

# 108 年度原住民族語言研究發展中心 (案號：107129)

## 附件九

### 108 年數位族語資料庫資料建置報告

(網路版)

研究員：Akiw 徐中文

原住民族語言研究發展中心  
國立臺灣大學

本文件屬原住民族語言研究發展中心版權所有

## 目錄

壹、前言	1
貳、數位族語資料庫概述	3
一、「數位族語資料庫」的概念	3
二、國內外語言資料庫與語言典藏相關網站	4
參、數位族語資料庫建置內容說明	17
一、數位族語資料庫建置目的	17
二、計畫執行方式	17
(一)前置工作	18
(二)辦理研商會議	19
(三)訪談及其他活動	20
(四)系統維運情形	21
(五)目前頁面狀況	26
三、語料建置工作說明	31
四、數位族語資料庫系統建置內容	32
(一)典藏資料	32
(二)族語新聞	33
五、口語語料及影音修復	34
肆、數位族語資料庫階段性任務分析	37
一、階段性建置任務及目標分析	37
二、資料庫建置與未來規劃	38
伍、結論與建議	41
陸、參考文獻及網站	43
附錄一、計畫相關會議簽到統計及會議記錄	48
附錄二、諮詢訪談及重要會議相關	49
附錄三、2019年建置之內容	50
附錄四、錄音帶修復建議	51

## 「108 年數位族語資料庫資料建置報告」

Akiw 徐中文 研究員

### 壹、前言

聯合國教科文組織（UNESCO）2009 年公布了「世界瀕危語言地圖」，其中臺灣有 5 種語言被列入極度瀕危，10 種語言被列入滅絕等級。而原民會「108 年度原住民族瀕危語言復振補助計畫」，則將卑南語、撒奇萊雅語、噶瑪蘭語、邵語、拉阿魯哇語、卡那卡那富語、賽夏語、茂林魯凱語、萬山魯凱語、多納魯凱語等語言列入，可見族語離「死亡」很近，我們必須嚴正看待。

為了避免族語「死亡」而無「復活」的機會，積極的保存語言是很重要的工作，也是為語言的復振做準備。原住民族語言發展法第 9 條即載明「中央主管機關應建置原住民族語言資料庫，積極保存原住民族語料」，而 David Crystal（2010:349）曾提到「保存語言的方式，是從語言記錄開始，以錄音（聲音與影像）、分析、文字記錄來保存語言，也能夠用於教育」。因此建立一個保存族語資料的平臺，分類收入語料，非常急迫也有其必要性。

有鑑於此，2017 年語發中心開始規劃「數位族語資料庫」的建置方向，規劃初期先查詢國內外族語相關資料庫時，整理出許多可參考的網站架構，像是國外 AILLA、PARADISEC 的網站平臺，都是以提供多媒體語言資料及檔案保存的服務，一般人都可以將所持有的媒體語料提供給該站，檔案就會以數位化的方式被建置，甚至部分網站有分級訪問權限的制度，因此吸引許多研究人員或是族人願意提供資料。而這些資料只要是公開的，也歡迎其他研究者協助分析、轉譯再反饋到資料庫中，讓資料庫收錄的越來越豐富。因此，這也影響了我們規劃族語資料庫的方向。

2018 年語發中心依前一年的規劃，開始建置小型的「數位族語資料庫」及收集語料。而今（2019）年則是保持系統的維運及調整，進行錄音帶語料的數位化修復、建置詮釋資料，並挑選 5 卷語料及共 17 篇族語新聞進行文字轉寫。最後則對於未來因應原語發法應建置的「族語資料庫」提出建議。未來，如何

以語發中心建置的族語資料庫基礎上提升規模，進行資源、計劃及人力的整合，並思考跨單位的合作方式，皆是要持續努力的目標。

語言保存刻不容緩，如何應用現代技術，將已留存的族語文獻資料，特別是將早期的族語資料彙整，並以數位化的方式保存，是首要之務。而將影音的部分加以轉譯及轉寫成文字，以利後續的再應用亦是語言保存中重要的工作。如 David Crystal (2001) 提到語言復振的策略之一，是將瀕危語言書寫成文字，與我們正在努力方向不謀而合。

族語資料庫除了收納之外，也應以原住民族知識體系來分類建置語料，汪明輝 (2009) 以鄒族為例提出知識體系的架構，大致上包含生活空間、時間、神話、歷史、政治司法、宗教信仰、社會、文化及經濟等面向。若能以這樣的知識架構建構語言的樣貌，語言的保存才能更完整。衷心期盼未來原住民族語言資料庫的建置，能夠更全面且走的更長遠，得以為我們的後代留下最珍貴的寶藏。

## 貳、數位族語資料庫概述

### 一、「數位族語資料庫」的概念

數位族語資料庫顧名思義是以數位的方式存取族語資料，是以現階段尚未發展到語料庫。而根據 2017 年及 2018 年語發中心關於數位族語資料庫的規劃及建置報告中，也將「資料庫」和「語料庫」概念上作區別。

簡單而言，資料庫 (DataBase) 就是保存資料的地方。進一步來說，資料庫是由一群相關的資料集合起來並儲存在一起，可藉由資料庫管理系統 (DataBase Management System, DBMS) 提供的語法功能來管理資料庫裡的資料，而透過資料庫系統的建置，使用者就能檢索、搜尋所需資料。

Crystal (1991:117) 則認為語料庫 (corpus) 是對語言資料收集，資料來源可以是書面語或是口語的轉譯文本，它們可以被當做語言描述的基礎或驗證關於語言假設的方法。而陳浩然、潘依婷 (2017:2) 則提到語料庫是指透過大量且根據特定原則進行文本搜集，以電子格式儲存，也就是使用者可以將大量的語言使用情形，進行詞彙、句子的檢索和統計。而語料庫中所收集的語料通常具有真實性、代表性、規模性、結構性、可機讀性 (何婷婷, 2003:4)，即可知語料庫和資料庫的最大差異在於語料庫所使用的語料文本是曾經處理，甚至分析及標記，例如收到音檔，要先轉成文字再進一步分析。資料庫則是收到資料，進行數位化，輸入描述資訊檔編碼即儲存，未有語料標記及分析的工作。而現在也有很多技術，如 OCR 光學字元識別 (OCR, Optical Character Recognition)，可以幫助書面資料數位化後的字型辨識，再由人工確認，不必從頭開始打字，增加很多效率，但辨識度還是要視資料狀態而定 (程婉如, 2008)。

因此，我們認為「數位族語資料庫」就是各種族語資料 (包含書面、手稿、錄音帶、錄影帶、音檔、影音資料等) 經數位化後以電子資料形式儲存的空間 (如下圖所示)，除了典藏族語資料，也讓使用者可以進行檢索，並在合理的規範下，進一步將這些資料做其他教學、研究的用途。

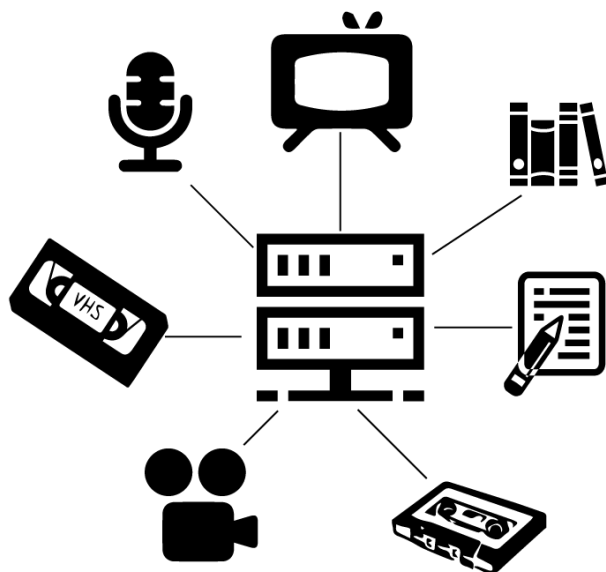


圖 1：資料庫的想像示意圖

當然，基於語料的保存及應用來說，進一步進行語料分析研究的工作也是十分重要的，雖然目前以我們之能還沒辦法做到真正嚴謹、有分析和標記的語料庫形式，但我們挑選口語語料及族語新聞轉譯成族語文字，即是希望先產出文字，以利未來進行族語語料庫的開發。

而在瞭解我們對於數位族語資料庫的定義後，我們接著介紹國內外語言典藏／語料庫等相關網站，皆是數位族語資料庫建置的參考基礎。

## 二、國內外語言資料庫與語言典藏相關網站

為了瞭解國內外相關的語言資料庫或語料庫，我們參考需多網站及系統，於 2017 年規劃、2018 年建置，其中，在版權及相關權益的訂定主要是參考 AILLA 網站<sup>1</sup>，而主要的系統概念則是參考 PARADISEC<sup>2</sup>。而未來則會視任務型態擴增及修改系統。

以下為 2017 年規劃主要參考的資料庫，主要是以表格的方式呈現重要參考網站的歸納與分析，另外有一些像是台灣原住民族數位典藏資料庫或


---

<sup>1</sup> AILLA (The Archive of the Indigenous Languages of Latin America)，拉丁美洲原住民族語言檔案館，網址：<https://ailla.utexas.org/>

<sup>2</sup> PARADISEC (Pacific and Regional Archive for Digital Sources in Endangered Cultures)，太平洋地區檔案數位資源瀕危文化庫，網址：<http://www.paradisec.org.au/>

中研院阿美族、雅美族口語傳說採錄翻譯資料數位典藏計畫網站、VCenter 數位典藏國家型計畫網站、文化部國家記憶庫網站等雖未列入下表，但對於數位族語資料庫的系統規劃及後續資料的存取考量，皆有莫大的助益，也希望未來資料庫亦能將族語相關資料納入其中，其他參考網站亦於文末「參考文獻及網站」羅列。

表 1：國內外語言資料庫及語言典藏相關網站分析

序	名稱（按筆劃）	形式	內容分析
1	AILLA (The Archive of the Indigenous Languages of Latin America)，拉丁美洲原住民族語言檔案館 <a href="https://ailla.utexas.org/">https://ailla.utexas.org/</a>	大型綜合資料庫	<ul style="list-style-type: none"> <li>➢ 支援：由澳洲爾本大學 Rachel Nordlinger 教授及阿德萊德大學 Ian Green 博士合作，並得到 ARC 卓越語言動力中心的資助。</li> <li>➢ 目的：AILLA 的使命是保存拉丁美洲原住民族語言的錄音、文本和其他多媒體的數位語言檔案，並將他們提供給原住民、研究人員和其他有興趣的朋友。</li> <li>➢ 特色與功能： <ol style="list-style-type: none"> <li>1. 網站必須註冊及通過審核。</li> <li>2. 資料有分級限制，保存的語言資料是有分析過的也有原始資料，他們很歡迎使用者在合理的規範下使用，並協助轉譯成文字或進一步的語言分析。</li> <li>3. 主要收藏是錄音、影音視頻，內容包含敘事、詠唱、演講、對話、歌曲。</li> <li>4. 另外也有很多語言文本資料，例如語法的分析、字典或民族誌的現場筆記。</li> </ol> </li> <li>➢ 分析：這個網站的相關權利益務很清楚，資料也有分級，也在乎族人的隱私及知情同意權。智慧財產權制度都可以做為參考。</li> </ul> <p>檢索頁面</p>  <p>The screenshot shows a search page titled 'Languages' with a navigation bar and a grid of folder icons representing different languages: Achi, Achele, Aho Anale, Aguaruna, Awa, Akmer Osham, Algonquin, and Amharic.</p>

序	名稱（按筆劃）	形式	內容分析
2	Daly Languages (AUSTRALIA), 澳洲 Daly 語語料庫 <a href="http://www.dalylanguages.org/">http://www.dalylanguages.org/</a>	小型語料庫	<p>➤ 支援：由澳洲爾本大學 Rachel Nordlinger 教授及阿德萊德大學 Ian Green 博士合作，並得到 ARC 卓越語言動力中心<sup>3</sup>的資助。</p> <p>➤ 目的：針對研究人員和社區提供 Ian Greens 在 1980 年至 1996 年間於澳大利亞北部戴利地區的許多語言的廣泛的實地記錄，記錄和分析。另外還有一個目的是提供 Daly 達利語言的入門資訊以及相關的可用資訊。</p> <p>➤ 特色與功能：</p> <ol style="list-style-type: none"> <li>1. 需有 PARADISEC 帳號。</li> <li>2. 語料庫包含大量的田野筆記，以及字典材料，未出版的手稿，其中包含大部分語言的語法描述和語言分析，以及歷史重建。</li> <li>3. 11 種語言共 157 小時的錄音，部分包含語法、語音、詞彙等文字。所有錄音已經數字化，並在 PARADISEC 中提供。</li> <li>4. 大量的現場筆記本將掃描後增加至 PARADISEC 中。</li> </ol> <p>➤ 分析：網站資訊清楚，是有目的性的將現有的語料數位化，但資料是儲存在外部，即 PARADISEC 網站。</p> <p>存有之資料</p> <ul style="list-style-type: none"> <li>• Marrihiyel, 32 hours, including texts and elicitation (grammatical, phonological, lexical)</li> <li>• Marri Tjevin, 6 hours, grammatical elicitation</li> <li>• Magati Ke, 3 hours, grammatical elicitation</li> <li>• Marri Ngarr, 36 hours, elicitation (grammatical, phonological, lexical)</li> <li>• Marramaninjsji, 6 hours, elicitation (grammatical, phonological, lexical), some texts</li> <li>• Batjamalh, 2 hours, elicitation</li> <li>• Merranunggu, 10 hours, including texts and elicitation (grammatical, phonological, lexical)</li> <li>• Menthe, 7 hours, elicitation</li> <li>• Emmi, 2 hours, elicitation</li> <li>• MalakMalak, 36 hours, elicitation (grammatical, phonological, lexical)</li> <li>• Matngele, 17 hours, elicitation (grammatical, phonological, lexical)</li> </ul> <p>語言介紹</p>

<sup>3</sup> <http://www.dynamicsoflanguage.edu.au/>



序	名稱（按筆劃）	形式	內容分析
			<p><b>Marri Tjevin</b></p> <p>also called: <a href="#">Tanggural, Ngangityaben, Murinjaban, Murindjabin</a></p> <p>dialects with: <a href="#">Marrithiyel, Marri Ammu, Marri Dan</a></p> <p>various spellings: <a href="#">[?]</a></p> <p><b>WESTERN DALY GROUPING</b> <a href="#">[?]</a></p> <p>Marri Tjevin and <a href="#">Marri Ammu</a> are two closely related dialects whose country lies along the coast north of the mouth of the Moyle River, extending up towards the Dashwood Plains. Marri Tjevin occupies the southern part of this area, with <a href="#">Magali Ke</a> and <a href="#">Marri Ngarr</a> to the south. Marri Ammu then has <a href="#">Menthe</a> and <a href="#">Merranunggu</a> to its north and north-east respectively. Their sister dialect <a href="#">Marrithiyel</a> lies to their east. There are a number of active outstations on Marri Tjevin and Marri Ammu country proper, with Marri Tjevin and Marri Ammu communities also resident at Wadeye.</p> <p>While neither Marri Tjevin nor Marri Ammu have been extensively documented, there is general agreement that there are only minor linguistic differences between them, with Tryon (1974) suggesting that their grammars are effectively identical and that they are only distinguished through vocabulary. The community language experts that we have worked with have indicated that there may also be minor phonological differences. Tryon (1968, 1970, 1974) and Kofod (1986a, 1986b), together with sections of Ford (2007), are the only published sources on the dialects, although publications on plant and animal terminology (Northern Territory Government), and a learner's guide (Batchelor College), appear to be under development.</p> <p>It had been thought that no full speakers of either dialect remained. However, two senior Marri Tjevin men, Atu Frank Dumoo (sadly now deceased) and Malbiyin Claver Dumoo, began work in Wadeye with Dr John Mansfield (University of Melbourne) in the development of an extensive set of audio and visual recordings providing detailed documentation of Marri Tjevin language and material culture. The products of this work can be found under the Recordings and Field Notes tab, with additional work being added as it becomes available.</p> <p>Ian Green's corpus of Marri Tjevin consists of 6 hours of recordings of lexical, and grammatical elicitation. The first recording has some textual material in Marri Tjevin (Ambrose Piallum), <a href="#">Marrithiyel</a> (Bill Parry) and <a href="#">Marri Ngarr</a> (Jimmy Nambatu). The other recordings were made with the following speakers:</p> <ul style="list-style-type: none"> <li>• Les Kuntjul</li> <li>• Colin Worumbu</li> <li>• Edward Nimarluk</li> </ul> <p><b>檔案內容，需連結到 PARADISEC 才能下載</b></p>  <p>Recordings and Field Notes</p> <p>These media files are stored externally at <a href="#">paradisec.org.au</a>. To view them, you need to either <a href="#">sign in</a> or <a href="#">create a free account</a>. <a href="#">Hide this message</a></p> <ul style="list-style-type: none"> <li><a href="#">John Mansfield (collector), 2012; Marri Tjevin (JMS), Digital collection managed by PARADISEC. [Other Access] DOI: 10.4225/72/5796DA8836497</a></li> <li><a href="#">G2-001, Marri Tjevin/Marrithiyel, audio (as spoken by: Nambatu, Parry, Piallum)</a></li> <li><a href="#">G2-002, Marri Tjevin, 10/11/93, audio (as spoken by: Worumbu, Kuntjul)</a></li> <li><a href="#">G2-003, Marri Tjevin, 10/11/93, audio (as spoken by: Worumbu, Kuntjul)</a></li> <li><a href="#">G2-004, Marri Tjevin, 10/11/93, audio (as spoken by: Worumbu, Kuntjul)</a></li> <li><a href="#">G2-005, Marri Tjevin, 13/11/93, audio (as spoken by: Worumbu, Nimarluk)</a></li> <li><a href="#">G2-006, Marri Tjevin, 13/11/93, audio (as spoken by: Worumbu, Nimarluk)</a></li> <li><a href="#">John Mansfield (collector), 2012; Marri Tjevin (JMS), Digital collection managed by PARADISEC. [Other Access] DOI: 10.4225/72/5796DA8836497</a></li> </ul> <p>Key Published Resources</p> <p>Learning Resources</p> <p>Other Resources</p>
3	<p>Glottolog，世界語言參考文獻網</p> <p><a href="http://glottolog.org/">http://glottolog.org/</a></p>	<p>大型資料庫</p>	<p>➤ 此為協作性的工作，哈拉爾·哈馬斯特羅姆（HaraldHammarström）收集許多書目變成主書目，以及關於世界各種語言證明的家譜關係的大量訊息。</p> <p>➤ 目的：Glottolog 旨在提供一個語言學家需要能夠識別的 languoids（家庭，語言，方言）的綜合列表。每個 languoid 都有一個獨特而持久的名為 Glottocode 的標識符，由四個字母和四個數字組成[abcd1234]。並組織大型書目導入，是一個很大的書目、語言分類查詢網站。</p> <p>➤ 特色與功能：</p> <ol style="list-style-type: none"> <li>1. 識別語言分類，至 2017 年 11 月 22 日止，有 7384 種口語語言（指第一語言的</li> </ol>

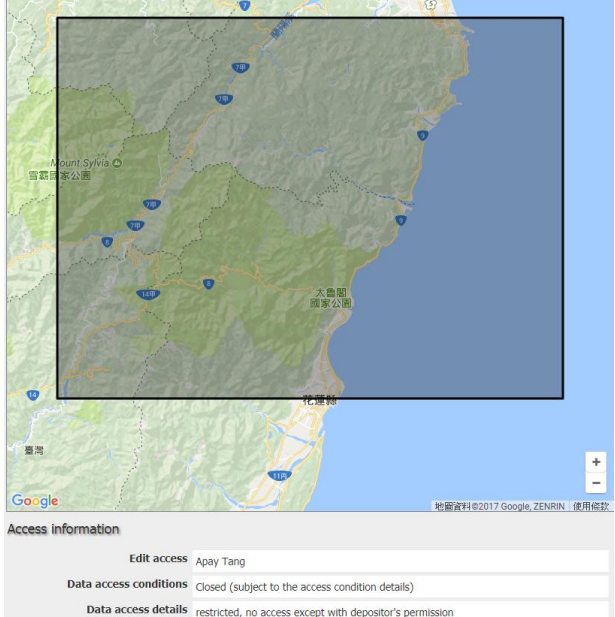
序	名稱 (按筆劃)	形式	內容分析																																																																		
			<p>口語)。</p> <ol style="list-style-type: none"> <li>2. 搜尋結果會以樹狀結構產生，點繫語別下面會出現相關的書目。</li> <li>3. 書目有 citation 功能可直接引用。</li> <li>4. 收集世界語言多樣性遺產。</li> </ol> <p>➤ 分析：這個網站主要是分類語言及語言的關係，再導入相關的書目，並結合地理資訊系統。</p> <p>分類方式</p> <p>Language: Paiwan 📌 ✎</p> <div style="border: 1px solid #ccc; padding: 10px; margin: 10px 0;"> <p><b>Classification</b> 📌</p> <ul style="list-style-type: none"> <li>— ▼ <b>Austronesian (1275)</b> <ul style="list-style-type: none"> <li>▶ Atayalic (2)</li> <li>▶ <b>Bunun</b></li> <li>▶ East Formosan (6)</li> <li>▶ Malayo-Polynesian (1255)</li> <li>▶ Northwest Formosan (2)</li> <li>▶ <b>Paiwan</b> ●</li> <li>▶ Puyuma</li> <li>▶ Rukai</li> <li>▶ Tsouic (3)</li> <li>▶ Western Plains Austronesian (3)</li> </ul> </li> </ul> <p><b>Comments on subclassification</b></p> <p><a href="#">Robert A. Blust and Victoria Chen 2017</a> <a href="#">Robert A. Blust 2009</a></p> </div> <p>書目呈現</p> <table border="1" style="width: 100%; border-collapse: collapse; font-size: small;"> <thead> <tr> <th>Details</th> <th>Name</th> <th>Title</th> <th>Any field</th> <th>CS</th> <th>Year</th> <th>Pages</th> <th>Language</th> <th>CS</th> <th>Provider</th> <th>Alt</th> </tr> </thead> <tbody> <tr> <td><a href="#">Details</a></td> <td>DeGroot, James Maxwell 1993</td> <td>The island of Formosa, past and present: history, people, resources, and commercial prospects. Tai, Chinese, English, Greek, Latin, Malaya, Occidental, Greek, and other projections.</td> <td>✓</td> <td></td> <td>1993</td> <td>800</td> <td>ethnographic, historical, political</td> <td></td> <td>IS</td> <td></td> </tr> <tr> <td><a href="#">Details</a></td> <td>Chang, Hsueh-ling 2002</td> <td>Paiwan Introduction: Paiwan-Deutsch, Deutsch-Paiwan</td> <td>✓</td> <td></td> <td>2002</td> <td>418</td> <td>introduction</td> <td></td> <td>IS, IS-2</td> <td></td> </tr> <tr> <td><a href="#">Details</a></td> <td>Yang, Robert and Oldeman, John 2003</td> <td>One Spelled Paiwan-Isle</td> <td>✓</td> <td></td> <td>2003</td> <td>323</td> <td>lex, grammar, syntax</td> <td></td> <td>IS, IS-2</td> <td></td> </tr> <tr> <td><a href="#">Details</a></td> <td>Harrell, George 1992</td> <td>Paiwan Dictionary</td> <td>✓</td> <td></td> <td>1992</td> <td>300</td> <td>dictionary, grammar, syntax</td> <td></td> <td>political, postwar, IS, IS-2</td> <td></td> </tr> <tr> <td><a href="#">Details</a></td> <td>Chang, Anna-Hsiang-Hsueh 2006</td> <td>A Reference Grammar of Paiwan</td> <td>✓</td> <td></td> <td>2006</td> <td>488</td> <td>grammar</td> <td></td> <td>IS</td> <td></td> </tr> </tbody> </table> <p>地理資訊</p>	Details	Name	Title	Any field	CS	Year	Pages	Language	CS	Provider	Alt	<a href="#">Details</a>	DeGroot, James Maxwell 1993	The island of Formosa, past and present: history, people, resources, and commercial prospects. Tai, Chinese, English, Greek, Latin, Malaya, Occidental, Greek, and other projections.	✓		1993	800	ethnographic, historical, political		IS		<a href="#">Details</a>	Chang, Hsueh-ling 2002	Paiwan Introduction: Paiwan-Deutsch, Deutsch-Paiwan	✓		2002	418	introduction		IS, IS-2		<a href="#">Details</a>	Yang, Robert and Oldeman, John 2003	One Spelled Paiwan-Isle	✓		2003	323	lex, grammar, syntax		IS, IS-2		<a href="#">Details</a>	Harrell, George 1992	Paiwan Dictionary	✓		1992	300	dictionary, grammar, syntax		political, postwar, IS, IS-2		<a href="#">Details</a>	Chang, Anna-Hsiang-Hsueh 2006	A Reference Grammar of Paiwan	✓		2006	488	grammar		IS	
Details	Name	Title	Any field	CS	Year	Pages	Language	CS	Provider	Alt																																																											
<a href="#">Details</a>	DeGroot, James Maxwell 1993	The island of Formosa, past and present: history, people, resources, and commercial prospects. Tai, Chinese, English, Greek, Latin, Malaya, Occidental, Greek, and other projections.	✓		1993	800	ethnographic, historical, political		IS																																																												
<a href="#">Details</a>	Chang, Hsueh-ling 2002	Paiwan Introduction: Paiwan-Deutsch, Deutsch-Paiwan	✓		2002	418	introduction		IS, IS-2																																																												
<a href="#">Details</a>	Yang, Robert and Oldeman, John 2003	One Spelled Paiwan-Isle	✓		2003	323	lex, grammar, syntax		IS, IS-2																																																												
<a href="#">Details</a>	Harrell, George 1992	Paiwan Dictionary	✓		1992	300	dictionary, grammar, syntax		political, postwar, IS, IS-2																																																												
<a href="#">Details</a>	Chang, Anna-Hsiang-Hsueh 2006	A Reference Grammar of Paiwan	✓		2006	488	grammar		IS																																																												


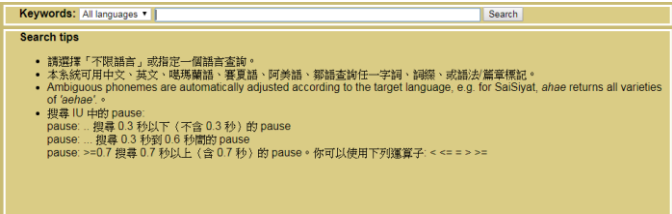

序	名稱（按筆劃）	形式	內容分析
			
4	<p>PARADISEC (Pacific and Regional Archive for Digital Sources in Endangered Cultures)，太平洋地區檔案數位資源瀕危文化庫</p> <p><a href="http://www.paradisec.org.au/">http://www.paradisec.org.au/</a></p>	大型資料庫	<ul style="list-style-type: none"> <li>➤ 支援：由澳洲雪梨大學、墨爾本大學、澳洲國立大學三所大學組成的，接受捐款。</li> <li>➤ 目的：將世界上小語種的語言以數位方式保存。</li> <li>➤ 特色與功能： <ol style="list-style-type: none"> <li>1. 目前超過 1000 種語言。</li> <li>2. 主要專注於數位化錄音磁帶，但另外開放可以存放音檔衍生而成的文本材料（論文、字典、文章等）。</li> <li>3. 需註冊。</li> <li>4. 提供語言與音樂數據管理，錄音及數據連接的培訓。</li> <li>5. 檢索功能可點地圖，或從搜尋欄打關鍵字，或從語言別、國家、前 100 的收藏或是可以自己創建目錄。</li> <li>6. 結合地理資訊系統。</li> <li>7. 標準化檔案，亦可在 OLAC（開放語言檔案社群網）中搜尋資料。</li> <li>8. 與當地機構合作（如所羅門群島博物館）提供備份和數據保存服務。</li> <li>9. 付費磁帶修復及數位化<sup>4</sup>。</li> <li>10. 有各式協助指南<sup>5</sup>（歸檔工作、數據儲存、入門介紹、Metadata 說明等）。</li> </ol> </li> </ul>

<sup>4</sup> <http://www.paradisec.org.au/deposit/costing-a-deposit/>

<sup>5</sup> <http://www.paradisec.org.au/resources/downloads/>

序	名稱 (按筆劃)	形式	內容分析																																					
			<p>➤ 分析：網站資訊清楚，資料庫訊息一目瞭然，結合地理資訊系統，檢索方式多元也很方便。</p> <p>檢索頁面</p>  <p>資料明細</p> <p>Collection details</p> <table border="1" data-bbox="710 884 1412 1265"> <tr> <td>Collection ID</td> <td>AIT1</td> </tr> <tr> <td>Title</td> <td>Recordings of Taroko (Taiwan)</td> </tr> <tr> <td>Description</td> <td>Recordings of narratives in Taroko (Taiwan).</td> </tr> <tr> <td>Archive link</td> <td><a href="http://catalog.paradisec.org.au/repository/AIT1">http://catalog.paradisec.org.au/repository/AIT1</a></td> </tr> <tr> <td>Collector</td> <td>Apay Tang <a href="#">Find similar</a></td> </tr> <tr> <td>Operator</td> <td></td> </tr> <tr> <td>Originating university</td> <td>University of Hawaii at Manoa</td> </tr> <tr> <td>Countries</td> <td>Taiwan - TW <i>To view related information on a country, click its name</i></td> </tr> <tr> <td>Languages</td> <td>Taroko - trv <i>To view related information on a language, click its name</i></td> </tr> <tr> <td>Region / village</td> <td></td> </tr> <tr> <td>Cite as</td> <td>Apay Tang (collector). 1997; <i>Recordings of Taroko (Taiwan)</i> (AIT1). Digital collection managed by PARADISEC. [Closed Access] DOI: 10.4225/72/56E7A74E33F87</td> </tr> </table> <p>所存資料 (圖為未開放)</p> <p>Items in Collection (4)</p> <table border="1" data-bbox="710 1355 1412 1512"> <thead> <tr> <th>Item</th> <th>Title</th> <th>Actions</th> </tr> </thead> <tbody> <tr> <td>001 <span>closed</span></td> <td>traditional Truku stories</td> <td><a href="#">View</a></td> </tr> <tr> <td>002 <span>closed</span></td> <td>daily stories</td> <td><a href="#">View</a></td> </tr> <tr> <td>003 <span>closed</span></td> <td>Daily and traditional stories</td> <td><a href="#">View</a></td> </tr> <tr> <td>004 <span>closed</span></td> <td>Tama visitation</td> <td><a href="#">View</a></td> </tr> </tbody> </table> <p>地理資訊</p>	Collection ID	AIT1	Title	Recordings of Taroko (Taiwan)	Description	Recordings of narratives in Taroko (Taiwan).	Archive link	<a href="http://catalog.paradisec.org.au/repository/AIT1">http://catalog.paradisec.org.au/repository/AIT1</a>	Collector	Apay Tang <a href="#">Find similar</a>	Operator		Originating university	University of Hawaii at Manoa	Countries	Taiwan - TW <i>To view related information on a country, click its name</i>	Languages	Taroko - trv <i>To view related information on a language, click its name</i>	Region / village		Cite as	Apay Tang (collector). 1997; <i>Recordings of Taroko (Taiwan)</i> (AIT1). Digital collection managed by PARADISEC. [Closed Access] DOI: 10.4225/72/56E7A74E33F87	Item	Title	Actions	001 <span>closed</span>	traditional Truku stories	<a href="#">View</a>	002 <span>closed</span>	daily stories	<a href="#">View</a>	003 <span>closed</span>	Daily and traditional stories	<a href="#">View</a>	004 <span>closed</span>	Tama visitation	<a href="#">View</a>
Collection ID	AIT1																																							
Title	Recordings of Taroko (Taiwan)																																							
Description	Recordings of narratives in Taroko (Taiwan).																																							
Archive link	<a href="http://catalog.paradisec.org.au/repository/AIT1">http://catalog.paradisec.org.au/repository/AIT1</a>																																							
Collector	Apay Tang <a href="#">Find similar</a>																																							
Operator																																								
Originating university	University of Hawaii at Manoa																																							
Countries	Taiwan - TW <i>To view related information on a country, click its name</i>																																							
Languages	Taroko - trv <i>To view related information on a language, click its name</i>																																							
Region / village																																								
Cite as	Apay Tang (collector). 1997; <i>Recordings of Taroko (Taiwan)</i> (AIT1). Digital collection managed by PARADISEC. [Closed Access] DOI: 10.4225/72/56E7A74E33F87																																							
Item	Title	Actions																																						
001 <span>closed</span>	traditional Truku stories	<a href="#">View</a>																																						
002 <span>closed</span>	daily stories	<a href="#">View</a>																																						
003 <span>closed</span>	Daily and traditional stories	<a href="#">View</a>																																						
004 <span>closed</span>	Tama visitation	<a href="#">View</a>																																						

序	名稱 (按筆劃)	形式	內容分析
			
5	<p>臺大臺灣南島語多媒體語料庫</p> <p><a href="http://corpus.linguistics.ntu.edu.tw/index_zh.php">http://corpus.linguistics.ntu.edu.tw/index_zh.php</a>。</p>	<p>中型語料庫</p>	<ul style="list-style-type: none"> <li>➤ 支援：由臺灣大學語言學研究所建立，原為國立臺灣大學資訊電子科技整合研究中心「多媒體整合實驗室」子計畫之一(2001-2003)，由臺灣大學語言學研究所黃宣範、蘇以文及宋麗梅教授共同主持。</li> <li>➤ 目的：以臺灣南島語為主軸，藉由資訊科技運用，建構起語言資料典藏及後設資料之標準，有系統彙整、轉寫、保存彌足珍貴的口述語料。</li> <li>➤ 特色與功能： <ol style="list-style-type: none"> <li>1. 語料庫中建置賽夏語、噶瑪蘭語、鄒語及阿美語資料庫。有聲音或影像檔。</li> <li>2. 未來展望，希望為臺灣南島語建置語言資料典藏及後設資料之標準。並以半自動化為目標。</li> <li>3. 不論是語言學者或是一般大眾，只要是對南島文化及語言有興趣者，藉由本語料庫，即可應用最便捷、多元的介面，取得豐富而珍貴的語料。</li> <li>4. 語料庫的基本原則，普及性、口語性、一致性、跨語言性、交叉性、相容性、便利性、多元性。</li> <li>5. 語料文本基本上有 Pear Stories 和 Frog</li> </ol> </li> </ul>

序	名稱 (按筆劃)	形式	內容分析
			<p>Story，故有範圍限制，較能夠做跨語言的比較分析。</p> <p>6. 語料呈現方式為族語句子、註解（英）、註解（華）、完整翻譯（英）、完整翻譯（華），有族語、華語、英語翻譯，讓國際人士都方便查詢。</p> <p>➤ 分析：網站資訊清楚，資料庫訊息一目瞭然。</p> <p>依語別檢索</p>  <p>關鍵字搜尋 (可跨語言)</p>  <p>語料呈現方式</p> 
6	<p>臺灣南島語數位典藏 (目前已停用)</p> <p><a href="http://languagearchives.sinica.edu.tw/cht/index">http://languagearchives.sinica.edu.tw/cht/index</a></p>	<p>中型語料庫</p>	<p>以下內容整理自 Zeitoun et al. (2003)；齊莉莎 (2013)；齊莉莎、余清華(2004, 2007)。</p> <p>➤ 支援：是國科會 NDAP 機構計畫下分項主題「語言典藏」的子計畫之一，整體建置作業之規劃暨主持人為齊莉莎女士，計畫分二期 2002-2006、2007-2011，現已停用。</p>

序	名稱（按筆劃）	形式	內容分析
	<a href="http://php.html">php.html</a>		<ul style="list-style-type: none"> <li>➤ 目的：建立一數位圖書館(digital library)，期能收集、保存、編輯及透過網路散播語言資源，供使用者存取已錄製及記音的南島語。</li> <li>➤ 特色與功能：               <ol style="list-style-type: none"> <li>1. 語料庫中經語料分析的有魯凱語、賽夏語、泰雅語、布農語、排灣語、卑南語。其他語料輸入有布農語、鄒語、巴宰語、卡那卡那富語、西拉雅語、魯凱語、排灣語、阿美語。</li> <li>2. 另外也建置語言學相關書目的資料庫。</li> <li>3. 有語言地理資訊系統了解，臺灣南島語的分布圖，並觀察同源詞與非同源詞的分布情形。</li> <li>4. 有進階檢索功能，也提供詞彙用字索引，讓使用者檢視詞彙的分布。</li> <li>5. 提供原始語料與重新分析語料做對照。</li> <li>6. 有特殊符號代碼，方便查詢時使用。</li> <li>7. 因各族書寫不一致，故使用 IPA 記音，也較不容易混淆。</li> <li>8. Metadata 包含語料的側寫（如：題目、語言及方言、文體），田野調查（如：發音人、田野工作者/謄寫者/編輯者/翻譯者、資料收集/編輯/翻譯的日期）以及管理策略（版權宣告）。</li> <li>9. 開發 AnnoTool（提供分析用縮寫，華語英語縮寫可以翻譯）及 Chkgloos（檢，查不同語言間的詞彙數量）兩種軟體，協助分析者能提供更正確及一致性高的語料。</li> </ol> </li> <li>➤ 分析：網站功能很完整，語料資源豐富，又結合地理資訊系統及建置書目，是確實可以產出學術研究的多功能語料庫。但很可惜現在計畫沒有繼續因此停用中。</li> </ul> <p>檢索頁面</p>



序	名稱 (按筆劃)	形式	內容分析																																																																																																														
			 <p>中央研究院 - 台灣南島語數位典藏 Academia Sinica Formosan Language Archive - Windows Internet Explorer</p> <p>http://formosa.sinica.edu.tw/</p> <p>中央研究院 ACADEMIA SINICA</p> <p>國立臺灣大學國語學系 台灣南島語數位典藏 Academia Sinica Formosan Language Archive</p> <p>語言典藏</p> <p>中文版 ENGLISH</p> <p>You are visitor 011830 since May 5, 2003.</p> <p>Institute of Linguistics All Rights Reserved</p> <p>檢索功能</p> <p>語料庫查詢</p> <p>典藏內容 使用說明 回首頁</p> <p>§ 自訂語料庫搜尋範圍 §</p> <p>設定語言別搜尋範圍：  <input checked="" type="radio"/> 全部  <input type="radio"/> 設定範圍：語言別： <input type="text" value="魯凱語"/> 方言別： <input type="text" value="高山"/></p> <p>設定文體搜尋範圍：  <input checked="" type="radio"/> 全部  <input type="radio"/> 設定範圍：<input type="checkbox"/> 傳說 <input type="checkbox"/> 口述 <input type="checkbox"/> 對話 <input type="checkbox"/> 歌謠</p> <p>設定語料呈現方式：  <input checked="" type="radio"/> 全部  <input type="radio"/> 設定範圍：<input type="checkbox"/> 原始語料 (Original data) <input type="checkbox"/> 重新分析語料 (Re-analyzed data)</p> <p>設定語料尋找方式：          請選擇： <input type="text" value="瀏覽模式 (course mode)"/></p> <p>執行 重置 輔助說明</p> <p>句子顯示</p> <p>搜尋範圍 - 語言別: 魯凱語/高山 文體: 全部 語料呈現: 全部 語料尋找: 瀏覽模式</p> <p>文本排: 1. 我們的話 語料資料</p> <p>檢視排: 顯示</p> <table border="1"> <thead> <tr> <th>Lang/Dial</th> <th>魯凱語/高山</th> <th>Genre</th> <th>全部</th> <th>Text</th> <th>1 我們的話</th> <th>Chp</th> <th>全部</th> <th>Phn</th> <th>國際</th> </tr> </thead> <tbody> <tr> <td>Format:</td> <td>raw</td> <td>Audio:</td> <td>on</td> <td>Sentence translation:</td> <td>TL</td> <td>Using Chinese</td> <td>Word Tran</td> <td></td> <td></td> </tr> <tr> <td>01.001.a</td> <td>onaŋi ʔaamaɓala-a-nai ta-piʔa-aa-na-ɓa po-aɟatja ʔoponoho m-ia.</td> <td></td> <td></td> <td></td> <td></td> <td></td> <td></td> <td></td> <td></td> </tr> <tr> <td>01.001.b</td> <td>Donaŋi a-kaavaŋi-ŋa-ɓa ʔaomo mani aɟa ʔitsaʔotsaʔo ʔina vaha-nai ʔoponoho i</td> <td></td> <td></td> <td></td> <td></td> <td></td> <td></td> <td></td> <td></td> </tr> <tr> <td>01.001.c</td> <td>mami poa a-pa-rama-ramao kapa-ka-ɓao-ɓao ta-ʔimjao ʔi tali-ɟakai ʔina vah</td> <td></td> <td></td> <td></td> <td></td> <td></td> <td></td> <td></td> <td></td> </tr> <tr> <td>01.001.d</td> <td>mami aɟ-inama patoʔo "maʔ-itolu-toloɟo-nom" m-ia.</td> <td></td> <td></td> <td></td> <td></td> <td></td> <td></td> <td></td> <td></td> </tr> <tr> <td>01.001.e</td> <td>mami ihoʔo-nai ʔina ʔoponoho "ka-iʔa-ta aanaŋi" m-ia-nai.</td> <td></td> <td></td> <td></td> <td></td> <td></td> <td></td> <td></td> <td></td> </tr> <tr> <td>01.001.f</td> <td>Donaŋi a-kaavaŋi-na-ɓa ʔaomo, o-ihoʔo-ka-nai maʔ-itolu-toloɟo-mita ɟa</td> <td></td> <td></td> <td></td> <td></td> <td></td> <td></td> <td></td> <td></td> </tr> <tr> <td>01.001.g</td> <td>kasa-ta ʔoponoho vaha-ta om-ia-nai.</td> <td></td> <td></td> <td></td> <td></td> <td></td> <td></td> <td></td> <td></td> </tr> <tr> <td>01.002.a</td> <td>ʔinaŋi vaha-nai ʔi ʔoponoho toramoro ka ma-kotsuŋai.</td> <td></td> <td></td> <td></td> <td></td> <td></td> <td></td> <td></td> <td></td> </tr> <tr> <td>01.002.b</td> <td>tsaala ʔinaŋi ia-a-laka-nai ta-ovaŋi-ʔja: o-ihoʔo-ka-i ka pa-itolobo ʔinaŋi ta-</td> <td></td> <td></td> <td></td> <td></td> <td></td> <td></td> <td></td> <td></td> </tr> </tbody> </table>	Lang/Dial	魯凱語/高山	Genre	全部	Text	1 我們的話	Chp	全部	Phn	國際	Format:	raw	Audio:	on	Sentence translation:	TL	Using Chinese	Word Tran			01.001.a	onaŋi ʔaamaɓala-a-nai ta-piʔa-aa-na-ɓa po-aɟatja ʔoponoho m-ia.									01.001.b	Donaŋi a-kaavaŋi-ŋa-ɓa ʔaomo mani aɟa ʔitsaʔotsaʔo ʔina vaha-nai ʔoponoho i									01.001.c	mami poa a-pa-rama-ramao kapa-ka-ɓao-ɓao ta-ʔimjao ʔi tali-ɟakai ʔina vah									01.001.d	mami aɟ-inama patoʔo "maʔ-itolu-toloɟo-nom" m-ia.									01.001.e	mami ihoʔo-nai ʔina ʔoponoho "ka-iʔa-ta aanaŋi" m-ia-nai.									01.001.f	Donaŋi a-kaavaŋi-na-ɓa ʔaomo, o-ihoʔo-ka-nai maʔ-itolu-toloɟo-mita ɟa									01.001.g	kasa-ta ʔoponoho vaha-ta om-ia-nai.									01.002.a	ʔinaŋi vaha-nai ʔi ʔoponoho toramoro ka ma-kotsuŋai.									01.002.b	tsaala ʔinaŋi ia-a-laka-nai ta-ovaŋi-ʔja: o-ihoʔo-ka-i ka pa-itolobo ʔinaŋi ta-								
Lang/Dial	魯凱語/高山	Genre	全部	Text	1 我們的話	Chp	全部	Phn	國際																																																																																																								
Format:	raw	Audio:	on	Sentence translation:	TL	Using Chinese	Word Tran																																																																																																										
01.001.a	onaŋi ʔaamaɓala-a-nai ta-piʔa-aa-na-ɓa po-aɟatja ʔoponoho m-ia.																																																																																																																
01.001.b	Donaŋi a-kaavaŋi-ŋa-ɓa ʔaomo mani aɟa ʔitsaʔotsaʔo ʔina vaha-nai ʔoponoho i																																																																																																																
01.001.c	mami poa a-pa-rama-ramao kapa-ka-ɓao-ɓao ta-ʔimjao ʔi tali-ɟakai ʔina vah																																																																																																																
01.001.d	mami aɟ-inama patoʔo "maʔ-itolu-toloɟo-nom" m-ia.																																																																																																																
01.001.e	mami ihoʔo-nai ʔina ʔoponoho "ka-iʔa-ta aanaŋi" m-ia-nai.																																																																																																																
01.001.f	Donaŋi a-kaavaŋi-na-ɓa ʔaomo, o-ihoʔo-ka-nai maʔ-itolu-toloɟo-mita ɟa																																																																																																																
01.001.g	kasa-ta ʔoponoho vaha-ta om-ia-nai.																																																																																																																
01.002.a	ʔinaŋi vaha-nai ʔi ʔoponoho toramoro ka ma-kotsuŋai.																																																																																																																
01.002.b	tsaala ʔinaŋi ia-a-laka-nai ta-ovaŋi-ʔja: o-ihoʔo-ka-i ka pa-itolobo ʔinaŋi ta-																																																																																																																



序	名稱（按筆劃）	形式	內容分析																																							
			<p>01.001.b  日本人來了以後就（開始）學我們高山和屏東縣（排灣及魯凱族）的話。          ðonaʔi a-kaavaʔi-rja-ða      ʔaomo mani a a      ʔitsaʔotsaʔo ʔir          那      分句名物化-動態-非限定 亦-已囉-他-屬 日本人      就      動態-非限定 幸 動態-處謂式 學 話</p> <p>01.001.c  他們比較這兩種語言後，（就發現）我們的話有許多（和）魯凱族是一樣的。          mani poa      a-pa-rama-ramao      kapa-ka-ðao-ðao ta-ʔini ao          歌      動態-非限定 弄 動態名物化-動態-非限定 相同-重疊-相同      狀態-非限定-重疊-步 主語名物化-動</p> <p>段落顯示</p> <p>1.           onaʔi ʔaamaðalaa-nai ta-piʔa-aa-na-ða po-alaʔa ʔoponoho m-ia.          ðonaʔi a-kaavaʔi-rja-ða ʔaomo mani a a ʔitsaʔotsaʔo ʔina vaha-nai          ʔoponoho ia tali- ao- ao. mani poa a-pa-rama-ramao kapa-ka-          ðao-ðao ta-ʔini ao ʔi tali-ðəkai ʔina vaha-nai. mani a a-inamə patoʔo:          "maʔ-itolo-to o o-nomi" m-ia. mani  ihoʔo-nai ʔina ʔoponoho "ka-iʔa-          ta aana " m-ia-nai. ðonaʔi a-ki-kaavaʔi-na-ða ʔaomo, o- ihoʔo-ka-nai          maʔ-itolo-to o o-mita ðəkai m-ia. kasa-ta ʔoponoho vaha-ta om-ia-          nai.</p> <p>我們的祖先高山自稱是高山人。日本人來了以後就（開始）學我們高山和屏東縣（排灣及魯凱族）的話。他們比較這兩種語言後，（就發現）我們的話有許多（和）魯凱族是一樣的。然後他們告訴我們說：「你是阿宗（的）。」我們高山人（才）知道「原來咱們是（屬於魯凱族）。」日本人還沒來之前，我們不知道我們和魯凱族同宗。我們只（認為）咱們的語言（就）是高山話。</p> <p>內容檢索</p> <p>內容檢索</p> <p><a href="#">回上頁</a> <a href="#">使用說明</a> <a href="#">回首頁</a></p> <p>■ 搜尋範圍-語言別: 魯凱語/高山 文體: 全部 語料搜尋: 搜尋模式</p> <div data-bbox="742 1003 1380 1243"> <p>關鍵詞示例</p> <p>1. 關鍵詞以原文輸入時,若使用到 IPA音標,請點選對應的代碼如下:</p> <table border="1"> <tr> <td>IPA</td> <td>ð</td> <td>ŋ</td> <td>ɔ</td> <td>l</td> <td>ʔ</td> <td>ɪ</td> <td>θ</td> <td>ʌ</td> <td>ɤ</td> </tr> <tr> <td>代碼</td> <td>240</td> <td>331</td> <td>598</td> <td>601</td> <td>621</td> <td>660</td> <td>616</td> <td>952</td> <td>654</td> <td>630</td> </tr> </table> <p>IPA</p> <table border="1"> <tr> <td>ɸ</td> <td>ʃ</td> <td>æ</td> <td>œ</td> <td>β</td> <td>ɰ</td> <td>ʏ</td> <td>ʈ</td> </tr> <tr> <td>代碼</td> <td>216</td> <td>643</td> <td>230</td> <td>339</td> <td>946</td> <td>649</td> <td>404</td> <td>607</td> <td>648</td> </tr> </table> <p>2. 關鍵詞以註解輸入時,請輸入可能出現的字(如:酒、芒草、書)</p> <p><input type="button" value="執行"/> <input type="button" value="清除"/></p> </div>	IPA	ð	ŋ	ɔ	l	ʔ	ɪ	θ	ʌ	ɤ	代碼	240	331	598	601	621	660	616	952	654	630	ɸ	ʃ	æ	œ	β	ɰ	ʏ	ʈ	代碼	216	643	230	339	946	649	404	607	648
IPA	ð	ŋ	ɔ	l	ʔ	ɪ	θ	ʌ	ɤ																																	
代碼	240	331	598	601	621	660	616	952	654	630																																
ɸ	ʃ	æ	œ	β	ɰ	ʏ	ʈ																																			
代碼	216	643	230	339	946	649	404	607	648																																	
7	<p>蘭嶼達悟語口語資料典藏</p> <p><a href="http://yamiproject.cs.pu.edu.tw/yami/yami_ch/c_h_index_flash.htm">http://yamiproject.cs.pu.edu.tw/yami/yami_ch/c_h_index_flash.htm</a></p>	<p>小型語料庫</p>	<ul style="list-style-type: none"> <li>➤ 支援：為靜宜大學團隊執行計畫。</li> <li>➤ 目的：有系統的使用資訊科技來保存達悟語。</li> <li>➤ 特色與功能：             <ol style="list-style-type: none"> <li>1. 有基本族群的介紹。</li> <li>2. 依主題劃分文章，有音檔案語言分析，含族語、標記、翻譯。</li> <li>3. 有上課教學影片可以下載。</li> <li>4. 數位典藏部分收錄教材、語料、聖經可以查詢。</li> <li>5. 另有連結應用功能的蘭嶼達悟語線上學習網，有課程和字典查詢功能。</li> </ol> </li> <li>➤ 分析：網站分析語料是以 XML 方式，結合線上學習網，是研究和學習達悟語很好的工</li> </ul>																																							

序	名稱（按筆劃）	形式	內容分析																																
			<p>具。</p> <p>語料分類</p> <div data-bbox="703 414 1398 667"> <p>[達悟語料]</p> <p>故事篇,生活篇,祭典與傳統文化篇,政見發表篇,女性議題篇,動畫篇,其他篇</p> <table border="1"> <thead> <tr> <th>主分類</th> <th>次分類</th> <th>篇序</th> <th>達悟篇名</th> <th>中文篇名</th> </tr> </thead> <tbody> <tr> <td></td> <td>意外經驗</td> <td>01</td> <td>niomahas a longtsang si Kawcincong</td> <td>越獄犯高連忠</td> </tr> <tr> <td></td> <td>學校生活</td> <td>02</td> <td>mapakazakazakat a yalalaw so sinsi</td> <td>裝死嚇老師</td> </tr> <tr> <td></td> <td rowspan="4">部落生活</td> <td>03</td> <td>aney o hango ko!</td> <td>還我林投果!</td> </tr> <tr> <td></td> <td>04</td> <td>man'acing so kols</td> <td>圈豬腳</td> </tr> <tr> <td></td> <td>05</td> <td>miarap</td> <td>打群架</td> </tr> <tr> <td></td> <td>06</td> <td>teymangangasien ko ya</td> <td>我也是個有愛心的人</td> </tr> </tbody> </table> </div> <p>語料分析</p> <div data-bbox="703 712 1398 1093"> <p>1. yaken rana am, mina apia dang o kakawan a.</p> <p>我 主格 已經 呢 曾經 很好 當時 主格 天氣 助</p> <p>1.S.NOM already PAR PA good then NOM weather PAR</p> <p>k-aro namen a ka-k-anak-an</p> <p>故-很多 我們 屬格 除 繫 疊-名綴-小孩子-名綴</p> <p>CON-many 1.PGEN.EXCL LIN RED-NF-child-NF</p> <p>a mi-al-alam do Jicinikikan ito a.</p> <p>繫 主無-疊-玩耍 處格 地名 那個 助</p> <p>LIN AF-RED-play LOC PLN that PAR</p> <p>【我呢，那天天氣很好，所以我們很多小孩子就都去Jicinikikan遊玩】</p> <p>As for me, one day the weather was very good, and a lot of us children went to Jicinikikan to play.</p> </div> <p>關鍵字查詢</p> <div data-bbox="703 1144 1398 1568"> <p>關鍵字查詢</p> <p>查詢方式</p> <p><input checked="" type="radio"/> 達悟語查詢 <input type="radio"/> 中文查詢 <input type="radio"/> 英文查詢</p> <p>查詢來源</p> <p><input checked="" type="radio"/> 所有來源 <input type="radio"/> 四冊教材 <input type="radio"/> 二十篇語料 <input type="radio"/> 聖經</p> <p>查詢種類</p> <p><input checked="" type="radio"/> 部份查詢 <input type="radio"/> 全字查詢</p> <p>請輸入關鍵字:</p> <p>開始查詢</p> </div>	主分類	次分類	篇序	達悟篇名	中文篇名		意外經驗	01	niomahas a longtsang si Kawcincong	越獄犯高連忠		學校生活	02	mapakazakazakat a yalalaw so sinsi	裝死嚇老師		部落生活	03	aney o hango ko!	還我林投果!		04	man'acing so kols	圈豬腳		05	miarap	打群架		06	teymangangasien ko ya	我也是個有愛心的人
主分類	次分類	篇序	達悟篇名	中文篇名																															
	意外經驗	01	niomahas a longtsang si Kawcincong	越獄犯高連忠																															
	學校生活	02	mapakazakazakat a yalalaw so sinsi	裝死嚇老師																															
	部落生活	03	aney o hango ko!	還我林投果!																															
		04	man'acing so kols	圈豬腳																															
		05	miarap	打群架																															
		06	teymangangasien ko ya	我也是個有愛心的人																															

## 參、數位族語資料庫建置內容說明

根據前一章的介紹，我們瞭解了語發中心數位族語資料庫的規劃及建置的概念，及大致瞭解國內外語言典藏網站的呈現方式，接下來則回顧本計畫的工作目的、執行方式、語料建置工作說明，並彙整現階段資料庫建置內容，最後則提出未來轉譯工作及資料修復需注意的幾點建議。

### 一、數位族語資料庫建置目的

建置數位族語資料庫的核心目的，就是要進行語言資料的保存及應用。2017年6月14日公布的原住民族語言發展法第九條亦明訂「需建置原住民族語言資料庫，積極保存原住民族語料」。

而因數位族語資料庫未來應是大型的語言資料庫，故2017年至2019年主要是預先研擬資料庫的小型模型，規劃、建置及資料存取測試。換句話說，語發中心建置「數位族語資料庫」即是希望可以作為基礎，未來可連結各計畫及單位的族語資料、培育轉寫人才並推廣供民眾使用，促進語言的保存之外，更能讓語言工具的使用更便利。目前族語資料庫網站含語料典藏、族語新聞轉譯及原住民族語言研書目等，可供民眾檢索使用。

### 二、計畫執行方式

本計畫今（2019）年主要的工作目標，即是數位族語資料庫的維運工作，除了持續收集語料並進行數位化轉檔修復外，亦針對長篇口語語料及族語新聞請族人協作轉寫及翻譯。

而主要執行的方式包含田調訪談、辦理研商會議（語料轉譯、資料庫維運及修正工作、系統修正），另外則是請族人協助文字轉寫、轉譯及詮釋資料確認等工作，並積極參與相關的活動及會議。而本年度執行架構如下圖：

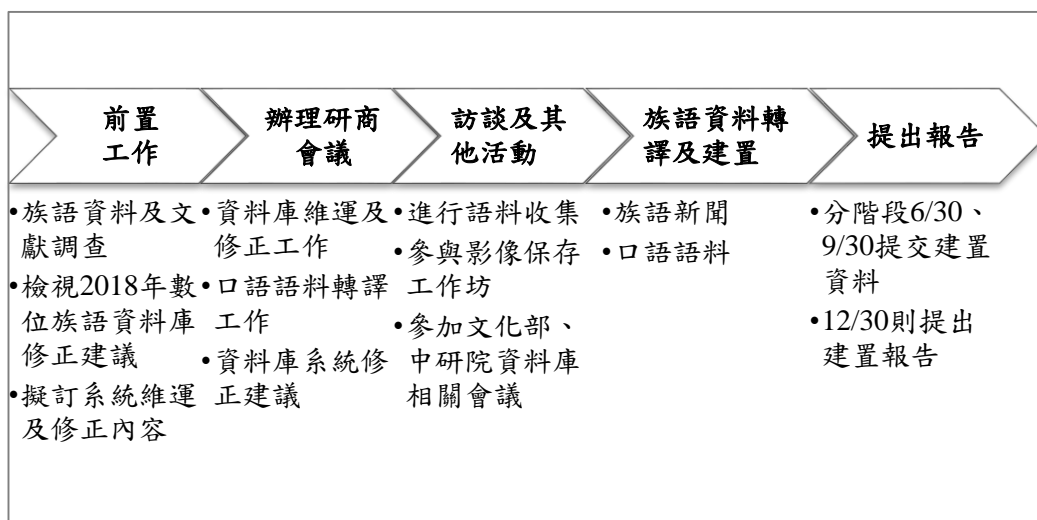


圖 2：數位族語資料庫—2019 年計畫執行架構

以下我們就分項說明計畫執行的方式：

#### (一) 前置工作

為將珍貴的族語文獻資料彙集並進行數位典藏，以利族語資料之保存、研究及應用，因此 2017 年我們透過國家圖書館「臺灣書目整合查詢系統」、「臺灣原住民族圖書資訊中心」、「臺灣原住民族數位典藏資料庫」等系統性檢索，並以及田野調查的方式，初步彙整相關的語料清單<sup>6</sup>。

而語料清單分為書面語料、口語語料及其他影音語料，根據語料清單調查的結果，教材的數量最多有 375 筆，而詞典類也有 51 筆，其餘類型如文學創作、口傳文學等也不多。依族群來看，阿美族有 109 筆是數量較多，卡那卡那富族及拉阿魯哇族則是僅列 6 筆，也反應了多方言、單方言及族群人數多寡之差距，因此在未來資料的建置時程上，我們必須加快腳步，也一起配合搶救瀕危語言計畫，盡量保存我們的語言資料。

經過 2017 年田調訪談，計畫面臨到一個困境，即無法於現階段取得私人族語文獻（語料）持有相關的資料清單，訪談的結果一致都是希望網站能先建立起來，並且確立智慧財產權的保障，讓資料安全且

<sup>6</sup>數位族語資料庫語料清單見原住民族語言研究發展中心 2017 年「原住民族語言研究發展中心數位族語資料庫建置評析報告」。

是有限制的被使用，另外也提到若能提供資料檔案的修復，亦是一大吸引。2018年語發中心在訪談調查的過程中，太魯閣族楊盛涂耆老不吝分享，授權提供40卷錄音帶，故2019年我們優先進行數位化及修復工作，經族人協助進行音檔確認及整理後，可修復及完整的音檔我們在完成詮釋資料後建置於資料庫中，另挑選5卷語料進行文字化及轉譯工作。

另外，我們也針對以中研院數位典藏技術分項所建置的「VCenter」中的影音做初步的檢索，發現若以「阿美族」為關鍵詞，至少有30則以上的影音是可以轉寫的族語口語語料素材，未來應考量合作的方式及資料使用的可能性。

而除了前述基礎的語料整理之外，我們也依2018年的系統修正建議做部分的修正及微調，並維持系統的運作。

## (二) 辦理研商會議

為了進行系統修正、維護及轉譯工作，2019年共辦理9場資料庫修正研商會議、5場轉譯工作研商會議、1場資料庫修正建議研商會議，會議統計及記錄詳見附錄一。會議重點建議整理如下：

項目	綜合建議整理
轉譯工作	<ol style="list-style-type: none"> <li>1. 目前族語新聞逐字稿及編輯系統對於年紀較長的族人操作起來較為不便。</li> <li>2. 未來需將其他平臺逐字稿檔案匯入的功能考量，或是資料庫平臺整合時應注意整體規劃。</li> <li>3. 書寫規範尚未完成，譯者需保持書寫的一致性，有利於未來調整。</li> <li>4. 應開發口語語料轉譯系統，並需參考youtube平臺字幕功能，有利於其他計畫人員銜接。</li> </ol>
系統修正	<ol style="list-style-type: none"> <li>1. 未來需專職瞭解南島語言結構之工程師協助規劃及溝通，有利資料庫平臺的整合。</li> <li>2. 應跨機關單位合作，應用現有技術、借鑑其他語言資料庫建置情形再調整。</li> </ol>

(三) 訪談及其他活動

為持續收集語料並確認口語語料內容，2019年以田調訪談方式收集語料，訪談相關記錄見附錄二。另一方面，也積極參與相關活動，如中研院「國家語料庫規劃與建置論壇」、文化部「規劃國家語言資料庫及推動國家語言相關事務諮詢會議」、南藝大音像紀錄研究所「影像修復工作坊」，重點整理如下表：

表 2：參與其他活動／會議重點整理

日期	重點
4/18	<p>中研院「國家語料庫規劃與建置論壇」</p> <ol style="list-style-type: none"> <li>(1) 過去語料庫接為分散經營且計畫不持續。</li> <li>(2) 技術人員缺乏或流動性大，需要共通格式及 POS 詞類標記等。</li> <li>(3) 版權問題應有公平合理應用的規範，因此著作權法、授權同意書等設計都是很重要的議題。</li> <li>(4) 國家語料庫應可跨語言檢索，方言也可以被納入，會聯繫到共同詞彙詞，是一種語意網的概念。</li> <li>(5) 需要有共通的核心詞表，包含核心固有詞及當代詞彙，以及例句，並有完整的話語材料，需考量詞序不同或是應配合各族文化知識的呈現。</li> <li>(6) 應包含日常口語語料的收集，需先分類主題，再做口語的平衡語料庫。</li> <li>(7) 一開始的資料應先多元收集，保存文化的多樣性，整個過程是滾動性的，如愛台語、新詞等。</li> </ol>
4/25	<p>文化部「規劃國家語言資料庫及推動國家語言相關事務諮詢會議」</p> <ol style="list-style-type: none"> <li>(1) 主要是針對文化部主責規劃建置國家語言資料庫而召開的會議，語發中心列席說明目前數位族語資料庫的建置情形。</li> <li>(2) 瞭解目前跨部會（教育部、科技部、客委會、原民會、中研院）及學術界語界建置相關語料庫的規劃或是情形。</li> <li>(3) 因目前仍為諮詢階段，故大多建議為，國家語料庫的建置需設置專責且永久性的單位，而非任務性導向，才能作長遠且宏觀的規劃，需有決策權，以利跨部會之分工及整合。</li> </ol>
8/31-	<p>南藝大音像紀錄研究所「影像修復工作坊」</p>

日期	重點
9/1	(1) 瞭解磁帶影音媒材的特性及歷史。 (2) 修復影音資料及數位化的重要性和與急迫性。 (3) 實際進行影帶修復工作。 (4) 進行修復後的影像放映，並敘說背後的故事。

#### (四) 系統維運情形

本年1月至12月，數位族語資料庫網站已持續委請工程師進行系統維護工作，並保持運作正常，相關證明如下：

##### 1. 伺服器租用情形<sup>7</sup>

The screenshot shows a server management page for a server named 'promo1' with ID 'BV550190900002'. The interface includes a '資訊' (Information) section with the following details:

運作狀態	● 已開機
申租狀態	✔ 已申租
登入帳號/預設密碼	root <a href="#">取得預設密碼</a>
產品名稱	標準型 Mini-Promotions(Linux)
作業系統	Ubuntu 16.04 eng, 64bit
規格	1 virtual core [等同於1GHz運算能力], 1 GB 記憶體
系統硬碟空間	一般硬碟 30 GB
區域	第一區域
Zone	CaaS第一區
雲伺服器監控週期	5分鐘 (免費) <input checked="" type="radio"/> 1分鐘 (進階版需額外付費)

<sup>7</sup> 伺服器租用情形為2019年12月6日於管理後臺截圖。

訂單資訊	
訂單編號:	2965717
名稱:	promo1
申請日期:	2018/11/14 16:31:14
起租日期:	2018/11/14 16:38:37
已使用:	387 天 0 小時 32 分
網路技術/管理聯絡人:	國立台灣大學

## 2. 伺服器流量<sup>8</sup>

自 2019 年 1 月 1 日至報告完成前 1 日 (12 月 5 日)，由伺服器寄送流量統計截圖，證明伺服器正常運行。未免太冗常，縮圖處理。

雲端服務編號	起始日期	起始時間	終止日期	終止時間	Public IP	流量
BV550190900002	2019/01/01	00:00:00	2019/01/01	23:59:59	210.61.46.35	26 MB
BV550190900002	2019/01/02	00:00:00	2019/01/02	23:59:59	210.61.46.35	6 MB
BV550190900002	2019/01/03	00:00:00	2019/01/03	23:59:59	210.61.46.35	19 MB
BV550190900002	2019/01/04	00:00:00	2019/01/04	23:59:59	210.61.46.35	59 MB
BV550190900002	2019/01/05	00:00:00	2019/01/05	23:59:59	210.61.46.35	46 MB
BV550190900002	2019/01/06	00:00:00	2019/01/06	23:59:59	210.61.46.35	37 MB
BV550190900002	2019/01/07	00:00:00	2019/01/07	23:59:59	210.61.46.35	33 MB
BV550190900002	2019/01/08	00:00:00	2019/01/08	23:59:59	210.61.46.35	26 MB
BV550190900002	2019/01/09	00:00:00	2019/01/09	23:59:59	210.61.46.35	50 MB
BV550190900002	2019/01/10	00:00:00	2019/01/10	23:59:59	210.61.46.35	55 MB
BV550190900002	2019/01/11	00:00:00	2019/01/11	23:59:59	210.61.46.35	19 MB
BV550190900002	2019/01/12	00:00:00	2019/01/12	23:59:59	210.61.46.35	28 MB
BV550190900002	2019/01/13	00:00:00	2019/01/13	23:59:59	210.61.46.35	38 MB
BV550190900002	2019/01/14	00:00:00	2019/01/14	23:59:59	210.61.46.35	47 MB
BV550190900002	2019/01/15	00:00:00	2019/01/15	23:59:59	210.61.46.35	38 MB
BV550190900002	2019/01/16	00:00:00	2019/01/16	23:59:59	210.61.46.35	22 MB
BV550190900002	2019/01/17	00:00:00	2019/01/17	23:59:59	210.61.46.35	43 MB
BV550190900002	2019/01/18	00:00:00	2019/01/18	23:59:59	210.61.46.35	27 MB
BV550190900002	2019/01/19	00:00:00	2019/01/19	23:59:59	210.61.46.35	43 MB
BV550190900002	2019/01/20	00:00:00	2019/01/20	23:59:59	210.61.46.35	66 MB
BV550190900002	2019/01/21	00:00:00	2019/01/21	23:59:59	210.61.46.35	77 MB
BV550190900002	2019/01/22	00:00:00	2019/01/22	23:59:59	210.61.46.35	12 MB
BV550190900002	2019/01/23	00:00:00	2019/01/23	23:59:59	210.61.46.35	7 MB
BV550190900002	2019/01/24	00:00:00	2019/01/24	23:59:59	210.61.46.35	6 MB
BV550190900002	2019/01/25	00:00:00	2019/01/25	23:59:59	210.61.46.35	9 MB
BV550190900002	2019/01/26	00:00:00	2019/01/26	23:59:59	210.61.46.35	6 MB
BV550190900002	2019/01/27	00:00:00	2019/01/27	23:59:59	210.61.46.35	12 MB
BV550190900002	2019/01/28	00:00:00	2019/01/28	23:59:59	210.61.46.35	22 MB
BV550190900002	2019/01/29	00:00:00	2019/01/29	23:59:59	210.61.46.35	30 MB
BV550190900002	2019/01/30	00:00:00	2019/01/30	23:59:59	210.61.46.35	38 MB
BV550190900002	2019/01/31	00:00:00	2019/01/31	23:59:59	210.61.46.35	9 MB
總計 (網路計費是以申請的所有服務之對外流量加總後合併計收)						956 MB

中華電信 < 雲端服務 >

<sup>8</sup> 此為伺服器營運商提供之伺服器流量記錄，以佐證伺服器有正常運作。



雲端服務編號	起始日期	起始時間	終止日期	終止時間	Public IP	流量
BV550190900002	2019/02/01	00:00:00	2019/02/01	23:59:59	210.61.46.35	7 MB
BV550190900002	2019/02/02	00:00:00	2019/02/02	23:59:59	210.61.46.35	17 MB
BV550190900002	2019/02/03	00:00:00	2019/02/03	23:59:59	210.61.46.35	29 MB
BV550190900002	2019/02/04	00:00:00	2019/02/04	23:59:59	210.61.46.35	39 MB
BV550190900002	2019/02/05	00:00:00	2019/02/05	23:59:59	210.61.46.35	14 MB
BV550190900002	2019/02/06	00:00:00	2019/02/06	23:59:59	210.61.46.35	13 MB
BV550190900002	2019/02/07	00:00:00	2019/02/07	23:59:59	210.61.46.35	18 MB
BV550190900002	2019/02/08	00:00:00	2019/02/08	23:59:59	210.61.46.35	13 MB
BV550190900002	2019/02/09	00:00:00	2019/02/09	23:59:59	210.61.46.35	19 MB
BV550190900002	2019/02/10	00:00:00	2019/02/10	23:59:59	210.61.46.35	6 MB
BV550190900002	2019/02/11	00:00:00	2019/02/11	23:59:59	210.61.46.35	11 MB
BV550190900002	2019/02/12	00:00:00	2019/02/12	23:59:59	210.61.46.35	31 MB
BV550190900002	2019/02/13	00:00:00	2019/02/13	23:59:59	210.61.46.35	10 MB
BV550190900002	2019/02/14	00:00:00	2019/02/14	23:59:59	210.61.46.35	14 MB
BV550190900002	2019/02/15	00:00:00	2019/02/15	23:59:59	210.61.46.35	12 MB
BV550190900002	2019/02/16	00:00:00	2019/02/16	23:59:59	210.61.46.35	11 MB
BV550190900002	2019/02/17	00:00:00	2019/02/17	23:59:59	210.61.46.35	6 MB
BV550190900002	2019/02/18	00:00:00	2019/02/18	23:59:59	210.61.46.35	18 MB
BV550190900002	2019/02/19	00:00:00	2019/02/19	23:59:59	210.61.46.35	26 MB
BV550190900002	2019/02/20	00:00:00	2019/02/20	23:59:59	210.61.46.35	11 MB
BV550190900002	2019/02/21	00:00:00	2019/02/21	23:59:59	210.61.46.35	29 MB
BV550190900002	2019/02/22	00:00:00	2019/02/22	23:59:59	210.61.46.35	18 MB
BV550190900002	2019/02/23	00:00:00	2019/02/23	23:59:59	210.61.46.35	18 MB
BV550190900002	2019/02/24	00:00:00	2019/02/24	23:59:59	210.61.46.35	21 MB
BV550190900002	2019/02/25	00:00:00	2019/02/25	23:59:59	210.61.46.35	43 MB
BV550190900002	2019/02/26	00:00:00	2019/02/26	23:59:59	210.61.46.35	37 MB
BV550190900002	2019/02/27	00:00:00	2019/02/27	23:59:59	210.61.46.35	29 MB
總計 (網路計費是以申請的所有服務之對外流量加總後合併計收)						520 MB
中華電信 < 雲端服務 >						

雲端服務編號	起始日期	起始時間	終止日期	終止時間	Public IP	流量
BV550190900002	2019/03/01	00:00:00	2019/03/01	23:59:59	210.61.46.35	9 MB
BV550190900002	2019/03/02	00:00:00	2019/03/02	23:59:59	210.61.46.35	23 MB
BV550190900002	2019/03/03	00:00:00	2019/03/03	23:59:59	210.61.46.35	15 MB
BV550190900002	2019/03/04	00:00:00	2019/03/04	23:59:59	210.61.46.35	17 MB
BV550190900002	2019/03/05	00:00:00	2019/03/05	23:59:59	210.61.46.35	29 MB
BV550190900002	2019/03/06	00:00:00	2019/03/06	23:59:59	210.61.46.35	13 MB
BV550190900002	2019/03/07	00:00:00	2019/03/07	23:59:59	210.61.46.35	46 MB
BV550190900002	2019/03/08	00:00:00	2019/03/08	23:59:59	210.61.46.35	33 MB
BV550190900002	2019/03/09	00:00:00	2019/03/09	23:59:59	210.61.46.35	21 MB
BV550190900002	2019/03/10	00:00:00	2019/03/10	23:59:59	210.61.46.35	6 MB
BV550190900002	2019/03/11	00:00:00	2019/03/11	23:59:59	210.61.46.35	43 MB
BV550190900002	2019/03/12	00:00:00	2019/03/12	23:59:59	210.61.46.35	36 MB
BV550190900002	2019/03/13	00:00:00	2019/03/13	23:59:59	210.61.46.35	32 MB
BV550190900002	2019/03/14	00:00:00	2019/03/14	23:59:59	210.61.46.35	18 MB
BV550190900002	2019/03/15	00:00:00	2019/03/15	23:59:59	210.61.46.35	27 MB
BV550190900002	2019/03/16	00:00:00	2019/03/16	23:59:59	210.61.46.35	49 MB
BV550190900002	2019/03/17	00:00:00	2019/03/17	23:59:59	210.61.46.35	46 MB
BV550190900002	2019/03/18	00:00:00	2019/03/18	23:59:59	210.61.46.35	365 MB
BV550190900002	2019/03/19	00:00:00	2019/03/19	23:59:59	210.61.46.35	30 MB
BV550190900002	2019/03/20	00:00:00	2019/03/20	23:59:59	210.61.46.35	36 MB
BV550190900002	2019/03/21	00:00:00	2019/03/21	23:59:59	210.61.46.35	40 MB
BV550190900002	2019/03/22	00:00:00	2019/03/22	23:59:59	210.61.46.35	13 MB
BV550190900002	2019/03/23	00:00:00	2019/03/23	23:59:59	210.61.46.35	33 MB
BV550190900002	2019/03/24	00:00:00	2019/03/24	23:59:59	210.61.46.35	31 MB
BV550190900002	2019/03/25	00:00:00	2019/03/25	23:59:59	210.61.46.35	15 MB
BV550190900002	2019/03/26	00:00:00	2019/03/26	23:59:59	210.61.46.35	13 MB
BV550190900002	2019/03/27	00:00:00	2019/03/27	23:59:59	210.61.46.35	81 MB
BV550190900002	2019/03/28	00:00:00	2019/03/28	23:59:59	210.61.46.35	27 MB
BV550190900002	2019/03/29	00:00:00	2019/03/29	23:59:59	210.61.46.35	32 MB
BV550190900002	2019/03/30	00:00:00	2019/03/30	23:59:59	210.61.46.35	25 MB
BV550190900002	2019/03/31	00:00:00	2019/03/31	23:59:59	210.61.46.35	20 MB
總計 (網路計費是以申請的所有服務之對外流量加總後合併計收)						1224 MB
中華電信 < 雲端服務 >						

雲端服務編號	起始日期	起始時間	終止日期	終止時間	Public IP	流量
BV550190900002	2019/04/01	00:00:00	2019/04/01	23:59:59	210.61.46.35	22 MB
BV550190900002	2019/04/02	00:00:00	2019/04/02	23:59:59	210.61.46.35	106 MB
BV550190900002	2019/04/03	00:00:00	2019/04/03	23:59:59	210.61.46.35	30 MB
BV550190900002	2019/04/04	00:00:00	2019/04/04	23:59:59	210.61.46.35	19 MB
BV550190900002	2019/04/05	00:00:00	2019/04/05	23:59:59	210.61.46.35	47 MB
BV550190900002	2019/04/06	00:00:00	2019/04/06	23:59:59	210.61.46.35	37 MB
BV550190900002	2019/04/07	00:00:00	2019/04/07	23:59:59	210.61.46.35	38 MB
BV550190900002	2019/04/08	00:00:00	2019/04/08	23:59:59	210.61.46.35	31 MB
BV550190900002	2019/04/09	00:00:00	2019/04/09	23:59:59	210.61.46.35	76 MB
BV550190900002	2019/04/10	00:00:00	2019/04/10	23:59:59	210.61.46.35	53 MB
BV550190900002	2019/04/11	00:00:00	2019/04/11	23:59:59	210.61.46.35	36 MB
BV550190900002	2019/04/12	00:00:00	2019/04/12	23:59:59	210.61.46.35	34 MB
BV550190900002	2019/04/13	00:00:00	2019/04/13	23:59:59	210.61.46.35	24 MB
BV550190900002	2019/04/14	00:00:00	2019/04/14	23:59:59	210.61.46.35	25 MB
BV550190900002	2019/04/15	00:00:00	2019/04/15	23:59:59	210.61.46.35	27 MB
BV550190900002	2019/04/16	00:00:00	2019/04/16	23:59:59	210.61.46.35	19 MB
BV550190900002	2019/04/17	00:00:00	2019/04/17	23:59:59	210.61.46.35	31 MB
BV550190900002	2019/04/18	00:00:00	2019/04/18	23:59:59	210.61.46.35	31 MB
BV550190900002	2019/04/19	00:00:00	2019/04/19	23:59:59	210.61.46.35	28 MB
BV550190900002	2019/04/20	00:00:00	2019/04/20	23:59:59	210.61.46.35	38 MB
BV550190900002	2019/04/21	00:00:00	2019/04/21	23:59:59	210.61.46.35	20 MB
BV550190900002	2019/04/22	00:00:00	2019/04/22	23:59:59	210.61.46.35	11 MB
BV550190900002	2019/04/23	00:00:00	2019/04/23	23:59:59	210.61.46.35	58 MB
BV550190900002	2019/04/24	00:00:00	2019/04/24	23:59:59	210.61.46.35	335 MB
BV550190900002	2019/04/25	00:00:00	2019/04/25	23:59:59	210.61.46.35	60 MB
BV550190900002	2019/04/26	00:00:00	2019/04/26	23:59:59	210.61.46.35	41 MB
BV550190900002	2019/04/27	00:00:00	2019/04/27	23:59:59	210.61.46.35	34 MB
BV550190900002	2019/04/28	00:00:00	2019/04/28	23:59:59	210.61.46.35	9 MB
BV550190900002	2019/04/29	00:00:00	2019/04/29	23:59:59	210.61.46.35	117 MB
BV550190900002	2019/04/30	00:00:00	2019/04/30	23:59:59	210.61.46.35	43 MB
總計 (網路計費是以申請的所有服務之對外流量加總後合併計收)						1480 MB
中華電信 < 雲端服務 >						

雲端服務編號	起始日期	起始時間	終止日期	終止時間	Public IP	流量
BV550190900002	2019/05/01	00:00:00	2019/05/01	23:59:59	210.61.46.35	16 MB
BV550190900002	2019/05/02	00:00:00	2019/05/02	23:59:59	210.61.46.35	14 MB
BV550190900002	2019/05/03	00:00:00	2019/05/03	23:59:59	210.61.46.35	41 MB
BV550190900002	2019/05/04	00:00:00	2019/05/04	23:59:59	210.61.46.35	14 MB
BV550190900002	2019/05/05	00:00:00	2019/05/05	23:59:59	210.61.46.35	14 MB
BV550190900002	2019/05/06	00:00:00	2019/05/06	23:59:59	210.61.46.35	27 MB
BV550190900002	2019/05/07	00:00:00	2019/05/07	23:59:59	210.61.46.35	25 MB
BV550190900002	2019/05/08	00:00:00	2019/05/08	23:59:59	210.61.46.35	47 MB
BV550190900002	2019/05/09	00:00:00	2019/05/09	23:59:59	210.61.46.35	49 MB
BV550190900002	2019/05/10	00:00:00	2019/05/10	23:59:59	210.61.46.35	43 MB
BV550190900002	2019/05/11	00:00:00	2019/05/11	23:59:59	210.61.46.35	15 MB
BV550190900002	2019/05/12	00:00:00	2019/05/12	23:59:59	210.61.46.35	23 MB
BV550190900002	2019/05/13	00:00:00	2019/05/13	23:59:59	210.61.46.35	51 MB
BV550190900002	2019/05/14	00:00:00	2019/05/14	23:59:59	210.61.46.35	16 MB
BV550190900002	2019/05/15	00:00:00	2019/05/15	23:59:59	210.61.46.35	29 MB
BV550190900002	2019/05/16	00:00:00	2019/05/16	23:59:59	210.61.46.35	29 MB
BV550190900002	2019/05/17	00:00:00	2019/05/17	23:59:59	210.61.46.35	40 MB
BV550190900002	2019/05/18	00:00:00	2019/05/18	23:59:59	210.61.46.35	25 MB
BV550190900002	2019/05/19	00:00:00	2019/05/19	23:59:59	210.61.46.35	33 MB
BV550190900002	2019/05/20	00:00:00	2019/05/20	23:59:59	210.61.46.35	33 MB
BV550190900002	2019/05/21	00:00:00	2019/05/21	23:59:59	210.61.46.35	19 MB
BV550190900002	2019/05/22	00:00:00	2019/05/22	23:59:59	210.61.46.35	41 MB
BV550190900002	2019/05/23	00:00:00	2019/05/23	23:59:59	210.61.46.35	23 MB
BV550190900002	2019/05/24	00:00:00	2019/05/24	23:59:59	210.61.46.35	34 MB
BV550190900002	2019/05/25	00:00:00	2019/05/25	23:59:59	210.61.46.35	15 MB
BV550190900002	2019/05/26	00:00:00	2019/05/26	23:59:59	210.61.46.35	42 MB
BV550190900002	2019/05/27	00:00:00	2019/05/27	23:59:59	210.61.46.35	53 MB
BV550190900002	2019/05/28	00:00:00	2019/05/28	23:59:59	210.61.46.35	28 MB
BV550190900002	2019/05/29	00:00:00	2019/05/29	23:59:59	210.61.46.35	31 MB
BV550190900002	2019/05/30	00:00:00	2019/05/30	23:59:59	210.61.46.35	53 MB
BV550190900002	2019/05/31	00:00:00	2019/05/31	23:59:59	210.61.46.35	22 MB
總計 (網路計費是以申請的所有服務之對外流量加總後合併計收)						945 MB
中華電信 < 雲端服務 >						

雲端服務編號	起始日期	起始時間	終止日期	終止時間	Public IP	流量
BV550190900002	2019/07/01	00:00:00	2019/07/01	23:59:59	210.61.46.35	31 MB
BV550190900002	2019/07/02	00:00:00	2019/07/02	23:59:59	210.61.46.35	23 MB
BV550190900002	2019/07/03	00:00:00	2019/07/03	23:59:59	210.61.46.35	14 MB
BV550190900002	2019/07/04	00:00:00	2019/07/04	23:59:59	210.61.46.35	23 MB
BV550190900002	2019/07/05	00:00:00	2019/07/05	23:59:59	210.61.46.35	22 MB
BV550190900002	2019/07/06	00:00:00	2019/07/06	23:59:59	210.61.46.35	14 MB
BV550190900002	2019/07/07	00:00:00	2019/07/07	23:59:59	210.61.46.35	10 MB
BV550190900002	2019/07/08	00:00:00	2019/07/08	23:59:59	210.61.46.35	17 MB
BV550190900002	2019/07/09	00:00:00	2019/07/09	23:59:59	210.61.46.35	18 MB
BV550190900002	2019/07/10	00:00:00	2019/07/10	23:59:59	210.61.46.35	23 MB
BV550190900002	2019/07/11	00:00:00	2019/07/11	23:59:59	210.61.46.35	32 MB
BV550190900002	2019/07/12	00:00:00	2019/07/12	23:59:59	210.61.46.35	15 MB
BV550190900002	2019/07/13	00:00:00	2019/07/13	23:59:59	210.61.46.35	19 MB
BV550190900002	2019/07/14	00:00:00	2019/07/14	23:59:59	210.61.46.35	18 MB
BV550190900002	2019/07/15	00:00:00	2019/07/15	23:59:59	210.61.46.35	13 MB
BV550190900002	2019/07/16	00:00:00	2019/07/16	23:59:59	210.61.46.35	77 MB
BV550190900002	2019/07/17	00:00:00	2019/07/17	23:59:59	210.61.46.35	41 MB
BV550190900002	2019/07/18	00:00:00	2019/07/18	23:59:59	210.61.46.35	25 MB
BV550190900002	2019/07/19	00:00:00	2019/07/19	23:59:59	210.61.46.35	9 MB
BV550190900002	2019/07/20	00:00:00	2019/07/20	23:59:59	210.61.46.35	12 MB
BV550190900002	2019/07/21	00:00:00	2019/07/21	23:59:59	210.61.46.35	16 MB
BV550190900002	2019/07/22	00:00:00	2019/07/22	23:59:59	210.61.46.35	48 MB
BV550190900002	2019/07/23	00:00:00	2019/07/23	23:59:59	210.61.46.35	121 MB
BV550190900002	2019/07/24	00:00:00	2019/07/24	23:59:59	210.61.46.35	55 MB
BV550190900002	2019/07/25	00:00:00	2019/07/25	23:59:59	210.61.46.35	22 MB
BV550190900002	2019/07/26	00:00:00	2019/07/26	23:59:59	210.61.46.35	27 MB
BV550190900002	2019/07/27	00:00:00	2019/07/27	23:59:59	210.61.46.35	9 MB
BV550190900002	2019/07/28	00:00:00	2019/07/28	23:59:59	210.61.46.35	14 MB
BV550190900002	2019/07/29	00:00:00	2019/07/29	23:59:59	210.61.46.35	52 MB
BV550190900002	2019/07/30	00:00:00	2019/07/30	23:59:59	210.61.46.35	40 MB
BV550190900002	2019/07/31	00:00:00	2019/07/31	23:59:59	210.61.46.35	23 MB
總計 (網路計費是以申請的所有服務之對外流量加總後合併計收)						883 MB
中華電信 < 雲端服務 >						

客戶名稱：國立臺灣大學文學院語言學研究所  
 HN 編號：HN55019090

雲端服務編號	起始日期	起始時間	終止日期	終止時間	Public IP	流量
BV550190900002	2019/12/01	00:00:00	2019/12/01	23:59:59	210.61.46.35	19 MB
BV550190900002	2019/12/02	00:00:00	2019/12/02	23:59:59	210.61.46.35	34 MB
BV550190900002	2019/12/03	00:00:00	2019/12/03	23:59:59	210.61.46.35	295 MB
BV550190900002	2019/12/04	00:00:00	2019/12/04	23:59:59	210.61.46.35	23 MB
BV550190900002	2019/12/05	00:00:00	2019/12/05	23:59:59	210.61.46.35	29 MB
總計 (網路計費是以申請的所有服務之對外流量加總後合併計收)						400 MB
中華電信 < 雲端服務 >						

(五) 目前頁面狀況

數位族語資料庫現行狀況及說明如下<sup>9</sup>：

■ 典藏資料檢索

- a. 目前能只設定基本檢索關鍵字
- b. 搜索結果顯示在下方

典藏資料檢索

搜索

主類(華語)	主類(族語)	族語別	方言別	所在部(華語)	所在部(族語)	備註
一、 齋祭品	simyus	太魯閣族	太魯閣語			<a href="#">編輯</a>
七、 漢語與阿太魯閣事件		太魯閣族	太魯閣語			<a href="#">編輯</a>
三、 房屋	Mnaa bəayu	太魯閣族	太魯閣語			<a href="#">編輯</a>
九、 齊樂或唱巴	Maa! Mgangah ka Hing	太魯閣族	太魯閣語			<a href="#">編輯</a>
二、 彩虹	Hakaw Unux	太魯閣族	太魯閣語			<a href="#">編輯</a>
五、 灰燼	Qbala Tahut	太魯閣族	太魯閣語			<a href="#">編輯</a>
八、 傳魯閣孩子的成長歷程	knshilan laq troku	太魯閣族	太魯閣語			<a href="#">編輯</a>
六、 手工藝	mnsusa	太魯閣族	太魯閣語			<a href="#">編輯</a>
十、 恩那阿威和恩那舞立	sknawo ni sknawo	太魯閣族	太魯閣語			<a href="#">編輯</a>
四、 樂成唱吧的詞	Maa Mgangah ka hing	太魯閣族	太魯閣語			<a href="#">編輯</a>

顯示 10 結果 共 12 項

上頁 1 2 下頁

<sup>9</sup> 進一步瞭解請至數位族語資料庫 <http://dailt.ilrdc.tw/>

## ■ 族語新聞檢索

族語新聞檢索

搜索

共傳2語資料	顯示 [10] 結果								
族語別	序號	視頻	標題	族語別	操作				
噶瑪蘭族(2)	1		海祭是噶瑪蘭族最大的祭典	噶瑪蘭族	<a href="#">查看</a>				
	2		年齡階級職務分工 使豐年祭順利進行	噶瑪蘭族	<a href="#">查看</a>				

顯示第 1 至 2 項結果,共 2 項 上頁 1 下頁

a. 播放影片一旁有逐字翻譯族語和中文的功能

### 族語新聞資訊

**年齡階級職務分工 使豐年祭順利進行**

**qangian na pakasa kelawkaway na sabasabasay**



zau lipay nani, ngid iku  
paqesupaR timaimuan tu waway  
niaq.

這個星期我要讓你們了解我們的祭典。

b. 可比對閱讀文字

文章
資訊

**族語文章**

zau lipay nani, ngid iku paqesupaR timaimuan tu waway niaq. paqesupaR tu zau qenabalang zau damu na, zau qatabanan ti, pitasawan kebalan nani, qawman melaut qatabanan qaya. matiwi ti ya zau sabasabasayan niaq ta kikinilan ay damu na, kudis. paqesupaR tu zau lanas na damu, qataban pa ti, zin na semanu timaimian.

nazau qadamuan na zau 'nay pateRungan qataban nani, qazuan na zau masang ay zau zana sabasabasayan zau, zana sabakes ay zau izip na, sining zau mekaway ay tu nazau ay lanas. maqezi ita nazau utani tasaw na zau zana sabasayan, qaniyau pasaya tu zau saqizuanan na meqataban ay, nazau ay kelawkawayan ay seRay salektiawan na sayza, nengi lanas ya sabasayan zaniaq zau, sikining ay zau mekaway. titan ta zau nininguan ku tu zau zana qataban na kebalan.

qataban na kebalan nani, mai sasetaR tu zuma ay Raqit. qinunangan na tu masang ay waway, ya pusia seRay metiyu mazas tu zau Raqit ta zau siangatuan na qataban, temunuR tu baqibai tu qanengian. ta ngayawan na qataban nani, paqesupaR ya zau pusia, temuluz tu zau nengi ay, sabakes ay izip na sining, mekaway ay tu lanas zau, paqesupaR tu zuma ay Raqit ta damuan ay baRqian nani, yau ti ta siwRatanan, melana tu temanay ay zau sabasabasayan.

izan na tu zana masang ay sabasay zau, mekassa tu Rakelawkawayan na. azuan na tasaw na zau. mekassa tu

**中文文章**

這個星期我要讓你們了解我們的祭典，告知附近的部落，豐年祭的時間到了，一年一度的噶瑪蘭豐年祭，請再來參與。我們的青年們去到附近的部落，立德，報訊息，我們即將要舉行豐年祭了。整個新社部落動員舉辦的豐年祭，是依循著過去的年齡階級，由身體強壯的青年負責這個事情。大家根據青年的年齡，由他們來籌備豐年祭的會場，諸如此類的工作以及舞蹈，我們的青年很會做事，所以由青年籌辦，我們來看我紀錄的噶瑪蘭族豐年祭影片吧。

噶瑪蘭族的豐年祭，與其他族群不同，是遵循著過去的習俗，由頭目及祭司帶領族人在豐年祭開始前，向祖先進行祈禱儀式。在豐年祭舉辦之前，頭目會告知、派選身體強壯的青年，負責這件事，報訊息給其他(部落)的族人。而在部落裡的耆老則會在聚會所中，等待青年們回來。

如同過去的年齡階級，各自進行分工，年齡相同的，分派各自的工作，籌備豐年祭的各項事宜，就這樣建置會場，教授我們族語的老師很辛苦，負責教小孩子們，學習歌謠以及舞蹈，孩子們在唱歌跳舞時，看起來非常可愛快樂。

比較年輕的青年，帶領來參加的族人學習歌曲，也讓大家在豐年祭當天一起配合歌唱。他們一起唱歌跳舞，仍然如同過去噶瑪蘭族的豐年祭。

c. 資訊呈現方式

文章 資訊

新聞標題(華語)	年齡階級職務分工 使豐年祭順利進行
新聞標題(族語)	qangian na pakasa kelawkaway na sabasabasay
新聞類別	教育
來源	IPCF-TITV原文會 原視
來源網站	<a href="https://youtu.be/gdWWH7LQdYQ">https://youtu.be/gdWWH7LQdYQ</a>
新聞日期	2017-08-17
最新修改時間	2019-05-21
轉譯者	杜佰克·瑪蘇笠

■ 臺灣原住民族語言研究書目檢索

臺灣原住民族語言研究書目檢索

或
領域
語言學領域 (一般語言學、應用語言學)
-

您搜索 **阿美**，共有 **334** 筆資料

顯示  結果
搜索:

序號	篇名	作者	年份	操作
1	東海岸阿美語的語調重音研究	劉大裕	2002	<a href="#">閱讀</a>
2	近代國家體制下學校教育對台東和平郡阿美族人語言使用變遷之研究	唐瑾信	2007	<a href="#">閱讀</a>
3	中部阿美語的否定詞研究	張裕龍	2007	<a href="#">閱讀</a>
4	阿美語詞綴與疑問詞研究	黃天來	2007	<a href="#">閱讀</a>
5	再論阿美語的u標記	吳宜娟	2009	<a href="#">閱讀</a>
6	花蓮地區客語阿美語接觸研究	廖致苡	2009	<a href="#">閱讀</a>
7	阿美語的聲學元音特性及其英文發音教學上的啟示	林麗誠	2012	<a href="#">閱讀</a>
8	中部阿美語方言詞義比較	林華美	2013	<a href="#">閱讀</a>
9	從不同的詞序探討阿美族幼兒於華語理解之差異研究	宥仕·少多宜	2015	<a href="#">閱讀</a>
10	阿美族語的語言活力	陳麗誠	2016	<a href="#">閱讀</a>

顯示第 1 至 10 項結果,共 334 項

 上頁  2 3 4 5 ... 34 下頁

a. 檢索結果顯示在下方



## ■ 族語新聞檢索

族語新聞檢索

搜尋

共有2篇資料

族語別	選擇編號(2)	序號	視頻	標題	族語別	操作
		1		海馬港噶瑪蘭族大的祭典	噶瑪蘭族	查看
		2		年齡階級職務分工 使豐年祭順利進行	噶瑪蘭族	查看

顯示 10 結果

顯示第 1 至 2 項結果,共 2 項

上頁 1 下頁

### d. 播放影片一旁有逐字翻譯族語和中文的功能

族語新聞資訊

年齡階級職務分工 使豐年祭順利進行

qangian na pakasa kelawkaway na sabasabasay

zau lipay nani, ngid iku paquesupaR timaimuan tu waway niaq.

這個星期我要讓你們了解我們的祭典。

文化小辭典

qangian na pakasa kelawkaway na sabasabasay

0.02 / 3:12

YouTube

### e. 可比對閱讀文字

文章 資訊

族語文章

zau lipay nani, ngid iku paquesupaR timaimuan tu waway niaq. paquesupaR tu zau qenabalang zau damu na, zau qatabanan ti, pitasawan kebalan nani, qawman melaut qatabanan qaya. matiwi ti ya zau sabasabasayan niaq ta kikinilan ay damu na, kudis. paquesupaR tu zau lanas na damu, qataban pa ti, zin na semanu timaimian.

nazau qadamuan na zau 'nay pateRungan qataban nani, qazuan na zau masang ay zau zana sabasabasayan zau, zana sabakes ay zau izip na, sining zau mekaway ay tu nazau ay lanas. maqezi ita nazau utani tasaw na zau zana sabasayan, qaniyau pasaya tu zau saqizuanan na meqataban ay, nazau ay kelawkawayan ay seRay salektiawan na sayza, nengi lanas ya sabasayan zaniaq zau, sikining ay zau mekaway. titan ta zau nininguan ku tu zau zana qataban na kebalan.

qataban na kebalan nani, mai sasetaR tu zuma ay Raqit. qinunangan na tu masang ay waway, ya pusia seRay metiyu mazas tu zau Raqit ta zau siangatuan na qataban, temunuR tu baqibai tu qanengian. ta ngayawan na qataban nani, paquesupaR ya zau pusia, temuluz tu zau nengi ay, sabakes ay izip na sining, mekaway ay tu lanas zau, paquesupaR tu zuma ay Raqit ta damuan ay baRqian nani, yau ti ta siwRatanan, melana tu temanay ay zau sabasabasayan.

izan na tu zana masang ay sabasay zau, mekassa tu Rakelawkawavan na. azuan na tasaw na zau. mekassa tu

中文文章

這個星期我要讓你們了解我們的祭典，告知附近的部落，豐年祭的時間到了，一年一度的噶瑪蘭豐年祭，請再來參與。我們的青年們去到附近的部落，立德，報訊息，我們即將要舉行豐年祭了。整個新社部落動員舉辦的豐年祭，是依循著過去的年齡階級，由身體強壯的青年負責這個事情。大家根據青年的年齡，由他們來籌備豐年祭的會場，藉如此類的工作以及舞蹈，我們的青年很會做事，所以由青年籌辦，我們來看我紀錄的噶瑪蘭族豐年祭影片吧。

噶瑪蘭族的豐年祭，與其他族群不同，是遵循著過去的習俗，由頭目及祭司帶領族人在豐年祭開始前，向祖先進行祈禱儀式。在豐年祭舉辦之前，頭目會告知、派選身體強壯的青年，負責這件事，報訊息給其他(部落)的族人。而在部落裡的耆老則會在聚會所中，等待青年們回來。

如同過去的年齡階級，各自進行分工，年齡相同的，分派各自的工作。籌備豐年祭的各項事宜，就這樣運籌會場。教授我們族語的老師很辛苦，負責教小孩子們，學習歌謠以及舞蹈。孩子們在唱歌跳舞時，看起來非常可愛快樂。

比較年輕的青年，帶領來參加的族人學習歌曲，也讓大家在豐年祭當天一起配合歌唱。他們一起唱歌跳舞，仍然如同過去噶瑪蘭族的豐年祭。

f. 資訊呈現方式

文章 資訊

新聞標題(華語)	年齡階級職務分工 使豐年祭順利進行
新聞標題(族語)	qangian na pakasa kelawkaway na sabasabasay
新聞類別	教育
來源	IPCF-TITV原文會 原視
來源網站	<a href="https://youtu.be/gdWWH7LQdYQ">https://youtu.be/gdWWH7LQdYQ</a>
新聞日期	2017-08-17
最新修改時間	2019-05-21
轉譯者	杜佰克·瑪蘇笠

■ 臺灣原住民族語言研究書目檢索

臺灣原住民族語言研究書目檢索

或

領域

語言學領域 (一般語言學、應用語言學)

-

您搜索 **阿美**，共有 **334** 筆資料

顯示  結果
搜索:

序號	篇名	作者	年份	操作
1	東海岸阿美語的語調重音研究	劉大裕	2002	<a href="#">閱讀</a>
2	近代國家體制下學校教育對台東和平郡阿美族人語言使用變遷之研究	唐瑾信	2007	<a href="#">閱讀</a>
3	中部阿美語的否定詞研究	張裕龍	2007	<a href="#">閱讀</a>
4	阿美語詞綴與疑問詞研究	黃天來	2007	<a href="#">閱讀</a>
5	再論阿美語的u標記	吳宜娟	2009	<a href="#">閱讀</a>
6	花蓮地區客語阿美語接觸研究	廖致苙	2009	<a href="#">閱讀</a>
7	阿美語的聲學元音特性及其英文發音教學上的啟示	林威誠	2012	<a href="#">閱讀</a>
8	中部阿美語方言詞義比較	林華美	2013	<a href="#">閱讀</a>
9	從不同的詞序探討阿美族幼兒於華語理解之差異研究	宥仕·少多宜	2015	<a href="#">閱讀</a>
10	阿美族語的語言活力	陳詠誠	2016	<a href="#">閱讀</a>

顯示第 1 至 10 項結果,共 334 項

上頁 1 2 3 4 5 ... 34 下頁

b. 檢索結果顯示在下方



### 三、語料建置工作說明

由於早期族語口語資料保存多以磁帶式媒材來紀錄，因此很可能已屆年限無法讀取，需盡早轉以數位化方式保存，也便於未來的應用。我們將去（2018）年已收集之一批太魯閣錄音帶語料進行數位化及修復工作，有斷裂或是轉速過快問題，也請影像修復專業人員協助進行修復。

而將錄音帶完成數位化後，我們進行初步的整理及資料詮釋建檔，另外以田調的方式進行訪談、敘事語料的收集，並從這些已收集的口語語料中挑選音質較佳、內容清楚的 5 卷音檔，請專業的族人進行族語文字轉寫、翻譯及確認後，建置於資料庫中，並於第二期（6 月 30 日）及第三期（9 月 30 日）繳交。而轉譯流程如下：

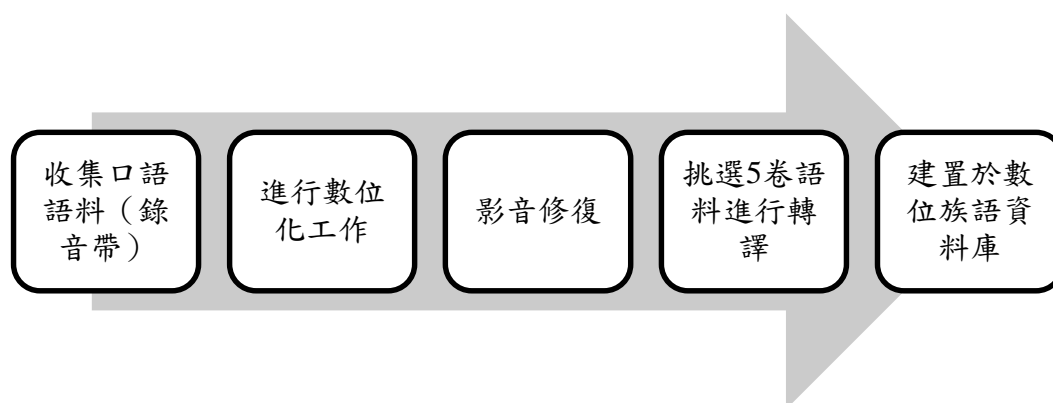


圖 3：太魯閣口語語料轉譯工作流程

除了上述的語料，我們也持續進行族語新聞的轉譯。2017 年我們已和原文會進行研商會議討論，現行族語新聞翻譯模式主要是以已上傳至原視 youtube 網站的族語新聞，挑選出適合的新聞進行文字轉寫，挑選的原則為與該族群文化相關且內容盡量都是族語的新聞。而轉譯人員則是優先邀請主播或是請已有經驗者協助。族語新聞本年度共完成 17 則族語新聞轉譯，亦以於前二期報告中提交，族語新聞的轉譯工作流程如下：

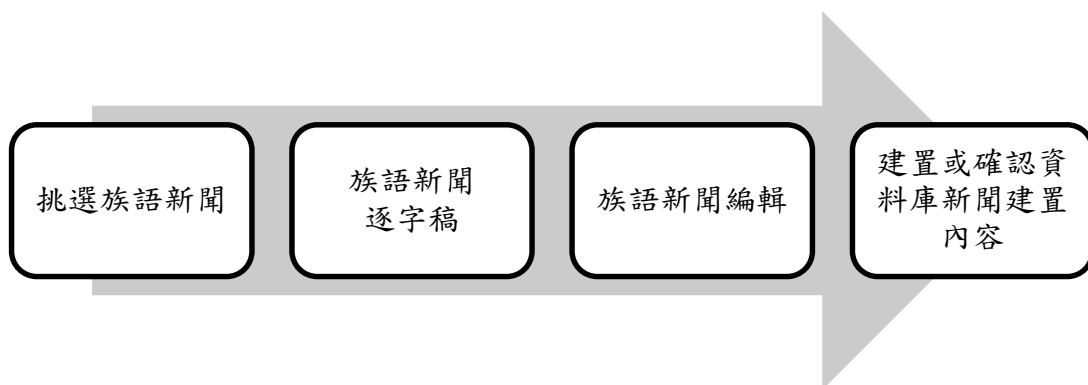


圖 4：族語新聞轉譯工作流程

#### 四、數位族語資料庫系統建置內容

「原住民族語言研究發展中心數位族語資料庫」網址為 <http://dailt.ilrdc.tw>。本年度試辦測試語料已全數建置於網站中，並已於前二期完成提交，詳細建置內容移至附錄三供參閱。目前已建置典藏資料 39 則（書面資料 10 則，其中 2 則有音檔，並重新檢視轉寫及翻譯）、族語新聞 16 族 22 則，研究書目 3,994 筆<sup>10</sup>，口語語料（含新聞）共計 16 小時，已完成轉譯約 4 小時。書面語料及口語語料因考量語料版權、轉譯人力及書寫規範的共識，目前以太魯閣族語料做建置的示範，清單羅列如下：

##### （一） 典藏資料<sup>11</sup>

##### 1. 書面語料（楊盛涂，2000）

序	標題	格式
1	一、獻祭品	文字辨識 pdf
2	二、彩紅	文字辨識 pdf
3	三、狩獵	文字辨識 pdf
4	四、變成啞吧的狗	文字辨識 pdf
5	五、灰燼	文字辨識 pdf
6	六、手工藝	文字辨識 pdf
7	七、漫談民初太魯閣事件	文字辨識 pdf
8	八、德魯固孩子的成長歷程	文字辨識 pdf
9	九、狗變成啞巴	文字辨識 pdf
10	十、思跟阿威和思跟希立	文字辨識 pdf

<sup>10</sup> 此為 2019 年 12 月 3 日之資料，仍在持續更新中，內容另於研究書目報告中呈現。

<sup>11</sup> 典藏資料文字、影音檔皆另附於光碟中。

## 2. 口語語料（有影音檔）

序	類別	標題	音檔時間	轉寫
1	敘事	青蛙的故事-金清山	10:44	
2	敘事	梨子的故事-金清山	06:09	✓
3	敘事	青蛙的故事-楊盛涂	07:27	
4	敘事	梨子的故事-楊盛涂	24:05	✓
5	敘事	祖先起源	12:45	✓
6	敘事	講述人名的命名以及連名後等音	32:19	
7	敘事	書寫、拼音及詞綴使用的問題	22:21	
8	敘事	談論過去的生活經驗	31:25	
9	敘事	講述歷史和對日戰役(1)	47:01	
10	敘事	講述歷史和對日戰役(2)	47:25	
11	敘事	講族人英雄事蹟(1)	33:41	
12	敘事	講族人英雄事蹟(2)	48:14	
13	敘事	調查部落的事情(1)	47:42	
14	敘事	調查部落的事情(2)	47:50	
15	敘事	講述以前部落的生活	1:18:20	
16	敘事	見晴部落歷史以及生活經驗	1:10:57	
17	敘事	田信德牧師、許通益牧師講解詞綴、數字、時間等用法	28:38	
18	敘事	母語的使用、動物說法	19:56	
19	敘事	母語的使用，詞根和詞綴的應用	19:01	
20	對話	日常對話：太魯閣族語發展與復振(1)	03:42	
21	對話	日常對話：太魯閣族語發展與復振(2)	17:04	
22	對話	日常對話：太魯閣族語發展與復振(3)	17:04	
23	對話	日常對話：太魯閣族語發展與復振(4)	17:04	
24	對話	日常對話：太魯閣族語發展與復振(5)	06:50	
25	對話	日常對話：生活經驗與族語現況	20:55	
26	對話	編織煮食農耕弓箭器具	01:02:46	✓
27	對話	五、灰爐（口語轉譯及檢視）	31:15	✓
28	對話	六、手工藝（口語轉譯及檢視）	31:15	✓
29	演講	朗讀稿示範	22:40	✓

### （二） 族語新聞

序	族語新聞標題	族語別	時間
1	阿美婚喪喜慶結束 Pakelang 完工答謝	阿美	04:32

序	族語新聞標題	族語別	時間
2	擔心沒獵物 阿美"Masetik"避免尷尬	阿美	03:44
3	Taeman 祭典落幕女性主導送靈祭	阿美	04:41
4	獵熊需抵命，泰雅獵人視為禁忌。	泰雅	04:16
5	排灣族祭儀祈福，小米梗扮演重要角色。	排灣	04:15
6	講求"文如其分" 排灣族文身手法解祕！	排灣	06:16
7	命名儀式前，Bunun 嬰兒祭祈新生兒平安成長。	布農	06:34
8	布農族主要有 6 個氏族，以玉山為中心建立族群。	布農	04:59
9	利嘉小米收穫祭，卑南族人暫停狩獵行動。	卑南	03:34
10	魯凱花環插飾規範多，各植物代表意義不同。	魯凱	03:28
11	鄒族小米祭儀，不可喧譁、禁止惡言。	鄒	01:36
12	敬重大自然是獵人必有的規範。	鄒	01:31
13	賽夏一年兩祭儀，感謝家族祖靈帶來豐收。	賽夏	04:58
14	雅美族人海水澆船，防止拼板舟裂開。	雅美	02:50
15	印象水沙連 呈現道光年間邵族部落生活。	邵	03:24
16	海祭是噶瑪蘭最大的祭典。	噶瑪蘭	02:16
17	年齡階級職務分工 使豐年祭順利進行。	噶瑪蘭	03:13
18	夢卜被視為祖靈 對未來運勢吉凶徵兆。	太魯閣	04:01
19	部落珍貴食材藤心，具有多種功效。	撒奇萊雅	02:34
20	賽德克族的狩獵袋、夢占及鳥占。	賽德克	04:19
21	搶救瀕危族語，原民會公告首創師徒制學習。	拉阿魯哇	04:07
22	那瑪夏的地名源自卡那卡那富族的傳說故事。	卡那卡那富	04:26

## 五、口語語料及影音修復

在建置族語資料的過程中，藉由訪談、諮詢、辦理會議、參與工作坊等方式，我們彙整口語語料轉譯及在收集、轉譯及修復等需要注意的問題，整理如下：

- (一) 轉譯人才不足：瀕危語、單方言僅少數人有能力轉譯，人力吃緊，故未來應培育相關人才，以利後續族語文字化的工作進行。未來應與書寫規範工作坊共同推廣辦理。
- (二) 系統操作門檻：由於族語能力可轉寫的老師通常年齡較高，對於學習逐字稿系統的操作通常有困難，需要年輕人從旁協助。因此建議除了語推人員外，培力年輕人搭配耆老轉譯，可藉由這樣的工作模式讓年輕人接觸、學習族語。

- (三) 轉譯費用落差：有譯者反應其他單位亦有全族語轉寫及翻譯的工作，提供的價格較高，一般計畫無法負擔。經詢問原文會「'a'iyalacho:開會了」節目字幕聽打及翻譯，即屬較高預算規格支給，費用大概會有至少 4 倍的落差。
- (四) 語料亟需積極搶救：澳洲國家電影及聲音檔案館（NFSA）認為，在 2025 年這個不久的將來，未被數位化的磁帶可能會面臨永久消失的問題，需積極搶救，因此，目前 NFSA 也積極收集及保護磁帶收藏品，並在網站<sup>12</sup>發起徵集的活動。未來台灣也應該積極進行大規模磁帶影音資料的數位化工作。
- (五) 影音數位化及修復技術：目前南藝大影音修復所正努力推動影像修復工作，未來原住民族語資料庫也應共同合作。相關技術人才亦需要培育。而關於錄音帶修復技術的步驟，則另於附錄四呈現。
- (六) 語料版權及研究倫理：早期收集的語料礙於現行法規限制，往往可能會觸及著作權或是否獲得受訪者知情同意等問題。而未來採及的語料亦應進一步規範，以妥善取得授權，以利語料的多元應用。目前語發中心已草擬資料提供授權書（如下圖），可供參考。

---

<sup>12</sup> 活動網頁 <https://www.nfsa.gov.au/corporate-information/publications/deadline-2025>

原住民族語言研究發展中心數位族語資料庫

資料提供授權書

本人同意授權原住民族語言研究發展中心（以下稱語發中心）保存及數位化本人提供的族語資料，授權範圍及權利關係如下：

1. 提供之資料為本人持有，並為記錄者或參與創作，或因為繼承關係被賦予決定此資料使用及授權之權利。
2. 提供之資料已獲得參與者的知情同意。
3. 授權語發中心將資料置於數位族語資料庫
  - 全部公開
  - 僅供語發中心研究使用，不直接公開資料
  - 有條件公開
4. 若同意公開，則同意語發中心和他人可以運用這些資料進行研究、教學或其他合理、合法且不為商業目的之使用。
5. 授權後，本人依然保留資料之相關著作權利，基於授權範圍，語發中心有變更使用方式之權利。
6. 若提供實體資料，不幸遭毀損、丟失或無法修復，本人放棄向語發中心追訴相關責任，但語發中心需盡保管之責任。
7. 若提供實體資料，經語發中心數位化後，實體資料檔案處理方式為
  - 交由語發中心保存
  - 交還給著作權人

\*未來語發中心資料可能轉移至原住民族語言發展基金會，相關授權亦一併移交，不另再簽署授權同意書。

資料清單如下：

本人\_\_\_\_\_已清楚及了解上述內容及授權範圍，並按上述約定授權語發中心使用上列資料清單之材料。

日期：\_\_\_\_\_年\_\_\_\_\_月\_\_\_\_\_日

收件人員：

圖 5：語發中心數位族語資料庫—資料提供授權書

## 肆、數位族語資料庫階段性任務分析

我們以 2017 年「數位族語資料庫系統的第一階段的規劃」為基礎，並彙整 2018 年及 2019 年的研商修正會議意見，提出未來資料庫的藍圖，並修正階段性建置任務及目標。以下依階段性建置任務及目標分析，再提出未來發展藍圖，以瞭解資料庫建置的現況與未來發展的方向。

### 一、階段性建置任務及目標分析

按原住民族語言振興第 2 期六年計畫的規劃，語發中心預期任務是到 2019 年，而未來財團法人原住民族語言研究發展基金會（下稱原語發基金會）成立後，相關計畫應移交由該會繼續執行。因此我們將階段性任務羅列如下圖：



圖 6：數位族語資料庫階段性建置任務及目標

如上圖所呈現，2019 年主要是繼續資料庫系統功能的修正，並建置不同形式之族語語料，並將問題彙整一併提出修正建議。而 2020 年後，則應整合相關資源，規劃大型整合建置計畫及研擬語料庫系統架構。

## 二、資料庫建置與未來規劃

瞭解前述階段性任務後，我們必須清楚資料庫現階段已完成之內容及未來發展的規劃藍圖，經整理後，以下圖呈現：



圖 7：數位族語資料庫建置及未來發展藍圖

根據上圖所示，2017 至 2019 年我們完成資料庫的規劃建置，並開始進行語料的數位化、轉譯及建置於資料庫中，另外也收集台灣原住民族的研究書目。在族語資料收錄時，亦需考量智慧財產授權及研究倫理問題。最後依使用者操作意見和專家建議，提出未來系統修正的建議。

依未來的發展而言，首要任務應評估如何成為整合性平臺，擴增系統功能、提升資料庫的位階，整合收集跨單位、平臺的族語資料。另外，亦需結合地方人力資源的工作，系統性的收集語料，並應以原住民族知識系統分類為基礎來建置語料，才能更全面的保存語言的樣貌。盧谷砾樂、吳



美美（2015）亦整理探討臺灣原住民族知識分布，其中提到澳洲原住民發展協會 Institute for Aboriginal Development (IAD)於 2000 年繪製出澳洲原住民族夢時代圖（The dreamtime chart）（轉引自盧谷碩樂、吳美美，2015:141），將原住民族知識體系分為人類、物質及神聖世界，再接軌到如天空、土地、儀式、法律、行為規範等面向（圖 8），表達了原住民族的知識觀。另外，汪明輝（2009：12-14）則以鄒族知識體系為例，提出知識體系之架構，大致上包含生活空間、時間、神話、歷史、政治司法、宗教信仰、社會、文化及經濟等面向（圖 9）。

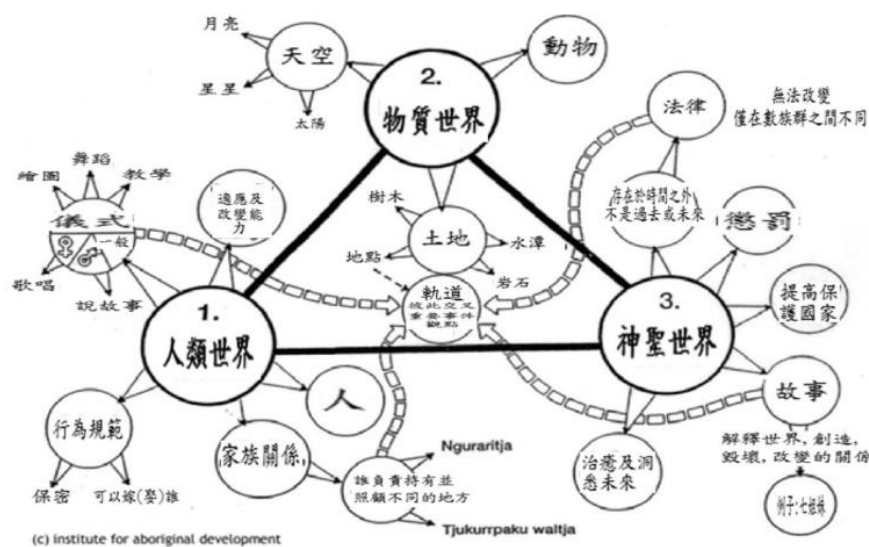


圖 8：澳洲原住民族夢時代圖（The dreamtime chart）

資料來源：“The Dreamtime Chart,” by Institute for Aboriginal Development, 2000. Retrieved from <http://aboriginalart.com.au/culture/dreamtime3.html> 轉引自盧谷碩樂、吳美美(2015:141)

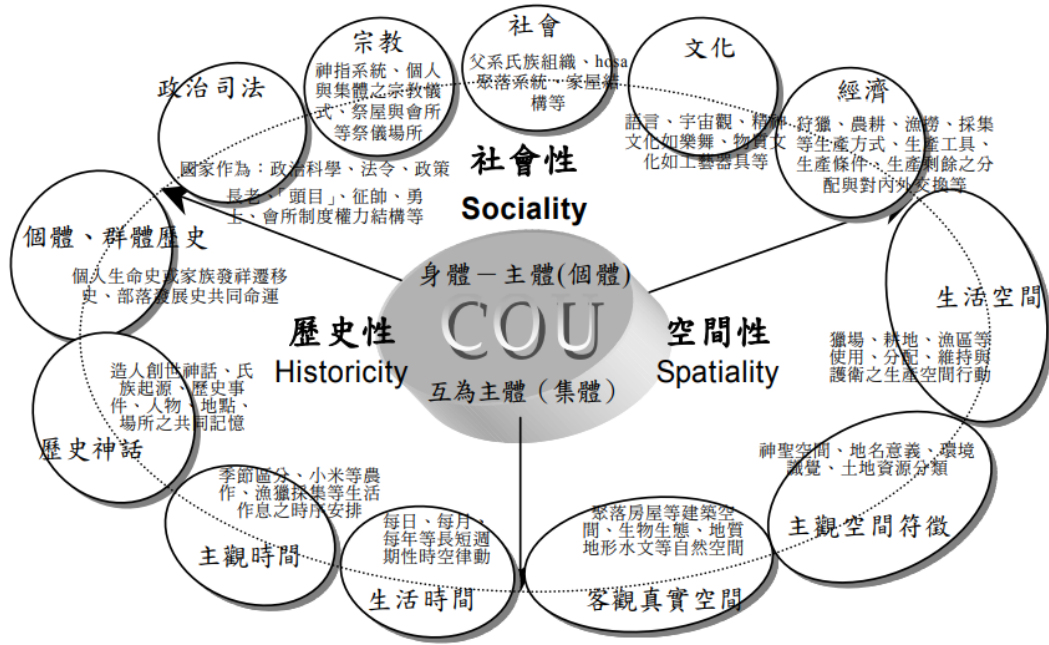


圖 9：鄒族存有及知識本體架構

若我們將上述範疇為基礎，在語料收集及建置時有清楚的目標，配合語推人員語料調查及轉寫的工作，或是各族語推組織對於族群內耆老及其傳統技能的盤點，都能找到擅長某個領域知識的人，記錄他的知識，這個部分是最急迫的，因為調查保存的工作，往往趕不上耆老的離開速度。除此之外，當代生活的自然語言也是需要收集的，例如在部落商店買東西、家門口的聚會等，往往能收集到不一樣的語料，像是委婉語、禁忌語、玩笑等，這些可能是一般學術調查中比較缺乏的。

而除了「收集」口語的語料，更需要將語料轉寫成文字，以發展語料庫，語料才能進一步的應用。因此完成各族的書寫規範及文字修訂，並且推廣族語文字，培育更多人能參與族語文字轉寫的工作，我們的族語資料就能更豐富，未來可以進行語用查詢、教材編輯、測驗試題，也能研究語言的改變，甚至是發展 AI 人工智慧的功能，如 google 即時翻譯功能，可提升族語的能見度及學習的便利性。未來族語資料庫要面臨的挑戰仍有很多，要仰賴後繼執行者從更高的角度整體的檢視，整合更多資源和力量一起前行。

## 伍、結論與建議

提供語言資料的保存，就是「建置數位族語資料庫」最核心的目的。本年度計畫重點項目為：一、將 2018 年收集的太魯閣語料錄音帶修復及數位化，並挑選 5 卷口語語料轉寫文字及翻譯；二、完成 17 則族語新聞逐字稿及文章編輯；三、資料庫系統維運及調整；四、辦理研商會議共 18 場。以上工作項目及內容皆已完成，並提出資料庫建置內容及階段性任務分析，報告中典藏資料之音檔、文件檔、網站程式另燒錄光碟備查。

目前數位族語資料庫的口語語料（含新聞）共計 16 小時，已完成轉譯約 4 小時，可知在語料收集、建置及轉寫翻譯等工作上，仍有很多需要努力的空間。未來若可結合其他計畫如語推人員、語推組織或與其他機關單位合作，才能更快速累積語料，並發展語料庫，更有利於族語的研究與語料的應用。

對於未來數位族語資料的工作發展與延續，提出一些建議：

### 一、應整合資料庫平臺，以利族語資料彙整

1. 擴大盤典及整合語料資源，如教育部（朗讀稿、族語教材、族語文學創作等）。
2. 未來各單位補助計畫或研究成果，事關族語語料部分，應於事前取得成果應用於資料庫之授權。

### 二、應以各族知識體系進行語料收集

1. 跨族討論主要的核心知識體系分類，有計畫的進行語料的收集。
2. 應結合其他計畫人員，如語推人員、語推組織或是原住民大專生獎勵計畫等，鼓勵知識語料的收集及文字化。

### 三、跨單位及跨計畫合作

1. 若資源人力尚不足，應考慮與其他專業單位合作，如國家檔案管理局（語料數位化修復及典藏）、國教院華語編譯中心（語料庫建置）、文化部（建置國家語言資料庫）、原文會（族語節目及新聞）或其他學術單位。
2. 未來政府補助之經費產出之族語語料，應需考量可納入資料庫中使用。
3. 由於收集語料及轉譯皆需大量的人力，應盡可能與現有計畫人力合作，以達事半功倍之效。

#### 四、發展語料庫

1. 可擇一族進行單語（或雙語）語料庫的建置，再評估若要建置各族語料庫所需的經費、人力、時程、需要語料的數量，以及若要跨機關單位合作，可依循之合作方式等。
2. 另一方面，也需徵詢語言學者，以評估進階、經語料分析的語料庫建置，其所需經費、人力及時程等，如客委會現行建置之客語語料庫，亦可借鑑中研院語言過去建置之臺灣南島語數位典藏資料庫。

#### 五、辦理語料轉寫工作坊，鼓勵民眾參與

1. 可結合書寫規範辦理語料轉寫工作坊，主要說明轉寫的工作、資料庫系統的運用。
2. 鼓勵族人於線上提供族語資料或協助轉寫。

數位族語資料庫的建置，往往會面臨很多的問題，收集資料會有授權、知情同意的問題，而重新田調訪談或是轉寫成文字，更是需要大量人力資源的投入。因此，整合資料庫平臺並以知識體系有計畫、有系統性的收集語料，進行跨單位的合作，發展語料庫及推廣使用，都是下一階段努力的方向，期待族語資料庫的發展走得更深、更遠。

## 陸、參考文獻及網站

### 一、參考文獻

Crystal, David. 1991. *Dictionary of Linguistics and Phonetics*. 3rd edition.

Cambridge, MA: Basil Blackwell.

Zeitoun, Elizabeth, Yu Ching-hua and Weng Cui-xia. 2003. The Formosan

Language Archive: development of a multimedia tool to salvage the

languages and oral traditions of the indigenous tribes of Taiwan. *Oceanic*

*Linguistics*, 42.1:218-232.

大衛·克里斯托 (David Crystal) 著，宋偉航 (譯)。2001。《語言的死亡》(Language Death)。臺北市：貓頭鷹出版。

朱清義。2013。〈阿美語書寫系統：o 及 u 的爭議〉。《原教界》。期 49，頁 62-64。

何德華、董瑪女。2016。《雅美 (達悟) 語語法概論》。新北市：原住民族委員會。

克里斯托 (David Crystal) 著，蔡淑菁、謝儀霏 (譯)。2010。《語言的秘密》(How Language works)。臺北市：貓頭鷹出版。

盧谷碩樂、吳美美。2015。〈從臺灣原住民族出版相關文獻探討臺灣原住民族知識分布〉。《圖書資訊學刊》。第 13 卷第 2 期。頁 135-160。

汪明輝，2009。〈台灣原住民族知識論之建構—以鄒族傳統領域資源管理知識為例〉發表於台灣原住民族教授學會、東華大學原住民族學院主辦「第一屆原住民族知識體系研討會」。壽豐，東華大學原住民族學院國際會議廳，5 月 15-16 日。

原住民族委員會。2013。「原住民族語言振興第 2 期六年計畫」(103-108 年)。網址：<https://goo.gl/qs6LPe> (2019 年 12 月瀏覽)。

原住民族委員會。2018。「107 年度原住民族瀕危語言復振補助計畫」。網址：<https://goo.gl/orrWbS> (2019 年 12 月瀏覽)。

原住民族語言研究發展中心。2017。「原住民族語言研究發展中心數位族語資料庫建置評析報告」。網址：<https://goo.gl/MiVhr5> (2019年12月瀏覽)。

原住民族語言研究發展中心。2018。「數位族語資料庫第一期建置內容及規劃修正報告」。網址：<https://goo.gl/MiVhr5> (2019年12月瀏覽)。

陳浩然、潘依婷。2017。〈語料庫與華語教學〉。收入於陳浩然編《語料庫與華語教學》。臺北：高等教育。

程婉如。2008。拓展臺灣數位典藏計畫：淺談光學字型辨識技術(OCR, Optical Character Recognition)。網址：  
<http://content.teldap.tw/index/?p=529> (2019年12月瀏覽)。

齊莉莎、余清華。2004。〈臺灣南島語數位典藏：詞類與標記原則〉。  
<http://linganchor.sinica.edu.tw/data/file/LC040907LC10.pdf> (2019年12月瀏覽)。

齊莉莎、余清華。2007。〈臺灣南島語數位典藏〉。《原住民族語言發展—理論與實務研討會》，臺北。

齊莉莎。2013。〈臺灣南島語典藏〉。發表於《兩岸瀕危語言典藏及語言振興訓練國際研討會》，臺中。

## 二、參考網站

AILLA (The Archive of the Indigenous Languages of Latin America)，拉丁美洲原住民族語言檔案館，網址：<https://ailla.utexas.org/>

Alilin 原住民族電子書城，網址：<http://alilin.apc.gov.tw/tw/>

ASEDA (The Aboriginal Studies Electronic Data Archive)，澳洲原住民研究電子數據檔案庫，網址 <http://aseda.aiatsis.gov.au>

Corpus of Contemporary American English，美國當代英語語料庫，網址：  
<https://corpus.byu.edu/>

Daly Languages (AUSTRALIA)，澳洲 Daly 語語料庫，網址：  
<http://www.dalylanguages.org/>

Glottolog，世界語言參考文獻網，網址：

<http://glottolog.org/langdoc/langdocinformation>

Kevin Scannell on Corpus-based language technology for indigenous and

minority languages，網址：

<https://www.youtube.com/watch?v=W05YmEl093o>

LDC (Linguistic data Consortium, University of Pennsylvania)，美國賓州大

學語言資料庫，網址：<https://www ldc.upenn.edu/>

OLAC: Open Language Archives Community，開放語言檔案社群網，網址：

<http://www.language-archives.org/>

PARADISEC (Pacific and Regional Archive for Digital Sources in Endangered

Cultures)，太平洋地區檔案數位資源瀕危文化庫，網址：

<http://www.paradisec.org.au/>

UNESCO Interactive Atlas of the World's Languages in Danger 聯合國教科文

組織世界瀕危語言地圖，<http://www.unesco.org/languages->

[atlas/index.php](http://www.unesco.org/languages-atlas/index.php)

中央研究院語言學研究所「語言典藏計畫」之臺灣南島語數位典藏（目前

已停用），網址：

<http://languagearchives.sinica.edu.tw/cht/index.php.html>。

臺大臺灣南島語多媒體語料庫，網址：

[http://corpus.linguistics.ntu.edu.tw/index\\_zh.php](http://corpus.linguistics.ntu.edu.tw/index_zh.php)。

臺灣原住民族數位典藏資料庫，網址：

<http://dore.tacp.gov.tw/tacp/pingpu/index.html>

臺灣聖經公會聖經網站，網址：<http://cb.fhl.net/>

臺灣語言學學會（語典及語料庫），網址：<http://linguist.tw/zh->

[tw/links/%E7%9B%B8%E9%97%9C%E7%B6%B2%E8%B7%AF%E8%B3%87%E6%BA%90%E9%80%A3%E7%B5%90/dictionaries%20and%200corpus](http://linguist.tw/zh-tw/links/%E7%9B%B8%E9%97%9C%E7%B6%B2%E8%B7%AF%E8%B3%87%E6%BA%90%E9%80%A3%E7%B5%90/dictionaries%20and%200corpus)

花蓮縣國小本土語文領域輔導團（太魯閣字典），網址：

<http://teacher.hlc.edu.tw/imax9.asp?id=81&sid=259>

阿美族、雅美族口語傳說採錄翻譯資料數位典藏計畫，網址：

<http://aya.ioe.sinica.edu.tw/>

阿美語萌典，網址：<https://amis.moedict.tw/>

阿美族兒童文學，網址：<http://www.amis.org.tw/2018/>

政大原民中心電子書城，網址：<http://ebook.alcd.tw/>

原住民族委員會臺灣原住民族圖書資訊中心，網址：<https://tiprc.apc.gov.tw/>

原住民族語言線上詞典（原住民族委員會），網址：

<https://e-dictionary.apc.gov.tw/>

國民記憶庫—臺灣故事島，網址：<https://storytaiwan.tw/default.html>

國家圖書館臺灣書目整合查詢系統(SMRT)，網址：

<http://metadata.ncl.edu.tw/blstkmc/blskm#tudorkmtp>

教育部朗聲四起（原住民族語朗讀文章），網址：[https://alcd-](https://alcd-con.tw/alr/lobby.php)

[con.tw/alr/lobby.php](https://alcd-con.tw/alr/lobby.php)

族語 E 樂園（原住民族委員會），網址：<http://web.klokah.tw/>

創用 CC 授權條款，網址：<http://creativecommons.tw/license>

達悟語線上學習網，網址：

<http://yamiproject.cs.pu.edu.tw/elearn/?&lang=zhtw>

數位典藏國家型計畫—臺灣社會人文電子影音數位博物館計畫，網址：

<http://proj1.sinica.edu.tw/~video/index.html>

蘭嶼媒體與文化數位典藏，網址：<http://lanyu.nctu.edu.tw/database/>

蘭嶼達悟語口語資料典藏網，網址：

[http://yamiproject.cs.pu.edu.tw/yami/yami\\_ch/ch\\_index\\_flash.htm](http://yamiproject.cs.pu.edu.tw/yami/yami_ch/ch_index_flash.htm)

蘭嶼達悟語線上學習網，網址：

<http://yamiproject.cs.pu.edu.tw/elearn/search.php>

NFSA（NATION FILM AND SOUND ARCHIVE OF AUSTRALIA），澳洲國



家電影及聲音檔案館，網址：<https://www.nfsa.gov.au/>

## 附錄一、計畫相關會議簽到統計及會議記錄

2019 年辦理之數位族語資料庫建置、轉譯工作、修正研商會議，相關資訊未於網路版本報告公開。

## 附錄二、諮詢訪談及重要會議相關

諮詢、訪談及重要會議清單，相關資訊未於網路版本報告公開。

### 附錄三、2019 年建置之內容<sup>13</sup>

今（2019）年完成 5 卷太魯閣族語料及族語新聞轉譯 17 則，另賡續擴充語料則，皆已建置於資料庫中<sup>14</sup>，目前權限設定為會員可查詢但尚未開放直接下載，相關資訊未於網路版本報告公開。

---

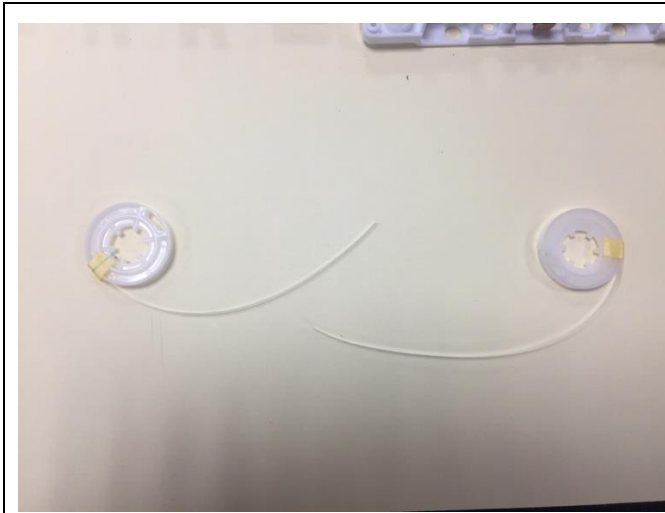
<sup>13</sup> 整理自 2019 年語發中心數位族語資料庫提交之第二期及第三期報告。

<sup>14</sup> 數位族語資料庫網址：<http://dailt.ilrdc.tw/>

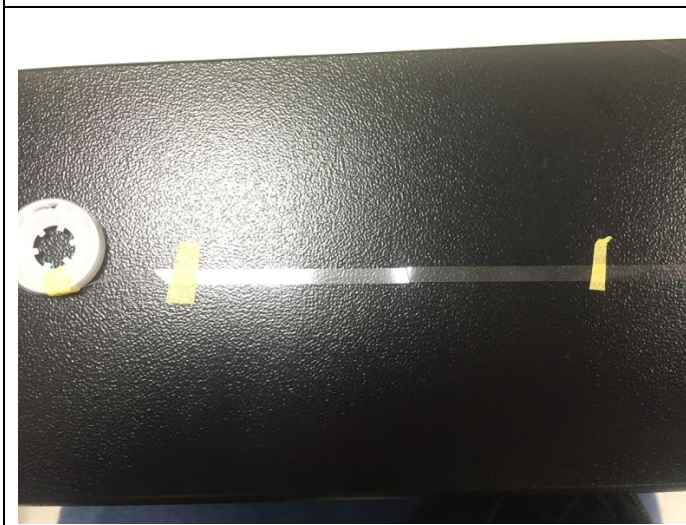
## 附錄四、錄音帶修復建議

臺南藝術大學 Uki 潘昱帆提供本中心錄音帶斷帶修復圖解，亦供未來錄音帶修復工作參考使用：

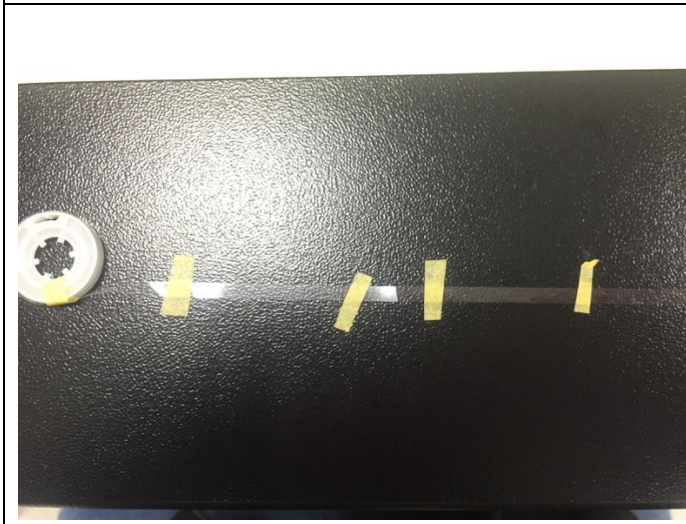
圖片	說明
	<p>先將錄音帶使用螺絲起子轉開，將錄音帶塑膠掀蓋打開。圖下黑色塑膠為內部保護蓋，多為塑膠材質，有些會是金屬材質。</p>
	<p>將磁帶捲頭取出。(錄音帶磁帶的材質膠軟，取出時須小心)</p>



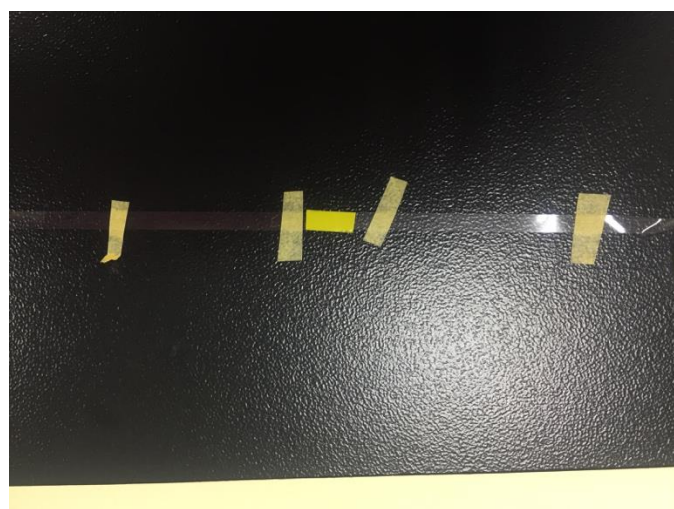
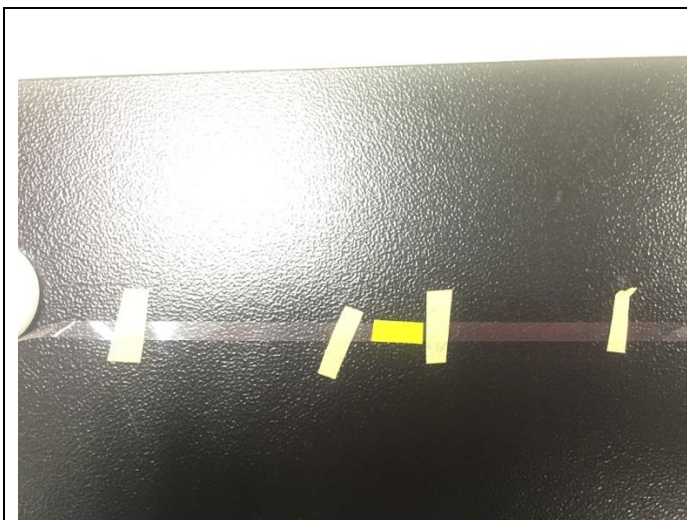
將斷帶取出，因錄音磁帶質地較軟容易卷饒成一團，所以用黃色膠帶固定兩軸。



用黃色膠帶固定兩端後，將斷帶的部分對齊。



接近斷裂處再貼膠帶固定，使其較好對齊。



亮黃色膠帶貼處以固定對其好的斷裂處，將斷裂兩端年處固定。**注意：膠帶黏貼處須在捲軸內側面。**





兩塑膠捲軸的固定膠帶先不用拿起來，將其固定，將黏接起的錄音帶捲軸放為塑膠殼內。



固定後將固定塑膠捲軸的黃色膠帶撕起，將塑膠保護片放回蓋上。



1. 最後將外塑膠殼蓋回。
2. 將外殼的螺絲轉回固定。
3. 錄音磁帶要海保護海綿之上。
4. 即可播放。



