

Multilingual, Multiperson, Multimedia: Linking Audio-Visual with Text Material in Language Documentation.

Patrick McConvell

AIATSIS

1. Introduction

'Language documentation' for endangered and Indigenous languages has been rapidly moving towards a more holistic view of what is to be captured, including a range of genres, conversation as well as narrative. Most of the languages concerned also exist in a multilingual, multivariety language ecology, in which different age groups may speak, and switch between, different varieties. This inevitably becomes part of what is being recorded and is crucial in the understanding of language shift and maintenance. Added to this is the growing realisation of the importance of paralinguistic elements such as gesture even to the basic interpretation of utterances. For proper documentation, what is required now is a system that can handle video, audio, transcription, translation and other annotation, linked by time codes. In this paper I will investigate the functionality of the CLAN system of a/v-transcript linking, widely used for child language and multilingual studies, and briefly compare this to other available alternatives.

As for archival holdings of a/v and transcriptions, most of what already exists cannot be immediately moved into such a/v-text linking systems, because of the amount of work involved. There is a need however, for some standard system for preliminary digital linking of a/v with existing transcripts, translations and annotations, which may be separated from each other physically and institutionally. From this, more robust linking for analysis and multimedia presentation can be developed. This paper reviews some of the systems being used and the extent to which the metadata element 'Relation' can be refined to carry out this task.

Finally the points made above, which are of general applicability, are reviewed and applied to the specific case of Australian Indigenous languages. An outline of how a 'two way documentation program' might be built is briefly presented.

2. Language documentation

Tony Woodbury in a keynote address to the Linguistic Society of America (reprinted as Woodbury 2003) outlines the newly emerging field of 'documentary linguistics'. Actually the concept and practice is not new, as Woodbury makes clear, but follows in the tradition of the kind of work carried

out by Franz Boas and colleagues at the birth of American descriptive linguistics as a subdiscipline of anthropology. However this broad ethnographic and textual approach to recording languages in the field, although continued by anthropological linguists and linguistic anthropologists, had largely become eclipsed by a more theoretical and less hands on approach, and in parallel with it, a focussing on words and sentences often taken in isolation from their context. The call for a new 'documentary linguistics' is not just for a return to Boasian practice but also importantly recognises that we have entered the digital era and this makes possible recording, archiving, analysis and distribution in very different ways and potentially with much greater efficiency.

Documentary linguists (like 'corpus linguistics') makes the corpus of recordings and texts central. According to Woodbury, a good corpus is (1) *diverse*; and (2) *large*.

The production of the corpus is (3) *ongoing*; (4) *distributed* (in the sense that several or even many different people may contribute to its compilation); *opportunistic* (taking every opportunity which presents itself to record) and (5) *ethical* (especially taking full account of the wishes of the language community in how to carry out the work and present the results). Materials built from the corpus should be: (6) *transparent* (understandable and usable by a wide range of people); (7) *preservable* (capable of being archived, with metadata, and retrieved) (8) *portable* (this refers mainly to the ability of materials to be read using different software and hardware - although literal portability would not be a bad idea and is becoming more feasible).

While the main focus of documentation is the language, with such a broad scope many other aspects of a culture or way of life are also documented with the added benefit that they are spoken about and enacted in the first language(s) of the local people, rather than is often the case in anthropological documentation, some lingua franca which may not be well known by local people or reflect local concepts accurately.

The 'Boasian trilogy' of Grammar, Dictionary and Texts (Darnell 1998:???) remains a desideratum in this 'documentary linguistics' approach, but the 'texts' element is broader than what is often offered in older text collections, and using modern technology, can be a multimedia product rather than just a written volume. The documentary linguistics approach would aim at recording a wider range of speakers than may have been the previous practice (in terms of gender, age, and variety of speech such as dialect); and a wider range of genres (not just the typical individual narrative of many text collections) including

1. Conversation
2. Multilingual, multidialectal speech
3. Context of the speech and the action accompanying it

4. Cultural commentary by speakers of the language on other recordings, artefacts etc.

Point 3 above points strongly to the need to use video in documentation rather than just audio. Point 4 points to the need to play back earlier recordings and add 'meta-recordings' to that as annotations. These points are further elaborated below.

It is useful to be able to compile lists of resources available for each language and quantify those to provide *indicators* of the level of resource need which exists. This can be combined in a formula with a measure of endangerment of the language, which can serve as a guide to where the most urgent work needs to be done. We may need to modify these as the new broader styles of documentation are adopted, in order for instance to assess how many different genres and types of speech have been recorded. One attempt was that of McConvell and Thieberger (2001) which proposed a set of indicators of documentation which yielded a score of 0-20 for languages, based on information recorded by them in the Indigenous Languages Database (a resource which is now to be upgraded and mounted on a website by AIATSIS). Each of six types of documentation is scored out of 3 or 4 and added up to yield the total documentation index (see below). In this scheme, 'ethnolinguistic information' includes a broad range of anthropological linguistic work, but there are no guidelines to assign scores when for instance there is a detailed thesis on songs but no information on other genres and modes.

Indicators of documentation (McConvell & Thieberger 2001)

Dictionaries: Detailed dictionary (e.g. Arrernte, Kayardild) (4); Medium dictionary (3); Small dictionary/ wordlist (e.g. Warnman) (2); Simple wordlist (e.g. Bates, Curr) (1).

Texts: Extensive text collection (3); Several texts (<10) (2); Elicited/example sentences (1).

Grammar: Detailed grammar (e.g. Gooniyandi, Kayardild) (4); Middle-sized grammar (eg. Handbook) (3); Grammar sketch or many technical articles (2); Few technical articles only (1).

Ethnolinguistic information: Substantial ethnolinguistic work (e.g. thesis) (3); Ethnolinguistic description (2); Some ethnolinguistic information (1).

Audio recording: More than several hours of audio (3); Less than several hours of audio (2); Less than an hour of audio (1); No audio recorded (0).

Other: Literature (including school readers and religious translation) in the language - more than 1000 words (2); more than 100 words (1); video or film with more than 100 words spoken or subtitled **or** multimedia with more than 100 words spoken and/or written (1).

It is obviously beneficial if metadata categories used in widespread standard schemes harmonised with the categories used for measurement here, or in the next updated version. Here the OLAC element *type.linguistic* is particularly relevant. There are three primary types in the latest proposed recommendation (<http://www.language-archives.org/REC/type.html>, 2002)—*language description*, *primary texts* and *lexicon*— each of which have a number of subtypes. Of these, only *lexicon* is a clear direct equivalent of the category *dictionary* in the list above, although *primary texts* is similar to *texts*; and *language description* is similar to *grammar*.

If the OLAC types (or some combination of their metadata elements) mapped more clearly on to resource descriptors, a search on a comprehensive database of languages would automatically yield documentation indices which were largely equivalent.

Audio- (or perhaps we should say audio-visual) documentation has a special place in this documentation list because it provides a direct link to what the fluent speakers had to say.

3, How can documentary linguistics contribute to language maintenance?

Documentary linguistics of the type outlined above clearly provides a richer storehouse of the linguistic and cultural heritage of an ethnic group than other narrower types of research. This is appreciated by the succeeding generations of scholars, and of descendants of the language group (whether the language is still spoken or not). The categories of scholar and speaker or descendant of a language group are not mutually exclusive, as we shall discuss below.

Another advantage of this broad type of data-gathering is that it can give a fuller picture of the *language ecology* of a group – which languages and varieties of languages are spoken and for which purposes in the bilingual or multilingual situations which are the typical sociolinguistic context of endangered languages. The language ecology can tell us which are the risk factors for languages, when we compare different languages, but we need a broad picture of the social and cultural embedding of languages to work this out. On the positive side it can tell us which are the elements of a language ecology which have been shown to be conducive to language maintenance, and thus provide models for interventions which are more likely to work.

Some of the more specific advantages of the documentary linguistic approach for language maintenance include the following:

- **As a source of user-friendly information on the old language which can assist and inspire the community to learn and maintain it.**

Full documentation enables people who would otherwise be partial speakers or non-speakers (if the language group has undergone language shift after documentation) to learn a more comprehensive and fluent form of the language. The existence of such a corpus can inspire people to make more effort both in their own research and in establishing language learning programs. Access to direct spoken language with helpful annotation is much less daunting than technical grammars and dictionaries which often need linguistic experts to interpret them.

- **As a repository of more natural kinds of speech**

Traditional grammars, dictionaries and even texts often do not contain the most common everyday ways in which people communicate with each other in language. Grammars may be based on elicitation, so the choice of sentences represents the linguist's choice of elements to test grammatical hypotheses rather than the way the speakers might naturally express themselves. Even the texts gathered, while immensely valuable might be skewed towards particular genres eg myths and legends, because the speakers and perhaps the linguist thinks these are important cultural material. This is probably correct but it may mean that more everyday styles of speaking are not recorded. In the worst case scenario all the texts may be in a special oratorical style used for such narratives which is quite different from ordinary language. Where care is taken to include all major styles in the corpus, this should not happen.

- **As a repository of special registers which may be important in language revival**

On the other hand, some special registers like speech making may be particularly important for language revival because it is a public activity invested with a lot of prestige. Among Maori, learning oratory was maintained longer than use of everyday language at home and has formed the basis for people re-expanding their knowledge of other genres. In the Kaurua language revival in Adelaide the making of speeches at funerals, festivals etc has been a key element (Amery 2000). Recording of such genres before they are forgotten provides a platform for later learning of them by descendants.

- **This information and learning can be a key element in land claims**

In Indigenous land and native title claims, both in Australia and overseas, knowledge encoded in the old language of the area concerned has proved decisive in providing a strong case. It is unfortunately the case that judges in Australia have sometimes viewed loss of a language as a symptom of a 'break in continuity of tradition' and this has been instrumental in the applicants' case failing. However whether or not the authorities have not taken such a hard-line stance the ability of the descendants of language to access from records detailed information about placenames (Hercus, Hodge & Simpson eds 2002), kinship (McConvell, Dousset & Powell eds); language used for land

matters (McConvell 2000) including invocations of spirits resident at places, and other environmental and general vocabulary, has been crucial to presenting cogent evidence. Not only that but the land claim process itself provides a means by which the descendant applicants learn or relearn about these things, with growing pride and confidence.

● **Community researchers get involved and find that they can discover more than outsiders**

One type of community research already mentioned is that stimulated by land claim and native title cases, but there is also a growing body of Indigenous researchers separate from this, collecting information on old language and cultural practices from written and recorded sources and from those elders who remember. Some of these are undertaking formal education and using the techniques of linguistics and other disciplines; others are less engaged with the western modes of data collection and analysis but proceed in their own way. In Australia, many of these people are attached to or working through the key Institutions which I mention in the last section of this paper – the Regional Aboriginal language centres; Batchelor Institute (CALL) and similar training centres; and AIATSIS. Others may work with University and college departments, in local communities or independently as individuals.

It is important to harness this vital force of Indigenous researchers in the task of documentary linguistics. The mode of operation of documentary linguistics and its typical products is generally much more acceptable than the academic scope, methods and products, but does not preclude the production of academic books and papers. The key here is ‘two-way’ or *Garma* research in which the Indigenous researcher, who is typically younger and more educated in the western sense than the elders with more traditional knowledge, forms a team with a linguistic technical expert and a knowledgeable elder as mentors (cf. Hinton & ??).

This process can be empowering for the Indigenous community researcher, one reason being that they may find that because of their prior knowledge and relationships, they can make more discoveries than a non-Indigenous researcher. So for instance, Lizzie Ellis (Ellis 2000) a Western desert language speaker with some training in linguistics undertook research with old people on words and expressions relating to fauna and discovered a great deal more detailed vocabulary than had ever been recorded before. Raymattja Marika, a Yolngu woman from North-east Arnhem Land who has a Masters degree has been studying both the clan languages (like her own, Rirratjingu) and the newly evolving koine Dhuwaya in her community, and has the great advantage of understanding all the varieties and the social circumstances in which they are used.

● **Local education projects can stimulate recovery of endangered languages affected by attrition**

As people engage in team projects with endangered languages, it has been observed that people begin to remember more detail of the old language. This seems to be particularly the case where there is a concrete outcome for the community, for instance educational resources. June Oscar (pers..comm) Chairperson of the Kimberley Language Resource Centre thinks that this occurred during the intensive community work on the production of the Bunuba CD-ROM. Helen Harper (2000) also reports that older people working on educational projects in Northern Cape York Peninsula, where the languages are far down the road towards being lost, were beginning to recall more and more as they engaged in the activity.

● **Some outsiders place value on this work and this increases its prestige in the community**

As well as local prestige and pride gained from projects in language documentation, local Indigenous people are also aware that other Indigenous people and non-Indigenous people are interested in and impressed by their efforts. This also provides positive feedback to drive projects along.

● **Information on the *linguistic ecology* helps to develop plans for language maintenance projects**

As mentioned above, wider documentation of all the languages and varieties spoken, and under which circumstances helps to build a picture of the linguistic ecology which aids language planning and 'reversing language shift' (Fishman 1995, 2000). The aspect of documentary linguistics which requires all types of speakers to be recorded logically also involves recording children and adult-child interactions and this can give direct pointers about language change and language shift, and help to plan what kind of language is appropriate for language learning materials at different ages.

Crucial questions in language maintenance include which young people keep knowledge of the old language and when and why do they use it? Can more people reproduce these conditions? In the case of Gurindji the children have been learning and speaking a 'mixed language' with a Kriol (English based creole) matrix and significant Gurindji vocabulary and morphology, for about 30 years (Dalton et al 1995), and have not been learning traditional Gurindji, or at least I thought not. Recently I noticed that a few adolescents at least still controlled a fairly standard traditional Gurindji but hardly ever spoke in it. This is the kind of issue that needs further study. The Aboriginal Language Acquisition Project is also investigating these kinds of issues in a number of Central Australian communities.

4. Why use video not just audio for our documentation research?

The Aboriginal Language Acquisition Project just mentioned is using digital video to record interactions between children and between adults and children. It has been the practice for some time for some language acquisition

researchers to use film or video. However there seems to be a good case now to use video (with good audio of course) for language documentation more generally. Digital video recording equipment is relatively inexpensive, easy to use and, increasingly easy to transcribe with ready-made programs like CLAN (described further below). The major objection has been the enormous amounts of digital space which video takes up, far more than audio and dwarfing text and still images. However as also noted below, this problem is decreasing as technology improves.

As to the advantages in terms of the linguistic and cultural phenomena captured, video has the following advantages:

- **Identifying speakers in multiperson conversation.** In natural conversation some people talk loudly, some softly; speech overlaps and background noise makes transcription difficult if not impossible for some passages. The visual image of the speakers makes everything easier: their body position and the movement of their lips shows who is talking, and to whom, and who is silent. Attaching voices to faces makes it easier to track what different individuals are saying.

- **Relation to environment, objects** Speakers frequently refer to objects in view and point to them (Hanks). Sometimes long passages may be descriptions of a painting or an artefact, with the pointing an integral part of the description. A purely audio record of such speech events is very hard to follow. If the recordist interrupts with description of referents or gestures it can ruin the recording; noting down in a notebook can be equally hard to follow later and is a distraction to the recordist; asking participants to recreate the full sense from audio can be posing as exacting task to them as to the recordist. Video is able to record all these essential parts of the speech event for later viewing and rich transcription.

- **Paralinguistics** This includes the study of gaze, proxemics (spatial relationship between bodies), gesture, facial expression and other forms of non-verbal expression. This is all lost in an audio recording, yet it can be an essential part of the overall meaning of communication in an interaction. This is all part of a communicative culture and should be recorded too. In Australian Aboriginal societies, for instance if someone sits as far away as possible from another individual with his or her back turned to him or her, it may not mean that they do not like each other, but that they are in an 'in-law' avoidance relationship.

- **Sign language.** The extreme example where most researchers accept that video is needed is where the conversation is wholly or partially in a manual sign language. Use of elements of such a language in general conversation (eg kinship terms) is fairly common in Indigenous Australia even where there are not deaf signers present, and would be entirely missed in an audio recording.

● **Sign altering propositional meaning** Signs or gestures commonly accompany spoken language and may have quite serious effects on the meaning of the utterance, undetectable in the audio-record. David Wilkins gave an example of this kind in his keynote address to the Australian Linguistic Society meeting in Canberra in 1997. The common hand-sign for 'nothing, no, not' was used by an Arrernte speaker while verbally saying a positive sentence, reversing the meaning to a negative sentence. A transcriber working only with audio would have missed the gesture and produced something meaning the opposite of what the speaker intended.

● **Gesture elucidating force**

There are many gestures and body postures which affect the pragmatic interpretation of utterances which are not as extreme as the example above but nevertheless form an important part of the overall meaning and which are lost in audio-only documentation. One example which is used commonly by Aboriginal people and some others in Australia – the tongue protruding briefly indicating 'just joking' - sometimes visible only to some participants.

● **Preferred by community as record**

The above points mainly concern making the role of video in making sure that as full a record of communication in speech events is made. There are other reasons, including the fact that communities where the language is spoken generally prefer video to audio records of people and events. This is in part because of the same reasons of improved interpretability, but also because of the generally more natural and appealing quality of the video record.

● **Memory for video costs less and less as technology improves**

As mentioned before, the main objection to use of video for digital archiving has been the large volume of memory needed and the consequent cost and difficulty involved. The cost of medium scale portable storage suitable for working documentation eg large hard disks of 400-500 GB capacity and DVD burners has decreased dramatically over the last few years and is likely to continue this trend. For larger archival storage there are several options. The Aboriginal Language Acquisition Project for instance has obtained Access to APAC mass storage at the Australian National University, and the associated links at little cost. This is 'near-line' rather than 'on-line' – there is some delay in downloading archived materials to work on. The CLAN archive has recently announced that researchers will be able to work 'on-line' on their archived material including video, although what kind of equipment is needed to make this feasible and efficient is not clear.

5. Linking Audio-visual files to text files

'Documentary linguistics' tends to reverse the traditional way that documents are seen. Traditionally, the high level analysis of the material in its published

textual form – the grammar and dictionary, comes first, with texts – the written record of spoken word usually – next, and finally, and often altogether out of the picture and inaccessible, the primary materials on audio (or occasionally video) tape. In the new order, these primary materials are central and the other products secondary ‘annotations’. Within the annotations, also, the basic transcriptions and translations are very important. If those exist, then grammars and dictionaries can be produced later.

Linguists generally across the world tend to use a quite similar set of conventions for transcription known as ‘interlinear text’ format. This similarity or convergence in the tradition is a very handy thing as we have moved into the digital era as it has proved quite amenable to digital representation. There are several programs which have been used for a few years commonly by linguists. In Australia, for endangered Indigenous languages, probably the most commonly used program is SHOEBBOX (<http://www.sil.org/computing/shoebox>) As the site states: “It is especially useful for helping researchers build a dictionary as they use it to analyze and interlinearize text”. For the Aboriginal Language Acquisition Project, on which I am working with a team of researchers, we have decided to use the program CLAN (<http://childes.psy.cmu.edu/clan>) for transcription for various reasons which will be alluded to but not fully discussed below. CLAN does not ‘build a dictionary’ as SHOEBBOX does but that is a less important feature for our particular project. Both these programs share common features in the design of the text transcriptions inherited from the traditional linguistic interlinear design.

The essential core of such interlinear text formats are four tiers for each line of the text, which are aligned with each other, 1 and 4 at the line level, and 2 and 3 at the morpheme level.

1. **The text** (usually a transcription in a standard or practical orthography)
2. **A morphological break up of text**
3. **Gloss line – one gloss or morpheme code per morpheme**
4. **Free translation**

A number of other tiers can be added - further detail of the sounds using eg phonetics; indications of the paralinguistic elements discussed above; cultural notes etc. One of the differences between programs is the amount of freedom they allow in types of alignment between parts of tiers. Obviously metadata needs to be added at the text level about recording date and context, participants etc. In CLAN there is a well worked scheme for such information which is placed in the header of the text.

What CLAN and a number of other programs eg ELAN (Eudico Linguistic Annotator; <http://www.mpi.nl/tools/elan.html>) allow is the insertion of timecodes taken from the audio or video of the text into the transcription in a fairly user-friendly way, at least at the end of each line. This allows easy return to the part of the audio/video one wants to review, and other functions which mimic

analogue transcribing machines, as well as easy selection and copying of segments of audio/video with aligned transcription/translation.

One other type of linking of text and video/film image which has been common in commercial and to some extent research applications is *subtitling*. Computer programs for subtitling digital video are available, but as far as I know, no one has developed software which takes the text and/or free translation output of a transcription program like CLAN or ELAN and transfers it to the video screen as subtitling. This would be a useful extension of the functioning of audio-visual/text linking.

CLAN is part of CHILDES, a project dedicated to systematic recording initially of Child Language, but which has branched out from there. It is also used for other conversation analysis, discourse analysis and studies of bilingualism and code-switching.

It includes a set of computer tools called CLAN (Child-Language Analysis) which can search a corpus of transcriptions and answer a lot of complex questions about the corpus. CLAN works with CHAT (the transcription system), is very well documented in the manuals downloadable free from the website mentioned above.

Researchers can also join the CHILDES network and contribute their data to, and look at other data deposited in, the very large database of child-language there. The network also runs email discussion lists which in the experience of our Aboriginal Language Acquisition Project team so far have proved very useful.

Given the importance of a video record, discussed above, one immediately useful aspect of CLAN for us is that it provides a way of **transcribing video** and linking the transcription to the video-audio via timecodes.

Another reason for choosing CLAN for our work is the fact that so many analytical tools have been developed using this framework which can answer questions like 'how many times does a child between 2 and 3 years use an inflected form of a verb in the corpus' and list those examples. However the way in which CLAN does this by combining data from text headers and information in the text tiers brings with it a drawback. What is really needed is one central database of information about participants etc., not bits of this information scattered across many text headers, which then have to be updated separately when the need arises. We are developing a solution to this problem.

5. Digital A/V Archives

Having settled on a way of documenting languages and cultures by digital audio-visual means, there is then the issue of how these products are to be archived. As mentioned above our research group are going to use mass

storage at a university in Australia for copies, which are generally minimally edited and we will be investigating placing another copy in the CHILDES database in the USA. For the original mini-DV tapes we will store these at the AV Archives at AIATSIS if this format is accepted, or otherwise at Melbourne University. Issues of appropriate metadata to accompany these materials are being worked out, and we are hoping that the database we are developing, referred to above will provide ready made metadata which will also be easily mapped on to the standards required by the archiving institutions and national and international standards (such as OLAC).

The technical issues are far from the only ones in audio-visual archiving. Rights including access rights of various parties including the speakers of the language are serious issues which can be the source of trouble if neglected. Doug Whalen in a keynote address to the International Congress of Linguists (2003) rightly stressed that language documentation is becoming much easier through computerization and the Internet. While the technology gap between the academic linguists and many of the speakers of endangered languages is great, it will probably narrow in the future and allow language communities to share in the fruits of these advances. However there is a need for language stakeholders to understand and agree to the system which is being put in place, and that means that they must see the advantages of the system of transfer of knowledge.

In my experience one of the key questions is: how can speaker community people see advantage in outsiders accessing their cultural and linguistic heritage? Appeals to the advancement of science, or the benefits of a shared national approach to heritage are not necessarily persuasive to people who may be disenchanted after years of oppression and neglect. There is a need for a 'two-way' process between the local language communities and the community of scholars, a free and equal exchange which can be mediated through regional bodies under Indigenous control (see following section). In Australia many Indigenous communities are happy to share information with the world at large (apart from restricted secret-sacred and other private matters) but there is a disturbing trend towards closing off access to most language materials to outsiders in a few places (Newry & Palmer 2003). On the other hand, however, I have just attended a workshop of the newly-constituted New South Wales Aboriginal Language Research & Resource Centre in which a large group of Indigenous people looked closely at the implications of web-delivery of language resources, and were not generally of a mind to be overly restrictive about access. An important element here is the sense engendered by such meetings among Indigenous stake-holders that they do not have to lose control of decision-making when entering the world of new technologies.

6. Bundling related resources

Returning to a more technical question about Archives, one of the problems facing us now is how to put in place the right kinds of links between different

resources or objects that need to be viewed together or work together as a multimedia output. This goes far beyond the notion of 'multimedia' as a specially-produced one-off assembly of vision, sound, text and other files, such as on a ready-made web page or CD-ROM. What we need now is mechanisms to put such elements together in the normal course of work as we search for and analyse data. I will examine first how to link different files if they are within one database, and then move on to discussing how this might be done if these are in different databases, even in different locations and institutions. This is written from the point of view of a researcher and user of such tools without much understanding of what particular computational solutions might come into play to implement them.

A simple and everyday kind of application of this idea of linking files is the need to use a video or audio file and a transcription file, synchronised with each other, when analysing basic linguistic documentation. CLAN will do this in a simple fashion, and ELAN working with its companion metadata framework IMDI will do it in a more sophisticated way. As noted above the ALRA project is engaged in building a database which will contain metadata about the a-v files, the participants etc, so this will need to be linked as needed too to feed information into CLAN.

In the IMDI approach, different files which have as a common element the 'session' at which the primary data was recorded are linked together as a 'bundle', which has a unique code. As the IMDI documentation describes it (<http://www.mpi.nl/IMDI/tools/>)

A session or resource **bundle** contains different files. The following four types of files exist:

(1) Metadata files

These files contain information about the session, e.g., its date and location, its content and its participants. They are of the IMDI Editor format.

(2) Media files

These files contain the audio or video recordings. They are usually digitized in one of the following formats: MPEG (*.mpg), Cinepak-Quicktime-Movies (*.mov), WAVE (*.wav).

(3) Annotation files

These files contain the transcripts, codings and annotations. Their format varies (e.g., ELAN, Media Tagger, Shoebox, CHAT, etc.)

(4) Info files

These files contain further background information on specific topics. They are in PDF or HTML formats.

The Archive for the Indigenous Languages of Latin America (AILLA; <http://www.ailla.org>) has adopted the IMDI scheme. Heidi Johnson of AILLA has pointed out

'A problem for the DC[Dublin Core]/OLAC model: How can we keep related resources together & make sure users get all the parts they need? IMDI protocols support bundling, a key consideration for AILLA'
(<http://www.language-archives.org/events/olac02/presentations/imdi.ppt>)

'An AILLA [archive] resource is actually a *bundle* of closely related *items*, or *files*. That means that there are several files associated with the "One-eyed Grandmother" resource, including audio files in different formats and a transcription and translation file in PDF format'

While this arrangement (with the file coding system which implements it) achieves the objective of being able to relate different files that are about one session, transcripts and translations in PDF format would not presumably allow synchronised linking like CLAN or ELAN.

Both IMDI and the AILLA scheme which is derived from it are also heavily based on the notion that the common component is the recording session. This is probably a good principle to work on to introduce bundling order into an archive. However one can easily imagine other types of important links between files that would need to be available in the metadata, and realised as actual links in some relational database or hypertext structure. For instance, a name of a participant in a speech event could be linked to that person in a genealogy, or a text dealing with places could be linked to gazeteer or map.

Another issue is that these schemes assume that the linked resources in a bundle will be in the same digital archive in the same location. It is quite frequently the case however, that parts of a virtual 'bundle' or related resources will be in one place and others in another place. This may be unfortunate in a sense but given institutional inertia it is not going to change fast, but it does not mean that links cannot be made. Given that we are fast moving towards compliance with international metadata and other technical standards and hence interoperability, it is quite likely that researchers will be able to work on separated elements of resource bundles as if they are together.

How this could be achieved is beyond my ability as a non-computational linguist to say. However it has struck me that the Dublin Core metadata element *relation* is somewhat under-used and could be used to link files in different ways, once suitable refinements were introduced. The current set is like this:

Relation metadata element (Dublin Core/Dspace)

relation		Catch-all for references to other related items.
relation	isformatof	References additional physical form.
relation item.	ispartof	References physically or logically containing
relation	ispartofseries	Series name and number within that series.
relation item.	haspart	References physically or logically contained
relation	isversionof	References earlier version.
relation	hasversion	References later version.
relation	isbasedon	References source.
relation	isreferencedby	Pointed to by referenced resource.
relation function,	requires	Reference resource is required to support
		delivery, or coherence of item.
relation	replaces	References preceding item.
relation	isreplacedby	References succeeding item.
relation related	uri	References Uniform Resource Identifier for
		item.

Documentary linguistics could surely add to this list with some very basic types of relation such as x is a transcript of y; p is a translation of q. These would be an important means of linking remote related files via metadata.

More ambitiously one can imagine dynamic research projects that create multimedia linkages by means of other kinds of metadata links. Initially the projects would have to create such links but as knowledge networks converged on ways of handling relations, these could be at least partially in place. One (at this stage imaginary) based on work I have been carrying out jointly with archaeologists and material culture specialists (McConvell & Smith 2003; Akerman & McConvell 2002) is diagrammed below.

The dynamic bundling or linkage revolves around a particular type of artefact – a ‘muller’ or top grindstone. It may begin with a video of an Indigenous person demonstrating and talking about how seed-grinding was done – a highly endangered knowledge today. This digital or digitised video is transcribed and the transcription is linked by standard relation metadata (as well as by timecodes etc) to the video audio track. However additionally other links could be made whereby a user could jump to other objects which need not be in the same archive eg (a) a dictionary of the language in question at the entry for the particular word for this artefact; (b) a map showing locations referred to in the text such as quarries; (c) an image or images of such artefacts on a Museum website or some database of Australian Aboriginal artefacts.

LINKING DIGITAL RESOURCES REMOTELY

