Planning multimedia documentation

David Nathan

This paper describes some of the issues and processes in initiating and planning an interactive multimedia CD-ROM for an endangered language (EL). It provides some guidance for dealing with issues that are not normally faced by linguists. It is beyond the scope of the paper to detail the skills required for preparing data or creating usable applications (such as media editing, data processing, or multimedia authoring). However, often the greatest difficulty faced by linguists is to identify *which* skills are needed—skills they may need to acquire or get assistance with. The focus here is on the *Paakantyi* CD (Hercus and Nathan 2000), and the interweaving of planning, community participation, and a team approach in its construction. Paakantyi is the language of the lower Darling River, south-western NSW, Australia. While Paakantyi no longer has any fluent speakers, the input of Paakantyi people had a profound influence on shaping the CD's design and content.[1]

Documentation, archiving and multimedia

Many linguists have become aware of the deficiencies in current practice in handling endangered languages data. These deficiencies fall under three broad headings:

1. *documentation* of a broad range of linguistic phenomena (cf Himmelmann 1998)
2. *archiving*, including the preservation and access to data (cf Bird and Simons 2003)
3. *mobilisation* of materials into usable resources for practical language support (cf Nathan, 2003)

There is considerable confusion and debate about how to successfully conduct these three types of activities. Firstly, at least, it is crucial to distinguish them, as they have different and even opposed aims. For example, we often hear researchers proposing to "archive the data on the web", perhaps looking for a way to simultaneously achieve all of (1), (2), and (3). However, the web uses data formats that are not optimal for archiving; more generally, the web is probably the most unstable information system in use today.[2] Communities and language consultants may resist the idea of using the web to provide free and universal access to their EL data. Since most ELs are endangered due to social, political, military or economic reasons, many communities are sensitive about releasing or exposing what may be some of the last vestiges of their distinguishing identity. In today's EL research and documentation, community participants must be able to understand the potential outcomes of the work; if they know that materials will be freely and unqualifiedly disseminated, it may discourage

[1] I am grateful to Luise Hercus for comments on a draft of this paper.

[2] For example, typical current estimates for the lifespan of an average web page range from 44 to 100 days. Even pages that persist are commonly subject to undocumented changes in content or changed URLs.

collaboration or reduce the amount of language material that they allow to be collected (Nathan 2004b).[3]

Mobilisation of linguistic material has received less attention than either documentation or archiving, probably because it requires linguists to play an active role in the invention and evolution of genres for presentation of linguistic knowledge, something that linguists are, paradoxically, generally poor at. Such new genres are difficult to work in because they involve team participation by people with a range of skills (Csató and Nathan 2004), exacerbated by the tendency of linguists to work alone.[4]

Mobilisation is even more urgent than the other two needs where languages are ceasing to serve social and cognitive functions, and if communities have justifiable demands for linguists' support in shoring up languages while some last full speakers remain healthy. Actual efforts by language experts to counteract language endangerment and (impending) loss reflect the sincerity of their claims. One of the best ways to mobilise language materials is through multimedia delivered on CD or the web (Warschauer 98, Csató and Nathan 2004).

Multimedia development provides opportunities for language documentation as well. Its advantages over static, written materials, such as the ability to present sound as an integral component of the materials and results, ought to be profound for linguistics. Multimedia helps to focus project design on the variety, quality, and authenticity of language events and performances, thereby supporting a framework for creating rich language documentations that can support a variety of purposes (Himmelmann: 1998). Multimedia typically means multi-skills and therefore requires the participation of several people; it puts linguistic work into a team situation and exposes it to the expertise of others such as designers, teachers, and programmers. Using multimedia can provide the motivation and contexts for encouraging community participation in producing language documentations. Existing "legacy" materials can be given new life by using them as assets in a multimedia product.

On the other hand, working with multimedia poses many challenges. The diversity of inputs can create problems in dealing with intellectual property rights, as well as a range of issues to do with the complexity of construction, interaction design, asset file formats, product performance, and high demand for storage space and transmission bandwidth.

Multimedia materials are not easily archived. Because multimedia typically requires delivery via specialised (often proprietary) software, there are specific and possibly short-lived requirements for particular data and file formats, software versions,

---

[3] Current ethical principles governing much linguistic work include informed consent of language project participants. In addition, there is growing appreciation of rights to privacy. See, for example, AIATSIS 2000.

[4] Linguists do work collaboratively on some activities, such as the publication of papers and books. Here, the maturity and transparency of the roles and relationships in traditional publishing make the complexities of the collaboration between linguist, software, editor, designer, typesetter etc. almost invisible.

operating systems, and platforms. Nevertheless, arguments against producing multimedia on these grounds are weak. There are fundamental challenges to be faced in *any* contemporary use of digital media for documentation: the instability of open-standard media file format specifications, a lack of conventions for describing and building interfaces, and the fact that it is not yet known how to represent and archive abstract content such as navigation, layout, links and interactivity. These limitations mean that choosing to work with multimedia is a result of recognising its advantages for language documentation and language strengthening projects, rather than being a broad strategy that can satisfy various other needs such as long-term data preservation.

Conception

Multimedia projects are typically more time consuming and expensive than other activities; compared, for example, to producing an edited collection of papers, one would have to add other factors such as that the participants' backgrounds, skills and materials vary widely, and their contributions are to be intertwined in the product. Therefore, a clear conception of the type and scope of the product are required.

Community initiation, support, and ongoing participation are important ingredients. See AIATSIS 2000 for a good general ethical framework for working with Indigenous communities. Ethical and protocol considerations are highlighted in multimedia because it provides a more direct channel between the information providers and its audience.

Here are some of the factors to be considered when writing a plan or a funding application:

- community initiation of project or level of support for project's aims
- researcher's understandings of the language situation, including amount of usage, its resources, and the level of interest in the language in the community and education sector
- the available community contacts, relationships with them, and their accessibility
- the availability of professional colleagues and their interest/skills in multimedia
- other resources in community, e.g. sources of skills or materials for art/graphics, and music; links to education institutions
- the availability of relevant data e.g. dictionary, texts, recordings, photographs

The *Paakantyi* CD (Hercus and Nathan 2002) was produced by the author and Dr Luise Hercus in response to perceived language needs within the Paakantyi community and following discussions with community members. This took place in a fertile context; language revitalisation was beginning in the Paakantyi community and was already gaining momentum across New South Wales, and this was coupled with promotion of access-enhancing language projects by the Aboriginal and Torres Strait Islander Commission (ATSIC), the peak national Indigenous organisation, as a result of attention to issues of Aboriginal languages and identity following the Royal

Commission in to Aboriginal Deaths in Custody (Commonwealth of Australia: 1991).[5] Figure 1 shows our original statement of aims for the Paakantyi CD as submitted to ATSIC.

Fig 1. Paakantyi CD: the original rationale for the project in the funding submission

Aim
To produce, by collaboration between community, linguist and software developer, an interactive multimedia CD for the Paakantyi language that will:
- Provide a resource for Paakantyi language teaching
- Record and archive existing language resources
- Generate new language resources
- Generate enthusiasm for language activities
- Provide an introduction to multimedia production for community participants

The language community should clearly be designated as a major audience for a multimedia product. Not only will it probably be a valuable resource for them, but also community participation is crucial to the construction of meaningful and authentic multimedia materials. Funding and other available resources should be used to make a *new* contribution to language and community resources, rather than use a new format to dress up something that is already available.

How can you decide on the content of a multimedia resource, especially when it may be influenced during the course of the project by evolving ideas and a changing scope of collected materials? The first decision to make is whether you should use existing materials as the project's primary assets, or generate new materials. Generating new materials is preferable; it allows resources to be distributed more broadly (because the project and participants will probably have intellectual property rights in materials), and enables planning and development work to take place within the community, thus providing more opportunities for input, interaction, and the development of a "biography" of the project within the community (Csató and Nathan 2004). Newly recorded sound material is likely to be more consistent and of better quality.

Some (e.g. Rood 2004) have argued that the richness of older language sources places them above new recordings of lesser-accomplished speakers. Considered in a vacuum, this might be valid; however, multimedia production has to take into account socio-political realities in communities (which will largely determine how accepted and effective the eventual product will be), and accepted media production methodologies (where quality and consistency of presentation also determine the effectiveness of the product).

Planning

I discuss here some aspects of the planning and evolution of the Paakantyi CD. In initial consultations, members of the Paakantyi community expressed interest in a "talking dictionary". The value placed on dictionaries as symbols of a language's importance is well-known; in addition, following a number of approaches by

[5] ATSIC was established in 1989 and has recently been abolished by the Federal Government led by John Howard.

Aboriginal people following the publication of the text-only Kamilaroi/Gamilaraay Web Dictionary (Austin and Nathan 1996), we also knew that many Aboriginal people simply wanted to *hear* the words.

The emphasis on a dictionary was also attractive because Luise Hercus' published dictionary of Paakantyi (Hercus 1993) was in need of updating and could be revised and distributed through the project's CD. However, details of the content and design of a talking dictionary were unknown at the beginning.

Before we set out on our first fieldwork trip, we discussed the methodology for eliciting the pronunciations of words for a talking dictionary. In my role as multimedia author, I was concerned that it may not be possible to record enough words to provide sufficient coverage without asking language consultants to make prompted recordings. A debate centred on the authenticity of such "staged", pronunciations by limited, second-language speakers of ELs (or "rememberers"). Luise, on the other hand, as a linguist and researcher, wanted to learn more about Paakantyi through traditional elicitation methods, but also did not think that we would collect much data due to the severe state of language loss in the community. Nevertheless, the eventual outcome surprised all the participants: as a result of a creative tension between approaches, the unmitigated enthusiasm of participating community members, and, perhaps, partly through luck, the number of words (and other materials; see below) recorded was way beyond what any of us had expected. During the project's three fieldwork visits, the language consultants became comfortable with our evolved working styles and found it easier to recall—and produce—language that they had not heard or used for many years. And as other community members saw drafts of the emerging product taking shape, more people offered to do recording sessions with us.

We recorded many more words than anyone had thought possible. In addition, interesting patterns in the consultants' responses provided important hints and sources for the design of the CD, beyond the simple template that had been the basis for initial discussions. Here is one example: our consultants (Renie Mitchell, Lottie Williams, John Mitchell and Badger Bates) would often follow the pronunciation of Paakantyi words with English glosses and explanations, and possibly some example usages. For example, John Mitchell provided the recorded entry for the word *murarta* "fast, quick" (as in Hercus 1993). What he said was "mura-mararta 'hurry up'… mura-mararta thikalanaapa 'hurry up, I'm going home.'" Several non-trivial linguistic and design decisions needed to be made in order to accommodate such material:

- Paakantyi speakers sometimes used forms that were different from those in the published dictionary. In some of these cases, the new data led to revision/correction of the dictionary; however, in most cases we simply juxtaposed the published (upper part of screen—see Figure 2) and the speaker's (lower part of screen) forms. It is up to the user of the CD to choose the form that appeals to them.
- Speakers often followed a word by its gloss or translation in English. Originally we had planned to include only the speakers' Paakantyi productions. But not only was the "mixed" pattern pervasive across recordings, we also realised that it had important interface design and pedagogical advantages for the CD, since

it makes sound content independent of written material, and accessible to preliterate children, or people with poor eyesight or sitting away from the computer.

- For some words, several usage examples were provided. To handle this we used a database to record the assignments of examples to entries,[6] and designed an interface system of numbered access buttons that, when clicked, reveal the usage example text and play its sound (in Figure 2, there is only one usage example).

The usage examples became a main asset of the CD. They help address a pervasive problem encountered in language revival situations: the vast, mostly empty space between dictionary and grammar, where people no longer know how to express ordinary, everyday matters (cf Pawley and Syder 1983).

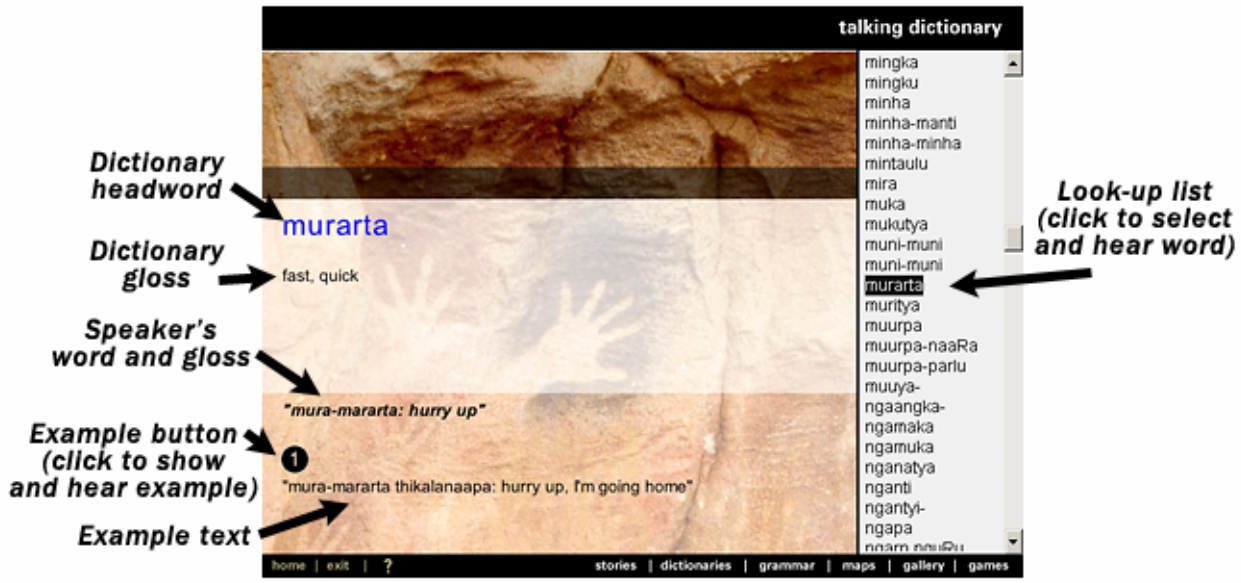


Fig 2. Paakantyi CD: design of the "Talking Dictionary" interface

Thus the type of materials recorded led to the design of a dictionary interface to accommodate and provide access to them. In fact, the CD included three dictionaries: a full Paakantyi dictionary (the update of Hercus 1993), an English to Paakantyi dictionary (recognising that for most people the access to Paakantyi words is via English), and the talking dictionary. The talking dictionary was presented as a separate dictionary (rather than, say, only having sounds for a subset of words in the main dictionary), for two reasons: so that community members could easily find the resource that they had expressed their strong interest in, and for consistency, so that within the talking dictionary, *any* word can be clicked on to predictably hear its sound.

The talking dictionary was more richly structured and populated than we had initially expected, and when we finally presented the CD to community members they told us that it was exactly what they had wanted! The reasons: because we had listened in the planning stage to their views, had involved as many people as we could in the

[6] Described further later in this paper.

development, and had regularly produced working drafts to show people how the product was evolving. This result, at least, was not a matter of luck.

The team

Multimedia means multi-skills and usually a project will be run by a team of people with a variety of roles and skills. There are four main roles. Firstly (because this is probably you, the reader), there is the linguist, someone with skills and experience in relevant areas of linguistics including language description and documentation, sociolinguistics, lexicography etc.. The linguist should have a good background not only in the language, but also with the community, and its local etiquette—in other words, the linguist ought to be a specialist, not a “generic” linguist. Secondly, community members should take various roles—as speakers/consultants, artists, recordists, advisors, and liaison. The community category may also, where appropriate, include people “married-in” to the community, and may be extended to include local enthusiasts, and clerics etc.. Thirdly, there will be a software developer or “IT” person, typically a multimedia developer rather than a business application programmer—this person may work closely with the linguist, or may coordinate the project.

Finally, you should engage a graphic designer. One of the pitfalls in conceiving multimedia materials is to imagine that merely adding in some sound or graphics can transform what you ordinarily do into a multimedia product. In deciding to make a multimedia product you are entering the realm of an entirely different genre, one that is populated with carefully designed and often richly interactive titles such as games. Therefore, do not attempt to make a significant multimedia product without a graphic designer. Professional graphic designers know how to interpret the project brief and your materials in order to create screens that will enable the product to achieve its aims. Furthermore, a designer will bring style and individuality to the project, and will most likely make valuable suggestions about interactivity and other matters surprisingly relevant to the linguistic content.

Try to find a graphic designer who is recommended by someone you know. It is not necessary to find someone who has a lot of experience designing CDs; it is more important to find a good, professional designer who is motivated toward your project than one who is widely experienced in digital media. Feed your designer with overviews, examples, and a large amount of assets such as photos, artwork, and ask them to come up with a design concept.

As with other members of the team, you will need to “go halfway” with your designer, for example, by learning some of the specialist terminology used in each area, using software in common, and efficiently exchanging graphic materials. Be open and be honest about your skills and aims, but give in very easily to the designer’s advice on design issues. On the other hand, take care to avoid over-design at the expense of depth of content and interactivity—a telltale sign of this is that the deeper you navigate, the less content and interactivity there is. Make sure that the designer does not bring

stylised notions of the language community to the project.[7] It is easy to avoid inappropriate representations by having community members' input on art and design, and by providing regular draft versions of the product for evaluation. Typically, you will have to work together to design the interactivity, and to solve problems in handling text, such as choosing or creating fonts or deciding whether screen text should be active (selectable/clickable) or graphic.

In some circumstances you may have the resources or the luck to have access to other skills, such as instructional or pedagogical design.

In a good team, there will be differences of opinion, but members will find it possible to defer to the views of the member primarily responsible for a particular area. Typically, the linguist and language consultants should have final say on linguistic issues, and the graphic designer should make the aesthetic decisions.

The cost

People often ask how much it costs to make a multimedia CD-ROM. This question is impossible to answer generally because CDs vary enormously in scope. Practitioners' costs vary wildly too; design and multimedia authoring can be expensive but on the other hand many projects are completed through the unpaid efforts of people dedicated to completion of projects where the funding is limited.

Budgets should take into account payments to community members for language work, artwork licensing, liaison, and other contributions. Remember too that CDs are "media hungry"—they typically require many images, which need to be sourced and paid for (to photographers and artists, or for royalties/licences), and that these images still require significant graphic design work to adapt them for the product. Projects may also include activities such as in-community workshops which may be expensive to organise and run.

Nevertheless, here are rough figures for some project budgets. The Paakantyi CD received about $A50,000 (£20,000) funding from ATSIC, which was mostly used for fieldwork, salary, language and art input from the community, software, and graphic design. Most of the content and time supplied by Luise Hercus was unpaid, and a significant proportion of the funds were retained by the host institution AIATSIS. The *Yolngu Languages and Culture: Gupapuyngu* CD (Christie et al 2001) cost a similar amount to develop; again, not including the time spent by academic colleagues developing materials over several years prior to the project. The *Spoken Karaim* CD (Csató and Nathan: 1998) was funded through approximately one year's salary for the developer (Nathan) and a direct grant to the graphic designer. Much of the material, such as recordings of speakers, transcriptions etc. was also created by the linguist prior to the project. However, it was a complex, innovative software project, with research and development taking well over two years.

---

[7] For example, some designers seemed to have believed that Indigenous people need graphics of landscape objects such as rocks and animals for navigation buttons. However, multimedia has its own conventions (some of them closely related to standard software conventions) and users are not helped by attempts to simulate realities using absurd associations.

Storyboarding

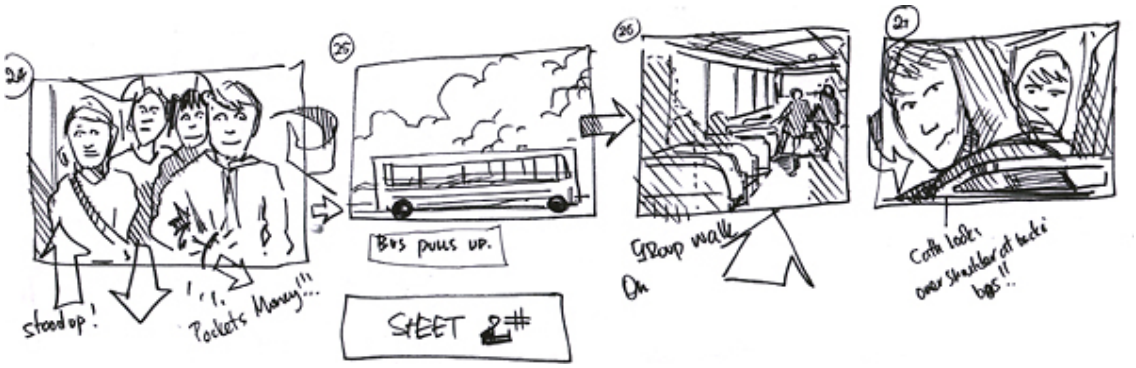


Fig 3. Cinematic storyboard fragment from 'What Goes Around', (First Light) http://www.firstlightmovies.com/storyboard.php

Storyboards originated in cinema production; they basically describe graphic change over time. Storyboarding is even more important for multimedia projects, because the conventions of multimedia are nowhere near as developed as in cinema.

Storyboarding is typically done by the software developer and the linguist, with later input from community members and the designer. Although it is a time consuming process, storyboarding fulfils many functions. Perhaps most importantly, it provides a gentle, non-threatening way to begin the design process and to solidify and flesh out ideas. It also provides a forum for negotiation of the project's scope and boundaries/limits, for identifying omissions or problems, and for bringing together disciplines. It is the phase of the project where participants check interpretations of terms and develop shared understandings of concepts and terminology. In fact, it is a good strategy to use the storyboarding phase to develop a shared nomenclature for the project's objects, sections etc.

Later, the storyboard becomes the actual blueprint that the designer and multimedia author will follow.

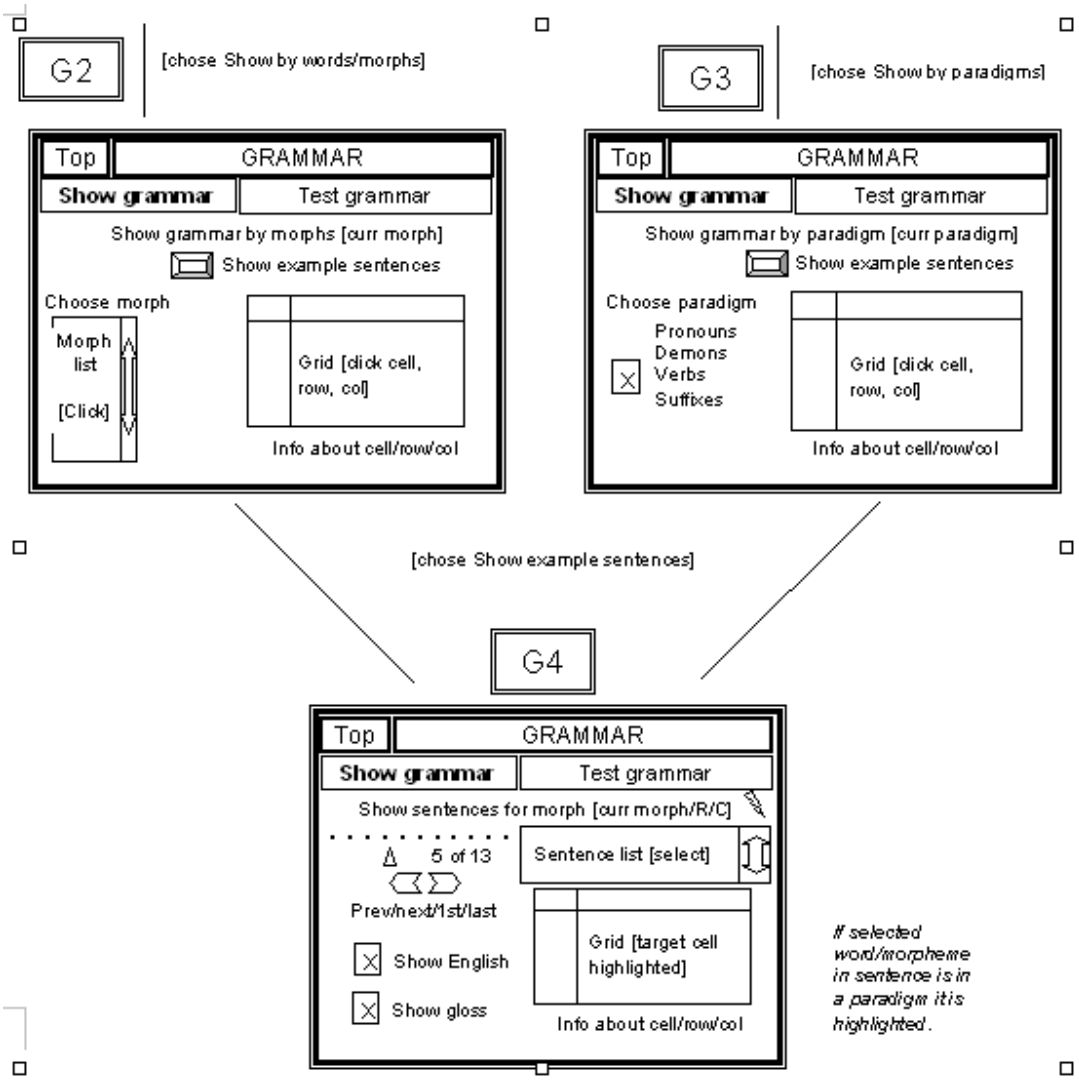
Here are some hints for getting started and creating storyboards:

- first, review products that are similar to what you imagine developing; view them in a team situation and say what you like and dislike about each one
- don't take storyboarding too seriously at first; just start drawing out screens on large sheets of paper, starting with the "splash" or opening screen
- your storyboard should describe the *functional* content and *behaviour* of the product, not the graphic design and layout. Your graphic designer will interpret the storyboard to make a design
- you need to deal with every single functional object you want to see on the screen and every conceivable navigation that the users will make—fudging or hand waving will cause difficulties later
- make sure that the following aspects are fully described: number of major areas/divisions, menu structure, main active data objects (e.g. scrolling clickable lists, other controls). Typically you will have a menu that corresponds to the major areas in the product; it usually also includes Help, Home,

Settings/Preferences, Quit, and perhaps Forward and Back (the latter are challenging to implement in interactive multimedia)

- remember to cover sounds—when and how they start and stop playing
- typically you will underestimate the amount of linking and interactivity that good multimedia should have
- there should be *more than one way* of doing most things (e.g. navigating, issuing commands, making choices)
- work alternatively top-down and bottom up—don't spend too long on either perspective
- many types of changes may still be made later. These can be made in reference to the original storyboard. Typically, some components will be added, and some not implemented
- ensure that the developer does not divert the linguist from linguistic aspects; on the other hand, encourage the developer to alert the linguist to possibilities that he/she may not have thought of
- beware the temptation to have a set of screens that are static or appear in a fixed order; this is a sign that you are making a training video rather than an interactive resource

Fig 4. Example—part of storyboard for another CD, *Gupapuyngu*



Recording sound

Methods for handling and presenting sound remain a major gap in the emerging field of language documentation. Sound is challenging to acquire, process, present, distribute, and archive.

If you are making new recordings, pay the most attention to microphones, the "ears" of your project. Microphone selection and usage is an area where linguists have typically been under informed.[8] But even before plugging in a microphone, the fieldworker should carefully consider the physical environment and human context—these will affect not only the quality of the recordings but also the smooth running of the sessions and the happiness of the participants. The best place to record is in a recording studio. You should consider spending some of the project funds on travel and accommodation for language speakers so that they can work with you in a studio. You may find a local radio station not far from the fieldwork area that has suitable facilities and is willing to let you use them. Simpler strategies can also work: for the

[8] For guidance on microphone selection and usage, see http://www.hrelp.org/archive/advice/microphones.html

Paakantyi project we stayed at a motel in a centrally-located village. We rented an extra motel room so that the one of the main consultants who lived more than 100km away could stay together with us (she could also use the opportunity to visit family and do shopping). Recording in the motel meant that we could work flexibly, we could avoid some of the noise and interruption that inevitably occurs at people's houses, and we could better control the recording environment without intruding on the integrity of people's homes (for example, we were able to turn off the refrigerator and move the furniture).

Strategies for making good recordings are covered elsewhere; here it suffices to make three points that will determine your ability to make recordings suitable for interactive multimedia:

- multimedia will typically be used on computers which emphasise high-pitched sounds, so low-hiss, *digital* recordings are best
- ultimately, microphone selection and location is the most influential factor in determining the quality of recordings
- in interactive multimedia, sounds will be encountered in various orders, and in unpredictable juxtapositions, so consistency (of volume, quality etc.) is absolutely crucial and extraneous noises should be avoided

There are many other issues to consider in dealing with sound, such as recorders, backup, digitisation and data processing (described below), and designing and programming interfaces for access and control; however, observing the three points above should make sure your team's recording efforts result in suitable sound assets for a multimedia product.

Processing of recordings

This section describes what was subsequently done with the recordings made in the field for the *Paakantyi* CD's talking dictionary.

In the field or immediately after our return, the recordings (made on Sony minidisk) were redigitised using a desktop computer with a good quality sound card, to produce 44 KHz/16 bit/mono .wav files (which correspond to CD quality except mono).[9] Each session (minidisk track) corresponded to one sound file whose name identifies the original session (eg. PMD3-2.wav was the second session on the third minidisk). These files were then written to data CDs.

Next we needed to transcribe or at least classify the content of the sound files. This was a task for the linguist, Luise Hercus, and to save her time having to listen to more than 10 hours of recording, much of it with no useful material, I made a rough edit and condensed all the segments containing potentially useful material into four CDs. Each of the segments was named to retain a link to the original session (eg. PMD3-2_ed.wav was the edited/condensed version of PMD3-2.wav).

[9] It would have been slightly better and possibly easier to use a ful-sized MD deck to do direct digital recording (our portable minidisk only had analogue output). However, this is not worth taking trouble over for most purposes; redigitisation (ie DA-AD) results in imperceptible loss of quality if reasonable quality equipment is used carefully.

Luise transcribed the condensed material, also annotating time her transcription with time offset values and comments about potential usage of words and phrases—see Figure 5.[10]

Fig 5. Example of annotated transcription for *Paakantyi* CD's talking dictionary material

transcription	sound offset	type	gloss	speaker
nhumparka that's green	022	lex+gloss		IM
wiirpa	042 or 43	lex	cloth	LW/IM
yarra	043	lex	tree	LW
yarra is a tree and kamara means gum	052	lex+gloss		LW
kurkuru	054	lex	box tree	LW
kurkuru yarra, that is where you used to get the grubs from when you couldn't get them from the gum tree	055	lex+gloss		IM
thuluru kunti partaana	117	sentence	many mosquitoes are biting	IM-LW
muni-muni	125	lex	green ant policeman	LW
wiikilaana	129		?	IM
you can talk.. yarnta too, you got any yarnta	135	gloss		IM
kaankuru	138	lex	horse	LW
thartu-pulyki	145	cpd	head-hair	LW

Luise's annotated transcriptions were then processed and entered into a database which had already been created and seeded with data from her published Paakantyi Dictionary (Hercus 1993). Further tables and fields were added to the database as we worked on the data that would eventually support linking the dictionary text, sounds, and example phrases/sentences (see Figure 6).

Fig 6. Database in transition, assigning sound assets for use as entries or examples

ID	Section	file src	form	gloss	DicID	type	speaker	comment
16	S1	116	thuluru	a lot, many	1741	l.	LW	
17	S1	117	thuluru kunti partaana	many mosquitoes are biting		sentence 2094, 433, 1741	IM-LW	
18	S1	125	muni-muni	green ant, policeman		l. ALSO 724	LW	25 JUNE REPEATED IN 746
19	S1	129	wiikilaana	?			IM	
20	S1	132	yarnta	stone, money	2152		IM	
21	S1	135	you got any yarnta?			gloss 2152	IM	

The details of the database work are beyond the scope of this paper. To summarise, we evolved data of the type shown in Figure 6 into discrete, regularised tables that could be

[10] This could also have been done using software such as *Transcriber* (http://www.etca.fr/CTA/gip/Projets/Transcriber/); however, it is the data structures that are important, not the software used to create them, and Luise preferred to use a simple sound editor to get the time offsets and write the data into tables in a word processor—a process that was familiar to her and resulted in well-structured data for import into a database.

imported into the multimedia application and then support the kinds of linking and interaction we had planned. A fragment is shown in Figure 7—notice how some of the data in the second data row of Figure 6 (ID 17) has been represented as simple relationships that correspond to some association or action. In Figure 7, each row corresponds to a use of an item (identified by its SoundID in the second column) as an example for an entry (identified by its LexID in the third column) in the talking dictionary—see Figure 2.

Fig 7. Fragment of a derived, normalised table

ID1	SoundID	LexID
4	17	1741
5	17	433
6	17	2094
8	21	2152
9	24	1602
10	24	1404

The time offset data was used to locate the sound segments to appear in the CD. We created individual files for each item, naming them carefully according to a predesigned schema, and storing the filenames in the database.[11] In a small number of cases, detailed sound editing was done, e.g.. to replace syllables that had noise intrusion such as paper shuffling with a suitable syllable from another item. All sound files were then prepared as assets for interactive multimedia:

- fade in / fade out of each item (*high priority*: so that transitions and juxtapositions are always smooth, and items do not cause clicks or thumps when they start or end playing[12]
- normalise sound volumes of some items (*medium priority*: some sounds were rather quiet)[13]

The remaining steps involved exporting the sound linking data from the database into the multimedia application, programming the way that sounds were accessed and controlled, and integrating the lexical data and the dictionary graphics produced by the graphic designer.

Conclusion

This paper has discussed various steps in planning the *Paakantyi* CD. Many other activities were required to bring the CD to completion and are not described here; yet,

[11] Today, the computational environment for sound has changed such that it is no longer necessary to make individual sound files for each playable item, but rather to store the sound offsets for the start and end.

[12] The same result can now also be obtained in high-end multimedia authoring software such as Macromedia Director through scripted control of playback qualities of individual sound files or segments of sound files (see previous footnote).

[13] We also applied noise reduction processing to some items, but this was of low priority. Noise reduction is usually not worth doing unless very high-end equipment and a knowledgeable operator are available.

in many ways, the positive outcomes of the project were largely determined by the steps described here. Some of those outcomes were:

- a multimedia resource meeting some of the community's needs and requests
- a resource for assisting language teaching programs at local schools
- raised community awareness of and interest in the Paakantyi language
- creation of the first talking dictionary and talking crossword of an Australian language
- creation and publication of new Paakantyi materials (Mutawintyi story, Paakantyi grammar)
- documentation of previously unrecorded words and expressions; additions and corrections to the Paakantyi Dictionary
- development of new interfaces for accessing sounds and better understanding of interfaces for presenting materials to remote Aboriginal children
- assertion of the Paakantyi community's relationship to the Mutawintyi site
- financial rewards for language consultants and artists

Relationships between stakeholders, and the quality of the product, will be optimised if the planning and development process is concrete and open. By exposing product planning and evolution to the community you can not only get valuable feedback about design, content, and usability,[14] but also build a local "story"—a biography—for the eventual multimedia product. The community's relationship with it when it emerges will most likely enhance its acceptance and usage, thereby best supporting the aim of such documentations: supporting communities in their efforts to maintain/revive their languages.

References

AIATSIS 2000. *Guidelines for Ethical Research in Indigenous Studies*. http://www.aiatsis.gov.au/corp/docs/EthicsGuideA4.pdf

Austin, P. and Nathan, D. 1996. *Kamilaroi/Gamilaraay Web Dictionary*. http://coombs.anu.edu.au/WWWVLPages/AborigPages/LANG/GAMDICT/GAMDICT.HTM

Bird, S. and Simons, G. 2003. "Seven dimensions of portability for language documentation and description". *Language 79*

Christie, M., Gaykamangu, Waymamba, and Nathan, D. 2001. *Yolngu Languages and Culture: Gupapuyngu*. Faculty of Aboriginal and Torres Strait Islander Studies, NTU [Multimedia CD-ROM]

Csató, É. and Nathan, D. 1998. *Spoken Karaim* (L version). Tokyo University of Foreign Studies. [Multimedia CD-ROM]

[14] For example, we were able to get (positive) evidence about the usability of the crisp, text-based interface that our graphic designer had provided.

Csató, É. and David Nathan 2004. “Multimedia and documentation of endangered languages”, in P. Austin (ed.) *Language Description and Documentation*, pp 73-84

Commonwealth of Australia, 1991. *Report of the Royal Commission into Aboriginal Deaths in Custody* (Volumes I-V)

Hercus, L.A. 1993. *Paakantyi Dictionary*. Canberra: L.A. Hercus

Hercus, L.A. and Nathan, D. 2002. *Paakantyi*. Multimedia CD-ROM. Canberra: ATSIC.

Himmelmann, N. 1998. “Documentary and Descriptive Linguistics”. In *Linguistics* 36 (1998), pp 161-95

Nathan, D. 2003. “Endangered languages documentation: from standardisation to mobilisation”. Paper presented at Digital Resources for the Humanities, Cheltenham, England, September 2003.

Nathan, D. 2004b. “ELAR—Reconceptualising the Digital Archive as Ark-hive”. Paper presented at the Annual Conference of the Australian Linguistic Society, University of Sydney, July 2004

Pawley, A. and Syder, F. 1983. Two puzzles for linguistic theory: native-like selection and native-like fluency. In J. Richards and R. Schmidt (eds) *Language and Communication*. London: Longman, pp. 91-226

Rood, D. and Mirzayan A. 2004. “Are new digital recordings always helpful in the preservation of Severely Endangered Languages? The Case of Wichita” Paper presented at *A World Of Many Voices: Interfaces in Language Documentation*, University of Frankfurt, 4th - 5th of September, 2004

Warschauer, M. 1998. “Technology and indigenous language revitalization: Analyzing the experience of Hawai'i”. *Canadian Modern Language Review*, 55(1), 140-161.