Abstract

In this paper we present the Daly Languages Project (www.dalylanguages.org), funded by the ARC Centre of Excellence for the Dynamics of Language, and in collaboration with the Pacific and Regional Archive for Digital Sources in Endangered Cultures (PARADISEC), which has developed website landing pages for all of the languages of the Daly region of northern Australia. These landing pages provide a useful and usable interface by which a range of users can access primary recordings, fieldnotes, and other resources about the Daly languages; they are powered by a relational database which allows for easy updating, ensuring consistency across the website and allowing for an immediate response to community requests. Moreover, since the website is built with a commitment to open source, it is available for other researchers to adapt to their own projects and language groups. In this paper we discuss the goals and outcomes of the project, the design and functionality of the website landing pages, and advise readers on how they can access and adapt the open-source framework for their own purposes.

Keywords: Daly languages, language repatriation, landing pages, website, heritage recordings
Introduction

The development of sustainable, accessible corpora of small languages has been an increasing focus of language documentation in recent years (see Henke & Berez-Kroeker 2016 for a summary of developments). Central to this has been the discussion of best practice in making good language records, and in archiving them appropriately to ensure their longevity and availability for future generations (e.g. Johnson 2004; Thieberger & Berez 2012). One important goal of such archives is to provide access to these language records, most importantly, for language communities and their descendants, and for other researchers. However, the ways in which such accessibility is best achieved are not yet well understood and have only recently become a focus of language documentation research (see, for example, Holton 2014).

In this paper we present the Daly Languages Project (www.dalylanguages.org), funded by the ARC Centre of Excellence for the Dynamics of Language, and in collaboration with the Pacific and Regional Archive for Digital Sources in Endangered Cultures (PARADISEC [www.paradisec.org.au]; see also Thieberger 2014; Thieberger, Barwick & Harris 2015), which has developed website landing pages for all of the languages of the Daly region of northern Australia. These landing pages provide a clear and easy-to-use interface for the extensive corpora of Ian Green, including recordings, fieldnotes, and manuscripts of 11 languages across the region, most of which are no longer spoken fluently. The websites are powered by a relational database which allows for easy updating, ensuring consistency across the website and allowing for an immediate response to community requests about the information contained on the various pages.

We believe the Daly Languages Project provides a useful interface model for other researchers, showing how corpora can be packaged and presented in a way that is usable and accessible for both community members and other researchers. In this paper we discuss the goals and outcomes of the project, the design and functionality of the website landing pages, and advise readers on how they can access and adapt the open-source framework for their own purposes.

Daly Languages

For this project, we use the term ‘Daly Languages’ to cover a group of 22 language varieties of the Daly region of the Northern Territory of Australia, southwest of Darwin. These languages (excluding Murrinhpatha) were first discussed as a group and proposed as a single language family in Tryon’s (1974) volume, which provides brief sketches of each language and, for many, still represents the key published resource on the language. However, Tryon’s evidence

---

1 We would like to thank Nick Reid, Stefan Schnell, and Nick Thieberger for useful discussion of the issues involved in designing and building the Daly Languages website, and lots of invaluable advice. We also thank Brighde Collins, Chris Venning, Kate Charlwood, and Geneva Goldenberg for research assistance support in digitising materials and entering content into the database, and audiences of the Australian Linguistics Society 2016 conference and corpus workshops run by the ARC Centre of Excellence for the Dynamics of Language, as well as two anonymous reviewers for useful feedback and suggestions. We are grateful to PARADISEC for support in digitising and archiving the Daly language materials, and the ARC Centre of Excellence for the Dynamics of Language (Project ID: CE140100041) for funding the project and the 2016 repatriation trip.
for considering these languages to be a single family appears to be largely typological (Tryon 1974: 304) and subsequent work, most notably by Ian Green (e.g. 2003) and Mark Harvey (e.g. 2003a; 2003b), has revised this picture. The current understanding of the relationships between the languages of the Daly region (see Nordlinger 2017) is that the 22 language varieties can be grouped into 10 different languages belonging to five language families. It is not clear whether there are any superordinate relationships among these five families, apart from proto-Australian, and indeed Evans (2003) suggests that (some of) these Daly families may be early offshoots of the proto language itself. Harvey (2003a) suggests a possible remote connection between the Eastern Daly group and the Gunwinyguan languages of Arnhem Land to the east, but this has yet to be definitively established.

The languages of the Daly region (except perhaps for Murrinhpatha) are greatly under-represented in the Australianist literature, and, beyond a brief chapter in Tryon (1974), many have had only limited description. One of the goals of the Daly Languages Project is to collect all available materials on these languages and make them accessible wherever possible to help redress this dearth of available resources. Nordlinger (2017) provides references to all available linguistic materials on the Daly languages, many of which are also available on the Daly Languages website that is the focus of this paper. Unfortunately, Murrinhpatha is now the only language of the region that is still being acquired by children: it is still being used by around 2500 people for daily communication in the community of Wadeye (Port Keats) and is acquired by all Wadeye children as their first language. All other languages of the region are highly endangered; many have only a handful of elderly speakers left and some (for example, Marramaninjsji, Yunggurr) have no fluent speakers left at all.

**Ian Green’s materials**

During the period of 1980 through to 1996, the second author worked in the Daly region and amassed over 157 hours of recordings from 11 of the Daly languages. These recordings were accompanied by 23 books of handwritten fieldnotes and transcriptions. Many of the people Green worked with and recorded were among the last fluent speakers of their languages – languages now no longer spoken regularly in their communities – and all have long since passed away. Apart from Marrithiyel, which was the focus of Green’s PhD thesis (Green 1989), most of these languages have been the subject of very little description, and available materials on the languages range from limited to non-existent. It was therefore very difficult for community members to access any information about their heritage languages, making Green’s corpus a precious resource for communities. At the inception of the Daly Languages Project, the large majority of these materials were in analogue form, unarchived and unpublished, although well-maintained and safely stored in Green’s home office. The original Green corpus is made up of the recordings and resources listed below.

**Recordings**

There is a total of 157 hours of recordings for 11 languages, as listed below. All recordings have now been digitised and most are available on open access in PARADISEC (http://www.paradisec.org.au).
• Marrithiyel, 32 hours, including texts and elicitation (grammatical, phonological, lexical)
• Marri Tjevin, 6 hours, grammatical elicitation
• Magati Ke, 3 hours, grammatical elicitation
• Marri Ngarr, 36 hours, elicitation (grammatical, phonological, lexical)
• Marramaninjsji, 6 hours, elicitation (grammatical, phonological, lexical), some texts
• Batjamalh, 2 hours, elicitation
• Merranunggu, 10 hours, including texts and elicitation (grammatical, phonological, lexical)
• Menthe, 7 hours, elicitation
• Emmi, 2 hours, elicitation
• MalakMalak, 36 hours, elicitation (grammatical, phonological, lexical)
• Matngele, 17 hours, elicitation (grammatical, phonological, lexical)

Fieldnotes, transcriptions, and analyses
Accompanying these recordings are 23 books of handwritten fieldnotes, transcriptions, and grammatical analyses. Each book has been scanned and added to the PARADISEC archive as an individual item. Where these notebooks contain transcriptions of field recordings, these pages have additionally been added to the same item as the recording in PARADISEC, so that the recording and the associated pages of the field notebook can be accessed together. Unpublished manuscripts laying out linguistic analyses, wordlists, and historical analyses have also been scanned or uploaded as required and are archived in PARADISEC.

Goals of the Daly Languages Project
The overall goals of the Daly Languages Project are (i) to protect the Daly language resources and ensure their longevity through digitisation and archiving in PARADISEC and (ii) to ensure their accessibility to community members and other researchers by providing a simple and functional online interface to these archived resources. More specifically, we wanted a digital portal that would be clear and simple for all users, both academic and non-academic; that would provide the content relevant to different user groups, including community users and researchers; and that would be easy to maintain and update without requiring the assistance of a programmer, so that we would minimise the risk of it becoming a graveyard website over time. A further goal was to provide a site where all the available information about each language could be brought together in one place – recordings, fieldnotes, published sources, language learning materials, cultural information, photos, and more. So, while the initial impetus was to protect and make accessible Ian Green’s corpus, we have expanded the scope of the project to include any available materials on the Daly languages (subject to accessibility and copyright). Currently this amounts to another 60 resources, but we will continue to add more as they become available. Finally, in order to ensure maximal accessibility, it was important that the portal have functionality across multiple digital platforms (computer, mobile phone, tablet)
and also in offline mode, given that most community members live in remote locations where internet connections are not reliable or readily available.

A second aspect of the project was to repatriate Green’s recordings to the families and communities of the people that he worked with in the 1980s and 1990s and, in doing so, to inform the communities of the digital portal where they could access materials about their languages, discuss access conditions and permissions, and get input on the functionality of the portal and suggestions for improvement.

We developed the Daly Languages website (http://www.dalylanguages.org/) to achieve these goals. It consists of landing pages for all of the languages of the Daly region of Australia and links to digitised copies of Green’s recordings, his fieldnotes, dictionary materials, unpublished manuscripts containing grammatical description and linguistic analyses of most of the languages, as well as historical reconstructions. Our purpose in this paper is to discuss the design and functionality of this website and the associated database in case it may be of use for other researchers to adapt for other language areas. Firstly we discuss the design of the website landing pages, and then the administration of the website and its construction. In addition, Green and Nordlinger undertook a repatriation trip through the Daly region in July 2016. This trip and its outcomes are discussed in the final section.

The Daly Languages website

The primary functionality of the Daly Languages website (Figure 1) is to provide information about each of the languages of the Daly region, links to their primary and secondary resources (recordings, fieldnotes, etc), and a unified means of navigating the website as a whole. In this section we describe both the functionality and content of our website, including website navigation and the language pages.

---

2 We have not yet created a fully functional offline mode for the website, apart from the fact that all the resources are downloadable, but hope to develop this in the future.
Website navigation

Navigation toolbar

Every page on our website has the same navigation mechanism. This is a toolbar located at the top of each web page. The consistency of this navigation toolbar across all pages is intended to simplify engagement with the website and make it easier for the user to navigate around it.

![Figure 2. The navigation toolbar as it appears on a desktop](image)

This toolbar allows a user to access all areas of the public website. Note that the presentation of the toolbar (but not its functionality) changes depending on the device used to access the website. When viewed on a smaller device, the navigation toolbar is reconfigured to that shown in Figure 3:

![Figure 3. The navigation toolbar as it appears on a mobile phone](image)

Map

The typical means by which language pages are accessed is via the language map. This map is reached either from the ‘Map’ tab in the navigation toolbar, or by clicking on the ‘Languages’ button on the entry page of the website (see Figure 1). The map is a customised Google map that displays each language and its general location.

A nice feature is that this map is dynamically resized depending upon which device a visitor uses to access the website. For smaller devices, the map of Australia is moved above the main map to give the user more screen space to access the region containing the language names, as shown in Figure 5.
Language pages

The language pages contain most of the website’s content. For each language we provide basic information about its relationships with other languages, alternative names and spellings, information about the language’s location and speakers, and pictures of country and people, where possible. The layout of each language page is consistent, to ensure ease of navigation: the basic layout is shown in Figure 6. Below we discuss the various features in more detail.
Although the layout of the language pages appears clean, it hides a deceptively large amount of information. This is by design, as many of the language details will not be relevant to most users – as such they don’t appear unless clicked, reducing visual clutter. This allows us to focus on the most important information for most users – namely the name of the language, what material is available and who is recorded.

Figure 6 is the view a visitor would see when accessing the language page from a desktop. For smaller devices, the information is moved to be more suitable to the vertical scrolling they utilise, as shown in Figure 7.
Speaker information

Through our community consultations on the repatriation trip it was clear that many users are particularly interested in accessing recordings by speaker, rather than by language. In other words, they may be interested in finding all of the recordings involving a particular family member, rather than all of the recordings in a particular language. Thus, every language page lists the speakers that feature in the available recordings of that language. If a speaker’s name is clicked, the visitor is directed to that speaker’s web page, which includes links to their audio recordings, any other resources they worked on, and a list of other community members with whom they were also recorded, as shown in Figure 8.

![Figure 8. Example language consultant page](image)

Alternative names and various spellings

Many Australian languages are known by different names with various spellings. This causes difficulties in web searches, since available resources for a language may not be identified if the user searched with an alternative name or spelling. We therefore attempted to include all alternative names and spellings that we were aware of for each language variety so that the information in the website is more likely to be found on web searches.

This naming information is split into two groups, alternative names and various spellings. The ‘alternative names’ section is used to display alternative language names that bear no similarity to the name we have taken as standard. Some of these alternative names are ones that have contemporary currency in the Daly region – for example, Matngale is also known as Dakayu or Werat – and others are alternative names from the historical literature. These alternative names are always visible as we can’t assume that a visitor to the website will know a language by the name we use.

The ‘various spellings’ link lists all the known spellings of the language names. By default this information is hidden from a visitor as it is not usually needed. Should a visitor want to view this information, they must expand the associated list by clicking the ‘various spellings’ link:
Recording these various spellings and grouping them together allows search engines to associate these different spellings with the language web page. For example, a Google search on 'Madngela language' returns as a search result a link to the Matngele language page shown in Figure 6 above, despite the fact that Madngela spelling is not used as the standard spelling for the language name on that page. This functionality is important as it greatly increases the chances of someone discovering the website and its resources, even if they search on an alternative language name or spelling.

Dialectal information

Language varieties that are considered (by linguists) to be in a dialectal relationship with each other are linked in the ‘dialects with’ section. No one of these language varieties is given a privileged status in this respect – they are treated identically, with each being listed as a dialect with the other – which is important so that the website provides a neutral perspective equally acceptable to all heritage communities. For example, on the Matngele page, Matngele is listed as being a dialect with Yunggurr, and on the Yunggurr page Yunggurr is listed as being a dialect with Matngele.

Language group information

Our current understanding of the diachronic relationships between the various Daly language varieties is reflected in the ‘Group’ name on the language page, which can be expanded to show the details of the relevant group. This information is also accessible via the ‘Language Groups’ tab in the navigation bar at the top of the page.

![Western Daly Grouping](image)
Each language is associated with one of five language groups, based on our current understanding of how these languages are related to each other (Nordlinger 2017). As further historical comparative work is undertaken, this information may be revised. Distinct but related languages are listed horizontally (for example, Marramaninjsji, Merranunggu, Marri Ngarr above). Each vertical stack of language varieties represents those that we consider to be related as dialects (for example, Marri Ngarr and Magati Ke), although it is important to remember that this is a linguistic classification rather than a community-oriented one. Each language name is an active link, and so the user can easily move between the pages of related languages as they wish.

Information links
Overlaying the photo of country, we have five hyperlinks which link to major online sites which provide additional information and material about the languages: OLAC, Ethnologue, Glottolog, Austlang, and Wikipedia. Should one of these resources have no information about a particular language, the corresponding hyperlink does not appear.

Resource links
We have categorised the resources associated with a language into four groups:

- Recordings and fieldnotes
- Key published resources
- Learning resources
- Other resources

These titles are fairly self-explanatory, but these categories were created to best suit the different users of the website. Community members are generally more interested in hearing recordings and accessing learning resources (such as posters, learner’s guides, storybooks). Academics and other researchers may additionally want to access key published resources which reference the language.

An important aspect of the project design is that all primary resources are appropriately archived, and that nothing is stored only on the website. To ensure this, we link directly from the website to the archived item wherever possible. Thus, the items under the ‘Recordings and Field Notes’ button link directly to the associated items in PARADISEC (see Figure 11). The disadvantage of this is that it requires new users to undertake a two-step process to access the recordings: they must first sign up (or sign in) to a PARADISEC account before they can access the Daly language recordings, and this can make the access more complex for users unfamiliar with the PARADISEC interface. However, it has the advantage of ensuring that the recordings have a level of protection from spurious use that they wouldn’t have if they were immediately accessible by an internet search.
The 'Key published resources' includes 5–6 published linguistic resources for each language. We have not attempted to include every single publication here, but simply to identify the key published resources, and to make them available online wherever possible. For some languages, there are very few published resources, and so the information provided varies significantly for each language.

The full set of resources listed on the whole Daly Languages website can be accessed via the Resources tab in the navigation bar at the top of each page – along with the type of resource (pdf, recording, etc) and the language(s) to which it relates. Where the resource is an audio file, we also note the language consultant. This means that the recordings can be found by community members using a search engine. For example, when searching for Worumbu, a speaker of Marri Tjevin, the search results include links to every audio recording featuring Worumbu on our Daly Language website (although the user still needs a PARADISEC account before being able to access them, as discussed above). This functionality is very appealing to non-linguist users who may not know to search particular language archives (such as PARADISEC or AUSTLANG, for example) for resources, but might just use a general search engine such as Google. Thus, this significantly increases the visibility and accessibility of these Daly language materials for community stakeholders.

**Website administration**

The website was designed from the outset to be easy to administer and update so that it could be maintained easily without requiring programming or specialised web design skills. In the next section we discuss the design of the website and the nature of the relational database that drives it. In this section we briefly describe the administration side of the website.

The administration side of the website allows content to be added, edited, and deleted. The two main administration pages control information regarding languages and resources. The administration pages are only available to users who have a username and password to access this part of the website. We decided not to implement individual username/passwords, instead
using the functionality offered by web hosts to create password protected folders on a website.\(^3\) This simplified the design of the website, and still offers a practical and straightforward means to control access to the administration pages of the website.

On logging into the administration web pages, the user sees the following menu:

<table>
<thead>
<tr>
<th>Edit Resources...</th>
</tr>
</thead>
<tbody>
<tr>
<td>Language Families...</td>
</tr>
<tr>
<td>Languages...</td>
</tr>
<tr>
<td>People...</td>
</tr>
<tr>
<td>Clans...</td>
</tr>
<tr>
<td>Language Dialect Groupings</td>
</tr>
<tr>
<td>Other...</td>
</tr>
</tbody>
</table>

Figure 12. The administration control panel

Each of these links has the same functionality, allowing the administrator to add, edit, or delete information. For example, clicking on the ‘Languages...’ tab reveals:

<table>
<thead>
<tr>
<th>Edit</th>
<th>Delete</th>
<th>Batjamalh (Batjamalh)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Edit</td>
<td>Delete</td>
<td>Emmi (Emmi)</td>
</tr>
<tr>
<td>Edit</td>
<td>Delete</td>
<td>Kamu (Kamu)</td>
</tr>
<tr>
<td>Edit</td>
<td>Delete</td>
<td>Magati Ke (Magati Ke)</td>
</tr>
<tr>
<td>Edit</td>
<td>Delete</td>
<td>MalakMalak (MalakMalak)</td>
</tr>
<tr>
<td>Edit</td>
<td>Delete</td>
<td>Marramaninjji (Marramaninjji)</td>
</tr>
<tr>
<td>Edit</td>
<td>Delete</td>
<td>Marri Ammu (Marri Ammu)</td>
</tr>
<tr>
<td>Edit</td>
<td>Delete</td>
<td>Marri Dan (Marri Dan)</td>
</tr>
<tr>
<td>Edit</td>
<td>Delete</td>
<td>Marri Ngarr (Marri Ngarr)</td>
</tr>
</tbody>
</table>

Figure 13. Administration of languages

Languages

Each language area in the database contains the following information:

- Linguistic name (the most commonly used name for the language)
- Traditional name (if different to linguistic name)
- Language family

\(^3\) For a typical example of how this is accomplished see [https://www.a2hosting.com/kb/cpanel/cpanel-security-features/password-protect-directories](https://www.a2hosting.com/kb/cpanel/cpanel-security-features/password-protect-directories).
• URLs to information in OLAC, Ethnologue, Glottolog, Austlang, Wikipedia
• Alternative language names
• Alternative spellings
• Related dialects
• Photo of country URL and caption
• Public notes
• Private notes
• Geographical location (the latitude and longitude for where the language is spoken, for positioning on the map)

Administrators may enter this information if adding a new language to the website or edit the information for existing languages. We have discussed most of these categories in the section above. The public and private notes are textboxes in which the administrator can write whatever relevant material they require. The public notes appear on the web page associated with the language and can be marked up in HTML. The general information about Matngele visible in Figures 6 and 7 above, for example, comes from the public notes. The private notes never appear on the public website and may be used by administrators to record workflow notes or other housekeeping information that may be useful. The geographical location is used to locate the language on the map page.

Resources

We have developed a relatively extensive web page for managing resources, since this is the major purpose of the website. Beyond basic attributes such as the resource’s creation date, citation, and URL, we also categorise resources by type (fieldnotes, recordings, published works, and other) and how they relate to the different languages. A resource can relate to various languages in many different ways. For example, an article could be a key published resource for one or two languages, containing transcripts etc, whereas it might only tangentially mention another language (in which case it would appear in the ‘other resources’ section for the web page for that language). The resource page allows an administrator to specify how a resource relates to a language (Figure 14, left), and also to give a more specific citation for that language as well (such as a specific chapter). As the documentation of a resource can take some time, we have included an option which allows the administrator to create a resource in the database, but not to include its information in the public website. This feature is useful when waiting on a copyright holder’s permission to use a resource online.

The creator (author, etc) and language consultants are associated with the resource, drawing upon a list of people’s names. This ensures that all references to people are consistent in spelling (Figure 14, right).
Construction of the website and relational database

Although people commonly consider a website’s usefulness in terms of the content it contains, the design and construction of a website contribute significantly to its utility. A well-designed website will help users locate the content they want, when they want to find it. Furthermore, over the longer term, the utility of a website comes not just from the information it contains, but from the fact that this information is updated regularly so as to be current. In constructing the Daly Languages website, we had four major design goals, namely that the website be: maintainable, extendable, optimised for the user, and reusable.

Maintainability

For a website to be useful to a community it needs to be maintained. Users of websites typically consider maintenance to be principally concerned with the addition of new content – however, the maintenance of websites revolves around both content and infrastructure.

Maintaining website content

Beyond the addition of new material, content also needs to be updated and occasionally deleted. For static websites, it is the updating and/or deletion of material that is particularly time consuming, especially on larger websites. Consider what happens if a resource (say a book) is deleted from a static website. This requires that the book be removed not only from the list of resources, but also from every language page that may reference that book. Likewise, changing the spelling of an author’s name (or even one of the languages) requires every reference to that name in the website to be updated. In reality, in a large static website, mistakes are made and slowly accrue – authors’ names may have different spellings on different pages, or a web page may reference a resource that has been deleted. These types of errors are not harmless, as searches on the data contained within the website may return inaccurate information.

Database-driven websites offer a solution to this problem at the cost of more initial programming in the creation of the website where administration web pages are built to allow the website to be easily updated. Thus, we decided to create a relational database to store our content, and a specialised content management system to drive our website that would allow administrators to easily add and delete information and to maintain consistency across the various pages. The basic workflow is shown in Figure 15.
For example, as part of the administration web pages, an administrator can view all the languages entered into the database, and choose whether to add a new language, edit one, or delete one, as we saw above. If choosing to add a new language, they see a new web page, allowing them to add the information regarding the new language:

![Figure 16. Adding a new language to the database](image)

Once this information has been added, the new language is immediately added to the map and a new language page is created, as shown in Figures 17 and 18. Note that the new language page even includes a new family tree for the Northern Daly family (since we selected this as the Language Family when we entered information about the new language in Figure 16).
The Daly languages are spoken across the Daly River region of the Northern Territory of Australia, south-west of Darwin. The language varieties and their approximate locations are provided on the map below – clicking on the language name will take you directly to the associated language page.

Figure 17. Map with new language added

A New Language

**Northern Daly Grouping**

Northern Daly

A New Language: MeleKarak

Tetnity

This is the new language we just found out about.

Figure 18. New language page generated
Additional information about the language and the resources can then be added by the administrator as required, via the administrative tools discussed in the section on website administration above.

**Maintaining website infrastructure**

The foundations of websites also need maintenance: software packages need to be updated and have security updates installed and underlying software needs to be made compatible with new standards. The key to making this aspect of website design as simple as possible is to build a website upon widely used and supported software. We discuss this further below, but the core technologies we use (PHP and MySQL) are de facto standards – so much so that they are usually preinstalled, made available, and maintained by the website host company. This means that the most important updates (such as security updates) are usually installed automatically.

**Extendability**

There are three different ways that a website can be extended, by adding (i) more content, (ii) new functionality, or (iii) new types of content. The administrative functionality of our website allows anyone with the appropriate credentials to easily add, modify, and delete content, as described and demonstrated in the discussions above.

**New functionality**

We have designed the website with the view to making it extensible. This is accomplished, as much as possible, by separating content and presentation, and by not duplicating functionality. For example, the key way visitors to the website navigate to the different languages is by using the map. However, a future web-designer might prefer to have the languages also listed in the navigation toolbar. Although the toolbar functionality appears on every web page, it is actually only created once in a file called 'publicheader.php'. Consequently, any new functionality added to this file is automatically included over the entire website. Note, too, in this particular instance no new data is needed to create the functionality as the list of languages is already stored within the database. The creation of new functionality requires a programmer with some basic web skills in PHP, MySQL, and Javascript. However, advanced skills in these languages are not required and it is generally not difficult to find someone with enough knowledge to add this type of functionality.

**New types of content**

Adding new types of content to a database-driven website would be the most complex type of extension. For example, say a web-designer wants to add information about the communities that speak these languages. This requires three additions to the website:

- An extension to the database that stores information about the communities and links them to the language(s) they speak.
- An extension to the administration web pages allowing an administrator to add, edit, and delete information about the communities.
- An extension to the public web pages to display the community information.
The creation of this type of functionality would require a programmer. However, their job is simplified by the availability of a database schema (a formal description of the database and how the information is linked together) and existing administration pages that could be used as a basis for new functionality (see http://dalylanguages.org/feedback.php for details).

Optimising user experience

A primary goal of website design is that visitors to a website should be able to find the information they want easily, regardless of which browser and which device they use to access it. For example, whether a visitor uses Safari, Chrome, Firefox on a desktop, tablet, or mobile, the website needs to display information clearly on their device. We accomplish this by using Bootstrap – an open-source framework that helps web page designers create pages that automatically reposition content so that it appears appropriately on different devices. Another important aspect of the users’ experience with the website is that they should only be presented with information that is relevant to them. Rather than displaying all the possible information at once, we show only the most crucial or interesting information. For example, the languages pages display a photo of country, some basic information about the language, and who we have recorded speaking it. Should visitors want to view additional information (such as other resources, alternative spellings, language groupings, etc), they click a link to do so. This link does not redirect the visitor elsewhere, rather it is used to reveal (or hide) the additional information on the existing page. For example, language groupings are mainly of interest to linguistic researchers and so this information is initially hidden. However, if clicked, this information is revealed:

![Figure 19. Expansion of language grouping information](image-url)
Reusability

As noted above, the primary functionality of our website is to associate different languages to their relevant resources and provide a means of easily navigating amongst them. This means that potentially anyone with similar needs could make use of our code. This is possible because one of the design principles we followed was, as much as possible, separating content and presentation so that information about content (such as a language-specific information) and resources are stored in the database whereas how that information is presented is specified separately. Consequently, if someone wanted to create a similar landing page website for any other language group, they could do so with only minor modifications to our code. Such reuse would require a person with basic programming skills and take a day or two to have up and running.

Technologies used

As mentioned above, our database and website are constructed with open-source software. Open-source software is licensed in such a way that anyone may use it, and anyone may also alter it. Although there was no need for us to alter the software we used, the ability for others to do so is important. This is because widely used open-source projects (such as we used) have a large community of developers working on the software, creating updates where deficiencies have been identified and adding new features from time to time. Ultimately this means the foundations of our project are maintained by a community of altruistically minded developers. Most of the widely used open-source software for the web is supported by website hosts – including the automatic installation of security updates etc. This further simplifies the ongoing maintenance of the project.

As indicated, we made use of the following open-source software:

- MySQL – the database software
- PHP – a scripting language that allows web pages to access and modify data in a database and present it in a web page
- jQuery
- Bootstrap.

jQuery and Bootstrap are programming languages which run inside the browser. jQuery is a specially written programming language that is designed to work on many different browsers (such as Chrome, Firefox, Safari, etc), as well as older versions of these same browsers. Bootstrap (which itself is built on jQuery) is a software framework that allows web pages to dynamically resize themselves depending upon what device they are being viewed on. This means that a page will appear one way when viewed on a desktop, another on a tablet and yet another when viewed on a mobile, as we have demonstrated for our website in various places above. Utilisation of such sector-standard, open-source, regularly updated software, with an operational capacity that well exceeds the projected size of the site database, makes for page load speeds that easily meet industry benchmarks.

Readers interested in using our framework for their own projects can find all the relevant information about licensing, source files and contacts at http://dalylanguages.org/feedback.php.
Repatriation and community consultation

While it is important to archive primary data to ensure its longevity and future accessibility, of equal importance is how we ensure that target communities are aware of the archived material and its ability to be accessed. In recent years the field of documentary linguistics has moved towards participatory models for linguistic archiving which “break traditional boundaries between depositors, users, and archivists to expand the audiences and uses for archives while involving speaker communities directly in language documentation and archival processes” (Henke & Berez-Kroeker 2016: 428; see also Holton 2014 and many other papers in Nathan & Austin 2014). A key objective in the Daly Languages Project is the repatriation of language recordings back to community, so that people have ownership of their language and their family’s recordings, and so that these can be used in community-based language maintenance and revitalisation efforts. We also wanted to consult with these communities about the website and the accessibility of materials to ensure that our design was maximally usable. In July 2016, Ian Green and Rachel Nordlinger (accompanied also by fellow Daly linguist Nick Reid) travelled through the Daly region of the Northern Territory, visiting communities and returning language recordings. On this trip, we distributed to family members and community organisations more than 40 USBs containing language recordings Green made in the 1980s and 1990s. The USBs were also printed with the Daly Languages website address. A story about the trip was published by Melbourne University’s Pursuit magazine (https://pursuit.unimelb.edu.au/articles/preserving-precious-indigenous-languages), and photos are available on the Daly Languages website under the Photos tab.

Figure 20. Project USB worn by a family member (photo: Rachel Nordlinger)
A secondary purpose of the trip was to discuss the Daly Languages website with community stakeholders; confirm permissions and access conditions for the recordings to be archived in PARADISEC; and seek stakeholder input on the design and functionality of the website. This ended up being much more difficult to do than we initially anticipated, given the limited access to computers and internet in many of the remote communities we visited, although those who were able to access the website were generally very positive and found it generally easy to navigate. A number of the website features discussed above arose through this consultation process, since community members were primarily interested in hearing recordings of their own family members, seeing photos of country and people, and, given that these languages are no longer spoken, accessing learning materials that might help them relearn some of their heritage language. In order to respond to these user needs, we added language consultant names associated with each relevant resource to our database, so that users could search for a family member’s name to find all associated resources, as discussed above. We also categorised our resource lists and included a ‘Learning resources’ category for each language, so that learners could more easily identify the relevant materials for their needs from among the list of resources for each language. The fact that the website was easily navigable on mobile phones was crucial to many of these community stakeholders since access to computers is so limited in remote communities.

We also discussed access conditions for the recordings archived in PARADISEC with family members. All families, apart from one, were happy with recordings being on open access in PARADISEC to ensure that the recordings would be accessible to their descendants into the future. One family wanted open access for family members only, a condition that is very difficult for us to implement long term since we can’t know who the family members will be in the future. We are still in discussions with this family to negotiate a practical arrangement.
that they will be comfortable with, so for the time being these recordings are not accessible from the Daly Languages website.

On the other hand, community response to the repatriation of recordings was emotional and overwhelming, reinforcing the importance of researchers ensuring they return recordings to the communities they work with. In general, people were quite overcome to hear the recordings of their past family members; to hear languages that they had often not heard since childhood; and to obtain their own copies on the USBs. They were also astounded and appreciative of our efforts to find them in what were often very remote locations and return the recordings in person. For those of us engaged in language documentation, archiving our recordings is now an important step in our workflow. However, we cannot simply assume that our responsibility to return the recordings to the communities we work with ends there. Community stakeholders are often not aware of linguistic archives and how to access them, and they may not even be aware of the possibility that there are recordings of their languages available online. We therefore need to build repatriation trips into our language documentation workflow as well, if we are to complete our responsibility as researchers to the communities who share their languages and cultures with us.

Conclusion

The Daly Languages Project aims to provide a useful and usable interface by which a range of users can access primary recordings, fieldnotes, and other resources about the Daly languages of the Northern Territory. Moreover, since the website is built with a commitment to open source, it is available for other researchers to adapt to their own projects and language groups. We have attempted to create a resource that will be longlasting and adaptive to the needs of different user groups (community members, language learners, researchers), although we are of course limited by the fact that we can’t know what the demands for such resources will be in the future. The database-driven website provides a user-friendly interface to archival records and is easy to maintain and update in response to community requests. Details more likely to be of interest only to other researchers (such as language family relationships, typological features of the languages, academic linguistic resources) are included in clearly separate parts of the website so that each group of users can easily access exactly the information they are after without having to search through other information of less relevance. Combined with the repatriation trip to return heritage recordings to language communities throughout the Daly region, we believe the Daly Languages website provides a useful model for the mediation of archival language materials (Holton 2014) so that they may be readily accessible and usable by speaker communities as well as by other researchers.

References


