
- Austin, Peter K. "Data and Language Documentation." In *Essentials of Language Documentation*, edited by Jost Gippert, Nikolaus Himmelmann and Ulrike Mosel, 87-112. Berlin: Mouton de Gruyter, 2006.
- Nathan, David. "Proficient, Permanent, or Pertinent: Aiming for Sustainability." In *Sustainable Data from Digital Sources: From Creation to Archive and Back*, edited by Linda Barwick and Tom Honeyman, 57-68. Sydney: Sydney University Press, 2006.
- Nathan, David. "Digital Archives: Essential Elements in the Workflow for Endangered Languages Documentation and Revitalisation." In *Language Documentation and Description*, Vol. 5, edited by Peter K. Austin, 103-19. London: SOAS, 2008.
- Nathan, David. "Archives 2.0 for Endangered Languages: from Disk Space to MySpace." *International Journal of Arts and Humanities Computing* 4.1-2 (2010): 111-24.
- Austin, Peter K. 2012. "Language Documentation and Meta-documentation." In *Keeping Languages Alive: Documentation, Pedagogy and Revitalization*, edited by Sarah Ogilvie and Mari Jones, 3-15. Cambridge: Cambridge University Press, 2012.

**Outputs b and c were submitted to the RAE 2008.**

**Output f is submitted in REF 2.**

### 4. Details of the impact (indicative maximum 750 words)

ELAR was established through a £20-million sponsorship of the Endangered Languages Project by Arcadia in 2002. This enabled research and practice in documentary linguistics to provide direct benefit to a range of users through the various activities of the archive. ELAR is not only part of the forefront project to document and support endangered languages worldwide, but is also an acknowledged leader in applying new media technologies in support of the goals of language documentation (1, below).

ELAR's catalogue serves an average of over 400 unique visitors per day with 1.4 million page-views per year. ELAR has a registered membership of 1,040 people from 85 countries, with

## Impact case study (REF3b)

membership increasing by about 5% per month (2, 3). Along with linguists, professional members include a number of anthropologists, archivists, and language teachers. In addition, a growing number of language and culture enthusiasts use the archive. A prominent example is the well-known New York-based poet Bob Holman who has used ELAR in 2012 and 2013 to create *Lost Wor(l)ds: an Endangered Cento*, a collage poem in which each line reproduces a single line of poetry from one of the world's endangered languages (4, 5).

Holman learned of ELAR as a visitor to "Endangered Languages Week" (ELW) at SOAS. Initiated by ELAR and the Endangered Languages Academic Programme in 2007, the annual week-long series of outreach events - which includes demonstrations, talks, performances and films - attracts hundreds of participants and audience members (6).

Of ELAR's membership, endangered language community members comprise 10%; for ELAR, such engagement of members of endangered language communities is vital, both as documenters of their own languages and as users of the archive. Eli Timan, a native speaker of the highly endangered and now globally dispersed Jewish Iraqi dialect of Arabic, provides an excellent example (7). Jews lived in Baghdad for more than 2,500 years and by the early 20th century they comprised one-third of the city's population. Of 137,000 Jews in Iraq in the 1940s, 124,000 had fled by 1952 as victims of persecution following the 1948 Arab-Israeli war. When Timan began documenting his native language in 2005, it was virtually a language of the diaspora alone, with only 10 Jews remaining in Baghdad. The last seven Jews in Iraq were named by Wikileaks in 2011, applying further pressure on them to flee.

Having learned of the Endangered Languages Project from a newspaper article, Timan visited SOAS in 2005, motivated to preserve his spoken language and create oral documentation of the recent history of Iraqi Jews. He took early retirement from a career in accountancy to visit diaspora communities in the UK, Canada and Israel to record his language and culture. Having successfully applied for funding, Peter Austin, David Nathan and others provided training in historical linguistics, phonetics, phonology, morphology, language revival, database management, and the use of recording equipment and linguistic documentation software.

Timan writes: "My life has been transformed since then as I embarked on my most important and interesting project." He has completed much of his project, and has deposited 34 fully-documented stories and interviews in the language with ELAR. He confirms the importance of this archived material to his community and beyond:

"Other members of the community have commented favourably on the material I deposited. They do not see it as a language project, but rather as a "Living history project". Recently there has been a great interest in "Iraqi Jews" from fellow Muslim Iraqis challenging the taboos of 35 years of Saddam's reign where nothing about us was allowed to be published as we were branded "Zionist spies". It may be that the archive will prove more valuable to those Iraqi Muslims anxious to discover a vital element of Iraq past history."

Crucial to Timan is the knowledge that his work is properly archived at ELAR and will remain accessible regardless of changes in technology. He has deposited materials with other Jewish Iraqi centres in the diaspora, but because they lack technical expertise, that work remains inaccessible.

Endangered language community members who wish to reconnect with and learn more about the languages their parents and older family members spoke are also users of the archive. A poignant example of a request to access materials documenting the language of the Wasco Native American Indian tribe of Oregon state, who now number approximately 200, was received in May 2013 (8):

"My name is G. My Uncles are N, H and J who are Wascos. Their Uncle was the last fullblood Chief. As a young boy, I was taken fishing ... by a tribal elder T who was (Wasco/Chinese) and he taught me Wasco ... I have been gone away from home for the past 22 years ... only coming home for funerals and vacations. I now live in Q and although I am closer to home, I am not around

**Impact case study (REF3b)**

anyone who speaks the language. Speaker C is a close friend and she is one of the keepers of the Language along with K. I would like to be able to access the language so I can practice what I have forgotten and keep the language alive."

**5. Sources to corroborate the impact** (indicative maximum of 10 references)

1. ELAR website: <http://elar-archive.org> [Most recently accessed 20.11.13].
2. Selected collection statistics: <http://www.elar-archive.org/about/statistics.php> [Most recently accessed 20.11.13].
3. Data relating to ELAR user numbers supplied directly from ELAR's server data. Usage reports can be made available upon request.
4. Bob Holman, poet and endangered language enthusiast who has used ELAR.
5. Link to description of the *Endangered Language Cento* project: <http://citylore.org/grassroots-poetry/endangered-poetry-initiative/endangered-cento/> [Please copy and paste link into browser] [Most recently accessed 20.11.13].
6. Endangered Languages Week: [www.hrelp.org/events](http://www.hrelp.org/events) [Most recently accessed 20.11.13].
7. Eli Timan, native speaker of the Jewish Iraqi language who has deposited his own language documentation project materials and outputs with ELAR.
8. Anonymised ELAR access request from member of the Wasco Native American community, which can be provided upon request.