

【六年研究成果分享】族語研究與保存

Pisoped to sowal ato rayray no ikor

原住民族語言研究發展中心，自 103 年 6 月 17 日來，六年中已累積了非常豐碩的成果，整體研究以基本的五大核心為基礎，包含新詞與讀本研究、認證測驗研究、族語教學研究、族語政策研究與推廣、族語研究與保存。一共完成 36 項計畫、69 項次的研究成果。

在族語研究與保存的這個項目中，包含書寫系統修訂與書寫規範、數位族語資料庫以及族語人才資料庫和族語研究目錄的盤點等，大致介紹如下：

一、書寫系統修訂與書寫規範

自 104 年起，語發中心開始參與書寫系統修訂的工作，主要的歷程如下圖：



圖 1：104 年至 108 年語發中心書寫系統修訂與書寫規範研究計畫歷程

根據上圖可以看到書寫的相關研究在 104 年、105 年主要是在檢視書寫系統的問題，瞭解自 94 年書寫系統頒布來，使用的問題，並加以彙整。106 年則以 105 彙整的資料為基礎，辦理各族的書寫系統修訂公聽會，除了讓更多族人參與討論外，更希望能夠促進族人的共識。107 年則是因原住民族語言發展法於前年通過，為符合第二條第二項關於文字的修訂需與各族會商，故與各語推組織進行共識會議，以瞭解各族在書寫上共識的情形，到目前為止僅剩阿美族、魯凱族及雅美(達悟)族尚未完成共識確認。108 年則是與 Truku 的語推組織合作，進行書寫規範的研擬，提出 Truku 語書寫規範參考指南，供族人參考使用並進行滾動式修正，而整個建置的架構及流程，未來亦可推廣到各族來建置書寫規範。

而書寫規範的架構與建置流程如下：



圖 2：108 年 Truku 書寫規範計畫架構與建置流程

根據上圖可知，一開始的書寫規範架構，是先參考其他南島語系國家的規範，其中主要以毛利語的拼寫法為參考基礎，先初步調整為較適合台灣南島語的架構。之後我們邀請 Truku 族人協助以 Truku 語填充，再邀學者一起先檢視初稿。另一方面，也向語推組織說明，並尋求合作的方式，由語推組織進行地區、方言或是對族語書寫熟稔及對語言的掌握等多方考量後，來建立約 15 人的討論小組，其中有 7 位左右為核心成員，基本上每次都需出席。而整體討論的方式是由少至多，先由核心成員達成基本共識再拓展，最後辦理工作坊讓更多族人使用和給與建議。這樣的方式，討論上能夠更具體且聚焦，而有學者的加入，才能夠保持書寫規範的原則性。而目前 Truku 已將試行版公告在語推組織的官網，也推動希望讓更多族人瞭解，未來這樣由下而上並合作產出的規範，如何由政府協助推動，並實際應用在如教材教具、測驗試題中等，是需要努力的方向。

最後，對於未來書寫工作的延續，有幾項重點建議：

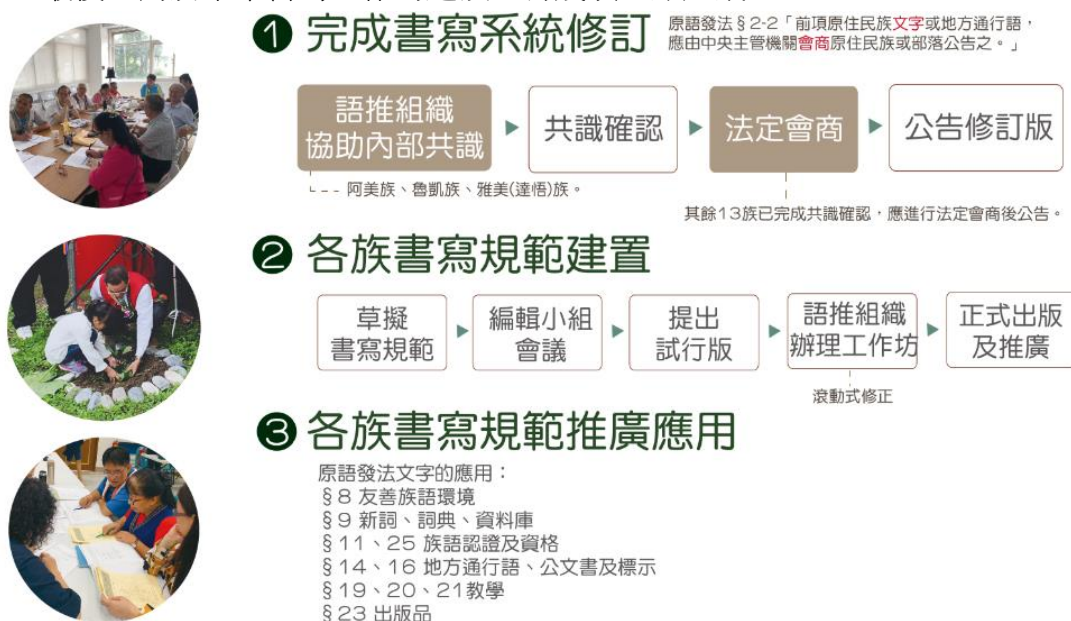


圖 3：書寫系統修訂與書寫規範未來工作的延續與建議

承上圖，書寫系統是族語文字的基礎，因此因早日完成修訂的工作。而書寫規範則有助於族語文字書寫具有原則性，規範完成且線上詞典、學習詞表、教材等皆配合修訂，對於後續文字的使用與推廣更有幫助。

二、數位族語資料庫

數位族語資料庫一直以來的核心概念就是「先保存、後應用」，建置的目的及歷程如下：



主要目的

先保存、後應用

研擬語料收集平臺，
希望以數位化的方式保存語料，
且族語資料能再被應用。



計畫歷程

- 106 規劃數位族語資料庫。
- 107 建置數位族語料料庫系統。
- 108 收集語料、數位化及修復語料、族語新聞及口語語料轉寫、階段性任務分析。



圖 4：數位族語資料庫的目的與計畫歷程

106 年的時候我們參考國內外的族語相關資料庫和網站，規劃數位族語資料庫。而 107 年即開始建置，並開發線上輸入逐字稿及編輯文稿的系統。108 年則將收集的語料進行數位化，建置於資料庫中，並挑選其中 5 卷為示範進行文字的轉寫與轉譯。除此之外，亦進行各族共 17 則族語新聞的轉譯，以及資料庫網頁的維運。

數位族語資料庫目前的網址：<http://dailt.ilrdc.tw/>，示意圖如下（左列表為後台功能列，另右上有語言別切換，首頁可切換 16 族語言）：

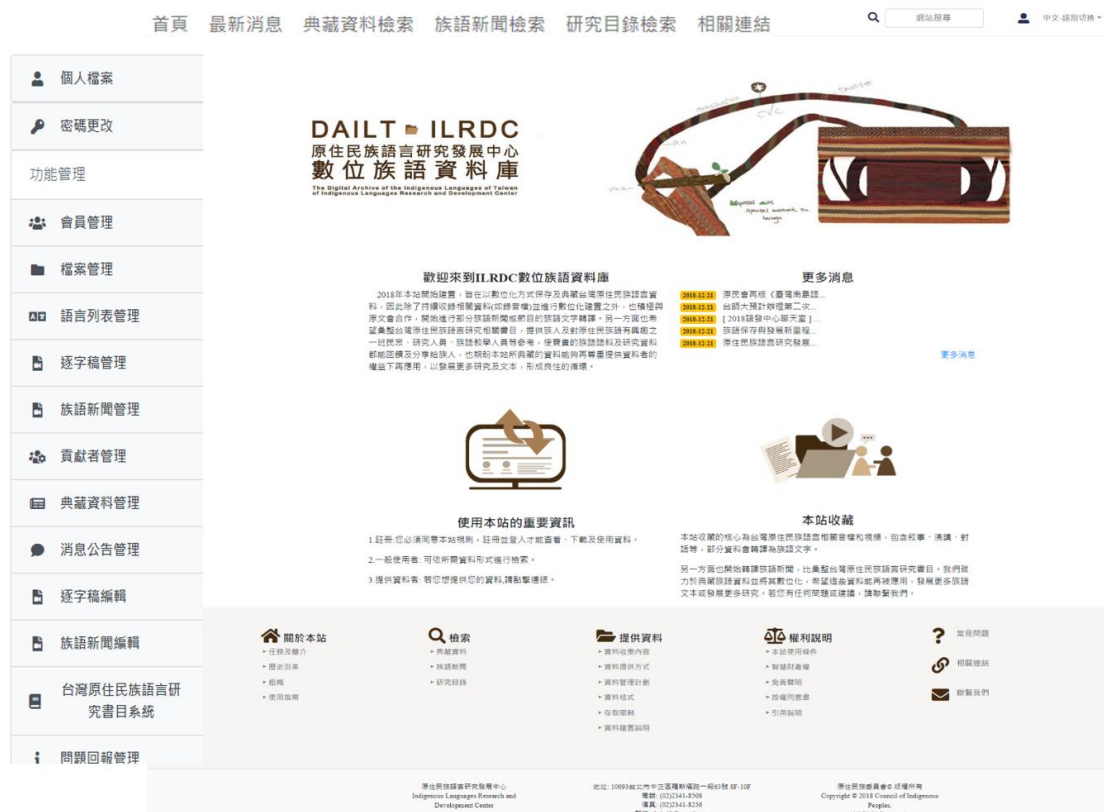


圖 5：數位族語資料庫網站示意圖

而數位族語資料庫的整個工作大致上包含什麼呢？我們整理如下圖：



圖 6：數位族語資料庫工作概述

也就是說，我們將記錄或收集族語語料，部分需要數位化或修復處理，並進行詮釋資料的建置，通常會需要族人的協助，完成後會儲存在資料庫中典藏。而部分語料會請族人轉譯，以利於未來若要開發「語料庫」或是「發展教材」來使用。目前資料庫典藏資料有 39 則、族語新聞 22 則，口語語料共約 16 小時，已轉寫成文字者約 4 小時。

對於未來數位族語資料庫的工作延續建議如下：

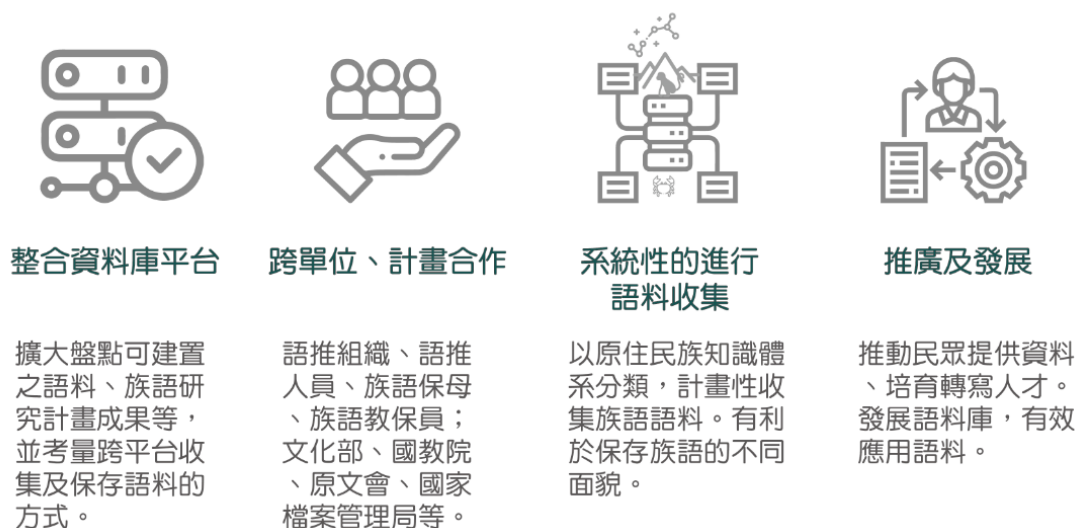


圖 7：數位族語資料庫未來工作的延續與建議

未來應該要整合資料庫存放的平台，並跨單位及計畫合作，才能將計畫的研究成果，有效的整合並能再應用。另一方面，也需要積極的進行系統性的語料收集，如結合語推組織的人力盤點，可以請具有各方專長的族人講述相關故事，

可以記錄下語言的不同面貌，更全面的「保存」我們的語言。最後，更應該積極推廣，鼓勵民眾提供早期的影音記錄，亦可推廣使用族語資料庫系統，也希望更多有能力的族人（尤其是年輕人）可以參與文字的轉寫，語料越多，未來訓練機器也更容易，能夠發展更多 AI 智慧的產品。

三、族語人才資料庫、研究目錄

1. 族語人才資料庫

族語人才資料庫目前共計有 4,180 筆，主要是提供各機關學校在查詢所需的族語人才使用，可依所在縣市、人才類別、關鍵字、所需族語及方言別等條件搜尋，主要類別包含教學類、翻譯類、創詞類、配音類、口譯類、族語推廣人員、族語專職教師、族語振興人員研習，另外測驗類及試題建置與命題的部分則為限制性查詢。未來應可開放部分資訊讓一般民眾、私人機構在需要族語人才時可供查詢。

The screenshot displays the web interface of the Indigenous Language Research and Development Center's Indigenous Language Talent Database. At the top, the center's name is written in both Chinese and English. Below this, the database title '原住民族語言人才資料庫' is shown with a count of 4,180 entries. The interface includes several dropdown menus for filtering: '族語別' (Language), '方言別' (Dialect), '人才類別' (Talent Category), '縣市' (City/County), and '鄉鎮市區' (Township/City Area). A search button labeled '關鍵字搜尋' (Keyword Search) is also present. Below the search filters, there is a section titled '資料庫使用說明' (Database Usage Instructions) which provides detailed information about the database's purpose and the types of talent profiles included.

圖 8：語發中心原住民族語人才資料庫示意圖

2. 臺灣原住民族語言研究目錄

研究目錄的部分，目前是建置於數位族語資料庫中的「[臺灣原住民族語言研究目錄](#)」，106 年至 107 年彙整過去已整理過之書目至線上平台，書目筆數約 3,994 筆。108 年書目文獻共收集到 1,075 筆文獻，包含學位論文、期刊論文、論文集論文、專書論文、會議論文集、專書、技術報告、線上語料庫、雜誌等分類。而經語發中心研究員戴佳豪及陳振助的整理，1906 年至 2019 年書目文獻中，學術期刊 145 筆、學位論文 130 筆、會議論文 269 筆、專書 (ISBN) 327 筆、專書 (無 ISBN) 10 筆、專書論文 (ISBN) 70 筆、專書論文 (無 ISBN) 25 筆、論文集論文 (ISBN) 13 筆、論文集論文 (無 ISBN) 15 筆、技術報告 23 筆、雜誌期刊 48 筆。截至目前為止共約 5,069 筆書目文獻，而這些書目資料仍持續滾動修訂，數據會依修訂後結果有所增減。

文獻類別／年份	1906-1986	1987-1996	1997-2018	2019	總計
學術期刊	28	24	76	17	145
學位論文	12	25	70	23	130
會議論文			222	47	269
專書(ISBN)	2	4	55	266	327
專書(無ISBN)	4	4	2		10
專書論文(ISBN)	3	3	58	6	70
專書論文(無ISBN)		20	5		25
論文集論文(ISBN)	1	1	11		13
論文集論文(無ISBN)		4	10	1	15
技術報告				23	23
雜誌期刊				48	48
總計	50	85	509	431	1,075

圖 9：語發中心 108 年新增臺灣原住民族語言研究目錄文獻統計

未來仍會陸續收錄，也希望若能取得授權，可直接保存或是連結文件資料。而每年也可依增加的資料進行族語研究方向的分析與評估，有利於族語的研究發展。另外也建議在《原住民族委員會原住民專門人才獎勵要點》之獎勵項目增加「族語研究專門著作」及「族語著作」列為重要獎勵項目，以符合現行相關法令推動族語發展，促進族人對族語之態度能透過獎勵的方式有正向之支持。

臺灣原住民族語言研究目錄檢索

圖 10：語發中心臺灣原住民族語言研究目錄檢索示意圖

以上是「族語研究與保存」項目六年成果的精華介紹，若需要更進一步瞭解歷年的研究成果，請至語發中心歷年成果參閱相關報告

<http://ilrdc.tw/data/out/10.php>。

原住民族語言研究發展中心 Akiw 研究員撰