科技輔助，人爲主體——原住民語AI系統如何打造師生的數位雙贏？

テクノロジーが支え、人間が主体に——原住民族語AIシステムはいかにして教師と生徒の双方有利を生み出すのか？

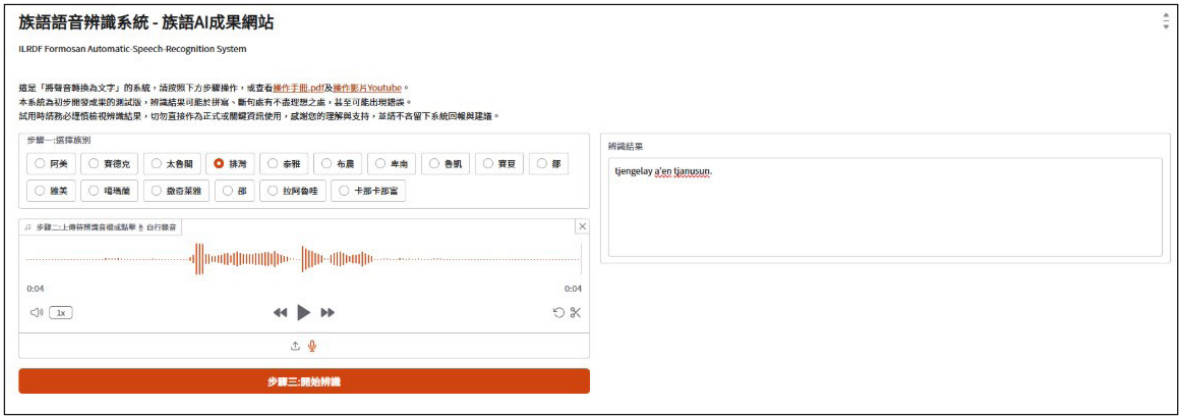
Technology-Assisted, Human-Centered: How Indigenous Language AI Systems Create a Digital Win-Win for Teachers and Students

文・圖—王凱弘（財團法人原住民族語言研究發展基金會 研究發展組研究員）

隨著人工智慧時代的來臨，AI 能為族語教育帶來什麼實質的幫助？根據聯合國教科文組織（UNESCO）於2010年發布的《世界瀕危語言地圖》顯示，台灣的原住民族語言過半數被列為「脆弱」等級，部分甚至面臨「確定瀕危」至「極度瀕危」的處境。當能流利使用族語的人口逐年流失，第一線教學資源也隨之吃緊。為此，原住民族語言研究發展基金會（原語會）正積極透過「族語AI系統」的建置，期盼科技能成為族語教學及學習上最得力的輔助工具。

賦予電腦「聽、說、譯」的能力——族語 AI 的三大核心技術

為了讓族人們在族語學習和教學上有更多元的選擇，原語會從2023年年底至2025年間，初步建置了「族語AI系統」。這套系統不僅僅是單一的工具，更是後續開發族語數位App或學習平台很重要的基礎，目前主要涵蓋以下三大技術：**(1) 族語語音辨識系統（ASR）：**系統的主要功能為，將接收到的語音自動轉化成族語文字。目前已支援16族語言的語音辨識。原語會針對秀姑巒阿美語、德固達雅賽德克語及太魯閣語等三語的抽樣評測結果顯示，三個語別的「字元錯誤率（Character Error Rate, CER）」皆低於6%。換言之，辨識準確度高達 94% 以上；**(2) 族語語音合成系統（TTS）：**此系統能讓使用者在輸入一段族語文字後，透過合成技術，產出一段擬真的語音。系統目前支援42語別的語音合成。在擬真度評測中，針對前述三個語別，原語會各邀請了10 位老師進行盲測，結果顯示三個語別的評測皆獲得了高達 4.5 的高分（滿分5分），打破過去機器發音死板的刻板印象；**(3) 基礎翻譯系統（MT）：**支援族語與華語間的雙向翻譯功能。目前系統在處理日常短句、簡單句時已有不錯的效果；雖然在面對長句、複雜篇章時，受限於現階段語料庫規模，翻譯精準度仍有改善空間。然而，這也提醒了我們：AI只是輔助工具，系統翻譯後仍需仰賴使用者的把關，人機協作才是最理想的模式。



族語語音辨識系統截圖。
（圖片來源：https://ai-labs.ilrdf.org.tw/sapolita-kaldi/）

化身教師的神隊友——族語 AI 在備課、互動與評量的應用

在教學端，族語AI系統能為教師在課程準備、學習活動及學習評量上，提供更多元的數位學習方向。在進行備課時，若須先進行文化性知識的語料採集，在逐字稿轉寫及教材的製作上往往曠日廢時，族語語音辨識系統的延伸應用「逐字稿辨識系統」此時便能省去許多備課時間，讓教師們能快速產出初步逐字稿，僅須做後續的文字校對工作，便能作為教材的參考資料。若教師想編寫對話教材，也可以透過基礎翻譯系統的雙向翻譯功能，先產出基本的對話文本，最後，再透過語音合成系統產出有聲音檔並置於教學簡報中，即可快速完成影音教材。

在學習活動上，為了實踐以學習者為中心的課程設計，教師可透過語音辨識系統設計發音挑戰賽，讓學生用麥克風唸出單詞或句子，如果系統能成功辨識學生的聲音，代表學生的發音清楚且正確。除此之外，在進行情境對話教學時，教師可以透過語音合成系統預先生成多個問句的音檔，讓學生與系統進行一對一的口語練習，並錄下學生每句答句的聲音。如此一來，便能減少課堂上因師生一對一互動，而讓其他學生空等的時間。

最後，在學習評量上，語音辨識系統可以初步比對學生在口說練習上的發音與標準文本的差異，透過辨識出來的族語文字，教師可以明確指出學生發音的偏誤，並提出具體發音的改善方法。以霧台魯凱語為例，「花」的魯凱語為「bengelray」，其中捲舌邊音 lr / ɭ/ ，常常是許多學習者較難正確發出的輔音。如果辨識系統產出的結果是「bengelay」，代表學生將捲舌邊音 lr / ɭ / 唸成齒齦邊音 l / l / ，教師可透過可視化的結果明確指出學生的錯誤並加以糾正。語音合成系統則能讓老師們快速產出聽力測驗的考題，省去句句親自錄音的時間，也更能標準化口說與聽力測驗的題幹語音。

打造無壓力的專屬學伴——族語 AI 在自主學習的應用

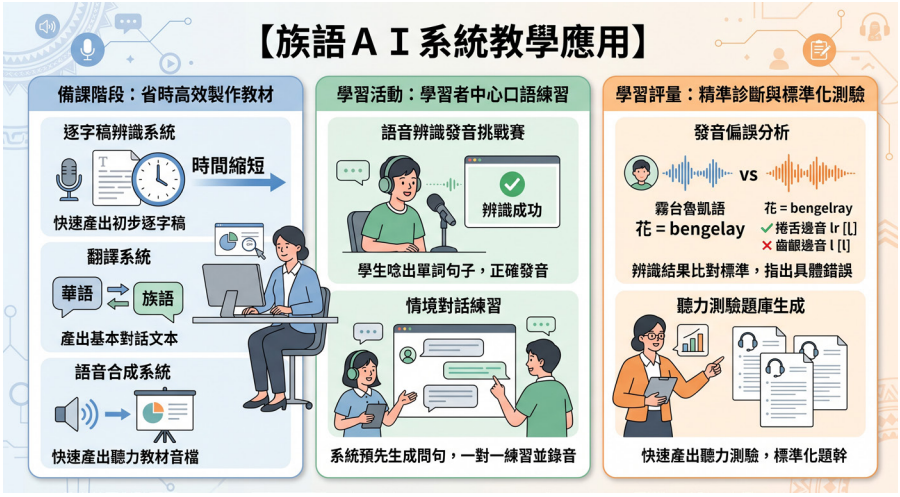
在學習端，族語AI系

雖然在面對長句、複雜篇章時，受限於現階段語料庫規模，翻譯精準度仍有改善空間。然而，這也提醒了我們：AI只是輔助工具，系統翻譯後仍需仰賴使用者的把關，人機協作才是最理想的模式。





族語語音合成系統介面截圖。（圖片來源：https://ai-labs.ilrdf.org.tw/hnang-kari-ai-asi-sluhay/）



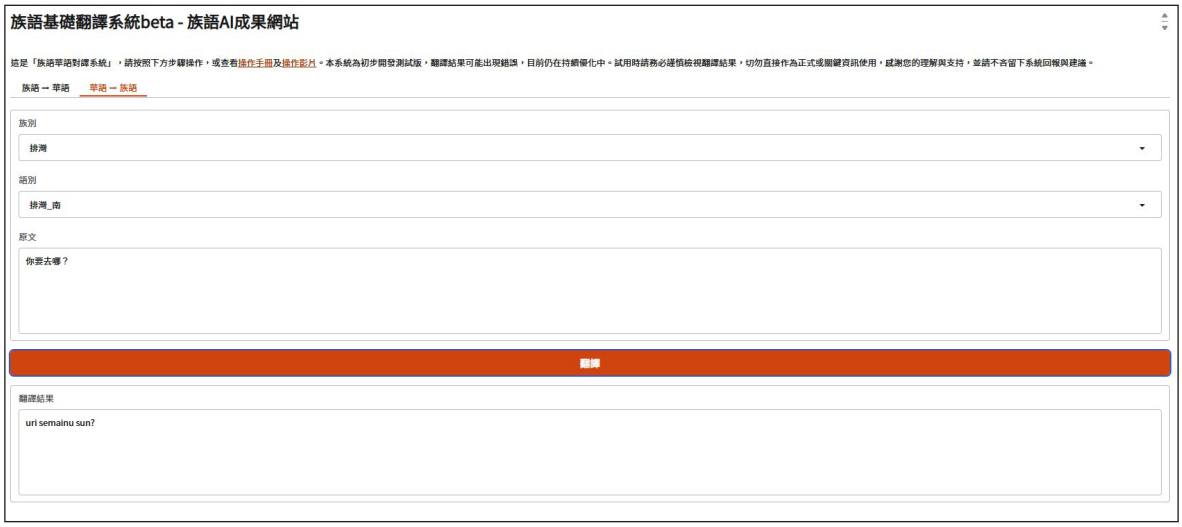
族語AI系統教學應用—備課階段、學習活動、學習評量。

統能成為學生在族語自主學習的入門磚。從學習環境來看，現階段國民中小學的族語課程，每週僅一堂，以筆者曾服務的部落小學為例，即便學生身處部落，學生在課堂外主動說族語的機會也少之又少。如果家中長輩不主動說族語，學生便難以在足夠的族語輸入下習得該語言，等到了國高中階段，離開部落的學校，學業壓力變重，更難以將心思放在族語學習上。因此，教師如何透過族語 AI 系統設計自主學習活動，讓學生懂得用數位工具「為自己的族語學習負責」，便成為十分重要的課題。

首先，族語 AI 系統創造了半擬真的族語環境。語音辨識系統可以成為課後永遠不嫌煩的傾聽者，教師能透過課後作業的指派，讓學生使用辨識系統反覆地練習，直到單字及句型辨識正確。系統的辨識機制讓學生能降低面對人說話的焦慮感，當口說練習成為反射動作，遇到真人時也更能自信地

即便是單一語別，不同部落與地區所使用的腔調與詞彙，也都蘊含著在地差異。如何在導入AI科技的同時，不抹煞原住民族語言的「多樣性」，是每一位AI系統開發者必須面對的課題。





族語基礎翻譯系統介面截圖。（圖片來源：https://ai-labs.ilrdf.org.tw/kari-seejiq-tnpusu-ai-hmjil/）

講出常用句型。除此之外，學生也可以透過語音合成系統的結果，強化語音輸入，當離開教室後，學生簡單輸入文字，就算不會的單字或句子，系統都會不厭其煩地唸出來。

此外，針對中級以上的族語二語學習者，教師可以引導學生透過自主學習任務，達到自主學習的效果。例如，教師指派任務讓學生輸入若干個族語單詞至合成系統中，並在下載音檔後隨機播放，自行練習聽寫，這種「讓自己成為自己出題的主動學習者」之模式，有助於強化單詞記憶及拼寫。除此之外，學生也可以透過基礎翻譯系統，比較中文和族語間的語法差異。以南排灣語為例，學生輸入「我喜歡香蕉。」，產出的結果為「tjengelay a'en ta veljevelj.」，學生可透過詞彙的擺放位置了解兩個語言語序上的差異，學習族語的語法知識。

AI 無法取代的溫度——科技爲輔、人類爲主的語言復振之路

回顧族語 AI 建置的前兩年，「低資源語言」語料的稀缺仍是最大的挑戰。儘管目前系統已涵蓋16族42語別，但各語別語料的規模與落差，確實直接影響了系統的效能。此外，即便是單一語別，不同部落與地區所使用的腔調與詞彙，也都蘊含著在地差異。如何在導入AI科技的同時，不抹煞原住民族語言的「多樣性」，是每一位 AI 系統開發者必須面對的課題。為此，原語會正持續優化族語 AI 系統，並積極開發多元的應用工具，期盼讓每一個區域的聲音都能被如實記錄。然而，使用者也必須時刻提醒自己：AI帶來的只是技術上的便利，科技始終是輔助工具，語言復振的希望，依然仰賴每一位具備智慧的族人與教育工作者來傳承。◆



王凱弘
新北市土城區人。1996年生。台灣大學語言學研究所碩士。現任財團法人原住民族語言研究發展基金會研究員，曾任台灣大學語文中心兼任華語教師、屏東縣獅子鄉楓林國小教師。目前在原語會負責族語資料庫、族語AI系統、語料處理人才培訓等業務，關注原住民族語言發展與復振議題。