
ELKL Conference, Central University of Jharkhand

Language documentation and language revitalisation: partners or just good friends?

Peter K. Austin

Department of Linguistics, SOAS, University of London

2017-02-25

© 2017 Peter K. Austin

Creative commons licence

Attribution-NonCommercial-NoDerivs

CC BY-NC-ND

Outline

- What is language documentation?
 - What is language description?
 - What is language revitalisation?
 - Relationships between these fields
 - Some current challenges and developments
 - Conclusions
-

Language documentation

- “concerned with the **methods, tools, and theoretical underpinnings** for compiling a **representative and lasting multipurpose** record of a natural language or one of its varieties” (Himmelman 1998)
 - Features:
 - *Focus on primary data*
 - *Accountability*
 - *Long-term storage and preservation of primary data*
 - *Interdisciplinary teams*
 - *Cooperation with and direct involvement of the speech community*
 - Outcome is **annotated and translated corpus** of archived representative materials on a language, cf. DoBeS, ELAR
-

Language revitalisation

- efforts to increase **language vitality** by taking action to:
 - increase the domains of use of a language and/or
 - Increase the number of speakers (often in the context of reversing language shift)
 - older than language documentation (serious work began in 1970s and 1980s among Maori, Native American groups and others)
 - Speech/language community members are often more interested in revitalisation than documentation
 - Often assumed revitalisation = formal language learning (school lessons, immersion)
-

Examples

- Amery, Rob 1998 WARRABARNA KAURNA: Reclaiming Aboriginal Languages from Written Historical Sources: A Kurna Case Study. PhD Thesis, Linguistics, University of Adelaide.
- Hinton and Hale 2001 'The green book of language revitalisation'
- Hinton 2002 'How to keep your language alive'
- Hinton 2013 'Bringing our languages home: revitalisation for families'



Components of documentation projects

- *Planning* – language, funding, fieldwork, equipment
 - *Recording* – of media and text (including metadata) in context
 - *Transfer* – to data management environment
 - *Adding value* – transcription, translation, annotation, notation and linking of metadata
 - *Archiving* – creating archival objects, assigning access and usage rights
 - *Mobilisation* – creation, publication and distribution of outputs
-

Documentation and description

- **language documentation:** systematic recording, transcription, translation and analysis of the broadest possible variety of spoken (and written) language samples collected within their appropriate social and cultural context
 - **language description:** grammar, dictionary, text collection, typically written for linguists
 - Ref: Himmelmann 1998, 2006, Woodbury 2003, 2011
-

Documentation and description

- documentation projects must rely on application of theoretical and descriptive linguistic techniques, to ensure that they are usable (i.e. have accessible entry points via transcription, translation and annotation) as well as to ensure that they are comprehensive
- only through linguistic analysis can we discover that some crucial speech genre, lexical form, grammatical paradigm or sentence construction is missing or under-represented in the documentary record
- without good analysis, recorded audio and video materials do not serve as data for any community of potential users. Similarly, linguistic description without documentary support risks being sterile, opaque and untestable (not to mention non-preserved for future generations and useless for language support)

As a consequence

- Documentation needs an epistemology for media capture – audio and video recording (see Nathan)
- Need to pay attention for good practices in recording – eg. microphone choice and spatiality in audio, framing-lighting-editing for video (“recording arts”)
- Concern for socio-cultural context (“ethnography of speaking”)
- Concern for data structuring and data management – eg. ‘portability’, relational modeling, XML
- Concern for ‘standards’ and cross-project comparability, especially typology and data mining
- Concern for ethics of research – documentation collects language use in “intimate” personal contexts, impacts on potential users and uses of documented speech events
- Changing models of research and relationships with people

Despite the rhetoric



- lone wolf linguists primarily focussed on language
- little real interdisciplinary interest
- the linguist decides what to deliver to communities (dictionaries, orthographies, story collections, etc.)

Issues in language revitalisation

- Revitalisation is bound up with politics, attitudes and ideologies
 - of speakers, activists, linguists ...
 - often poorly understood
 - ‘New uses’ and ‘new speakers’ may be contested or ideological
 - It has not attracted the same level of funding or recognition as documentation
 - excluded from many grants - seen as ‘unscientific’ praxis
 - often seen as a waste of time by mainstream linguists (‘linguistic social work’ – Newman 2003)
 - also fails to engage applied linguists (Cope 2012)
-

Issues in language revitalisation

- Under-theorised
 - Aims are often not articulated, activities are not evaluated
 - Poorly meta-documented – lack of good ethnographies of language revitalisation
-

Revitalisation and documentation – not a simple relationship

- ❑ ‘Revitalisation’ has been seen by some documentary linguists as a simple technical add-on
 - = orthography, dictionaries, videos, primers, multimedia
- ❑ But documentary methods and outputs are not always useful for revitalisation

“Work on language documentation to this point has tended to focus on what sorts of records are required to facilitate the creation of grammars, dictionaries, and texts, rather than, for instance, considering what kinds of records are required to adequately document patterns of variation in a community or to provide sufficient context to inform community efforts at language standardization.” (Childs, Good & Mitchell 2014)

Documentation and revitalisation

- most language documentation outputs are **unsuitable** for revitalisation:
 - inappropriate genres or topics
 - primarily speech of older fluent speakers (reflects linguists' ideology of "saving the language" or "getting the best language") – may be difficult for learners to process
 - no learner-directed speech (cf. Slow Italian website)
 - Observed language practices may not match perceived/stated ones
 - Some speakers/language activists may prefer 'folk linguistics' or purism to documentary evidence
 - Documentation which demonstrates low vitality, attrition, 'decline', variation and change may be unwelcome
-

Documentation **for** revitalisation

- what would language documentation look like if it was done with a goal of producing outputs for revitalisation?
 - *different genres*: conversation, not just narratives or rituals (Sugita, 2007; Amery, 2009), interactions (greetings, leave takings, ‘phatic communication’)
 - ‘*chunks*’ of language, from fixed/formulaic expressions to whole discourses (eg. ‘Welcome to Country’)
 - Dorian 1980 ‘semi-speaker’ – “a speaker of an endangered language who has a partial linguistic competence” but can sometimes appear ‘more competent’ because they can interact appropriately
 - research suggests proficient language users know a large number of formulaic sequences (e.g. Fillmore, Kay, and O’Connor, 1988; Wray, 2002), with Pawley and Syder (1983) suggesting that speakers know several hundred thousand of these sequences
-

From language loss to revitalisation – via documentation

- Documentation needs to be accessible and useful for revitalisation and teaching
 - e.g. conversation, child-directed language, functions, annotated for applied categories
 - Documentation and theorisation of revitalisation needs to be developed
 - and made accessible to activists and practitioners
 - More genuine collaboration is needed
 - including applied linguists and educators
 - Community and disciplinary ideologies need to be explored and taken into account
-

Thank you!

This talk reflects joint work with Julia Sallabank. I would also like to acknowledge the input of Lise Dobrin, Anthony Jukes, Yan Marquis, David Nathan, Candide Simard and other colleagues in discussions which informed the development of this presentation.

But address all criticisms to me!

Email: pa2@soas.ac.uk

References

Austin, Peter K. 2014. Language documentation in the 21st century. LIPP 3, 57-71.

Grenoble, Lenore. 2010. Language documentation and field linguistics: The state of the field. In Grenoble, Lenore A. and N. Louanna Furbee (eds.) *Language Documentation: Practice and values*, 289-309. Amsterdam: John Benjamins Publishing Company.

Good, Jeff. 2010. Valuing technology: Finding the linguist's place in a new technological universe. In Lenore A. Grenoble and N. Louanna Furbee (eds.) *Language Documentation: Practice and values*, 111-131. Amsterdam: John Benjamins Publishing Company.

Berge, Anna. 2010. Adequacy in documentation. In Grenoble, Lenore A. and N. Louanna Furbee (eds.) *Language Documentation: Practice and values*, 51-66. Amsterdam: John Benjamins Publishing Company, 2010.

Schultze-Bernd, Eva. 2006. Linguistic annotation. In Jost Gippert, Nikolaus P. Himmelmann and Ulrike Mosel (eds.) *Essentials of Language Documentation*, 213-251. Berlin: Mouton de Gruyter.

Woodbury, Anthony C. 2007. On thick translation in language documentation. In Peter K. Austin (ed.) *Language Documentation and Description*, vol. 4, 120-135. London: SOAS, University of London.

Woodbury, Anthony C. 2014. Archives and audiences: Toward making endangered language documentations people can read, use, understand, and admire. In David Nathan and Peter K. Austin (eds.) *Language Documentation and Description*, vol. 12. London: SOAS, University of London.
