

線上機器翻譯：東南亞新住民語言服務芻議

史宗玲

本文研究動機緣起於近年來東南亞新住民在臺灣人口快速成長，卻缺乏以其母語提供網路資訊的平台，另則線上統計式機譯系統翻譯品質已大幅改善，可以成為多語服務的媒介，因而思考如何使用線上機器翻譯來做為提供東南亞新住民語言服務的平台。本研究首先採用文獻考察及初探性調查，探討何以線上機譯系統可以成為移民線上資訊取得之輔助工具，接著採用訪談方式進行機譯品質評估，證實前編措施確實能夠幫助東南亞新住民對其母語（越南語及印尼語）機譯內容大意之了解。最後再探究機譯語言服務所彰顯的語言人權及民主意涵。本研究藉由訪談四位越南籍人士和三位印尼籍人士，就其對母語機器譯文大意理解度來做分析，結果顯示越南籍及印尼籍受訪人士之理解程度皆達到八成以上；受訪者表示，只要網站能提供八成以上理解度之機器譯文，他們願意利用機器翻譯工具來擷取更多臺灣社會與文化的相關資訊。有鑑於此，作者建議臺灣網頁文本可增加控制性中文書寫版本，以利提供東南亞新住民友善的線上機譯服務，此種語言服務在東南亞新住民已佔有相當比例的臺灣社會中，可彰顯其人權意涵，如語言平權、多元文化主義與庶民語言運動。

關鍵詞：東南亞新住民、機器翻譯語言服務、控制性中文、語言平權、多元文化主義

收件：2017年9月8日

修改：2018年2月7日、2018年6月21日

接受：2018年6月27日

史宗玲，國立高雄科技大學應用英語系教授，E-mail: clshih@nkfust.edu.tw。

作者感謝本文研究接受科技部專題計畫補助（編號 105-2410-H-327-030-），並感謝匿名審查者與本期主編提供寶貴意見，以精進其內容；此外，編譯論叢助理編輯細心校正，一併於此敬致謝忱。

A Probe into the Language Service of Online Machine Translation for Southeast Asian Immigrants

Chung-ling Shih

This article looks at the use of online machine translation (MT) to provide language service to Southeast Asian immigrants in Taiwan. Its motivation is twofold: (a) Taiwan's immigrant population has been growing in recent years and yet the immigrants still lack a web-based information service, and (b) there has been a significant improvement in the online statistical MT systems' capacity for automated translation. This research began with a review of the MT-relevant literature and a pilot investigation, followed by an interview-based survey to demonstrate how pre-editing using controlled Chinese helped Vietnamese and Indonesian audiences to understand the gist of MT outputs. It concluded with a probe into the implications of online-MT-enabled language services. The findings of the MT-based comprehensibility survey showed that an average score of more than 80% earned by Vietnamese and Indonesian interviewees supported the effectiveness of pre-editing in improving MT performance. Meanwhile, all the interviewees expressed that they would use the MT tool to access Taiwan's socio-cultural information, since the gist of MT outputs was more than 80% comprehensible. Thus, besides natural Chinese web-texts, controlled Chinese web-texts can be supplemented, and so Southeast Asian immigrants will be able to read the MT outputs for the purpose of information mining and knowledge acquisition. Despite the small samples used in this research project, the positive findings have shed light on some humanistic issues: equal language rights, multi-culturalism, and the "civic language" campaign. The importance of controlled-Chinese-enabled MT service cannot be overlooked in Taiwan's contemporary society as Southeast Asian immigrants now constitute a significant part of Taiwan's population.

Keywords: Southeast Asian immigrants, machine-translation-based language services, controlled Chinese, language rights, multiculturalism

Received: September 8, 2017

Revised: February 7, 2018, June 21, 2018

Accepted: June 27, 2018

壹、前言

近年來臺灣東南亞新住民人口愈形趨多，根據內政部移民署（2015）目前於網路公佈的最新資料，至 2015 年 8 月底統計，東南亞新住民在臺灣的總人數為 591,420 人，含印尼籍 220,688 人、越南籍 165,849 人、菲律賓籍 122,063 人、泰國籍 65,194 人及馬來西亞籍 17,626 人；而根據教育部（2017）針對新住民子女人數分布概況統計，民國 105 學年新住民子女就讀國中、國小一年級的學生總人數為 196,178 人，面對約 80 萬東南亞新住民人口，再加上政府如火如荼推動「南向政策」，更突顯東南亞語言、文化之重要性。目前臺灣有客家電台、原住民電台，但仍未見有東南亞相關語言如越南語、泰語、印尼語等的電台或電視節目。其次，網站的使用語言，除了主流之中文或英文，偶見德文、法文、日文、西班牙文、義大利文等之外，少見東南亞語文；透過網路搜尋也僅有中央廣播電台的網站提供少許泰文、越南文及印尼文的節目，故臺灣的東南亞新住民欲以母語閱讀臺灣網站資訊的機會微乎其微，相較於臺灣其他族群如客家人、原住民等，東南亞語言之使用率和呈現率更是相形見绌，顯然還有很大的改善空間。今日臺灣已是多語族群的社會，政府應該思考實施多語政策，俾使不同族群可以其母語閱讀或取得資訊，尤其是來自東南亞國家的新住民，更需盡快了解臺灣社會文化的資訊，期能早日融入當地社群。「語言人權之理想」在實踐過程中難免會遇到一些困難和挑戰，如政府經費補助不足或欠缺持續的推動等，以致於理想無法如期實現。為免經費不足之憾，本文提議使用免費線上機器翻譯（Machine Translation，簡稱 MT）來作為東南亞新住民語言服務的小幫手，然欲達成此目的，我們必須先探討機譯服務之可行性和機譯品質，並提出相關配套措施，以證實其應用效益。

Bowker (2009) 文章中曾提及政府預算有限，加上專業譯者不多，加拿大政府無法提供充裕的法語翻譯官方網頁，故加拿大

政府官方語言專員辦公室 (Office of the Commissioner of Official Languages) 建議政府機構使用科技來解決翻譯供需差距的問題 (Adam, 2013; Office of the Commissioner of Official Languages [OCOL], 2005)。OCOL (2005) 人員表示，政府應該探究如何經由新的管道 (explore new avenues) 來增加翻譯文件的效益與效率，語言科技工具 (technolinguistic tools) 即可列入政府考量政策之一，以強化其翻譯服務成效 (p. 58)。其實，加拿大氣象報告自 1976 年起，就開始應用 TAUM-METEO 系統來進行法英、英法翻譯，而 Systran 系統自 1978 年以來，就已應用在盧森堡歐洲共同體委員會的多語文件翻譯 (Somers, 2003a)。Bowker 在其論文亦中提到，根據 Brace 之研究顯示：

自 1990 年代以來，歐洲執行委員會的 MT 使用率已飆升。於 1996 年，MT 系統處理了 220,000 頁文件；到了 2000 年時，數量已倍增至 546,000 頁。(引自 Bowker, 2009, p. 124)¹

此外，英特爾 (Intel) 和微軟 (Microsoft) 等公司也提供了產品知識庫的自動翻譯，減少客戶來電要求技術支援的次數，大幅降低公司開銷 (Dillinger & Gerber, 2009)。

目前臺灣政府提供多語翻譯文件的經費有限，加上能夠勝任中越、中印、中泰翻譯的專業譯者不多，我們應可考慮使用線上 MT 系統提供東南亞新住民免費之語言服務。許多文獻已提出 MT 工具輔助專業譯者之效益，GALA Global 公司報告指出，使用 MT 的成本降低許多，平均 5%—20% 的機器譯文可做最終譯文，而 40% 左右的機器譯文經過後編輯後，可達到出版的程度 (Juskewycz, 2016)。SDL (2012) 公司報告亦指出，MT 減少許多專業譯者時間，Best Western 公司藉由 MT，交件的時間已從 24 個月減少至 8 個月；此外，CNH 公司使用 MT 多語資源

¹ 本文引用自英文文獻之中譯文，由作者自譯。

系統文件，節省了高達一千五百萬的成本；Chrysler 公司則節省了 35% 的成本。SDL 公司經理 Beaupre (2010) 更表示，全球化浪潮下，多語市場已逐漸擴大，翻譯成本逐漸上升，故使用 MT 可縮小語言差距，節省成本，進而提升商業利益。上述文獻皆已明白指出使用 MT 工具除了提高翻譯產能外，亦有頗高之經濟效益。

由於上述 MT 系統處理的是英文譯入法語、西班牙語或其它印歐語系語言，而非中文譯入東南亞語言，二者無法相提並論。細究其差異，一則英語與印歐語言特性之相似性較中文與東南亞語言高；另則，印歐語言與英語互譯的文件資訊量語料庫已累積甚鉅，其比中譯東南亞語之平行語料庫更為龐大，故印歐語言與英文互譯的機譯品質自然比中文與東南亞語言互譯的機譯品質為佳。然而，近年來線上統計式 MT 系統如谷歌翻譯 (Google Translate) 所產出的譯文品質已比多年前進步許多 (Brown, Pietra, Pietra, & Mercer, 1993)，這點雖讓人寬心許多，但仍需探究統計式 MT 系統是否也同時可提升中文翻譯越南語、泰語、印尼語之品質。其次，也要思考應該提供何種配套措施，方能使中譯東南亞語言的 MT 品質達到「可知可解」的目的，而非僅停留在「不知不解」的狀況。最後，MT 應用在多語社會中所呈現的民主意涵，如語言人權、高經濟價值之 MT 服務、多元文化平權與庶民語言運動，亦值得進一步討論。有鑑於此，作者提出三個問題作為本研究的基調與指引：(1) 何以 MT 系統可作為移民取得線上資訊之輔助工具？(2) 何種配套措施有助於提升中譯東南亞語機器譯文品質？(3) MT 應用在臺灣多語社會情境中呈現的語言人權及民主意義為何？

本文探討如何使用 MT 系統作為提供東南亞新住民語言服務之小幫手，分為三階段：(1) 證實線上 MT 系統輔助翻譯之成效；(2) 使用前編 (Pre-editing) 之配套措施，改進機器譯文品質；(3) 強調 MT 對於移民組成的多語社會，具有民主實踐、語言正義及文化平權之貢獻。此三個階段的研究目的，如圖 1 所示。

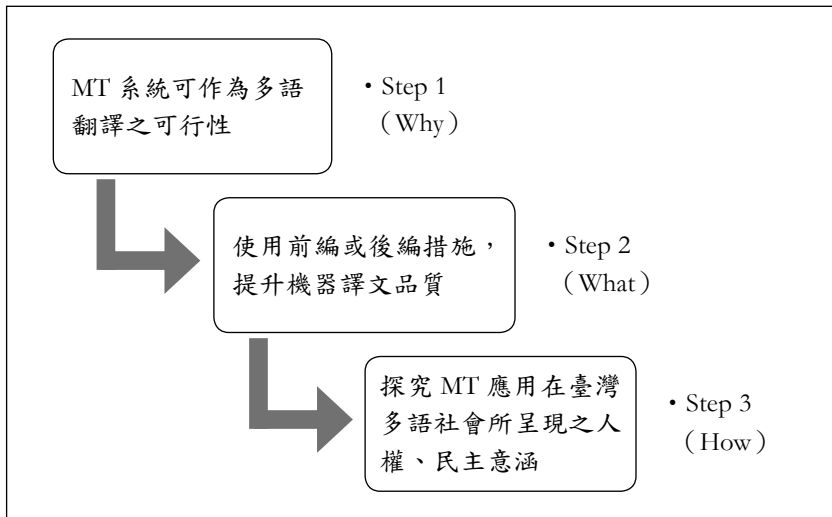


圖 1 本研究三大階段及目的
資料來源：作者自行整理。

本研究將援用文獻考察和訪談方式進行分析。首先，作者耙梳 MT 相關文獻，並使用前導訪談調查，以證實 MT 推動東南亞語翻譯服務之可行性。接著討論如何藉由書寫控制性語言作為前編輯來提升 MT 品質，並進行小規模實驗，提出受訪者評量 MT 品質的結果。最後再探討 MT 應用如何協助臺灣邁向一個平等、更多元之移民社會。

貳、機器翻譯定義及發展史

本文建議使用 MT 作為東南亞新住民翻譯服務之助手，故於此先行介紹 MT 定義與發展歷史。MT 顧名思義乃是指電腦或機器系統自動產出譯文，或藉由電腦程式將文字或演說從一種自然語言轉譯成另一種自然語言。換言之，MT 是透過電腦系統將一個自然語言轉換成另一個語言的一門學科。該主題乃是建立在語言學、數學及電腦科學這三門學科的基礎上（馮志偉，1995）。

毫無疑問的，MT 是一項極具前景的科技，自 1940 年代後期開始，MT 一直是人工智慧領域的重要研究項目 (Hutchins, 1986)。MT 發展已經歷半世紀之久，但其不斷地遭受質疑。從最早期的字典配對方式，接著是字典結合語言學的規則法 MT 系統，近年來則是語料庫統計式 MT 系統，甚至發展到神經 MT 系統，翻譯品質愈見優良，故作者提議利用線上 MT 系統，再配合前編網頁文本，作為東南亞新住民閱讀臺灣網站的 MT，作為認識臺灣之便捷途徑。

有關 MT 的發展歷史，根據 Somers (2003a) 所言，可區分為早期研究 (early research)、第二代研究 (second generation)、商業應用 (commercial application) 及最新研究 (the latest research) (pp. 4-7)。Shih (2006) 則將 MT 發展歷史分為四階段，包括：(1) 開創期 (pioneering stage)，時間從 1950 年代至 1970 年代；(2) 黑暗期 (dark stage)，時間從 1970 年代至 1980 年代初期；(3) 上市期 (marketing stage)，時間從 1980 年代中期至 1990 年代初期；以及 (4) 復興期 (revival stage)，時間從 1990 年代中期至今 (p. 13)。

二次世界大戰期間，洛克斐勒 (Rockefeller) 基金會副總裁瓦倫·韋弗 (Warren Weaver) 首次提出 MT 研究及發展之構想。於 1947 年至 1949 年間，他提出一項計畫，藉由電腦解碼的技術來發展 MT，此計畫引起同事的興趣 (Weaver, 1949)。這項創新的概念得到麻省理工學院的正面支持，故麻省理工學院在 1951 年成立了一個 MT 研究小組，並在麻省理工學院所舉辦的研討會中發表研究成果。這項研究成果普遍影響了日後其他國家的 MT 研究，包含了蘇維埃社會主義共和國、英國、加拿大等。此乃是早期 MT 研究的成果。

接著，美國科學院自動化語言處理諮詢委員會經過兩年的研究後，於 1966 年發佈一份名為《語言與機器》的報告。報告中全面否定了 MT 的可行性。正因如此，只剩下三個計畫可獲得美國政府的補助得以繼續研究。然而到了 1975 年，美國政府對於 MT 研究的補助全面停止。在美國中止 MT 研究後，加拿大、法國、德國等國家卻持續支持大學研究

MT。而且從 1976 年起，由加拿大蒙特婁大學與加拿大聯邦政府翻譯局聯合開發 TAUM-METEO 系統，開始應用在加拿大氣象局的英譯法及法譯英的雙向氣象報告，其譯文正確性高達 90% 以上。同時，歐盟組織亦開始使用 Systran 系統處理多國之間的公文 (Somers, 2003a)。此時期雖為黑暗期，但已逐漸露出曙光。

於 1981 年，由楊百翰大學團隊所研發的 MT 系統 Alps 上市。日本也在此時期研發出許多不同的 MT 系統，如 1984 年的 Medium-Pak、1985 年的 Micro-Pak J/E III、Rosetta、Pivot、Duet-E/J、1986 年的 X-EJ、1987 年的 ASTRANSAC、1988 年的 Pense。這些 MT 系統主要用於英日的科技以及商業文件翻譯 (陳子昂、黎偉權，1991)。此外，1993 年 IBM 的 Brown 和 Della Pietra 等人提出基於詞對齊的翻譯模型，代表著現代統計式 MT 系統的誕生 (科大訊飛 cobbyli、zawang, 2017)。由於許多 MT 產品開始在歐洲及日本市場推銷，故作者將此時期定義為上市期。

到了 1990 年代中期，MT 發展進入復甦期，此時許多中英互譯 MT 產品問世，而翻譯記憶系統也漸漸受人矚目。在中國以及臺灣地區，許多科技公司都已克服了科技以及語言學上的困難，發展出中英互譯的 MT 系統。在中國上市的 MT 系統有東方快車、華建、金山快譯以及中軟譯星。在臺灣則有翻譯小精靈、譯典通、黑色譯神、譯橋以及譯言堂 (Shih, 2006, p. 16)。同時，MT 研究人員致力於研發翻譯記憶系統，其概念則是由 Martin Kay 在 1980 年所提出。一開始翻譯記憶並沒有受到太多重視 (Somers, 2003b, p. 32)。後來，Alan Melby 在楊百翰大學的研究團隊將「重複作業」(repetitions processing) 的概念融入 MT 系統 Alps 之中，算是翻譯記憶使用之雛形 (Somers, 2003b, p. 33)。到了 1990 年代中期以後，許多科技公司開始販售翻譯記憶產品，並受到使用者們的歡迎。著名的翻譯記憶產品包括 Trados、Star Transit、SDL International 以及 WordFast。直至今日，MT 系統已經與翻譯記憶系統融為一體，並成為現今翻譯科技中的主流趨勢。同時隨著線上大數據增

加，谷歌（Google）公司 1991 年推出統計式 MT，讓 MT 的準確度大幅提升。統計式 MT 系統（statistical machine translation/SMT）乃是從許多資料中提取使用出現頻率最高的翻譯詞組，再串聯成建議譯句，其譯文比傳統規則法 MT 系統譯文更優良。上述 MT 四個重要發展時期可列表如下。

表 1

MT 四大發展時期

	開創期 (1950 年代— 1970 年代)	黑暗期 (1970 年代— 1980 年代初期)	上市期 (1980 年代中期— 1990 年代初期)	復興期 (1990 年代中 期—2010)
研究動機	自動破解敵人 軍事資訊的高 品質翻譯	著重於計算語 言學及人工智 慧的學術研究	商業利益	商業用途及跨 文化溝通
主要地點	美國	歐洲及加拿大	歐洲及日本	亞洲（中國大 陸及臺灣）、 歐洲和美國
資金來源	軍方、政府及 情報部門	私人企業	政府及私人企業	企業及政府
應用	官方用途於特 定的文件	大眾用途如天 氣預報及跨公 司之間的溝通	產業用途如科技 及商業文件翻譯	個人用途或是 多重商業用 途：如資訊獲 得、訊息傳遞 與本地化

資料來源：譯自 Shih（2006, pp. 17-18）。

近年來，隨著深度學習崛起，2014 年學界將深度學習應用於 MT，創造出神經網絡 MT（neural machine translation/NMT）系統，其運行原理是讓機器自動從更大規模的語言資料中學習規則，而非直接利用給予的規則。百度於 2015 年 5 月發佈其網路 NMT 系統，谷歌在 2016 年 9 月也發布第一款 NMT 系統，對外宣稱該機器譯文之理解性和正確性

平均可達 60%，接近中等雙語使用者的水準（南僑專欄，2016）。當然，某些語對翻譯成效比其他語對較好。2010 年之分析研究指出法譯英、義譯英之正確性極高（南僑專欄，2016）。媒體平台專欄作家謝宇程（2015）在網路上貼文發表其使用 MT 經驗，表示他曾將一篇義大利文放入谷歌翻譯譯成英語，沒想到譯文非常通順，可完全了解其訊息，他並指出歐洲語和英語對譯可產出極佳譯文，而阿拉伯語、漢語、日語及英語對譯亦可讓讀者猜出 50%—60%。隨著科技進展及日益擴增之語料庫，機譯前途將無可限量，難怪美國政府早在 2009 年就把 MT 列為未來最具影響力的十大技術之一。正因為線上 MT 品質（以谷歌翻譯為例）已大幅改進，若再配合簡易書寫風格的網頁文本，更能提升其機器譯文可理解性。

參、MT 可作為移民資訊取得之輔助工具

透過 MT 相關文獻考察，得知線上 MT 系統可作為外來移民取得資訊之輔助工具。根據 DePalma（2007）的研究，非本地語言的小眾社群，尤其是外來移民，很難獲得主流社會大眾語言的翻譯服務，故美國舊金山城市使用 MT 系統提供翻譯給當地英語能力不好的各種不同移民人士。該城市政府網站上開門見山寫到：「與其不提供任何譯文，我們寧可提供自動機器譯文，使得非英語人士至少可得知網頁訊息的大意」；原文為：“We prefer to provide automated translation rather than no translation at all in order at least provide speakers of other languages an overall sense of the information available on a web page”（Bowker, 2009, p. 124; DePalma, 2007, p. 46）。如果早在十年前美國舊金山當地政府就已使用 MT 系統來協助解決移民的閱讀障礙，十年後的今天，線上統計式機譯品質或許已大幅改進，我們當更有信心新住民可藉由 MT 系統來輔助其使用母語來閱讀網頁機器譯文，以取得各種資訊，這是 MT 系統帶給移民社會的契機與轉機。

舊金山政府的美意早已獲得 Hutchins (2001) 的支持：

雖然專業譯者不能接受品質低落的機器譯文，但多數民眾卻可接受；因為他們想要立即得到資訊，且線上閱讀環境往往需要立即取得資訊，並且立即處理資訊。(p. 8)

雖然仍有許多專業譯者及語言教師排斥 MT，但一般民眾更重視訊息溝通的便利性，MT 系統對他們而言，仍有其存在價值。以下兩例亦可提供佐證。Allied Business Intelligence 的報告指出：

MT 系統可以生產信件大意或其它訊息翻譯。在商業領域，把握時機傳遞訊息是很重要的，故商場愈來愈可接受低品質文件。一般使用者寧可犧牲品質問題以換取訊息溝通之便利性。(引自 Bowker, 2009, p. 144)

Gerber (2008) 也指出，自從英特爾公司 (Intel Corporation) 使用線上 MT 系統處理技術資訊文件英譯西班牙文，顧客明顯減少來電請求協助，甚至有些機譯品質已具一定水準，根本不需人工編輯。他更解釋道：

多數使用者不會要求高品質的文件，故英特爾公司僅要求資訊大意清楚，以減少民眾要求技術支援。當公司評量 MT 譯文之可行性時，語言學家往往是認為不行的，但其公司在中、南美區域的管理階層卻認為可行；MT 譯文可提供給西班牙語人士閱讀，因為他們得到西班牙機譯資訊的協助後，可減少致電求助。(Gerber, 2008, p. 16)

其實機器譯文被接受的程度，往往因使用目的不同而有所影響。學者認為若使用者只想了解文本大意，如個人資訊取得 (personal assimilation

of information) 等，則目前 MT 譯文是綽綽有餘的 (O'Hagan & Ashworth, 2002; Vasconcellos & Bostad, 1992; Wagner, Bech, & Martínez, 2002)。同樣的，對於不懂官方語言網頁資訊的移民而言，MT 譯文既然可協助他們快速取得資訊大意，那麼利用 MT 系統來作為移民線上資訊取得之輔助工具自是無庸置疑的。

上述兩個例子的文本內容皆屬於公司商業資訊，如客服、線上通訊及技術支援文件等，皆不含有地方俚語、文化用詞、隱喻表述等，故其 MT 譯文之可讀性自然比較好；但若為文化文本，則因含有地方俚語、文化用詞、隱喻表述等，其機器譯文之可讀性自然比較差；為了證實不同文類會影響 MT 譯文之可讀性，作者請一位國立高雄科技大學就讀國際管理碩士學位學程 (IMBA) 的越南籍學生閱讀兩篇中譯越南語的 MT 譯文，內容分別為花王公司簡介和七夕節日的介紹，他表示前者可讀性及正確性較後者高出很多；公司簡介越南語之 MT 譯文的理解度約 60%—70%，而七夕文章的越南語 MT 譯文大約只可了解 20%—30%，可見商業文件 MT 譯文之正確率比文化文本之 MT 高出許多。接著，作者又要求該越南學生閱讀評量一篇公司網站的越南語 MT 譯文，答案竟然出乎意料，其理解程度竟高達 80%。探討其原因，乃是此篇公司簡介的中文先行使用控制性中文進行前編輯，致使 MT 品質大幅上升。由上可知，公司網頁機器譯文的確可提供東南亞新住民大意資訊，但遇到語意、句型複雜之資訊內容時，宜先作前處理，方能提高其 MT 譯文之可讀性。至於含有文化用詞之文本，則更需要前處理，亦即使用控制性語言書寫，故本論文將進一步檢視是否文化文本亦適用控制性語言及東南亞語機器翻譯，以增進 MT 提供新住民語言服務之效能。

肆、提升機譯效能之配套措施

從上述文獻得知 MT 可作為移民取得線上資訊之輔助工具，但先決條件是以控制性語言書寫網頁文本，以下將介紹控制性中文書寫提升

MT 品質之效益，並調查越南籍及印尼籍人士對控制性文化書寫之 MT 譯文的理解程度。

一、控制性語言書寫

已有若干研究指出，控制性語言書寫的文本有助於提升 MT 譯文之可讀性與可理解性。OCOL (2005) 網站指出，泛美衛生組織 (Pan American Health Organization) 平日使用 MT 系統處理多國語翻譯文件時，必先經過自動修正或人工改寫，之後再送入 MT 系統自動處理 (p. 59)。經過前編，生產成本比人工翻譯每字平均節省 31% (OCOL, 2005)。高雄第一科大口筆譯碩士班研究生 Lo (2015) 針對 20 位泰語人士進行一項 MT 譯文問卷調查，結果顯示使用自然中文書寫原文時，其泰文 MT 譯文的理解度平均只有 37.75%，使用控制性中文書寫原文後，泰文 MT 譯文的理解度則可高達 89.7%。由此可見，使用控制性中文書寫原文確實對改進 MT 譯文品質有極大之助益。換言之，MT 系統使用合宜的配套措施，如前編，則有如虎添翼之效，將可全面提升 MT 品質，並節省成本。值得注意的是，作者僅提到前編作業，乃是網路翻譯提供即時大意翻譯，無法立即進行後置作業。此外，多數網路使用者閱讀 MT 譯文時，僅希望從中獲得全文大意及相關資訊，並非特別在意 MT 譯文流暢度及文法正確性。

「前編」意指將來源語文本改成「控制性語言」，以改進 MT 譯文可讀性。控制性語言又稱為人工語言 (artificial language) 或次語言 (sublanguage)。1970 年代該特殊語言便已出現，當時是用來書寫技術文件，純粹為了方便讀者了解文件內容，與翻譯毫無關聯。今日隨著全球化來臨，許多商業文件、使用者手冊、產品說明書等為了滿足全球客戶的需求，必須多語呈現以方便拓展市場。經由前編或使用控制性語言書寫，其多語 MT 文本大意的正確性及可讀性能夠大幅提升，基於此功能需求，控制性語言乃成為提升 MT 品質的重要書寫條件。如前所述，一位越南人士表示，自然中文公司簡介的越南文 MT 譯文可解讀性為

60%—70%，控制性中文公司簡介的越南文 MT 譯文則達 80% 以上。

我們無法針對各種不同的東南亞語言特性提出客製化的前編策略，況且 MT 系統技術日新月異、語料庫隨時變更內容，其譯文品質也會日益進步，所以僅提供一些控制性中文書寫的通用原則即可。此外，於先行研究階段時，作者使用中譯英的前編原則書寫兩篇控制性中文（花王公司簡介和七夕節日的介紹），再使用機器翻譯將其譯入越南語，發現越南籍讀者可了解 80% 以上的文本大意，故目前適用於中譯英的控制性中文書寫原則，也可適用於控制性中文譯入東南亞語。控制性中文書寫原則的特性，可歸納為：（1）避免使用一詞多義之詞彙；（2）避免使用分詞、片語及代名詞；（3）避免使用俚語俗語；（4）避免使用關代子句及語意不完整的句子；（5）盡量使用短句；（6）一句子應以傳達一個旨意為限；（7）將左側太長的修飾詞改寫成不定詞片語或其他句構；（8）每一子句使用 SVO 或 SV+ 其他詞性的語詞結構；（9）盡量使用被動式；（10）表時間或地點之介係詞片語，宜置於句首或句尾。语法方面需注意的秘訣和細則包括：（1）使用數詞及量詞；（2）使用「以」表不定詞，銜接兩個動詞；（3）使用所有格，如「她的」、「我們的」；（4）使用連接詞銜接兩個子句；（5）使用時態標示詞來說明各種動詞時態，如動詞後增加「了」、「過」以表示過去式；使用「將」或「將會」表示未來式；（6）使用「該」字表示指稱詞；（7）於動詞前加上助動詞，如「可以」、「將」，使其動詞機譯文法正確；（8）被動式有主事者，應置於句尾，且前方加「由」（史宗玲，2013，頁 55、63）。值得注意的是，若是地方文化用詞，不妨加上解釋，如：七娘媽（孩童的保護神），有助於來自不同文化背景的新住民了解其意涵。下例使用「新年禁忌」的兩種文本說明之，如表 2。

表 2
自然中文改寫成控制性中文之對照表

	自然中文	控制性中文	說明
例句 1	「我們不能清掃房子，不然會把運氣及財富掃出去」	「我們不能打掃房子；否則我們的運氣與財富會被掃走」	自然中文的附屬子句「會把運氣及財富掃出去」欠缺主詞，其使用句構為 O+V，故需要修正為 S「我們的運氣與財富」+V「被掃走」。
例句 2	「我們打破碗盤時，要說碎碎平安，這樣才不會有霉運」	「如果我們打破碗盤，我們必須說 Sui Sui Ping An。如此一來，我們可以解除壞運氣」	自然中文的附屬子句「要說碎碎平安」沒有主詞，且「碎碎平安」對於外國讀者無法了解此俗語之意涵，故添加主詞「我們」，把「要說」改成「必須說」，且將「碎碎平安」改為羅馬拼音「Sui Sui Ping An」。其次，自然中文「這樣才不會有霉運」也欠缺主詞，「霉運」是中國俗語，要避免使用之，故整句可編輯為「如此一來（連接詞），我們（主詞）可以解除（動詞）壞運氣」，意思較為清楚。
例句 3	「不要用刀子與剪刀，這樣才不會受傷」	「為了防範受傷，我們不應使用刀子與剪刀」	原自然中文使用祈使句，但附屬子句「這樣才不會受傷」欠缺主詞，仍不符合控制性中文句構 S+V 的原則，故我們必須修正為「我們（主詞）不應使用（動詞）刀子與剪刀（受詞）」。
例句 4	「出嫁的女子在新年第一天不能回娘家，不然她父母會變窮」	「在新年的第一天，出嫁的女子不能返回她的父母家。否則，她的父母將會很窮」	依照上述控制性中文書寫原則，介係詞片語「在新年第一天」必須移至句首，且避免使用俗語，「回娘家」可改成「返回她的父母家」。其次，依據使用短句及一句一旨意原則，我們將原句子改成兩個獨立句，並增加轉折詞「否則」，將更能清楚傳達旨意。

資料來源：作者自行整理。

目前臺灣市面上並無控制性中文書寫系統，故仍需採用人工作業，然而就像訓練公文書寫一樣，只需一段時間便能上手；歐美大公司內部的控制性技術性文件書寫者也是需要接受訓練。由於線上統計式 MT 系統之語料庫存量日益增多，MT 系統日益聰明，即使不遵守某些原則，其英文 MT 譯文仍可達到一定水準。反觀中文翻譯東南亞語文的語料庫存量，目前仍處低量，所以很多人會使用機器翻譯系統，將中文先譯入英文，再將英文譯入東南亞語文，故建議書寫控制性中文時仍應盡量遵守上述原則。若非網路翻譯，一般翻譯公司會採用後編，目前歐洲許多翻譯公司則使用 MT 再加上後編流程，但本文強調使用新住民線上即時資訊取得，可以改寫原文書寫風格，以提高 MT 譯文可讀性，另一種合作法則是建置東南亞語翻譯語料庫，此牽涉到系統程式設計及翻譯資料的來源問題，不在本文探討範圍之內。

二、越南文及印尼文機器譯文之理解度調查

原先受邀的一位越南籍人士僅能了解 20%—30% 的自然中文之 MT 譯文。根據鄒劍宇（2016），採用 Bilingual Evaluation Understudy（BLEU）作為 MT 譯文評量標準，發現用戶閱讀英中 MT 譯文之理解度約為 40% 左右。而在韓國世宗大學與韓國國際協會於 2017 年 2 月初聯合舉辦的一場翻譯比賽中，得知由谷歌翻譯處理之英韓 MT 譯文之正確度為 46.6%（60 分中得到 28 分）、由 Papago 處理之英韓 MT 譯文之正確度為 28.3%（60 分中得到 17 分）、由 Systran 處理之英韓 MT 譯文之正確度為 25%（60 分中得到 15 分）（Kroulek, 2017）。由此可見雖然不同系統會影響到 MT 譯文的可讀性，但自然語言之 MT 譯文的可讀性及正確性平均而言仍不到 50%。職是之故，本研究建議使用控制性中文以提升 MT 譯文的可理解程度。

為調查使用控制性中文改寫後是否可大幅提升其 MT 譯文的可讀性，作者進行下列實驗，先請四位口筆譯碩士班研究生依上述控制性中文書寫原則書寫四篇介紹臺灣文化之短文，然後使用線上谷歌翻譯

(SMT)譯為越南文及印尼文。為了確保中文控制性語言書寫短文品質，作者先放入 MT 系統譯成英文，發現其譯文皆可了解，且文法皆正確。接著邀請四位越南籍及三位印尼籍人士來閱讀 MT 譯文（越南文／印尼文），並請他們提出修正意見和回饋。上述三位印尼籍受訪者為大學生，平均年齡為 20 歲，其中兩位是印尼華僑，其父母早年移居至印尼作生意，她們在印尼出生、接受國小、中學教育，而現在回臺灣唸大學。四位越南籍受訪者平均年齡為 28 歲，他們來自越南，就讀高雄科技大學的國際管理碩士學位學程，其大學專長背景包括法律及商業等，他們對臺灣文化很有興趣。本文以全文大意理解作為評量標準，並非以文法及用詞錯誤數量為依據，原因乃是網路 MT 譯文提供大意資訊，並非作為教學或品質評量使用。

第一篇題目為「臺灣坐月子」之控制性中文的越南文機器譯文，請四位越南人士閱讀後，請他們以 100 分作為分母，然後針對所有機器翻譯的句子給予概括性的代表數字。藉由此計算方式，他們主觀對全文大意之理解程度分別為 95%、80%、70% 及 88%。針對同一篇文章之印尼文機器譯文，三位印尼籍人士閱讀後，使用同樣的計算方式，得知他們對於全文大意之理解程度分別為 80%、95% 及 70%，兩者相較之下，越南籍人士及印尼籍人士對於機器譯文理解程度差距不遠。

受訪者皆表示，由於中文與越南文、印尼文間句子結構及單字字義有所差異，故易產生一些翻譯錯誤。有位印尼籍受訪者指出，第一篇文章「產後照護」中，印尼文 MT 為 Sejarah dan sosial budaya latar belakang（直譯：歷史與社會文化背景）有誤，應修改成 Latar belakang（背景）sejarah（歷史）social（社會）dan（與）budaya（文化）（直譯：背景歷史社會與文化），其原因乃是中文為前位修飾，形容詞應放置於名詞之前，但印尼文則為後位修飾，形容詞需置於欲修飾名詞之後。再者，有一名越南籍受訪者指出「以免得到一個偏頭痛」之越南文 MT 為 như vậy là không để có được một chứng đau nửa đầu（直譯：因此，並不會避免得到一個偏頭痛），為錯誤之譯文。“không”應置於“để”之後。原因在

於“không dễ”意指「並不會避免發生」，但“dễ không”意指「以避免得到」，正確譯文應為 *như vậy là dễ không có được một chứng đau nửa đầu*（直譯：因此避免得到一個偏頭痛）。此外，另一名越南籍受訪者指出原文「在農業社會，生活環境很不好」的越南文 MT 為 *trong xã hội nông nghiệp vào thời điểm đó, môi trường sống là rất kém*（直譯：在農業社會當時，生活環境很不好）。此譯文有些許文法錯誤，須將其編修為 *tại thời điểm xã hội nông nghiệp môi trường sống rất là kém*（直譯：在農業社會環境很不好）。

就單字字義而言，一名印尼籍受訪者指出，原文「產後的婦女們不能碰冷水」表示產後的婦女不被允許觸碰冷水，於此「不能」表示「不允許」。然而，印尼文 MT *wanita postpartum tidak bisa menyentuh air dingin* 中的“tidak bisa”表示「不能」。由於此產後禁忌與個人能力無關，因此 MT 譯文中的“tidak bisa”需更正為“tidak boleh”（直譯：不應該），表示「產後的婦女們不應該碰冷水」。在越南文 MT 中則有同樣的問題，一名越南籍受訪者指出，*phụ nữ sau sinh không thể chạm vào nước lạnh* 中的“không thể”（直譯：不能）需更正為“không nên”（直譯：不應該）。

第二篇控制性中文是介紹「臺灣婚禮」，四位越南籍人士閱讀越南文 MT 後，全文大意之理解程度分別是 85%、85%、80% 及 84%，三位印尼籍人士對於同篇文章之印尼文 MT 的全文大意理解程度則分別為 89%、95% 及 80%。綜而論之，無論是越南籍或印尼籍人士對於其母語之 MT 譯文之接受度皆超出八成以上。細查原因，乃是印尼籍受訪者有二位為華僑，對於臺灣結婚習俗如鬧洞房及歸寧有較充足的背景知識，故對於機譯的理解遠大於純越南籍人士。

根據一名越南籍受訪者的意見，在越南文機器譯文的第一個句子中，「今日」被譯為“hôm nay”，意思為「今天」，與文意不符，建議改為“ngày nay”（「近日」之意）較為正確。第二句中的「扇子」被譯為“fan”，但是“fan”為扇子的英譯，而非正確的越南文 MT，應該

改為“*quạt tay*”，代表「手扇」。此外，一名印尼籍受訪者指出，印尼文 MT 中，“*Taiwan pernikahan*”（直譯：臺灣婚禮）的語序應修改為“*pernikahan Taiwan*”（“*wedding of Taiwan*”）才符合印尼語文法。另外，在語意方面，“*tidak bisa*”（直譯：不可以）用於此處有誤，應改為“*tidak boleh*”（直譯：不應該），較能讓印尼籍人士了解。印尼文 MT *beberapa kerabat dan teman-teman, tua atau muda, akan bermain game, dan pengantin baru untuk menciptakan suasana meriah* 中的“*dan*”（直譯：且）用於此處不適合，應該改為“*dengan*”（直譯：與……一起），句子的意思才會正確。

第三篇控制性中文是介紹「臺灣喪禮」，四位越南籍人士閱讀越南文機器譯文後，全文大意之理解度分別是 80%、85%、100% 及 80%，三位印尼籍人士閱讀同一篇文章之印尼文 MT 後全文大意之理解程度則分別為 93%、85% 及 70%。前一篇 MT 譯文讀者接受度調查中，印尼籍人士高於越南籍人士，但該篇讀者反應剛好相反，越南籍人士的接受程度遠大於印尼籍人士。詢問之下，方發現印尼籍人士平均年齡為 20 歲，她們較為年輕，故對於喪禮的背景知識稍嫌不足，但較年長的越南籍人士（平均年齡為 28）則有較多關於喪禮的背景知識。

針對臺灣喪禮的文章，受訪者表示 MT 錯誤最主要的原因是有時無法產出正確的譯文。其中一名越南籍受訪者表示，機器翻譯中出現的羅馬拼音“*shaoguiqian*”意指「燒鬼錢」，但在越南文中，這樣的習俗則稱為“*hóa vàng mã*”，線上 MT 系統並不能產出正確的越南譯文，竟然出現羅馬拼音。此外，MT 也無法正確翻譯地名，如中文的「臺南」無法正確譯為越南文的“*Dài Nam*”。另一名越南籍受訪者表示「她的靈魂」應譯為“*linh hồn của hợ*”（直譯：她的靈魂），而不是“*linh hồn của mình*”（直譯：他的靈魂）。“*mình*”代表「他自己」，而“*hợ*”則代表「她自己」。關於印尼文 MT 錯誤，一名印尼籍受訪者指出，「在四十九天內，死者的家人不能剪他們的頭髮和鬍子」之 MT 譯句為 *dalam 49 hari dari keluarga almarhum tidak bisa memotong rambut dan jenggot mereka untuk*

menunjukkan belasungkawa mereka (直譯：在四十九天內，死者的家人不能剪他們的頭髮和鬍子)應更正為 dalam 49 hari dari keluarga almarhum tidak boleh memotong rambut dan jenggot (直譯：在四十九天內，死者的家人不應該剪他們的頭髮和鬍子)。印尼文中的“tidak bias”意指「不能」，因此須更正為“tidak boleh”，代表「不應該」。

第四篇控制性中文的題目是「臺灣成年禮」，四位越南籍人士閱讀此篇文章的越南文 MT 後，全文大意之理解程度分別是 85%、75%、90% 及 75%，三位印尼籍人士閱讀此篇文章的印尼文 MT 後，全文大意之理解程度則分別為 85%、85% 及 85%。兩者相較之下，印尼籍人士對於此篇 MT 譯文之理解程度大於越南籍人士。部分原因乃是印尼籍受訪者為華僑，他們對於臺灣成年禮有較多背景知識，故他們對於 MT 譯文的理解度自然較高。

同上述三篇 MT 譯文錯誤類型，該篇文章的越南文或印尼文的 MT 錯誤仍是語序及字詞方面的錯誤。舉例來說，原文「十六歲的成年慶祝活動」之印尼文 MT 譯句為 perayaan dewasa enam belas (直譯：成人的慶祝活動給十六歲的人)，一名印尼籍受訪者建議將其修正為 acara perayaan enam belas tahun (直譯：給十六歲的慶祝活動程序)。該句的越南文 MT 為 Mười sáu lễ trưởng thành (直譯：十六歲的成人儀式)，其語序亦不正確，應修正為 Lễ trưởng thành mười sáu tuổi (直譯：成人儀式給十六歲的人)。就字義錯誤而言，一名越南籍受訪者指出，原文「十六歲以下兒童們皆被招募來幫忙扛貨品」之越南文 MT 為 Trẻ em dưới mười sáu tuổi đang bị tuyển dụng để giúp vận chuyển hàng hoá (直譯：十六歲以下兒童們皆被徵召來幫忙扛貨品)，此句含有負面涵義，讀者可能會認為該句子表示小孩似乎是被強迫去幫忙扛貨品，故他建議將“bị tuyển” (被徵召) 改成 “được thuê” (受雇用)，且全句修正為 Trẻ em dưới mười sáu tuổi đang được thuê để giúp vận chuyển hàng hóa。此外，一位印尼籍受訪者表示，“hôm nay” 代表「今天 / today」，需修正為 “ngày nay” (現今 / nowadays) 才符合原文意思。另一位印尼籍受訪者

表示，原文「該儀式有不同的方式及程序」的印尼文 MT 為 *upacara ada cara dan prosedur yang berbeda*（直譯：有該儀式不同的方式及程序），必須修正為 *upacara memiliki cara dan prosedur yang berbeda*（直譯：儀式具有不同的方式以及程序），方能正確表達原文意義。

整體而言，越南籍受訪者對於四篇文章之越南文 MT 之大意理解程度約為八成，而印尼籍受訪者對於四篇文章之印尼文 MT 之大意理解程度亦為八成。基本上，導致中文譯入越南文及印尼文之 MT 錯誤的一個重要原因，乃是中文、越南文及印尼文之語序差異，例如越南文受到法文影響，形容詞會放在名詞後，但中文偏好用前置形容詞。此外，有些詞彙語意錯誤，其原因乃是機器無法從上下文判斷其正確意涵，只從表面判斷，會有選錯詞彙的情形，如「有」直譯為“*ada cara*”，應修正為“*memiliki cara*”。無可諱言，當翻譯語料庫愈加龐大時，其 MT 品質會更好（Och, 2005），但作者也注意到，當讀者群不同以及他們的文化背景不同時，對於 MT 譯文大意之理解程度也會有所差異。上述兩名印尼籍人士之父母是臺灣人，所以他們雖然在印尼成長及接受教育，但從父母口中早已了解一些臺灣文化，故其背景知識幫助他們更易於了解介紹臺灣文化之 MT 譯文。至此，我們可推論若臺灣東南亞新住民在臺灣居住久一些，也會增加其文化背景知識，自然會幫助他們更了解 MT 譯文大意；但是閱讀商業及公司訊息，則不太需要許多的文化背景知識。

本研究受訪者數量有其侷限性，原因有二：一為臺灣一般東南亞新住民，雖能了解文章大意，但無法明確指出文法錯誤並建議如何改進；二則為作者與一般東南亞新住民在溝通上有所困難，他們的中文程度只達一般會話水準，無法明確指出並說明母語與中文之語言特性差異所造成機器翻譯錯誤之原因。有鑑於此，作者最後邀請就讀高雄科技大學國際管理碩士學位學程及應外系的學生參與 MT 實驗，除了使用英文溝通之外，他們也能清楚表達機器翻譯錯誤之原因。

有鑑於控制性中文的東南亞語 MT 譯文可達「可知可解」程度，作者建議臺灣政府單位、公司機構在建置網頁文本時，可使用控制性語

言書寫，並連線至線上 MT 系統以譯入東南亞語文，新住民則可選擇其熟悉的母語來閱讀網頁譯文。雖然網頁文本最好以控制性中文書寫，但原先的自然中文網頁不需撤除，只需添加另一個控制性中文文本供使用者選擇即可。有關推廣控制性中文書寫，政府可補助大專院校辦理工作坊，針對教師、社會民眾實施相關訓練；政府亦可以身作則，推出一些簡易中文書寫網站。另一方面政府亦可協助國家教育研究院或中央研究院及大專院校建置東南亞語之翻譯語料庫，並與谷歌公司合作，將資料導入線上谷歌翻譯系統，亦有助於提升東南亞語機器譯文之可讀性及可理解性。

伍、機器翻譯應用之語言人權及民主意涵

新移民來到臺灣，除了享有一般的法律權利外，我們也應考慮其語言人權、多元文化平權與庶民語言運動等議題，俾其能快速了解、擷取當地文化與資訊，以順利適應、融入在地社群。上述文獻考察和訪談研究，證實了控制性文本線上 MT 譯文可提供臺灣東南亞新住民母語閱讀的環境，使其得以快速了解臺灣在地社會、文化等資訊。此突顯了 MT 服務對塑造語言人權、多元文化平權與庶民語言環境的效益與重要性，也有助於「語言多樣化，紓解族群間的張力，並且推動包容性民族主義（inclusive nationalism）」（施正鋒，2007，頁 175）。

一、語言人權

MT 提供臺灣外籍移民或東南亞新住民在網路上無限制使用其母語，基本上是賦予新住民基本語言人權，可視為實踐語言正義之默許與期許。一些學者指出，目前國際潮流逐漸重視語言正義議題，因為語言權利即是一種基本人權，兩者合在一起乃是語言人權（language rights），此種權利允許個人有權利在公共場合中使用及學習其母語（施正鋒，2002；De Varennes, 2001；Skutnabb-Kangas, 2000；Skutnabb-Kangas

& Phillipson, 1994)，而線上 MT 服務提供東南亞新住民在網路上使用其各種不同的母語，正符合此語言人權潮流。

二、高經濟價值機譯服務

試想一個中文網站譯為七種外國語言，如英文、日文、越南文、印尼文、韓文、德文及法文，就算一字一元計算，成本花費很大，況且網頁文本又需不斷更新，且多數網路使用者只需取得文本翻譯之大意，故 MT 服務可成為另種替代之經濟方案；但為精準 MT 品質，仍需花費在前置控制性語言書寫，一次控制性語言書寫即可同時在網路上翻譯多種不同語言，當可節省大量花費與時間。實際上，由於 MT 系統語料庫日漸增長，加上人工配套措施，其多語服務帶來之高經濟價值不容小覷。

三、多元文化平權

語言離不開文化，故我們使用 MT 服務實踐語言人權時，也是落實文化平權的理念，創造多元文化傳輸、溝通之橋梁。孫春在（2007）曾執行以外籍配偶為對象的多元文化數位學習內容研發研究計畫，其目的乃是建置 16 個數位學習活動，供外籍配偶學習中文、電腦與生活適應等技巧，其語言種類涵蓋東南亞語言，但其內容缺少社會文化、民情風俗層面。作者認為，欲使新住民了解臺灣在地文化與風土民情，可透過 MT 學習不同主題，其知識層面會較為廣泛，故 MT 仍是數位學習之理想工具。一方面，臺灣新住民可使用其母語之機譯快速了解臺灣社會文化；另一方面，臺灣民眾可借助線上 MT 服務了解新住民母國文化及外國文化，而新住民的下一代亦可借助 MT 服務了解與認識其祖先母國文化。這些語言行為無形中表示對於不同族群之多元文化的認同，連帶促成由下而上、屬於大眾層次的文化平權之普遍意識，此可呼應臺灣政府制定國家語言發展法，以達到尊重國家多元文化發展及語言文化多樣性之目的（文化部，2017）。

四、庶民語言運動

臺灣過去數年來推行多語政策之決心明顯進步許多；從解嚴（1987年）之後，歷經李登輝擔任總統（1988年－2000年）推行臺語運動，又至陳水扁擔任總統（2000年－2008年）廣設不同族群電台，實施多語政策；現今臺灣政府有鑑於東南亞新住民的語言平權，也從2010年開始，將越南語、印尼語納入國中小義務教育的選修語言課程。只是官方政府語言政策常受到政治及國家預算等外來因素的干擾，會產生一些變數，故推動線上MT服務，輔以配套教育措施，並鼓勵新住民使用，此乃是從日常生活底層做起，替代國家語言政策的庶民語言運動。

陸、結論

過去，多數人對MT有諸多誤解，甚至認為它在跨語言和跨文化溝通上是敵非友。事實上，MT隨著語料庫日益增大，有如人工智慧，愈來愈進步、也愈來愈聰明，所以我們必須學習如何更有智慧地使用MT，使它能更快速、更完善地幫助我們達到語言服務目的。李怡平（1994）指出，澳門每年有數百萬遊客，且其經濟發展需仰賴從其他國家吸收技術與設備，所以翻譯服務需求十分迫切，因此他呼籲政府能研發葡—英—中的多語MT系統，幫助澳門推行辦公自動化。美國樂圖公司（Lotousa Company）也已成功研制一套發音專業翻譯系統「狂飆譯族」（SPEED EICQ）以提供中英文語言學習及雙向翻譯；此電腦輔助工具可有效地滿足語言發音學習和網上資料及文件翻譯的迫切需要（華美協進社，2003）。有鑑於此，作者建議教育部或科技部可補助國家教育研究院語文教育及編譯研究中心或大專院校教師申請計畫，以建置東南亞語文翻譯語料庫，或補助東南亞語言教師與資訊工程專家合作研發東南亞多語文機譯系統。其次，政府亦可補助大專院校教師推動MT教育，強化學生使用MT技巧，並訓練學生使用控制性中文來書寫臺灣在

地文化、社會、經貿等資訊，以利新住民使用線上 MT 系統譯入東南亞語文，使其能了解臺灣在地文化，並使用 MT 教導他們的孩子學習其祖先的母語。此乃是結合學術資源、大學教育及科技工具應用，以提升多元社會數位語言服務品質之學術貢獻。

在全球化的時代，一個國家只說一種語言、閱讀一種語文及擁有同一種族群身分是不可能的。全球移民數量不斷增加以及社會流動，已促成多語社群的形成。今日，我們所要面對的議題並非阻止移民人口增加，而是應著重如何提供新住民友善的資訊溝通環境，並允許其使用祖先后母語及保存其祖先文化，以創造多語、多文化環境，並豐富臺灣語言文化資產。然而在推動多語政策過程中，由於資源分配和語言競合等因素，往往會使東南亞新住民之弱勢母語停滯、凋零，故為落實語言政策文化平權理想，公司機構及學術文化團體可考慮添加控制性語言之網路文本，以利東南亞新住民及其他外語人士使用 MT 作為語言學習和資訊取得之工具。綜而言之，線上 MT 服務在移民和全球化社會情境中，可協助推動多語溝通，亦是開啟非國家層次之庶民語言運動之利基，更是展現政府面對多語族群溝通之誠意及正面態度。

參考文獻

中文文獻

- 內政部移民署（2015年8月27日）。外僑居留人數統計表。取自 <https://www.immigration.gov.tw/ct.asp?xItem=1301702&ctNode=29699&mp=1>
- 【National Immigration Agency. (2015, August 27). *Foreign residents by nationality*. Retrieved from <https://www.immigration.gov.tw/ct.asp?xItem=1301702&ctNode=29699&mp=1>】
- 文化部（2017年7月3日）。國家語言發展法草案。取自 https://www.moc.gov.tw/information_253_65745.html
- 【Ministry of Culture. (2017, July 3). *Draft national language development law*. Retrieved from https://www.moc.gov.tw/information_253_65745.html】
- 史宗玲（2013）。網頁書寫新文體：跨界交流「快譯通」。臺中：白象文化。
- 【Shih, C. L. (2013). *Alternative web textual writing: Effective communication across borders*. Taichung, Taiwan: Elephant White Cultural.】
- 李怡平（1994）。機器翻譯及澳門的研究。取自 <https://www.macaudata.com/macabook/book138/html/11801.htm>
- 【Li, Y. P. (1994). Machine translation & current research in Macau. Retrieved from <https://www.macaudata.com/macabook/book138/html/11801.htm>】
- 南橋專欄（2016年10月2日）。機器翻譯將優化並擴大翻譯市場。取自 <https://read01.com/GEQGNd.html>
- 【Nangqiao Columns. (2016, October 2). Jiqi fanyi jiang youhua bing kuoda fanyi shichang. Retrieved from <https://read01.com/GEQGNd.html>】
- 施正鋒（2002）。語言權利法典。臺北：前衛。

【Shih, C. F. (2002). *Yuyan quanli fadian*. Taipei, Taiwan: Avantguard.】

施正鋒（2007）。語言人權。載於鄭錦全、何大安、蕭素英、江敏華、張永利（編），*語言政策的多元文化思考*（頁 171—176）。臺北：中央研究院語言學研究所。

【Shih, C. F. (2007). Yuyan renquan. In J. Q. Zheng, D. A. He, S. Y. Xiao, M. H. Jiang & Y. L. Zhang (Eds.), *Yuyan zhengce de duoyuan wenhua sikao* (pp. 171-176). Taipei, Taiwan: Institute of Linguistics, Academia Sinica.】

科大訊飛 cobbyli、zmwang（2017 年 6 月 15 日）。機器翻譯的前世今生。取自 <http://mini.eastday.com/mobile/170615122116050.html>

【Iflytek cobbyli, zmwang. (2017, June 15). Jiqi fanyi de qianshi jinsheng. Retrieved from <http://mini.eastday.com/mobile/170615122116050.html>】

孫春在（2007 年 11 月 30 日）。以外籍配偶為對象的多元文化數位學習內容研發研究成果報告。取自 <https://ir.nctu.edu.tw/bitstream/11536/90372/1/942524S009004.PDF>

【Sun, C. T. (2007, November 30). *Report on developing multi-cultural digital learning content for immigrant spouses in Taiwan*. Retrieved from <https://ir.nctu.edu.tw/bitstream/11536/90372/1/942524S009004.PDF>】

教育部（2017 年 6 月）。新住民子女就讀國中小人數分布概況統計。取自 http://stats.moe.gov.tw/files/analysis/son_of_foreign_105.pdf

【Ministry of Education. (2017, June). *Xin zhumin zinu jiudu guozhongxiao renshu fenbu gaikuang tongji*. Retrieved from http://stats.moe.gov.tw/files/analysis/son_of_foreign_105.pdf】

陳子昂、黎偉權（1991）。機器翻譯系統現況與展望。CIO 資訊傳真，143，頁 166—171。

【Chen, Z. A., & Li, W. Q. (1991). Current status and prospects of machine translation system. *CW Infopro Weekly*, 143, 166-171.】

華美協進社（2003 年 2 月 14 日）。“狂飆譯族”可發音專業翻譯技術，

中英文語言學習及雙向翻譯的電腦輔助工具。取自 <http://www.chineselectures.org/021404b.html>

【China Institute. (2003, February 14). “SPEED EICQ” ke fayin zhuanye fanyi jishu, zhongyingwen yuyan xuexi ji shuangxiang fanyi de diannao fuzhu gongju. Retrieved from <http://www.chineselectures.org/021404b.html>】

馮志偉 (1995)。自然語言機器翻譯新論。北京：語文。

【Feng, Z. W. (1995). *Ziran yuyan jiqi fanyi xinlun*. Beijing, China: Language & Culture Press.】

鄒劍宇 (2016年3月31日)。「億級翻譯需求」存在嗎？FT中文網。

取自 <http://big5.ftchinese.com/story/001066862?full=y>

【Zou, J. Y. (2016, March 31). “Yi ji fanyi xuqiu” cunzai ma? *Financial Times*. Retrieved from <http://big5.ftchinese.com/story/001066862?full=y>】

謝宇程 (2015年11月4日)。翻轉學語言，該想清楚的三大運勢。親子天下新聞。取自 <https://flipedu.parenting.com.tw/article/1724>

【Xie, Y. C. (2015, November 4). Fanzhuanxue yuyan, gai xiang qingchu de san da yun shi. *Parenting News*. Retrieved from <https://flipedu.parenting.com.tw/article/1724>】

英文文獻

Adam, D. (2013, September 27). *Annual report of the Office of the Commissioner of Official Languages (2000-2001)*. Retrieved from http://www.ocol-clo.gc.ca/html/ar_ra_2000_01_e.php

Beaupre, J. (2010, February). Machine translation: Integrating automation into an effective localization strategy. Retrieved from <http://www.translationdirectory.com/articles/article2115.php>

Bowker, L. (2009). Can machine translation meet the needs of official language minority communities in Canada? A recipient evaluation. *Linguistica Antverpirnsia*, 8, 123-155.

- Brown, P. F., Pietra, S. A. D., Pietra, V. J. D., & Mercer, R. L. (1993). The mathematics of statistical machine translation: Parameter estimation. *Computational Linguistics*, 19(2), 263-311.
- De Varennes, F. (2001). Language rights as an integral part of human rights. *MOST Journal on Multicultural Societies*, 3, 1. Retrieved from <http://www.unesco.org/most/v13n1var.htm>
- DePalma, D. A. (2007). Limited English proficiency not a bar to citizen access. *MultiLingual*, 18(4), 46-50.
- Dillinger, M., & Gerber, L. (2009). Success with machine translation: Automating knowledge-base translation. *ClientSide News Magazine*, 9(1), 10-11.
- Gerber, L. (2008). Recipes for success with machine translation: Ingredients for productive and stable MT deployments. *ClientSide News Magazine*, 8(11), 15-17.
- Hutchins, J. (1986). *Machine translation: Past, present, future*. Chichester, UK: Ellis Horwood.
- Hutchins, J. (2001). Machine translation and human translation: In competition or in complementation? *International Journal of Translation*, 13(1-2), 5-20.
- Juskewycz, H. (2016, November 15). *The future of machine translation and how to use it properly* [web log comment]. Retrieved from <https://lingohub.com/blog/2016/11/the-future-of-machine-translation-and-how-to-use-it-properly/>
- Kroulek, A. (2017, February 28). *A translation showdown: Man vs machine translation* [web blog]. Retrieved from <http://www.k-international.com/blog/human-translation-vs-machine-translation-contest/>
- Lo, P. (2015). *A study of normative controlled language writing for Thai MT application* (Unpublished master thesis). National Kaohsiung First University of Science and Technology, Kaohsiung, Taiwan.
- O'Hagan, M., & Ashworth, D. (2002). *Translation-mediated communication in a digital*

- world*. Clevedon, UK: Multilingual Matters.
- Och, F. J. (2005, September 12). *Statistical machine translation: Foundations and recent advances*. Retrieved from <http://www.mt-archive.info/MTS-2005-Och.pdf>
- Office of the Commissioner of Official Languages. (2005). *Bridging the digital divide: Official languages on the internet*. Retrieved from http://www.officiallanguages.gc.ca/sites/default/files/lang_internet_2005_e.pdf
- SDL. (2012). Intelligent machine translation: The future is here. Retrieved from <http://www.sdl.com/download/intelligent-machine-translation/75989/>
- Shih, C. (2006). *Helpful assistance to translators: MT and TM*. Taipei, Taiwan: Bookman Books.
- Skutnabb-Kangas, T. (2000). *Linguistic genocide in education—or worldwide diversity and human rights?* Mahwah, NJ: Lawrence Erlbaum Associates.
- Skutnabb-Kangas, T., & Phillipson, R. (1994). Linguistic human rights, past and present. In T. Skutnabb-Kangas & R. Phillipson (Eds.), *Linguistic human rights: Overcoming linguistic discrimination* (pp. 71-110). Berlin, Germany: Mouton de Gruyter.
- Somers, H. (2003a). [Introduction]. In H. Somers (Ed.), *Computers and translation: A translator's guide* (pp. 1-12). Amsterdam, Netherlands: John Benjamins.
- Somers, H. (2003b). Translation memory system. In H. Somers (Ed.), *Computers and translation: A translator's guide* (pp. 31-48). Amsterdam, Netherlands: John Benjamins.
- Vasconcellos, M., & Bostad, D. (1992). Machine translation in a high-volume translation environment. In J. Newton (Ed.), *Computers in translation: A practical appraisal* (pp. 58-77). New York, NY: Routledge.
- Wagner, E., Bech, S., & Martínez, J. (2002). *Translation for the European Union institutions*. Manchester, UK: St. Jerome.
- Weaver, W. (1949). Warren Weaver memorandum. *MT News International*, 22, 5-6.