

Language Documentation and Description

ISSN 1740-6234

This article appears in: *Language Documentation and Description*, vol 5. Editor: Peter K. Austin

The role of digital video in language documentation

LOUISE ASHMORE

Cite this article: Louise Ashmore (2008). The role of digital video in language documentation. In Peter K. Austin (ed.) *Language Documentation and Description*, vol 5. London: SOAS. pp. 77-102

Link to this article: <http://www.elpublishing.org/PID/064>

This electronic version first published: July 2014



This article is published under a Creative Commons License CC-BY-NC (Attribution-NonCommercial). The licence permits users to use, reproduce, disseminate or display the article provided that the author is attributed as the original creator and that the reuse is restricted to non-commercial purposes i.e. research or educational use. See <http://creativecommons.org/licenses/by-nc/4.0/>

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

The role of digital video in language documentation

Louise Ashmore

1. Overview¹

For the purposes of language documentation, multimedia recordings (audio and video) have been recommended as the basis for any documentary corpus (Himmelmann 2002:12; Lehmann 2001:9; Wittenburg 2003:124; McConvell 2003, 2007:2). As audio-visual technology becomes more accessible in terms of cost, storage and delivery, language researchers are increasingly adopting video media for the collection and analysis of language data. Although there will always be situations where video is inappropriate for recording communicative events (e.g. for political, cultural, ethical and practical reasons), video media presents a number of opportunities for language documentation work.

In contrast to audio and text records video has the potential to visually record the situational use of language, including spatially-encoded linguistic information, gesture and eye gaze. Video material can also be an invaluable transcription aid for recording complex, multi-speaker events and linguistic annotation tools are available to assist with managing and analysing video data (e.g. ELAN²). Importantly, speech communities are reported to prefer video as a record and resource (McConvell 2007; Wittenburg 2007). As a result, video has been promoted as ‘the ideal recording device’ for language documentation (Himmelmann 2002:12).

However, the popularity of video as a research tool also raises ethical and methodological challenges for the emerging field of language documentation. Thieberger and Musgrave (2007) identify new ethical issues arising from digital documentation and the long-term storage of audio-visual data. Digital documentation presents particular questions about informed consent, ownership, rights of access and reproduction and the return of materials collected within language documentation projects (Thieberger and Musgrave 2007:27). Concerns about the costs and complexities of archiving large

¹ This paper is developed from research presented in my MA thesis (Ashmore, 2006). I am grateful to Peter Austin, David Nathan and David Harrison for feedback and comments on this topic and to David Harrison and Anthony Jukes for discussing their fieldwork experiences and providing examples of field recordings. Any errors, omissions or inconsistencies are my own. I would also like to acknowledge staff at ELAR, filmmaker Simon Atkins and participants of the ELDP Grantee Training Workshops (2006-7) who generously provided information and stimulating discussion about their experiences with film and video.

² See <http://www.lat-mpi.eu/tools/elan>

amounts of video data and a lack of agreed upon standards for recording and processing video have also been raised (Nathan 2007). In addition, as video technology continues to develop, new genres and representations will be produced and new working practices adopted. Developing goals, methods and evaluative criteria appropriate for the different ways video is used in language documentation work, requires ongoing discussion.

In part, the uptake of video is a response to the diverse aims and outcomes of language documentation. Language documentation as a theory and practice is still being actively discussed and contested from the different perspectives of the language documentation community including linguists, archivists and speakers of endangered and minority languages. A prevalent theme in the development of language documentation has been the availability of digital technology to produce more comprehensive records of the linguistic practices and traditions of a given language community, particularly in situations of language endangerment (Himmelman 2002, 2006). Emphasis has also been placed on how primary research data can be representative, lasting and multipurpose, in the sense of being able to be used by different users for different uses (Gippert, Himmelman and Mosel 2007:v). For example, data can be of use for the academic community for future descriptive work and for community-oriented language revival goals.

Digital language archives are playing an increasing role in generating discussion about the collection, management and long-term storage of data resulting from documentation work.³ Equally, the requirements of funding organisations are shaping the ways that projects are conceived and implemented. Dobrin, Austin and Nathan (2007) point out how ‘documentary research is now frequently framed around the archival materials to result from it, a development Nathan (2004) has called “archivism”’ (Dobrin, Austin and Nathan 2007:4). Dobrin et al (2007) have highlighted the tensions arising in language documentation through the ‘discourse of endangerment’ and a resulting commodification of endangered languages. As the management of digital data requires systems of standards and quantifiable outcomes on one hand, linguistic fieldwork continues to require methods that allow a ‘responsiveness’ to the individuality of language situations on both a social and formal level (Dobrin, Austin and Nathan 2007:7). How then, do we approach the use of video within the context of language documentation?

³ For example, Dokumentation Bedrohter Sprachen (DoBeS), The Hans Rausing Endangered Languages Project (HRELP), PARADISEC and collaborative frameworks such as DELAMAN and E-MELD.

2. Introduction

This paper discusses the role of video in language documentation. As more linguists take up video for language documentation work, recommendations have been made for researchers working in situations where languages are endangered to ‘make as many video recordings as possible to document as much as we can before it is too late’ (Wittenburg 2007:5). Drawing on recent discussion about the use of video from publications (Nathan 2007; McConvell 2007; Wittenburg 2007), in training courses and in the informal domain of linguists’ forums and blogs⁴ this paper focuses on the methodological implications of using video for language documentation rather than specific equipment and tools. In order to develop methodologies, training and priorities for video use that are ‘responsive’ to individual language situations while also planning for future archive deposits, it becomes important to consider why and how video is currently being used for language documentation, what types of video records are being produced and how the resulting records may be evaluated.

The following section presents a sample of the working practices of a number of linguists currently using video for documentation work (section 3) and highlights the particular challenges that are faced in obtaining a varied, lasting and multipurpose corpus of audio-visual language data (section 4). While practical information about the use of video equipment and recording in the field is available to assist linguists to ‘train yourself for making ‘good’ videos’ (Klass 2005; see also Cholin 2004), broader evaluative measures for what constitutes ‘good’ video for the purposes of language documentation have not been sufficiently addressed. Section 5 considers possible criteria for evaluating and discussing the quality of video records in terms of the wider concerns of language documentation.

3. The use of video FOR language documentation

3.1. Background

Linguists have previously been described as ‘digital pioneers’ in the uptake of new technology for their work (Bird cited in BBC News, 2003). The availability of technology able to record synchronous sound and image has changed the working practices of language researchers. Video has been used extensively for research in language acquisition, phonetic studies and conversation analysis, for the documentation and research of sign language, as well as in the development of language elicitation and stimuli tools.

⁴ For example, see *Transient Languages and Cultures* <http://blogs.usyd.edu.au/elac/>

Anthropological linguistics, with a focus on recording language in its cultural context, has a long history of using visual media. As a result archives hold footage in a variety of legacy formats that can be invaluable resources for languages that are endangered or have ceased to be spoken.

The focus of this paper is the current role of digital video within the context of language documentation. However, the recent uptake of video for research also reflects a more popular interest in the use of digital video media across a range of domains. Websites provide a place to publish videos of varying quality, content and length, from the use of high quality equipment to footage taken with mobile phones, surveillance cameras and webcams. Future industry-driven innovation to next generation formats will present new opportunities to produce, store and disseminate video data. As a result, what we understand as 'video' in terms of equipment, aesthetics, content and delivery continues to evolve. The use of video for language documentation presents a number of issues intrinsic to the medium: from assessing the appropriateness of video use, the selection of recording equipment, format and genre, to the way video data is managed, stored and analysed and how it may be disseminated and repurposed for future audiences.

The rate in which video technology is developing presents particular challenges for digital language archives in setting archive standards, recommending 'best practices' and ensuring the portability and longevity of video data. The amount of storage space and associated costs required for video material remains a topic of contestation (see Nathan 2007; Wittenburg 2007:4). Nathan (2007:3) highlights an 'archiving conundrum' presented by video data. Archives, ideally holding uncompressed, high-resolution language material are accepting compressed formats of MPEG2 and even MPEG4 for video. In many situations the original tapes are being retained in the expectation that a better process will be developed in time as storage capabilities increase. Nathan argues that costs for storing and managing uncompressed video records are still so high that archives with limited resources may only be able to archive a limited number of hours of video (Nathan 2007:3). In response, Wittenburg (2007) states that the costs of video lie predominantly in the costs of the researcher (e.g. transcription time) and that storage costs are rapidly decreasing (Wittenburg 2007:5). Wittenburg therefore recommends language documenters to make as many video recordings as possible, if necessary leaving some linguistic analysis until a later date and to find archiving and curation schemes appropriate to individual projects.

The development of training sessions in the use of video for language documentation, held at the Endangered Languages Archive, highlighted a

number of unresolved questions about recommendations for video practice.⁵ For example, what kind of video should language documenters produce in the context of different projects? Is this video for research or video for language revival work? Who exactly are the video records for? Should language documenters be recording as much as possible and selecting segments to archive, in which case who makes such selections and on what basis? And what processes are involved in transforming video footage into a linguistic record? The different ways that fieldworkers are approaching the use of video for documentation is discussed in the following section.

3.2. Language Documenters

Over the course of language documentation training workshops held at the Endangered Languages Archive during 2006-7 linguists and fieldworkers were asked to complete a questionnaire about their use of video for language documentation work.⁶ Out of 39 completed questionnaires, 9 participants responded that they always used video for fieldwork, 20 often or sometimes used video, 4 never used video and 5 indicated that although they had not yet used video they intended to use video in the future. One respondent indicated that they had tried video but had subsequently ‘dropped’ it as it had too great an effect on the speech situation.

McConvell (McConvell 2007:2) has cited the following reasons for why he considers the use of video as the best way of doing language documentation:

1. it identifies speakers in multiperson conversation;
2. it captures the environment and objects in it;
3. it renders paralinguistic expressions;
4. it records sign language;
5. it shows signs that alter propositional meaning;
6. it shows gesture elucidating force;
7. it is preferred by the community as a record;
8. it costs less and less to store as technology improves

⁵ The video training session was developed as part of the Endangered Languages Documentation Project (ELDP) Grantee Training Workshops (2006-7).

⁶ The Endangered Languages Documentation Project (ELDP) Grantee Training Workshops are held at the Endangered Languages Archive (ELAR), SOAS, and provide documentation training for recipients of a selection of ELDP's grants (see <http://www.hrelp.org>). The questionnaires asked participants to provide information about their experience with video for linguistic fieldwork. The initial questionnaire was prepared by David K. Harrison in preparation for a video training session at the June 2006 workshop and was distributed with his permission for the 2007 training.

When participants were asked to comment on the most and least satisfying aspects of using video for linguistic fieldwork, positive responses reiterated the points identified by McConvell. In particular, the contextually rich data, the ability to capture extralinguistic information and speakers' articulation and to play back footage in the field, alongside the largely positive response to video by language communities were all cited as reasons to use video for fieldwork. However, the least satisfying aspects of video use were described in terms of equipment (notably, portability, weight and set up), the time and complexity involved in processing video data, problems with power in some areas and the impact of video on the language community and speech situation.

McConvell has suggested that the intrusiveness of video continues to lessen as equipment becomes smaller and language communities become more familiar with the medium (McConvell 2007:3). Yet intrusiveness was still cited as a negative factor to using video in some responses to the questionnaire. One respondent highlighted tensions between the preservation ethos of language documentation and concerns of speakers in some language communities about the long-term preservation of their image after their death. Other concerns about video centred on technical recording issues concerned with lighting, sound and night recording. The post-production tasks of editing, capture and copying were highlighted as particularly time-consuming and encompassed a variety of software.

The range of experience with video reflected the variety of endangered language situations represented in the ELDP training workshops. Evidently priorities will be very different for projects with different aims and resources. Moreover, working practices for language documentation are likely to change as developing recording equipment presents new opportunities for data collection. The possibility of compromising image quality to record longer and from multiple viewpoints (for example using web cams)⁷ is now a reality. Likewise, smaller equipment that bypasses tape formats has the potential to record more unobtrusively. Yet video use remains specific to different language situations. While in some field settings linguists strive to minimise the effect of the video camera on the speech event, in other language situations the presence of the video is required to stimulate performance. In other fieldsites it is neither possible nor appropriate to use video. While some language researchers choose to use video for selective recording, others are using video as their primary tool for data collection. The following case studies present a summary from interviews with two linguists who described

⁷ See *Transient Languages and Cultures* <http://blogs.usyd.edu.au/elac/>

their different approaches to using video for linguistic fieldwork in situations where the language is endangered.⁸

3.3. Case study 1

Anthony Jukes⁹ has chosen to primarily use video in his fieldwork for several reasons. These include the positive community response to video materials, the potential of video and annotation tools to simultaneously collect, analyse and produce community language resources and because of the specific language situation of Toratán (Ratahan). Toratán is an endangered language spoken by the older generation in a small number of villages located in southern Minahasa, North Sulawesi, Indonesia. Jukes' decision to first switch to video was informed by Tartius Timpal, a language co-worker who found transcribing audio recordings uninspiring work. In comparison to video, Timpal described audio-only recordings as not 'enak' (literally, not *tasty*).

Unlike other fieldworkers' experiences of the camera producing self-conscious or 'flatter' performances from speakers, Jukes reported that the presence of the video camera acted as a stimulus for conversation providing 'a sense of occasion – of performance' which proved difficult to obtain with an audio recorder. In part, this is particular to the language situation of Toratán. No longer used in the domain of everyday spontaneous conversation, its use has a more formal and performative nature. Yet Jukes did not constrain his recording to narratives and storytelling with individual speakers. One fieldwork video, a recording of three men in animated conversation was recorded with video and two radio microphones. It is an engaging and lively instance of 'language in use', stimulated by the presence of the camera. The filmed gestures of the speakers add substantial meaning to the recording (see Figure 1 and 2). Although the speakers are not directly addressing the camera, in this instance the conversation would not have happened without the camera or Jukes' presence. In such endangered language situations, video is not being used to record natural and spontaneous conversation as such but to stimulate it.

⁸ The interview summaries are reproduced from Ashmore (2006).

⁹ Interview with Anthony Jukes, September 2006, and reproduced with permission.

Figure 1. Three Toratán speakers share a joke during the videotaping of a conversation (l-r: Willem Goniwala, Anes Sumangandow, Bert Hosang).



Figure 2. The same conversation seen in Elan transcription with translation into Indonesian and English.

Elan - RTH021_tatengsan.eaf

File Edit Search View Options Help

Grid Text Subtitles Controls

Annotation

Nr	Begin Time	End Time	Duration
1	00:00:01.000	00:00:02.670	00:00:01.670
2	00:00:04.040	00:00:05.200	00:00:01.160
3	00:00:06.820	00:00:08.230	00:00:01.410
4	00:00:10.053	00:00:11.373	00:00:01.320
5	00:00:13.493	00:00:15.673	00:00:02.380
6	00:00:17.335	00:00:19.175	00:00:01.840
7	00:00:19.625	00:00:20.855	00:00:01.230
8	00:00:23.065	00:00:23.655	00:00:00.590
9	00:00:25.035	00:00:25.730	00:00:00.695
10	00:00:28.230	00:00:28.730	00:00:00.500
11	00:00:34.548	00:00:35.758	00:00:01.210

Selection: 00:00:17.335 - 00:00:19.125 1790

00:00:17.335

00:00:11.000 00:00:12.000 00:00:13.000 00:00:14.000 00:00:15.000 00:00:16.000 00:00:17.000 00:00:18.000 00:00:19.000 00:00:20.000 00:00:21.000 00:00:22.000

ng
ia o're 'lurua nusuka'.
jadi kalau, oh kalau torang dua
well if, oh if we're both the sam
nusuka'
sama-sam
the same
ah, @@@ te ya' le natuari
ah, [tatawa] iju kita linggak adik
ah, i'm a younger brother

te@WG
In@WG
te@WG
te@AS
In@AS
te@AS
te@BH
In@BH
te@BH

Annotation

1 na' te apa.
2 anem lapan.
3 num pulo: walu.
4 inda nusuka' pei ikurua,
5 ah, @@@ te ya' le natuari
6 @@@@ 0
7 nuahi e @@
8 0
9 0
10 Oh, kami le
11 Sawun to Jepang bee,

Annotation

00:00:17.335 00:00:18.000 00:00:19.000 00:00:20.000 00:00:21.000 00:00:22.000

musasoma to da...
bakudapa yang sudah begini...
meseing like this
te
itu
that
@@@ 0
[tatawa] lo
no, sawun le'e.
no, waktu itu
well, that time
nuahi e @@
sudah begini [tatawa]
that's right

Importing, compressing and converting video data to archival and working formats is time-consuming. However, Jukes sees video as having the potential for simultaneously recording, eliciting and producing language resources. An example is the short film he produced about the process of going to the forest to gather palm sugar. The film served as an elicitation stimulus for his language consultant, Bert Hosang, who was no longer physically able to make the journey to gather palm sugar himself. Jukes filmed Hosang watching the palm sugar film on the computer and narrating in Toratán to explain the process (see Figure 3).

Using ELAN it was then possible to add subtitle streams in Indonesian and Toratán from the data tiers already entered to produce a narrated, essentially ethnographic film, which was then copied to VCD and distributed within the community. In this sense the data already has a dual use and can be mobilised quickly back to the community (see Nathan 2004). Yet Jukes acknowledges that he is rarely concerned about the image from a ‘filmic’ or aesthetic perspective. Rather than wanting the camera to be the main focus and interrupting events to alter framing or lighting, he conceives of it more as a static observer, able to be switched on and left running, employing an observational, seemingly objective representation of the speech event.

3.4. Case study 2

In contrast, David Harrison,¹⁰ a linguist working with language communities in Central Asia and Siberia, uses video recording selectively in terms of the amount of video he records (only as much as he can adequately annotate on his return) and what he records. His more recent video data have a particularly cinematic quality, with evident attention to framing, lighting and composition. This may be due to Harrison’s previous collaboration with filmmakers and the fact that he often works with a colleague, allowing one of them to focus on the equipment.¹¹ Harrison mainly restricts the use of video recording to performative genres where, like Jukes, he sees the camera as a stimulus or audience for the narrator or storyteller. In addition his video recordings are often essentially contrived events – the speaker may be positioned to get the best light, speakers may choose particular clothes to wear or onlookers may be asked to leave the room to achieve better quality recording results with minimum disruption.

¹⁰ Interview with David Harrison, June, 2006, and reproduced with permission.

¹¹ Documentary filmmakers have previously accompanied Anderson and Harrison on fieldwork to film material for the documentary, *The Linguists* (Ironbound films).

Figure 3. Bert Hosang (l) explains the process of palm sugar production to Tartius Timpal, prompted by a pre-recorded video segment (shown in second window).

The screenshot displays a video player window titled "Elan - RTH030_hulawatu.eaf". The video content is split into two panels. The left panel shows two men, Bert Hosang and Tartius Timpal, in a traditional setting. The right panel shows a close-up of a large metal pot on a fire, with steam rising from it. Below the video panels is a control bar with various playback buttons and a selection range of 00:06:51.372 to 00:06:53.088. To the right of the video is a subtitle editor with a "Controls" tab. The subtitle list includes:

- 00:06:51.372 - 00:06:51.372: ref@BH
- 00:06:50.000 - 00:06:50.000: nanalangka no drum.
- 00:06:50.000 - 00:06:50.000: pick up the drum
- 00:06:51.000 - 00:06:51.000: namake salangka no drum.
- 00:06:51.000 - 00:06:51.000: use something to pick it up
- 00:06:52.000 - 00:06:52.000: meral one:
- 00:06:52.000 - 00:06:52.000: so there's...
- 00:06:53.000 - 00:06:53.000: Empty
- 00:06:54.000 - 00:06:54.000: Empty

The subtitle editor also features a "Text" tab and a "Subtitles" section with a list of subtitle IDs: t@BH, fe@BH, and Empty.

Figure 4. Video session by K. Mukhayev and A. Mukhayeva, last speakers of Tofa (Siberian Turkic), annotated in ELAN. Filmed in 2001 by K. David Harrison, funded by the DoBeS project, Volkswagen Stiftung.

File Edit Search View Options Help

The screenshot displays the ELAN software interface. At the top left, there is a menu bar with 'File Edit Search View Options Help'. Below the menu is a video window showing a scene with two people. To the right of the video is a 'Grid' window containing a table of annotations. The table has columns for 'Nr', 'Annotation', 'Begin Time', and 'End Time'. The annotations are numbered 14 through 26 and describe actions like 'Then he cried and cried, and flew away.' and 'He cried a lot, his wife was not there, had vanished.' Below the grid is a control bar with buttons for 'S', '←', '→', '↑', '↓', 'Loop Mode', and 'Selection Mode'. The main area of the interface is a spectrogram showing the frequency spectrum of the audio over time. The spectrogram has a time axis from 00:01:12.000 to 00:01:18.000. Below the spectrogram, there are three text windows: 'rt' with the text 'kuhuun bit-qi-lyn-oen-oti-nun kara-on-ak-kuzul-lyme', 'm/s' with 'in.winter hunger-CY INCHPST he-POSS eyes-ACC-LOC red PN', and 'Egloss' with 'In the winter, when he began to get hungry, his eyes were red.'

Nr	Annotation	Begin Time	End Time
14	Then he cried and cried, and flew away.	00:01:05.730	...
15	In the winter, when he began to get hungry, his eyes ...	00:01:11.740	...
16	He cried a lot, his wife was not there, had vanished.	00:01:17.130	...
17	After crying and crying, he also went, flew away.	00:01:26.810	...
18	To find his wife, his mate.	00:01:31.640	...
19	He lived for a long, long, long time.	00:01:37.010	...
20	He grew old, he became feeble.	00:01:43.170	...
21	Now I will not find my wife.	00:01:54.370	...
22	Now I will also die.	00:01:48.450	...
23	We Tofa also have lived in this way.	00:02:00.480	...
24	So far, we're still living like this.	00:02:06.640	...
25	There's no work.	00:02:15.020	...
26	And there's no more now.	00:02:18.840	...

Selection: 00:01:11.740-00:01:17.110 5370

00:01:14.980

00:01:12.000 00:01:13.000 00:01:14.000 00:01:15.000 00:01:16.000 00:01:17.000 00:01:18.000

rt
kuhuun bit-qi-lyn-oen-oti-nun kara-on-ak-kuzul-lyme

m/s
in.winter hunger-CY INCHPST he-POSS eyes-ACC-LOC red PN

Egloss
In the winter, when he began to get hungry, his eyes were red.

berke wala-an-ur kor-qa-b-lic
much cry-PST-3 wife-3 NEG V
He cried a lot, his wife was no

In his recordings when dealing with one speaker, Harrison often frames the speaker with head and shoulders in frame. One recording is a full-face close-up as the speaker narrates a softly spoken story, looking just left of the camera (see Figure 5). In ethnographic filmmaking the close up has been described as unnatural and decontextualising (see Heider 1976:75), in contrast to the wide-angled view favoured by observational filmmakers, yet it remains a strong convention of film tradition. Harrison, rather than attempting to produce a record that is multipurpose, is making informed decisions about framing and angle choice, choosing to omit the rest of the body to get a better quality image in terms of the limited natural light available and within the context of the performance, where the hearer would be positioned close to the speaker.

Figure 5. A speaker of Tsengel Tuvan, Mr. Gansukh, recorded in Mongolia in 2002 by K. David Harrison



While Harrison acknowledges that the material would not be useful for the linguist concerned with conversation analysis and gesture, or contain much extra-linguistic material in frame of interest to the anthropologist, the results have a particular quality that is difficult to define, precisely because we lack the necessary evaluation measures to discuss video. Unlike the filmmaker who can use edits and cut-aways to create narrative and context, the linguist is largely concerned with uninterrupted, synchronised sound and image of a speech event. Without having two or more cameras to produce multiple viewpoints, linguists must therefore make selective framing decisions that may well affect how the record can be used once out of the field.

3.5. Native speaker documentation

The experience of video documentation from the perspective of language communities is an area that requires further investigation. Although linguists working with video have reported positive community responses to the medium, more research needs to be directed to how communities engage with video records and the potential of indigenous media production for language revitalisation (See Cotter 2001; Hale 2001) as well as situations where video has had an unforeseen negative impact. The development of more collaborative and participatory research frameworks that emphasise training speakers of endangered languages to document their language has been recommended (see Grinevald 2003; Woodbury and England 2004). Within the context of language documentation, interesting results are emerging from collaborative projects of this type (e.g. see the Kuikuro linguistic and ethnographic study, DoBeS)¹² and from reflexive approaches to video documentation (see Cash Cash 2007).

The field of visual anthropology provides extensive literature about community self-representation and a history of research into collaborative film projects (Ruby 2000:208–238; 2005), notably Worth and Adair (1997), Michaels (1986; 1994) and Turner (1990), that can inform current documentation work about issues of representation, commodification and media. Importantly, as more video records of communities' linguistic and cultural practices are made and preserved, documentation by native speakers supports community control of the types of visual and linguistic representations of the language and culture being produced.

¹² <http://www.mpi.nl/DoBeS/projects/kuikuro/project>

3.6. The role(s) of video for language documentation

For the purposes of language documentation a number of interrelated roles for video can be proposed:

1. to produce a record of a cultural and linguistic event (video-as-record)
2. as a tool to assist with transcription and analysis
3. to produce material to elicit commentary (e.g. Jukes' palm sugar film)
4. to produce resources for language learning and cultural heritage work (e.g. an edited story or multimedia resource). This could include the particular role video material may have to develop language resources without using a writing system (see Hinton 2001:265)
5. as a tool to stimulate performance (see Case Study 1)
6. to engage EL communities in language documentation work (e.g. on the basis that communities often seem to prefer working with and watching video - see Case Study 1)

A further role should also be mentioned. That is the use of footage taken during fieldwork and included in documentary films to raise public awareness of language endangerment and to communicate about language documentation work (see for example, Harrison and Anderson's *Enduring Voices* project with *National Geographic*).¹³ Such films clearly have a very different agenda to other types of video records and have generated much discussion within linguistic circles (see Dobrin et al. 2007:2).

To what extent different roles for video require different methods is another issue. Nathan (2007) has pointed out that 'different aims will govern the selection of techniques and methodologies for planning, shooting, editing and deploying video' (Nathan 2007:3). Of course, video does not **do** anything without someone to operate, analyse and to watch and engage with it in its many forms and genres. An understanding of the different roles that video plays, of different audiences, of the language community and existing media alongside the methods that go into producing varied types of video records, becomes particularly important in selecting *what* to record and *how* to record it within a language documentation project.

¹³ <http://www.nationalgeographic.com/mission/enduringvoices/>

4. The selection and type of video records

4.1. Corpus selection

Much discussion has revolved around the selection of communicative events for a language documentation corpus (see Himmelmann 2002, 2006; Lehmann 2001; Seifart this volume). Himmelmann (2002) proposes two approaches to the compilation of a comprehensive and representative corpus. These include an anthropological (ethnography of communication) approach and a linguistic approach that selects a range of events along the parameters of spontaneity, modality and frequency. The quality of data in Himmelmann's model is evaluated along a continuum of 'naturalness'. While the ideal goal would be natural data, in reality the majority of data that linguists will obtain are 'observed' linguistic data (Himmelmann 2002:27). Ideally, the corpus would reflect a wide range of speakers in terms of age, gender, social status, language variety etc. However, reconciling suggestions of an 'ideal' documentation that is representative, comprehensive and multipurpose with the diverse realities of fieldwork is rarely straightforward.

Resource and situational constraints mean that language documenters have to prioritise and make decisions about what to record and how to record it, dependent on the specific aims and outcomes of the documentation project. For cultural and ethical reasons some language situations may not be able to be recorded (e.g. taboo or secret language systems). Practical reasons may inform selections (e.g. no night recording). Priority may be given to specific genres on the basis of the topic under investigation (e.g. PhD topic). In McConvell and Thieberger (2001) a set of indicators were proposed to establish existing resources for a given language and to identify areas where urgent documentation work could be directed. Selection may be also be affected by levels of endangerment, for instance prioritising the recording of most 'endangered genres' or focusing on gaps in existing data (e.g. no audio-visual records in an existing archive). Frequently, many recordings are opportunistic and unplanned. In addition, communities may have very different ideas about what types of representative records to produce, preferring more formal, edited speech events than raw, unedited casual conversation (see Mosel 2004).

Video records present particular challenges for the representativeness of a corpus. As Seifart (this volume) highlights, 'the concept of representativeness is most apparent when considering the opposite case: a misrepresentation of a language by a heavily biased corpus'. A concern that more 'filmable' events may take on an exaggerated importance has been identified in the experience of ethnographic filmmaking (see MacDougall 1992:92). An overrepresentation of certain cultural representations recorded with video (e.g.

ceremonial events) and an absence of others (e.g. domestic conversation) may in the future lead to a particularly unbalanced representation of a language community. Given the long-term preservation goals of language documentation, Thieberger and Musgrave highlight ‘that time and the use of technology may reify and lend authority to what were originally ephemeral acts’ (Thieberger and Musgrave 2007:34). Although this is also applicable to audio and text-based records, the multimodal nature of video and the emotional weight it can convey may have consequences in lending prestige to a particular variety in the future or giving evidential weight to casual remarks made in conversation (Thieberger and Musgrave 2007:34).

Language documenters will not always be able to anticipate future interpretations and uses of the video material they produce. Nevertheless, taking a more reflexive approach to video documentation, providing explicit information about the selection of speech events and recording methods and about how linguistic knowledge is produced would be beneficial for future interpretations of the material (see Pink 2007:24). To do this requires establishing appropriate terminology to describe the types of video records produced and the different methods that go into their production. The emphasis in some language documentation literature of the aims to produce ‘multipurpose’ data, able to be repurposed for different future uses and users that may not yet be envisaged, raises particular questions about what types of video a language documentation should aim to produce or prioritise – video material that can be used for research, video that can be used for publication material and language learning resources, or the possibility of video material that can do both?

4.2. Types of video record

McConvell (2007) has proposed that training in the use of video for language documentation should ‘be aimed at true documentation video, not making short films or documentaries except as by-products’ (McConvell 2007:3). The distinction between the types of audio-visual record produced as a basis for research has been an ongoing issue in the field of visual anthropology and has produced extensive literature on the topic (Hockings 1995; Ruby 2000; Pink 2007).

MacDougall (1978:405, cited in El Guindi 1998) makes the following distinction between ethnographic research and presentation forms: ‘ethnographic footage (raw material that comes out of a camera, like fieldnotes, used for a variety of purposes including the making of films) and ethnographic films (structured works made for presentation to an audience)’ (El Guindi 1998:486). MacDougall further divides ethnographic footage into ‘research footage to serve specific scientific inquiries, and record footage

made to provide more general documents for archiving and future research' (El Guindi 1998:486). El Guindi proposes the following typology for research purposes:

(1) the visual medium as a recording tool of data for analysis and/or archival purpose; (2) the visual medium for elicitation and discovery; (3) the visual medium used for experimental culture reconstructions;¹⁴ and (4) the visual medium as ethnography – visual ethnography (El Guindi 1998:487).

A more straightforward distinction is made between an observational record, uninterrupted filming that is unedited with minimal manipulation of cinematic conventions, often using a wide-angle perspective, and 'creative' footage that may incorporate editing to construct a narrative, for example (Pink 2007:97). Such distinctions remain problematic. Even the most visually 'objective' material, where the camera is left running to record an event, is only a selective representation. Equally, as Pink notes in relation to ethnographic data 'ethnographic knowledge does not necessarily exist as observable facts' able to be captured with a video camera (Pink 2007:98). As with linguistic data, what transforms digital bits to a linguistic record is the accompanying contextual information and interpretation – knowledge produced through fieldwork and interaction, in which linguists and the speech community are constantly engaged in the field (see also Gnerre, this volume).

McConvell suggests 'the key is for would-be language documenters to understand that a fuller record is possible, from which they can subjectively select and edit elements for various purposes' (McConvell 2007:3). This could be taken to mean a variety of video material, including research footage and record footage. Likewise, a fuller record may include 'extra' footage of communicative events that may be drawn on later for cutaways of use in editing and additional contextual information of use in producing a wider range of resources. The level of contextual material and interpretation accompanying video data that turns it into a fuller record is discussed in the following section.

¹⁴ 'Experimental culture reconstructions' here refers to the practice of reconstructing sequences of traditional behaviour, for example, particularly used in situations of culture loss (See El Guindi 1998: 480–483).

5. Evaluating the video record

According to Austin and Grenoble (2007), one of the unresolved issues in the field of documentary linguistics is how the quality and effectiveness of documentation outcomes can be defined and assessed. Funding applications for documentation projects frequently refer to the quantity of archival objects as an outcome of documentation (e.g. how many hours of video recordings). Austin and Grenoble suggest other quality metrics may be more appropriate (e.g. the effectiveness for language maintenance and revitalization), although how this would be assessed remains an open question (Austin and Grenoble 2007:22). Equally, criteria for evaluating individual records, particularly video media, as part of a language documentation corpus remain undefined. In part this is due to the varied aims and outcomes specific to different documentation projects. Also the multimodality of video, a lack of agreed standards for archiving or clear methodological approaches makes developing appropriate descriptive frameworks to discuss video records problematic. In addition, the processes of documentation identified by Austin, of recording, capture, analysis, archiving and mobilisation (Austin 2006:89) means that data are undergoing transformation into varied records at different stages. At what point these records are assessed – the recording stage, the archival object, at the point of mobilisation or in terms of long-term goals concerning language endangerment – is not clear.

The following sections discuss three components of digital video material that may be taken into account in planning and evaluating video for language documentation aims. These include (1) the technical elements of the record, (2) the contextual information and analysis that transforms video data into a video record, and (3) video data in relation to the broader documentation aims of the field of language documentation.

5.1. Technical Aspects

I use the term ‘technical’ here to refer to the techniques available for recording as well as to refer to the technical elements relating to the process of documentation including choices of media, compression and conversion of data to different working formats. Given the current popularity of video and the availability of affordable consumer range cameras, video production can be perceived as relatively easy. Nevertheless, the benefits of some training in the use of video equipment cannot be understated for camera operators to be able to handle varied recording conditions and challenges such as lighting, night-recording and producing high quality sound. Moreover, ease with the technical elements of video recording allows more time to concentrate on the interview, interaction or content of the recording. Considering the many years that linguists commit to linguistic training, a short practical filming course

could make a significant difference to how the production of video records is practiced and theorised.

Technical tips from professional filmmakers about composition, eye line, filming cutaways and editing can be useful regardless of the type of record being produced.¹⁵ Some understanding of editing assists in planning what you record and how you record it. Knowing what might be needed to turn a research record into a more richly contextualised representation of an event will direct what types of ‘extra’ footage may be filmed. This becomes more important in situations where the option of returning to collect different types of data may not be possible. In terms of producing a ‘fuller’ record, drawing on professional techniques for filming cutaways, to join together different sections of an event and recording contextual opening and closing shots may prove a useful resource if repurposing footage for different audiences.

How the technical components of video could be evaluated is another challenge. For example, should technical elements of image and sound be assessed on digital quality, on aesthetic grounds, on the basis of content or on the potential for repurposing? Should the content of the data (in terms of the selection of the communicative event) be prioritised regardless of image and sound quality? And could choices of compression rates and recording formats be evaluated on the longevity and stability of data for future use? Realistically, these different dimensions of an individual record are not easily separated. Questions about the assessment and description of video materials require further discussion in the field of language documentation. However, one component of video material that is considered key to the aims of language documentation is the contextual information supplied with data that transforms raw footage into a linguistic record.

5.2. Contextual information and analysis

One of the most prominent reasons that video is cited as the ‘ideal recording device’ (Himmelman 2002:12) is its apparent ability to capture ‘language in context’ or to result in ‘context rich’ data. Of course, the context-richness of video records is not intrinsic to the media. Without some level of metadata, contextual information and preferably time-aligned annotation, locating, searching and accessing video data becomes problematic. While there are no established guidelines for what constitutes adequate metadata in relation to video data, Nathan (personal communication) refers to it as textual support that signals ‘commitment’ (for example, at minimum a time-aligned

¹⁵ Thanks to Michael Yorke, Alison Kahn and Maarten Roos at Oxford Academy of Documentary Film for training and discussion of these topics.

annotation and list of topics discussed in the footage). Himmelmann describes three components that comprise each piece of data:

i.e. the ‘raw’ data in various forms of representation (transcription, tape and/or video), a translation (word-by-word/interlinear and free), and a commentary ... (Himmelmann 2002:13).

The commentary could include the circumstances of the recording, notes on the linguistic aspects of the recording and metalinguistic commentary by speakers as well as relationships to other data. In the production of video records it could be argued that analysis begins as soon as the camera is turned on, selecting to record a particular event in a particular way, though this is equally applicable to research in general. For video recordings, an explicit statement of intent about the selection process, context and technical decisions made at the recording stage would assist in developing a clearer methodology for video documentation, even for unplanned, spontaneous recordings. However, providing information about context requires some consideration. Widlok (2005) points out that there are two aspects to consider when discussing context: how things/events are embedded in context and how contexts are constructed through interpretation (Widlok 2005:7).

How context is understood has implications for the organisation of data in terms of the type of commentary supplied and how larger lengths of video are carved up into individual records or sessions as well as how relations between data are built within a corpus. At least some basic editing is usually required. Vaughan, an editor of ethnographic film, notes that many observational filmmakers consider that ‘the minimum of structuring will afford the maximum of truth’. Instead Vaughan suggests ‘the antithesis of the structured is not the truthful, nor even the objective, but quite simply the random’ (Vaughan 1992:100 cited in Henley 2006:400). Contextual information presented in different forms (cataloguing metadata, commentary, structure and organisation of data) makes data useable.

However, constructing commentaries to accompany recordings involves a process of ‘gradual and cumulative understanding’ (Evans and Sasse 2007:62). Evans and Sasse have highlighted that despite new technology that allows us to record more and more of the physical context of language, ‘the other side of language – what these recordings mean – remains problematic’ (Evans and Sasse 2007:60). The process of how translations are arrived at, what they term as the search for meaning, ‘is best seen as a never-ending stringing together of hypertextual commentary, which gradually leads to a better understanding of the utterances under study’ (Evans and Sasse 2007:60). The interpretation necessary to decode video material involves a wider complex of social, linguistic and cultural understanding that extends

well beyond the edge of the frame. These are activities that linguists, working collaboratively with native speakers, have long been engaged in. As Evans and Sasse highlight, ways to make this process more transparent, able to be added to and built on is an ongoing issue in the archiving and organisation of digital documentation corpora.

5.3. Documentation aims

More general aims of language documentation have included the potential for primary data resulting from fieldwork to provide lasting, varied and multipurpose records. These records may then provide a basis for future descriptive and community-oriented work. The different resources and scope of documentation projects make evaluation of these aims difficult. In addition, is it realistic to consider that an individual video recording can be useful for the conversation analyst, the phonetician, the linguist concerned with eye gaze, the language researcher concerned with the ethnographic context of the speech act and the needs of the language community? The competing aims of fieldwork, of formal standards and the need for a flexible approach to language situations, means that decisions must be made about priorities in recording and analysis. While individual records are unlikely to serve all purposes, the corpus as a whole could certainly aim to provide a range of media and types of video record, a ‘fuller’ record able to be revisited as part of the ongoing process of interpretation.

The potential for documentary corpora to have an impact on language shift remains a focus for language documentation in many situations. However, ‘to date there is very little experience indeed with regard to actually working with digitally-stored multimedia corpora of lesser-known languages’ (Himmelmann, Gippert and Mosel 2006:vii). It is possible that many of the materials produced in language documentation projects may never be retrieved by anyone but the researcher. However, assessing the quality of video records in terms of these long-term goals requires ongoing investigation and discussion by language documenters, archivists and language communities as technology and working practices evolve. Specifically, research into the response of endangered language communities to video recording and publication needs to be conducted more widely, including the kinds of material that are important in current revitalization and maintenance projects and alternative outcomes in areas where digital technology is not appropriate or available.

6. Conclusions

Digital video can be useful and effective for language documentation. However its utility is dependant on individual language situations and necessitates some level of planning, consultation and commitment to provide contextual information. Recommending standards and methods for video use remains problematic because of rapidly changing technology and the diverse realities of fieldwork. Yet as more video documentation is carried out it becomes increasingly important to anticipate issues arising from the production and long-term storage of digital representations of languages and speakers. Discussion about the role of video in language documentation will undoubtedly continue as technology and the field of language documentation evolves, presenting new opportunities and challenges for language researchers and language communities.

References

- Ashmore, Louise. 2006. Reducing language to bits: Implications for the use of digital video in language documentation. Unpublished MA Thesis, Department of Linguistics, School of Oriental and African Studies.
- Austin, Peter K. 2006. Data and Language Documentation. In Jost Gippert, Nikolaus P. Himmelmann and Ulrike Mosel (eds.), *Essentials of Language Documentation*, 87-112. Berlin and New York: Mouton de Gruyter.
- Austin, Peter K. and Lenore A. Grenoble. 2007. Current trends in language documentation. In Peter K. Austin (ed.), *Language Documentation and Description, Volume 4*, 12-25. London: SOAS.
- BBCNews. 2003. Digital race to save languages. Retrieved 10 November, 2007, from <http://news.bbc.co.uk/1/hi/technology/2857041.stm>
- Cash, Phillip E. 2007. Documenting language, visualizing culture: shooting digital video in two endangered language communities. In Peter K. Austin, Oliver Bond and David Nathan (eds.), *Proceedings of Conference on Language Documentation and Linguistic Theory*, 13-19. London: SOAS.
- Cholin, J. 2004. Video recording in the field. *Language Archive Newsletter*, 1(4), 5-8.
- Cotter, C. 2001. Continuity and Vitality: Expanding Domains through Irish-Language Radio. In Leanne Hinton and Kenneth Hale (eds.), *The Green Book of Language Revitalization in Practice*, 301-312. San Diego and London: Academic Press.
- Dobrin, Lise M., Peter K. Austin and David Nathan. 2007. Dying to be counted: the commodification of endangered languages in documentary linguistics. In Peter K. Austin, Oliver Bond and David Nathan (eds.), *Proceedings of Conference on Language Documentation and Linguistic Theory*, 59-68. London: SOAS.

- El Guindi, F. 1998. From Pictorializing to Visual Anthropology. In H. Russell Bernard (ed.), *Handbook of Methods in Cultural Anthropology*, 459–512. London: Altamira Press.
- Evans, Nicholas and Hans-Jürgen Sasse. 2007. Searching for meaning in the library of Babel: field semantics and problems of digital archiving. In Peter K. Austin (ed.), *Language Documentation and Description Volume 4*, 58-99. London: SOAS.
- Gippert, Jost, Nikolaus P. Himmelmann and Ulrike Mosel (eds.). 2006. *Essentials of Language Documentation*. Berlin and New York: Mouton de Gruyter.
- Grinevald, Colette. 2003. Speakers and documentation of endangered languages. In Peter K. Austin (ed.), *Language Documentation and Description Volume 1*, 52-72. London: SOAS.
- Hale, Kenneth. 2001. Strict Locality in Local Language Media: An Australian Example. In Leanne Hinton and Kenneth Hale (eds.), *The Green Book of Language Revitalization in Practice*, 277-282. San Diego and London: Academic Press.
- Heider, Karl. 1998. [1976] *Ethnographic Film*. Austin: University of Texas Press.
- Henley, P. 2006. Narratives: the guilty secret of ethnographic documentary? In M. Postma and P. I. Crawford (eds.), *Reflecting Visual Ethnography: using the camera in anthropological research*, 376-401. Højbjerg and Leiden: Intervention Press and CNWS Publications.
- Himmelmann, Nikolaus P. 2002. Documentary and Descriptive Linguistics (full version). In Osamu Sakiyama and Fubito Endo (eds.), *Lectures on Endangered Languages 5*, 37-83 (Endangered Languages of the Pacific Rim, Kyoto, 2002).
- 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.), *Essentials of Language Documentation*, 1–30. Berlin and New York: Mouton de Gruyter.
- Hinton, Leanne. 2001. Audio-Video Documentation. In Leanne Hinton and Kenneth Hale (eds.), *The Green Book of Language Revitalization in Practice*, 265-271. San Diego: Academic Press.
- Hockings, Paul. (ed.) 1995. *Principles of Visual Anthropology*. Berlin and New York: Mouton de Gruyter.
- IronBoundFilms. (Nd). *The linguists*.
<http://www.ironboundfilms.com/ironwork.html>
- Klass, G. 2005. Making video recordings. DoBeS A4 guides Retrieved July 15th, 2006, from <http://www.mpi.nl/corpus/a4guides/>.
- Lehmann, Christian. 2001. Language Documentation: A program. In Walter Bisang (ed.), *Aspects of typology and universals*, 83-97. Berlin: Akademie Verlag (Studia Typologica, 1).
- MacDougall, David. 1992. Complicities of style. In Peter I. Crawford and David Turton (eds.), *Film as ethnography*, 90-98. Manchester and New York: Manchester University Press.

- McConvell, Patrick. 2003. 'Multilingual Multiperson Multimedia: Linking Audio-Visual with Text Material in Language Documentation'. In Linda Barwick, Allan Marett, Jane Simpson and Amanda Harris (eds.), *Researchers, Communities, Institutions, Sound Recordings*. Sydney: University of Sydney.
- McConvell, Patrick. 2007. Video - A Linguist's View (A Reply to David Nathan). *Language Archives Newsletter*, No. 10, 2-3.
- McConvell, Patrick and Nicholas Thieberger. 2001. *State of Indigenous languages in Australia - 2001*. Australia State of the Environment Second Technical Paper Series (Natural and Cultural Heritage), Department of the Environment and Heritage, Canberra <http://www.ea.gov.au/soe/techpapers/index.html>
- Michaels, Eric. 1986. *The Aboriginal Invention of Television in Central Australia*. Canberra: Australian Institute of Aboriginal Studies.
- Michaels, Eric. 1994. *Bad Aboriginal Art: Tradition, Media and Technological Horizons*. Minnesota: University of Minnesota Press.
- Mosel, Ulrike. 2004. Inventing communicative events: conflicts arising from the aims of language documentation. *Language Archive Newsletter*, 1(3), 3-4.
- Nathan, David. 2004. Planning a multimedia documentation. In Peter K. Austin (ed.), *Language Documentation and Description, Volume 2*, 154-168. London: SOAS.
- Nathan, David. 2007. Digital video in documentation and archiving. *Language Archives Newsletter*, No. 9, 3-4.
- Pink, Sarah. 2007. *Doing Visual Ethnography*. (2nd Ed.). London: Sage Publications.
- Ruby, Jay. 2000. *Picturing Culture: Explorations of Film and Anthropology*. Chicago and London: The University of Chicago Press.
- Ruby, Jay. 2005. The last 20 years of visual anthropology - a critical review. *Visual Studies*, 20 (2), 159-170.
- Thieberger, Nicholas and Simon Musgrave. 2007. Documentary linguistics and ethical issues. In Peter K. Austin (ed.), *Language Documentation and Description, Volume 4*, 26-37. London: SOAS.
- Turner, Terence. 1990. Visual Media, cultural politics and anthropological practice. *Commission on Visual Anthropology Review*. Spring.
- Widlok, Thomas. 2005. Cross-disciplinarity in the documentation work of anthropologists and linguists. In Peter K. Austin (ed.), *Language Documentation and Description, Volume 3*. London: SOAS.
- Wittenburg, Peter. 2007. A Technologist's View (A Reply to David Nathan). In *Language Archives Newsletter*, No. 10, 3-5.
- Wittenburg, Peter. 2003. The DOBES model of language documentation. In Peter K. Austin (ed.), *Language Documentation and Description Volume 1*,. London: SOAS.
- Woodbury, Anthony C. and Nora C. England. 2004. Training speakers of indigenous languages of Latin America at a US university. In Peter K.

Austin (ed.), *Language Documentation and Description, Volume 2*, 122-139. London: SOAS.

Worth, Sol and John Adair. 1997 [1972]. *Through Navajo Eyes*. Albuquerque: University of New Mexico Press