Current Trends in Language Documentation

PETER K. AUSTIN, LENORE GRENOBLE

_____

_____

EL Publishing

For more EL Publishing articles and services:

      Website: http://www.elpublishing.org
      Terms of use: http://www.elpublishing.org/terms
      Submissions: http://www.elpublishing.org/submissions

Current trends in language documentation

Peter K. Austin and Lenore A. Grenoble

1. Defining documentary linguistics and language documentation [1]

Documentary linguistics is a newly emerging field of linguistics that is "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" (Gippert, Himmelmann and Mosel 2006:v). Documentary linguistics has developed over the last decade in large part in response to the urgent need to make an enduring record of the world's many endangered languages and to support speakers of these languages in their desire to maintain them (Whalen 2003, Austin 2007). It is also fueled by developments in information, communication and media technologies which make documentation and the preservation and dissemination of language data possible in ways which could not previously be envisioned. In addition it essentially also concerns itself with the roles of language speakers in documentary projects and their rights and needs in ways not previously considered within linguistics (see Thieberger and Musgrave, this volume).

Himmelmann (2006:15) identifies the following as important new features of documentary linguistics:

- *Focus on primary data* – language documentation concerns the collection and analysis of an array of primary language data to be made available for a wide range of users;

- *Explicit concern for accountability* – access to primary data and representations of it makes evaluation of linguistic analyses possible and expected;

- *Concern for long-term storage and preservation of primary data* – language documentation includes a focus on archiving in order to ensure that documentary materials are made available to potential users into the distant future;

---

[1] This paper was originally written for the 2006 Georgetown University Round Table on Languages and Linguistics "Endangered and Minority Languages and Language Varieties: Defining, Documenting and Developing", and then circulated among colleagues for feedback. For detailed comments on earlier drafts we are grateful to Oliver Bond, Anthony Jukes, and David Nathan, none of whom can be held responsible for errors or insitencies that remain.

- *Work in interdisciplinary teams* – documentation requires input and expertise from a range of disciplines and is not restricted to linguistics alone;

- *Close cooperation with and direct involvement of the speech community* – language documentation requires active and collaborative work with community members both as producers of language materials and as co-researchers.

We use the term *language documentation* to refer to the activities carried out by researchers and communities engaged in work that adopts a documentary linguistic approach. The historical genesis of the field of documentary linguistics has meant that the term 'language documentation' is sometimes used loosely to refer to any kind of language record, but documentary linguistics uses it in a more specific way, to refer to an activity with much larger and more specific goals. In particular, language documentation strives "to provide a *comprehensive* record of the linguistic practices characteristic of a given speech community" (Himmelmann 1998:166; our emphasis). Putting aside for the moment questions of how 'comprehensive' is to be interpreted (see below), we note that language documentation differs fundamentally and critically from language description. Language documentation seeks to record the linguistic practices and traditions of a speech community, along with speakers' metalinguistic knowledge of those practices and traditions. This includes 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 (Austin 2006, HRELP 2006). Analysis within language documentation is aimed at making the records, or rather the language data recorded, accessible to a broad range of potential users. This group includes not only linguists but also community members, who may not have first-hand knowledge of the documented language. The record is thus intended for posterity, and so some level of analysis is required, in particular glossing and translation into one or more languages of wider communication (see Evans and Sasse, this volume, for some of the challenges that entails), and systematic recording of metadata to make the archived document(s) findable and usable (Nathan and Austin 2004).

Language description typically involves the production of grammars, dictionaries, and collections of texts[2]. In contrast, language documentation is discourse-centred: its primary goal is the direct representation of a wide range of discourse types (Austin 2005, Woodbury 2003, Himmelmann 1998). Although description relies on documentation, it involves analysis of a

---

[2] Historically, and in some cases currently, some linguists use the term 'language documentation' to refer to what we are calling 'language description'. We attempt to be consistent in our usage of the two terms.

different order: description provides an understanding of language at a more abstract level, as a system of elements, rules, constructions and so on (see again Himmelmann 1998, 2002:48). Description and analysis are contingent by-products of documentation and will change and develop over time as research progresses (Woodbury 2003, Austin 2005). The intended audience of such products is typically linguists, and they are sometimes written in frameworks accessible only to trained linguists.

We take the core of a language documentation to be a corpus of audio and/or video materials with time-aligned transcription, multi-tier annotation, translation into a language of wider communication, and relevant metadata on context and use of the materials. Woodbury (2003) argues that the corpus will ideally be *large*, cover a *diverse* range of genres and contexts, be *expandable*, *opportunistic*, *portable*, *transparent*, *ethical* and *preservable*. As a result, documentation is increasingly done by teams, including community members, rather than 'lone wolf linguists'; both the technical skills and the amount of time required to create this corpus make it difficult for a single linguist, working alone in the field, to achieve.

2. Uses and users of language documentation

The documentation of a language can provide an empirical basis for a wide range of activities. Two of the most obvious uses are linguistic research and language revitalization. By virtue of its comprehensive nature, the documentation can provide data for research in all subfields of linguistics, ranging from phonology, morphology, syntax and discourse to sociolinguistics, typology and historical reconstruction. In addition, because the documentation consists of a range of discourse types, it can provide a database for the analysis of oral literature and folklore, as well as poetics, and the metrical and musical aspects of oral literature. The content of the documentation corpus can also ideally service work in oral history and anthropology; it provides information about a range of aspects of culture, kinship relations, rituals and ceremonies, knowledge of the environment and so on (Himmelmann 1998, Austin 2005).

Language education and revitalization are often of great interest to the language community, and increasingly researchers are expected to pay attention to them in framing their research projects. The documentation of a language is aimed at producing the most comprehensive record of linguistic practices possible, and so for communities who have lost their language, or whose language is seriously endangered, the documentation must ideally provide all the information they need to revitalise it. This is often the factor which motivates community members to collaborate with language

documenters in the first place; the resulting product needs to be of use to them to pursue their own goals.

To a certain extent, the nature of the documentation corpus is driven by the anticipated users of that corpus. As we suggested earlier, documentary linguistics differs from traditional descriptive linguistics in that the resulting corpora are created with the explicit intention of being of interest to, and accessible to, multiple audiences. This means that the collection, analysis and presentation of data should be useful not only for professional linguists, but also for research into the socio-cultural life of a given community, as well as a means of providing support for languages where community members and others wish to do so.

Moreover, the multiplicity of uses means that the documentation corpus needs to be structured, encoded and represented in a format that ensures it will be accessible to a broad audience of users. It should be analyzed and processed in such a way that it can be understood by both linguists and researchers of other disciplines, with analytical categories and decisions being made transparent. At the same time, it needs to be in a format which will make it usable by members of the speaker community. Ideally then it should be organised (and archived) in a way which does not require any prior knowledge of the language in question to guarantee accessibility by all potential users. This requires clear transcription, annotation, and translation into at least one language of wider communication, along with clear information about the nature of the representation, e.g. the transcription and annotation scheme adopted. Although we have been careful to distinguish between documentation and description, it is important to note that annotation, such as morphemic glossing, requires both analysis and description; it is a type of description in its own right (called by Nathan and Austin 2004 "thick metadata" to distinguish it from the "thin metadata" typically serving cataloguing and discovery purposes). Because a fundamental part of language documentation is long-term usability, such annotation should rest on widely accepted categories and be as free as possible from framework-specific terminology. Ideally, the categories and their definition should be included as part of the documentation. We may compare this with, for example, a large number of descriptive analyses written in the 1960's and 1970's in a tagmemic framework (especially popular among SIL linguists) that makes them difficult to use now, especially for community members.

There is also a tension between the needs of different groups of potential users and how the documentation corpus may best meet those needs. For example, in terms of the translation language, at present it is clear that English is the single language of wider communication which can guarantee global accessibility by linguists and other academics. At the same time, the language of wider communication which is of greatest use to community members may

well be the majority language which is replacing their endangered language, e.g. Spanish in Latin America, or Mandarin, Portuguese, and Russian and other locally dominant languages elsewhere. For community members, then, documentation using English as an interface might require them to learn yet another language of wider communication to access their community language, or to hire an interpreter to make sense of the documentary materials. Time constraints on individual linguists or documentation teams may make it impractical to create a trilingual corpus, but that may often be the ideal. Resolving this issue is complicated. Commonsense considerations may suggest that any linguist working on, say, some variety of Zapotec, might be expected to know Spanish, and so translation into only this one language might suffice (the same would hold for Russian in the case of Siberian indigenous languages, and so on). But if the corpus is to be truly accessible to typologists or others doing any kind of comparative work (in anthropology or folklore, for example), then the expectation of such knowledge is less well-founded. In addition, the requirements of individual archives where the archival materials from a documentary corpus is to be housed might be a consideration. Those housed in English-speaking countries might expect a corpus to be translated into English and may require that all metadata be prepared in English, for practical considerations of the archive's own work (this should not be an a priori expectation however, and Munro 2006 describes an archive information management system in which annotation and metadata can be prepared in any language of the user's choice).

Finally, there are a host of issues surrounding access rights and the use of language documentation. Any documentation project should respect intellectual property rights, moral rights, as well as both individual and cultural sensitivities about access and use (Austin 2005, Dwyer 2006, AILLA 2006, Thieberger and Musgrave, this volume). While at first blush this may be an obvious statement, in fact many researchers have historically ignored the intellectual property rights and access needs of the communities themselves. As a result, a number of communities have felt that they have contributed to the careers of external researchers without themselves seeing any benefits. In the extreme, there is a sense of loss of control over one's own cultural heritage and intellectual property.

Our emphasis here on the importance of a variety of users and uses for documentation projects stems from the fact that there is a growing awareness that linguistics has crucial stakeholders well beyond the academic community (see also Dwyer 2006:35-37). Important stakeholders are to be found in the endangered language communities themselves, and beyond; the very design of the documentation project needs to take into account the needs of these many stakeholders and the ways in which they will or will not be able to access the corpus. This is not necessarily easy to achieve. The differing needs and desires of the linguistic community on the one hand and the speaker

community on the other mean that the two groups can strive toward very different outcomes. For this reason, it is current practice to include community members in a project from its very conceptualization, so that they are full collaborators in the documentation (see Grinevald 2003).

3. Documentation in the context of modern technology and linguistic theory

Language documentation has emerged at this moment in history due to a combination of factors (see also Woodbury 2003). These include advances in technology; an increased attention to linguistic data, along with a new attention to linguistic diversity; a growing interest in and concern for archives; the emergence of extensive funding resources; and recognition of the needs of other stakeholders, those outside of professional linguists, who are interested in language documentation as a necessary first step toward language maintenance and revitalization, or as a safeguard against complete language loss.

3.1 Advances in Technology

Relatively recent developments in information, communication and media technology have made it possible to make, process and distribute high-quality audio and video recordings more affordably and more easily than ever before. We have new technology which enables linguistic data recording, digital capture and manipulation, representation and maintenance at relatively low cost and with relatively low technical training (see Gibbon 2002 for an example). Moreover, the emphasis on digital recording and digital maintenance of archives means that such data are readily portable and transferable (in line with the recommendations of Bird and Simons 2003). In other words, although some have argued that the technology itself drives language documentation, in fact it is the needs of the language documenter and the community members which is driving the uses of technology (see Nathan 2006 on the development of 'thick interfaces' for multimedia to access documentary data, and Good, this volume).

3.2 Linguistic Diversity and Data

To an increasing extent contemporary linguistics, including descriptive, theoretical and applied linguistics, is paying attention to the diversity of the world's linguistic ecology and moving away from a focus on large languages and literary forms. The development of language typology, for example, has emphasised empirical research methods based on a wide sample of language

data, and the presentation of materials in a way that makes research transparent and replicable. We have also seen the publication of grammars with linked corpora, such as Heath (1984) in a more traditional book format but which includes hundreds of hypertextual links to a published text collection, or more recent developments such as Thieberger (2006) which includes a CD-ROM with recordings of all the example sentences in the grammar.

3.3 Archiving

Archives are crucial in any documentation project, both as repositories for data and for the provision of advice and technical support to language documenters. The recent development of digital archives has made it both cheaper and more feasible to process and store the data on lesser studied languages more thoroughly, and we have also seen the involvement of dedicated professionals who can maintain and provide access to the digital archives, and support language documentation projects through advice and training. Among new initiatives, we can identify:

- AILLA – the Archive of the Indigenous Languages of the Americas, based at the University of Texas and dedicated to archiving data with a focus on Latin American languages

- DoBeS archive at the Max Planck Institute for Psycholinguistics, Nijmegen, which is dedicated to supporting the DoBeS projects sponsored by the Volkswagen Stiftung

- ELAR – the Endangered Languages Archive based at SOAS which receives data from researchers funded by the Endangered Languages Documentation Programme, students of the Endangered Languages Academic Programme, and others

- PARADISEC – the Pacific and Regional Archive for Digital Sources in Endangered Cultures, based at the University of Sydney, Australia, which focuses on data from communities in the Pacific Islands and neighbouring areas

There are also a number of more local digital archives, such as the Austrian Academy of Sciences Phonogramarchiv, Alaska Native Languages Centre archive, Survey of California and Other Indian Languages archive, and the Australian Institute of Aboriginal and Torres Strait Islander Studies audio-visual archive, who serve researchers and community members for languages of those particular areas.

Since 2000 increasing attention has been paid to the necessities of collaborating between and managing digital archives world-wide, and this has lead to the establishment of OLAC (the Open Language Archives Community) and DELAMAN (Digital Endangered Languages and Musics

Archive Network). This latter group through their website, discussion lists and occasional meetings aim to stimulate interaction about practical matters that result from the experiences of fieldworkers and archivists, and to act as an information clearinghouse. The E-MELD (Electronic Metastructure for Endangered Language Data) project has also aimed at promoting good practice in documentation and archiving.

All documentation projects should be conceived with an eye toward the ultimate deposit of the recorded data and analysis in an archive. Archivists can help support this process, including providing assistance in digitizing language resources, training documentation teams in the use of digital tools to record language, and to advising them on how to make their data archive-ready. Moreover, archives are responsible for maintaining and updating the digital data in the face of ever-changing technology. They also have an important role for access to the data, allowing access by authorised users and prohibiting unauthorised access. This latter point is particularly important in the face of serious concerns by speakers and community members about intellectual property rights, and by issues surrounding the accessibility of culturally or personally sensitive material.

3.4 Funding

Just as academic research in general is driven, to a certain extent, by funding opportunities, it is also fair to say that the success (and indeed the realization) of any documentation project is dependent upon funding. At present, a number of organizations provide funding for individual linguists or teams to go to the field to collect data, or to work with existing (legacy) materials to meet the goals of language documentation. These include the following:

- Documenting Endangered Languages (DEL) grants, a joint collaborative effort between the National Science Foundation, and the National Endowment for the Humanities in the United States

- the Volkswagen-Stiftung and its DoBeS project; and

- the Hans Rausing Endangered Languages Project through its Endangered Languages Documentation Programme.

Small research grants are provided by the Endangered Languages Fund (based in the US), the Foundation for Endangered Languages (based in the UK) and the Gesellschaft für Bedrohte Sprachen (based in Germany). The DEL program, relying on government funding, requires that at least the primary investigator be a US citizen or resident; the other granting agencies fund internationally.

Our own association is with the Hans Rausing Endangered Languages Project (HRELP) and so we will describe it in more detail here (further

information is available at http://www.hrelp.org.) HRELP was established in 2002 with support from the Lisbet Rausing Charitable Fund (now Arcadia) and is based at SOAS, University of London. It consists of three interrelated programs, the Endangered Languages Documentation Programme (ELDP); the Endangered Languages Academic Programme (ELAP); and the Endangered Languages Archive (ELAR). ELDP aims to distribute approximately £1million per year in five types of grants (individual postdoctoral fellowships; major documentation projects; individual graduate studentships; pilot project grants and field trip grants). At present, there are 110 teams of researchers around the world documenting languages and cultures with support from ELDP. The varying categories of ELDP grants are designed to capture the differing needs of researchers, ranging from preliminary exploratory fieldtrips on one end of the spectrum to large multi-year, multi-phased collaborative efforts. The Academic Programme (ELAP) offers postgraduate degrees in Field Linguistics and Language Documentation and Description, along with workshops and training courses; it is notable for offering a one-year MA degree in documentation and description with no previous linguistic training required. This programme underscores the commitment of not only HRELP, but of the field of documentary linguistics as a whole, to the training of a much larger body of field workers to meet the pressing demands of language documentation in the face of imminent language endangerment. The Endangered Language Archive (ELAR) archives and disseminates language documentation materials collected by researchers from the other two programmes, as well as others. ELAR, together with ELAP, runs training courses for grantees funded by ELDP, and publishes a range of book and multimedia materials on the theory and practice of language documentation.

4. Some outstanding issues

In concluding this paper, we would like to outline some unresolved theoretical issues with regard to language documentation, in the hope that they will spark further research and debate. In this section, we will focus on just four of the major questions which face documentary linguistics today: (1) the definition of a comprehensive record of a language; (2) issues of the quality of the documentation itself; (3) the boundaries between documentation and description; and (4) interdisciplinarity and cross-discipline collaboration. There are other theoretical and practical issues which remain to be resolved and will undoubtedly emerge as practices of language documentation develop.

4.1 The 'comprehensive' record

Language documentation is defined as providing a "comprehensive record" of a language (Himmelmann 1998, 2002), but it is unclear how "comprehensive" is to be understood. On a theoretical level, once can define "comprehensive" documentation as the collection of representative texts of *all* discourse types, *all* registers and genres, from speakers representing *all* ages, generations, socioeconomic classes, and so on. On a practical level, however, there are concrete limitations to the range and amount of language data which can be collected. Most linguists cannot devote their entire careers to time in the field, which would be required for a truly thorough collection and analysis of data.

It is clear that the success of a documentation project rests on intimate collaboration with community members. In the ideal, they can be trained to be engaged in data collection themselves, thereby expediting the process (for an example see Florey 2004). Even if this is not possible, community members can direct (external) linguists to varying discourse types and to differing speech patterns. Himmelmann (2002:66) identifies five major types of communicative events ranged along a continuum from unplanned to planned, as follows:

Parameter	Major Types	Examples
unplanned	exclamative	'ouch!', 'fire!'
	directive	'scalpel!'
	conversational	greetings, small talk, chat, discussion, interview
	monological	narrative, description, speech, formal address
planned	ritual	litany
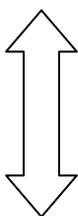
It is not clear to us that this typology is either truly comprehensive or that it is appropriate for all language situations (thus among some groups narrative is typically dialogical, and where does 'story telling' constantly interrupted by an audience responding to the story teller lie?). Further research to develop and expand this typology is needed.

4.2 Defining 'quality' documentation

As we have noted, the field of documentary linguistics is a young one and it remains unclear what the outcomes of language documentation can be or how quality of those outcomes is to be assessed. There is a tendency among some researchers to equate documentation outcomes with archival objects (part of what David Nathan has termed 'archivism'), that is, the number and volume of recorded digital audio and/or video files and their related transcription, annotation, translation and metadata. However, mere quantity of objects is not a proxy for quality of research. Equally, some would argue that outcomes which contribute to language maintenance and revitalization are the true measure of the quality of a documentation project (what better success of an endangered language project than that the language continues to be used?). Again, how to assess these is an open question at present.

4.3 Documentation versus description

Although Himmelmann (1998, 2002) has tried to delineate the different spheres of interest and research methods of language documentation and language description, it is unclear to us whether such a separation is truly meaningful, and even if it is where the boundaries between the two might lie. Documentation projects must rely on application of descriptive linguistic techniques, if only 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. It is only through linguistic analysis that we can 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 is sterile, opaque and untestable.

4.4 Interdisciplinarity

Himmelmann and others have pointed to the importance of taking a multidisciplinary perspective in language documentation and drawing in researchers, theories and methods from a wide range of areas, including anthropology, musicology, psychology, ecology, applied linguistics and so on (see Harrison 2005, Barwick 2005, Coelho 2005, Eisenbeiss 2005 for examples). However, in our experience, true interdisciplinary research, especially in teams carrying out fieldwork in remote locations, is difficult to achieve, both because of theoretically different orientations, and practical differences in approach (ranging from the trivial where linguists' and anthropologists' practices concerning payments for consultants traditionally

have differed, to more significant differences in academic paradigm that make communication and understanding fraught). Whether these problems can be resolved in meaningful ways remains open.

5. Conclusions

The past ten years has seen the emergence of a new field of linguistic research with the development of documentary linguistics and language documentation. For many researchers and communities, especially those speaking endangered languages, the focus of language work has shifted to a new attention to recording and analyzing language in use in ways that serves a wide range of constituencies, not least the speaker communities themselves. A number of influences have lead to the development of this new field, and it has benefited from developments in technology, a change in relations between researchers and those whose languages they study, and a change in the vision of what the goals and uses of linguistic research can be. A number of outstanding issues remain, and we can be sure that further challenges will appear, and be addressed as documentary linguistics as a field matures in coming years.

References

Archive of the Indigenous Languages of the Americas 2006. Intellectual property rights. Available at http://www.ailla.utexas.org/site/ipr.html accessed 15th May 2006

Austin, Peter K. 2005. Training for language documentation – the SOAS experience. Paper presented at the Workshop on Language Documentation, Summer Institute of Linguistics, Harvard/MIT, July 2005.

Austin, Peter K. 2006. Data and language documentation. In Jost Gippert, Nikolaus Himmelmann and Ulrike Mosel (eds.) *Essentials of Language Documentation* (*Trends in Linguistics. Studies and Monographs*, 178), 87-112. Berlin: Mouton de Gruyter.

Austin, Peter K. 2007. Survival of Languages. In Emily F. Shuckburgh (ed.) *Survival: Darwin College Lectures*. Cambridge: Cambridge University Press.

Barwick, Linda 2005. A musicologist's wishlist: some issues, practices, and practicalities in musical aspects of language documentation. In Peter K. Austin (ed.) *Language Documentation and Description, Vol 3*, 53-62. London: School of Oriental and African Studies.

Bird, Stephen and Gary Simons 2003. Seven dimensions of portability for language documentation and description. *Language* 79/3:557-82.

Coelho, Gail 2005. Language documentation and ecology: areas of interaction. In Peter K. Austin (ed.) *Language Documentation and*

Description, Vol 3, 63-74**.** London: School of Oriental and African Studies.

Dwyer, Arienne M. 2006. Ethics and practicalities of cooperative fieldwork and analysis. In Jost Gippert, Nikolaus P. Himmelmann and Ulrike Mosel (eds.) 2006. *Essentials of language documentation* (*Trends in Linguistics. Studies and Monographs*, 178), 31-66. Berlin: Mouton de Gruyter.

Eisebeiss, Sonja 2005. Documenting child language. In Peter K. Austin (ed.) *Language Documentation and Description, Vol. 3*, 106-140. London: School of Oriental and African Studies.

Florey, Margaret 2004. Countering purism: confronting the emergence of new varieties in a training program for community language workers. In Peter K. Austin (ed.) *Language Documentation and Description Vol. 2*, 9-27. London: School of Oriental and African Studies.

Gibbon, Dafydd 2002. Workable Efficient Language Documentation: a Report and a Vision, available at http://www.spectrum.uni-bielefeld.de/LangDoc/EGA/Docs/elsnews_WELD_oct2002_HTML/, accessed 15th May 2006.

Gippert, Jost, Nikolaus P. Himmelmann and Ulrike Mosel (eds.) 2006. *Essentials of language documentation* (*Trends in Linguistics. Studies and Monographs*, 178). Berlin: Mouton de Gruyter.

Grinevald, Colette 2003. Speakers and documentation of endangered languages. In Peter K. Austin (ed.) *Language Documentation and Description, Vol 1*, 52-72. London: School of Oriental and African Studies.

Hans Rausing Endangered Languages Project 2006. What is language documentation?
Available at http://www.hrelp.org/documentation/whatisit/, accessed 15th May 2006.

Harrison, K. David 2005. Ethnographically informed language documentation. In Peter K. Austin (ed.) Language *Documentation and Description, Vol 3*, 22-41. London: School of Oriental and African Studies.

Heath, Jeffrey 1984. *A Functional Grammar of Nunggubuyu*. Canberra: Australian Institute of Aboriginal and Studies.

Himmelmann, Nikolaus P. 1998. Documentary and descriptive linguistics. *Linguistics* 36:161-95.

Himmelmann, Nikolaus P. 2002. Documentary and descriptive linguistics (full version). In Osamu Sakiyama and Fubito Endo (eds.) *Lectures on Endangered Languages: 5.* Kyoto: Endangered Languages of the Pacific Rim, 37-83.

Himmelmann, Nikolaus P. 2006. Language documentation: What is it and what is it good for? In Jost Gippert, Nikolaus P. Himmelmann and Ulrike Mosel (eds.) 2006. *Essentials of Language Documentation* (*Trends in Linguistics. Studies and Monographs*, 178), 1-30. Berlin: Mouton de Gruyter.

Munro, Robert 2006. Multilingual metadata: between translation and coexistent annotations. Paper presented at ELAP workshop on Meaning and Translation in Language Documentation, 11th February 2006.

Nathan, David 2006. Thick interfaces: Mobilizing language documentation with multimedia. In Jost Gippert, Nikolaus P. Himmelmann and Ulrike Mosel (eds.) *Essentials of Language Documentation* (*Trends in Linguistics. Studies and Monographs*, 178), 363-379. Berlin: Mouton de Gruyter.

Nathan, David and Peter K. Austin 2004. Reconceiving metadata: language documentation through thick and thin. In Peter K. Austin (ed.) *Language Documentation and Description, Vol. 2*, 179-187. London: School of Oriental and African Studies.

Thieberger, Nicholas 2006. *A Grammar of South Efate: An Oceanic Language of Vanuatu*. Oceanic Linguistics Special Publication, No. 33. Honolulu: University of Hawai'i Press.

Whalen, Doug 2003. How the study of endangered languages will revolutionise linguistics. International Congress of Linguists, Prague.

Woodbury, Tony 2003. Defining documentary linguistics. In Peter K. Austin (ed.), *Language Documentation and Description*, *Vol. 1,* 35-51. London: School of Oriental and African Studies.

.