教育部教學實踐研究計畫成果報告
Project Report for MOE Teaching Practice Research Program

計畫編號/Project Number：PED1122127

學門專案分類/Division：教育　次領域：語言教育

申請年度/Project Period：112 年度 ☐一年期　☐多年期

執行期間/Funding Period：2023.08.01 – 2024.07.31

(計畫名稱/Title of the Project)：**人工智慧技術對提升大一英文學生中翻英句子口譯成效的探討**

(配合課程名稱/Course Name)：　英文(附二小時英聽)(二)

計畫主持人(Principal Investigator)：高照明

協同主持人(Co-Principal Investigator)：

執行機構及系所(Institution/Department/Program)：國立臺灣大學外國語文學系

成果報告公開日期：☐立即公開 ☐延後公開

繳交報告日期(Report Submission Date)：2024 年 9 月 20 日

(計畫名稱／Title of the Project)：人工智慧技術對提升大一英文學生中翻英句子口譯成效的探討

1. 研究動機與目的 (Research Motive and Purpose)

口語和寫作公認是台灣英語學習者語言技能中最弱的部分，而相關口語能力的研究並不多見，特別是運用人工智慧技術輔助口譯練習的研究幾乎完全空白。本研究希望能填補這個重要的研究缺口，開發一套能夠訓練中高級程度的大一學生中翻英文句子口譯能力的系統，並評估這個系統的實際成效以及學生對這套的系統的態度。

無論在國內外，對於提升外語口語能力研究，都相當有限，這是因為研究口語能力傳統上需透過人工方式將學習者的語音資料轉寫成文字，這樣的流程不僅費時費力，也無法用於智慧型電腦輔助語言教學，因為這類系統需在幾秒鐘內對學習者的口語表現給予自動評分或自動回饋。隨著語音辨識技術的進展，目前已經可以將使用者的語音資料自動轉寫成文字，並在正確率上達到可以接受的水準。

我們在109年的「大專校院教學實踐研究計畫」-探討語音科技對大一英文口語訓練的成效計畫中開發了一套具有語音辨識和語音合成功能，且能給予學習者自動回饋的英語線上覆誦系統，由於當時的線上錄音程式有一些問題，另外因為採用Google免費語音辨識的API導致有時無法得到語音辨識的結果，在這兩大問題的影響下，造成我們收集受試者的資料不如預期。109年的計畫還有一個問題是儘管覆誦練習(shadowing)在口譯界公認是一個非常有效提升口語能力的方法，但是在現今的英語教學界並不受重視，普遍被認為是行為主義學派外語學習audio lingual的老方法，與強調溝通功能的主流學派背道而馳。在經過數年的改良和反思後，我們決定重新設計當初的系統，將一部份口譯教學的方法納入大一英文口語教學訓練之中，以便實踐政府推動雙語教育的目標，並讓大一學生能夠具備初步的中英文口語轉換的能力。

2. 研究問題 (Research Question)

自動語音辨識（ASR）技術早已廣泛應用於國內外商業領域電腦輔助語言學習（CALL）系統，包括 Tell Me More、EnglishCentral、MyET、VoiceTube。這些系統幾乎都有句子覆誦練習，系統可以根據學習者複述的音檔自動給予回饋或顯示分數。這些系統即運用自動語音辨識技術（ASR）來衡量外語學者的回答與標準答案間的相似程度。最先進的 ASR 技術可以自動將語音轉寫成文字而將錯誤率控制在可以接受的範圍內。這樣的功能已被納入一些創新的線上測驗系統中，例如之前隸屬Pearson集團的Versant，提供包括英文、中文、法文、西班牙文、荷蘭文和阿拉伯文等多種語言的線上口語測驗。Versant技術的核心是使用ASR和自然語言處理對使用者的口語和書面回答進行自動評分，Versant使用包含母語和非母語人士的口說資料和發音特徵的統計模型。然後評估使用者的口語能力。和全民英檢（GEPT）口語測驗一樣，Versant有幾種不同的類型，包括朗讀、完成句子、回答問題、描述圖片和表達對某個主題的想法，不同之處是全民英檢是傳統的人工評分，而Versant則是採用電腦自動評分。

Versant是近年來人工智慧在語言處理技術長足進展的一個明證。目前AI不僅語音辨識和

語音合成技術已經廣泛使用於語言教育與商業應用，在語音辨識和語音合成技術的輔助下也可以語音作為輸入與輸出。著名的人工智慧研究機構Open AI在2022年推出一個功能強大的語音辨識開放軟體Whisper，另外Meta也推出免費的語音辨識和語音合成系統，從以上的科技進展可知，結合語音辨識，語音合成，以及其它人工智慧技術，開發一個具備AI的網路練習系統，來提升大學生的英語口譯能力的系統並不是遙不可及的。本研究主要的目的是利用人工智慧語言科技設計出適合大學中高級程度學生的中翻英句子口譯系統並評估其使用後的實際成效以及學習者對這套系統的態度。

有數篇論文討論了自動語音辨識技術（ASR）在外語學習和測驗的應用。例如，Coniam (1998) 探討ASR 成為英語測驗工具的可能性。 Dalby & Kewley-Port (1999) 討論ASR 技術在外語學習的應用，包括美國人學習西班牙語和中國人學習美語。他們的研究顯示，ASR 衍生的反饋可以改善發音。Luo et al. (2010) 探討音素、清晰度、和韻律流暢度等聲學特徵與學習者覆誦的分數之間的關聯性。Ginther et al. (2010) 討論評量英語口語流利度的方法。Franco et al. (2010) 開發應用於電腦輔助語言學習發音評分的語音辨識工具。Berstein et al. (2010) 提供證據顯示自動口語測試和口語能力面試兩種類型的測試具有很強的相關性。Streeter et al. (2011) 介紹了Pearson 公司在寫作、口語和數學三項測驗自動化評分的技術。這些研究顯示，自動語音識別技術在語言學習和測驗中所扮演的角色越來越重要。Bashori et al. (2022)研究顯示英語作為外語的學生對於英語口說大多有焦慮感，而在使用具有英文語音辨識功能的網站練習後對於英語口說的焦慮感有所降低。Evers & Chen (2022) 探討實驗組和對照組學生在使用SpeechNotes這個語音辨識軟體後，比較同儕互相給予回饋的實驗組和自行練習的對照組學生這兩組學生在朗誦課文發音的正確性上是否有顯著差別，實驗結果顯示同儕互相給予回饋的實驗組表現顯著比較好。

除了 ASR 之外，透過量化指標自動判斷學習者程度這個研究議題上，近年來也取得了重大進展。例如，Crossley & McNamara (2009) 使用計算語言學的方法在第二語言寫作中自動找出不同程度者在詞彙方面表現的差異。McNamara, Crossley, McCarthy (2010) 則提出了一套可以判定寫作程度的量化指標並進行了實驗。Crossley et al.(2011)針對語音資料提出一套可以預測詞彙程度的量化指標。Crossley & McNamara (2013) 依據 244 份 TOEFL口說測驗考生的文字稿，發現跟詞型和詞頻相關的量化指標可以預測人工評分約 61% 左右。 他們的研究清楚顯示，即使在沒有完全正確的語音、語調和重音訊息下，根據考生口說測驗文字稿的資料以自動化方式判定整體口語程度仍然是可行的。前面這幾項研究許多有賴於Coh-Metrix這個可以自動生成詞彙、句法和語義指標的系統。除了Coh-Metrix外，Lu (2010)提出的 14 個評估句法複雜性的量化指標，來評估英語作為第二語言的學習者所產出的作文的句法複雜性，同時利用這一套句法複雜性分析器軟體，來測試中國學習者書面英語語料庫的資料。結果顯示，所提出的指標可以可靠地區別不同程度學習者的作文。Lu (2012) 應用詞彙複雜性分析軟體探索詞彙豐富度與 ESL 英語學習者口語敘述能力兩者之間的關係。Ai & Lu (2013) 根據句子的長度、從屬結構數量、並列結構數量和詞組複雜程度，使用 10 種量化指標來比較母語和非英語母語者的作文。研究結果顯示，這些量化指標不僅可以區分母語人士和非母語人士的作文，還可以區分不同程度的作文。

本研究計畫的目的是研發一套可以評估學習者口譯能力的系統並評估這套系統對學習者學習的影響以及學習者對這套系統的態度與評價。

研究的問題如下：

研究問題 1： 受試者經過練習之後，在後測的表現是否顯著優於前測？

研究問題 2：受試者認為使用本計畫所建立的英語口說(視譯)訓練系統的效果如何？
研究問題 3：受試者是否認為可以在自主學習的情形下利用本計畫所建立英語口說(視譯)練系統與方法來提升英語視譯的能力? 對學習策略和後設認知的影響為何？

研究問題 4: 如何設計一套可以對視譯練習自動評分的AI系統？

3. 文獻探討 (Literature Review)

利用計算語言學的工具與方法自動辨識語言學習者的語法特徵已成為語言學習評估中的重要技術之一。Hawkins 和 Filipovic（2012）使用此技術對不同程度的英語學習者進行了分析，辨別學習者的正確和錯誤語法結構。研究顯示，根據大量測驗資料，能夠有效地計算出學習者掌握的文法結構與常犯錯誤，這對於評估語言能力尤為重要。Alexopoulou et al.（2013）則進一步開發了一個可視化工具，用於探索學習者的正面和負面文本特徵，並且提供了自動分級功能，便於不同程度的學習者進行個性化學習。Crossley 和 McNamara（2009）開發了一套專門用於分析學習者寫作程度的系統，通過計算詞彙多樣性、語義密度和句法複雜性等指標來進行評估。後續研究進一步拓展了該系統的功能，納入了連貫性和思路清晰度的分析，從而為教師提供了更加全面的學習者寫作評估工具（McNamara et al., 2010；Crossley et al., 2011）。這些技術能夠提供即時、個性化的反饋，促進學習者在寫作過程中的進步。事實上，目前的技術不僅輸入文本可以自動得到自動分析與回饋，口語也可以透過語音辨識以及對話機器人得到回饋。例如，Dizon（2017）針對日本大學英語學習者進行的實驗中，使用了 Alexia 智慧型語音助理來提高口語能力。研究表明，儘管學習者的發音影響了語音辨識的準確性，但系統通過間接的回饋機制，仍能有效促進學習者口語溝通能力的提升。Dizon & Tang（2019）探討了這種技術對英語自主學習的影響，雖然參與者對技術持正面態度，但實際使用率相對較低。後續研究（Fryer et al., 2020；Dizon, 2020）證實，長期使用該系統能顯著提升學習者的口語表現。

口譯和視譯訓練在全球化背景下成為語言學習的核心領域之一，是增強外語口說能力的一種重要的訓練方法。蔡小紅（2006）深入研究了中英口譯訓練中的文化因素，強調跨文化溝通對口譯員的重要性。劉敏華等人（2008）的研究則對比了國內外口譯考試命題和評分方法，提出了強化口譯員即時反應和記憶能力的重要性。Peng（2022） 探討了在大學課程中引入語音科技進行視譯訓練的可能性，並收集了學習者對這一技術的正面反饋，顯示技術創新能夠有效提高學習者的學習體驗。

近年來人工智慧（AI）技術逐漸被引入到口譯訓練中，為學習者提供了一種新的模擬環境。Cheung 和 Li（2018） 強調，語音辨識技術在同步口譯中的應用，能夠幫助口譯員即

時處理語音輸入，特別是在處理口音較重的講者時效果顯著。Fantinuoli（2018）編輯的專書彙集了多篇研究，探討了 AI 如何在高壓口譯環境下輔助譯員做出決策。Defranq 等人（2018）的研究顯示，AI 系統能夠顯著提升學習者在口譯訓練中的速度和準確性。Jiang 和 Lu（2021）則提出 AI 不僅能幫助口譯員進行記憶訓練，還能提供即時、個性化的反饋，從而改善學習效果。Pastor（2021）進一步研究了 AI 在未來口譯訓練中的應用，認為 AI 將成為輔助口譯員的核心技術，並有助於降低學習門檻，提高學習效率。

從以上的文獻回顧中可以發現，現代外語學習和口譯訓練已逐漸邁向AI化和自動化。自動語音識別（ASR）技術，語法分析以及AI和計算語言學的應用，可以有效提高外語學習效率和表現。

4. 教學設計與規劃 (Teaching Planning)

本研究的教學的目標是提升學生英語口語表達以及中翻英句子口譯的能力。方法採用自主學習法，學生透過電腦輔助語言學習系統練習。由於目標是發展學生自主學習英語口說（視譯）的能力，教學及訓練的部分都在課後以電腦輔助的方式實施，教師沒有介入。我們請學習者每一個視譯練習試5次。為了訓練學習者口語的流利度，每次錄音的時間設為10秒。超過10秒，就必須重新錄音。學習者上傳視譯練習錄音檔二次之後，即可看到參考答案，接著學習者需要唸出參考答案三次，音檔上傳時，系統會自動播放參考答案的音檔與學習者的錄音檔，讓學習者覺察其中的發音和用詞差異，以加深印象。

5. 研究設計與執行方法 (Research Methodology)

I. 研究架構

我們收集前後測以及每次練習的音檔資料並透過程式分析以及問卷調查和訪談的方式來收集質性和量化的資料。

II. 研究問題意識

我們的研究問題有以下幾項。

(1). 受試者在前測和後測資料有顯著的提升？

(2). 受試者認為使用本計畫所建立的英語口說(視譯)訓練系統的效果如何？

(3). 受試者是否認為可以在自主學習的情形下利用本計畫所建立英語口說(視譯)練系統與方法來提升英語視譯的能力? 對學習策略和後設認知的影響為何？

(4). 如何設計一套可以對視譯練習自動評分的AI系統？

III. 研究範圍目標

(1). 研究範圍
練習的範圍:中翻英句子口譯練習，且限制為中翻英視譯練習，也就是看到中文句子，以英語口譯。
(2). 使用的科技：網路播音以及錄音的元件，資料庫，語音辨識，語音合成，根據深度類神經網路語言模型建立之語義相似度計算，以及其它評估詞彙難度，語法結構，可讀性，流利度，以及詞串豐富度的量化指標。
(3). 歸屬的學科:電腦輔助語言教學。
(4). 研究目標：開發一套能夠訓練中高級程度的大一學生中翻英文句子口譯能力的系統，並評估使用這個系統練習後實際成效以及學生對這套的系統的態度。
影響學習成效的變數：學生的學習動機，每週投入的練習時間，每次練習時間長短，學生入學時的英文成績，學生課後是否有網路可以使用，學生是否有安靜的空間練習。

IV. 研究對象與場域

(1). 研究對象
本研究招募112學年度第二學期在台大修讀大一英文的大一學生24位及台北市立大學一年級同學54位。這些學生大部分受試者閱讀和聽力程度在CEFR B1以上。最後共收集到49位受試者練習(共1013個中翻英視譯練習音檔資料)和問卷資料。

(2). 研究場域: 線上練習的部分在課後實施。

V. 研究方法與工具

(1). 材料：由於系統在接近期末才完成，同學練習的時間只有短短的兩週，因此實驗的材料一共只有 8 個學術詞彙以及包含這 8 個學術詞彙的 40 句的中翻英口譯練習。學術詞彙依據 AWL 學術詞彙表中挑選其中的八個動詞，英文例句和中文翻譯則由 GPT 4.0 Turbo 產生。為了控制難度和結構，我們在提示中清楚說明產生 B2 等級難度的例句，且同一個動詞的練習的結構難度儘可能接近。

(2). 工具：我們針對109年的「大專校院教學實踐研究計畫-探討語音科技對大一英文口語訓練的成效計畫」中，建置的多媒體資料庫中和錄音程式進行修改。受試者登入系統後即可練習。圖 1 練習系統使用者介面。

圖 1 練習系統使用者介面

本研究的系統由三個子系統構成，子系統一語音合成系統負責將文字合成為語音。子系統二語音辨識系統負責將使用者的中翻英文句子口譯的語音轉成文字。子系統三則根據參考答案評估句子翻譯的正確性給予自動評分及回饋。由於自動評分及回饋的功能來不及在學期結束前完成，因此無法收集使用者的意見及資料。實驗未完成的部份，將在新學年度的課程中繼續收集資料。

依據我們參加大學入學考試中心學測和指考英文閱卷的經驗，大考中心採用的中翻英試題評分方式是由人工方式收集所有可能的答案，再根據以下原則評分。首先將一個句子分成幾部分，每個部分分別根據計分原則給分。由於這種計分方式，每題的計分原則都要寫成規則再轉成程式，複雜性頗高，不但很難自動化，且難以設計一個通用的模組，我們沒有採用這種方式。

翻譯和口譯的答案不只一個，只要文法和語義都對，不需要跟參考答案完全一樣，所以只靠比對參考答案的這種評分方法不夠周延，因為很容易漏掉沒有列在參考答案但卻是正確的答案。另一個經常被用來作為評估機器翻譯系統品質的工具是Bleu (Papineni, et al., 2002)。Bleu計算與參考答案的相似度主要是靠n連詞(ngram)的比對，n＝1時比對的單位就是個別的詞。如果n＝2時比對的單位是2連詞，n＝3時比對的單位是3連詞。Bleu以及衍生的算法會對不同長度的ngram給予不同的權重。例如，最簡單的算法是將1連詞，2連詞，3連詞的權重都設為1/3。之所以要考慮給予不同長度的n連詞具有一部份的權重的原因是，如果只考慮個別的詞出現與否，不考慮其它n連詞，那麼詞的順序不同，所造成句法或語義的錯誤就無法涵蓋。將不同長度n連詞納入計算，給予一部份權重等於把一部份有關於詞組或句法的資訊納入考量。假設 1(a)是標準答案，1(b)和1(c)是學生練習時回答的答案，根據NLTK裡面內建的sentence_bleu函數，1(b)和1(c)的分數分別為0.750和0.485。

(1).

(a). The quick brown fox jumped over the lazy dog.

(b). The fast brown fox jumped over the lazy dog.

(c). The fast brown fox jumped over the sleepy dog.

與minimal edit distance類似的是，Bleu只根據參考答案計算與答案之間的相似度，對於沒有列在參考答案但完全正確的答案，則會嚴重低估其分數。換言之，minimal edit distance和Bleu只適合計算與參考答案在字串層次的相似度，無法用來評估一個字串是否與參考答案語義上相似。目前在自然語言處理研究比較適合用來計算語義相似度的工具是使用paraphrase model或句向量模型來計算學習者輸出的口譯與答案的相似度，這必須仰賴近幾年深度類神經網路在自然語言處理領域的突破才有可能。現在AI的方法越來越接近通解，自然語言處理和影像辨識使用的理論和工具都很類似。換句話說，AI已經越來越偏重類神經網路數學模型的泛化能力，特定領域的專業知識已經不是AI難以突破的障礙。以前AI所謂的專家系統，現在被更廣泛抽象的知識表達方式（例如，深度類神經網路的語言模型）和演算法所取代。我們比較了三種目前常用計算語義相似的工具，分別是Huggingface sentence-similarity，spacy_sentence_bert函式庫以及sentence-transformers/paraphrase-distilroberta-base-v2這個paraphrase的模型。我們根據(2a)這個句子為基準，分別以三種模型比較5個不同的句子。實驗的結果(2)顯示最後一個模型最符合我們的直覺。經過多次的實驗，我們已經找出較佳的自動評分模型，可以輸入學習者的音檔，自動給予評分。評分標準則參考全民英檢中高級測驗口說能力測驗以及中翻英測驗的評分標準。

(2).

(a). He collapsed from loss of blood.

(b).

使用的模型	He lost consciousness.	He fainted due to his injury.	He lost his balance.	He lost his head.	He lost his way.
sentence-similarity	0.989	0.979	0.978	0.987	0.979
spacy_sentence_ber using en_stsb_roberta_large	0.494	0.722	0.529	0.463	0.347

sentence-transformers/paraphrase-distilroberta-base-v2	0.740	0.733	0.592	0.581	0.455

我們根據(2b) 可以發現透過深度類神經網路的模型將句子以向量表示的 sentence-transformers/paraphrase-distilroberta-base-v2這一個paraphrase模型可以大致顯示兩個句子之間的語義相似度關係。相似度值越高，兩個句子之間語義也越接近。我們將這一個工具做為計算兩個句子之間語義相似度的基本工具。

我們以sentence-transformers/paraphrase-distilroberta-base-v2這個深度類神經網路的語言模型計算得到的語義相似度作為和NLTK sentence_bleu函數得到的分數作為互補的工具，作為提供給學習者自動回饋的訊息。這些訊息包括語義相似度，與參考答案比較後得到最高bleu的分數，以及minimal edit distance演算法與參考答案比較後得到錯誤的部分。這些訊息可以提供學習者改善的方向。如前所述，由於自動評分系統來不及在學期結束前完成，我們的實驗並沒有包括自動評分及自動回饋這一部分。

我們利用一系列的軟體來評估給學習者的中翻英的練習材料，包括詞彙難度，語法結構，可讀性等各項量化指標是否大致在同一難度，以及學習者練習產出的資料是否在這幾個面向都顯示有顯著的進步。由於現有的工具，在測量中文句子的詞彙難度，語法結構，可讀性不如英文工具可靠，多元，且有系統，我們根據中翻英練習的參考答案，也就是以英文句子來評估詞彙難度，語法結構，和可讀性三個面向的量化指標，原則上，我們挑選詞彙難度，語法結構，和可讀性這三個面向的難易度量化指標比較接近的例句且適合B2學習者的材料讓學習者練習。

詞彙難度指標主要將以West (1953)編纂的General Service List (GSL)當中最常用的1000詞以及次常用1000詞出現的比例，作為判斷句子詞彙的最重要特徵。GSL雖是早在1950年代即已經編成，但是至今仍是學界公認英文最常用2000詞的參考詞表。包括Tom Cobb的Vocabulary Profiler (參見3b, 4b)，Tom Cobb 及Paul Nation 開發的RANGE程式，或是Laurence Anthony所開發的AntWordProfiler，都包含GSL詞表的2000詞，以及相關的衍生詞約4千餘個。以上這三個工具也都包含Academic Word List (AWL) 所涵蓋的570個詞，以及2千多個衍生詞。我們以這幾個常用詞表的指標作為判斷詞彙難度的最基本量化指標，同時也參考其它相關的工具，例如Lu (2012)提出的詞彙複雜度分析器Lexical Complexity Analyzer 裡面的25種量化的指標。句法結構我們以Berkeley Neural Parser作為主要的工具(參見3d, 4d)並參考Lu (2009，2013) 。其中D-Level Analyzer 可以根據句法結構區分8種程度，而L2 Syntactic Complexity Analyzer則有14種量化指標。我們用Tgrep這個擷取句法結構樹的工具判斷從屬子句，關係子句，名詞子句等較複雜的結構。至於可讀性，我們使用readability 0.3.1這個Python套件裡面有包含Kincaid、Flesch、GunningFogIndex、Coleman-Liau、SMOGIndex、LIX、RIX、ARI、

DaleChallIndex等9種可讀性公式(參見3d, 4d)。關於流利度的評估，我們使用readability 0.3.1這個Python套件裡面有關於詞彙數，音節數(參見3e, 4e)除以音檔的時

間作為計算流利度的指標。由於詞彙長短不一，我們以學習者口說練習的音檔裡面音節的總數除以時間長度，即可得到平均每秒說幾個音節。

(3).

(a). Sentence

The building collapsed after the earthquake.

(b\). Vocabulary Profile

	Families	Types	Tokens	Percent
K1 Words (1-1000):	4	4	5	83.33%
Function:	...	...	(3)	(50.00%)
Content:	...	...	(2)	(33.33%)
> Anglo-Sax	...	...	(2)	(33.33%)
K2 Words (1001-2000):				0.00%
> Anglo-Sax	...	...	()	(0.00%)
1k+2k			...	(83.33%)
AWL Words:	1	1	1	16.67%
> Anglo-Sax	...	...	()	(0.00%)
Off-List Words:	?			0.00%
	5+?	5	6	100%

Current profile	
%	Cumul.
83.33	83.33
0.00	83.33
16.67	100.00
0.00	100.00

Words in text (tokens):	6
Different words (types):	5
Type-token ratio:	0.83
Tokens per type:	1.20
Lex density (content words/total)	0.50
Pertaining to onlist only	
Tokens:	6
Types:	5
Families:	5
Tokens per family:	1.20
Types per family:	1.00
Anglo-Sax Index: (A-Sax tokens + functors / onlist tokens)	%
Greco-Lat/Fr-Cognate Index: (Inverse of above)	%

(c). Readability

Kincaid', 4.45), ('ARI', 11.40), ('Coleman-Liau', 16.51), ('FleschReadingEase', 73.86), ('GunningFogIndex', 9.07), ('LIX', 56.0), ('SMOGIndex', 8.48), ('RIX', 3.0), ('DaleChallIndex', 9.20)]

(d). Syntactic Structure

(S (NP (DT The) (NN building)) (VP (VBD collapsed) (PP (IN after) (NP (DT the) (NN earthquake))))

(e). Statistics

[('characters_per_word', 6.33), ('syll_per_word', 1.5), ('words_per_sentence', 6.0), ('sentences_per_paragraph', 1.0), ('type_token_ratio', 1.0), ('characters', 38), ('syllables', 9), ('words', 6), ('wordtypes', 6),

('sentences', 1), ('paragraphs', 1), ('long_words', 3), ('complex_words', 1), ('complex_words_dc', 2)])

---

(4)

(a). Sentence She collapsed after hearing about her husband's death.

(b). Vocabulary Profile

	Families	Types	Tokens	Percent
K1 Words (1-1000):	6	7	7	87.50%
Function:	...	...	(4)	(50.00%)
Content:	...	...	(3)	(37.50%)
> Anglo-Sax	...	...	(2)	(25.00%)
K2 Words (1001-2000):				0.00%
> Anglo-Sax	...	...	()	(0.00%)
1k+2k			...	(87.50%)
AWL Words:	1	1	1	12.50%
> Anglo-Sax	...	...	()	(0.00%)
Off-List Words:	?			0.00%
	7+?	8	8	100%

Current profile	
%	Cumul.
87.50	87.50
0.00	87.50
12.50	100.00
0.00	100.00

Words in text (tokens):	8
Different words (types):	8
Type-token ratio:	1.00
Tokens per type:	1.00
Lex density (content words/total)	0.50
Pertaining to onlist only	
Tokens:	8
Types:	8
Families:	7
Tokens per family:	1.14
Types per family:	1.14
Anglo-Sax Index: (A-Sax tokens + functors / onlist tokens)	%
Greco-Lat/Fr-Cognate Index: (Inverse of above)	%

(c). Readability

'Kincaid', 4.96), ('ARI', 6.62), ('Coleman-Liau', 10.31), ('FleschReadingEase', 75.50), ('GunningFogIndex', 8.04), ('LIX', 42.33), ('SMOGIndex', 8.478), ('RIX', 3.0), ('DaleChallIndex', 7.60)

Vocabulary Profile

(d). Syntactic Structure

(S (NP (PRP She)) (VP (VBD collapsed) (PP (IN after) (S (VP (VBG hearing) (PP (IN about) (NP (NP (PRP$ her) (NN husband) (POS 's)) (NN death)))))))) (. .))

(e). Statistics

'characters_per_word', 5.0), ('syll_per_word', 1.44), ('words_per_sentence', 9.0), ('sentences_per_paragraph', 1.0), ('type_token_ratio', 1.0), ('characters', 45), ('syllables', 13), ('words', 9), ('wordtypes', 9), ('sentences', 1), ('paragraphs', 1), ('long_words', 3), ('complex_words', 1), ('complex_words_dc', 2)]

如前所述，我們根據例句的詞彙難度，語法結構，和可讀性三個面向的量化指標，選出適合B2學習者且難度一致的題目作為練習。至於學生的答案，也以這些工具來評估他們的輸

出詞彙難度，語法結構，和可讀性三個面向的量化指標是否顯示顯著的進度。此外，也另外再搭配AI語義相似度計算的軟體，自動評估語義是否與答案類似。

VI 研究實施程序

步驟：

(1). 告知學生本研究的目的以及進行的方式，參與者的權益並請同意學參與這項研究的學生簽署同意書。 不同意者的資料不予收錄，但因為練習屬於課程的一部分，所以仍須做練習。

(2). 所有參與這項實驗的學生參與前測

(3). 介紹我們線上練習系統的使用方式，以及注意事項（避免在吵雜的環境使用系統）並讓學生建立並登入帳號練習。提供時間讓學生發問以及回答問題。

(4). 每位同學進行為期兩週的自主練習，使用本計畫所開發的系統練習中翻英視譯，共8個學術單字，40句中文翻譯。學術單字相同，練習的句子也刻意選擇結構及字彙難度類似的句子來翻譯。

(5). 進行後測。

(6). 評分者針對前測和後測進行人工評分。

(7). 系統根據前測和後測資料跑有關詞彙難度，語法結構，可讀性，詞串的豐富性，以及流利度等五個面向的量化指標，並進行統計分析。

(8). 根據人工評分量化指標，並進行統計顯著性檢定。

(9). 受試者問卷調查及訪談。

(10). 根據受試者問卷調查及訪談以及統計顯著性檢定的結果回答研究的問題。

6. 教學暨研究成果 (Teaching and Research Outcomes)

(1). 教學過程與成果

以下我們根據本計畫研究的成果依序回答本計畫四個研究問題。

研究問題 1： 受試者經過練習之後，在後測的表現是否顯著優於前測？

由於我們訓練的時間只有兩週且視譯的句子只有包含 8 個動詞學術詞彙的 40 句，受試者經

過練習之後，在後測的表現並沒有顯著優於前測。雖然如此，實驗的數據顯示，在練習的過程中，儘管視譯的難度稍微有增加，但是平均表現仍然非常類似，理論上難度增加，表現應該變差，但是實驗的結果顯示同一個動詞的視譯表現差異並不大。例如，表 1 是 distribute 視譯 5 個例句得分統計，可以發現差異並不大。背後的意義可能是因為我們產生的單字和結構較類似，受試者在反覆練習的過程中掌握了類似單字和結構的視譯。這一部分仍需要更進一步的實驗才能證實。

表 1 distribute 視譯 5 個例句得分統計

視譯句子編號	1	2	3	4
平均數	3.91	4.07	3.95	3.97
中位數	4.5	4.5	4.5	5.0
眾數	5.0	5.0	5.0	5.0
標準差	1.39	1.24	1.59	1.75

研究問題 2：受試者認為**使用本計畫所建立的英語口說訓練系統的效果如何？**

根據問卷調查的結果(圖 2)顯示超過 7 成的學習者肯定這個練習系統可以提升口語流利度。

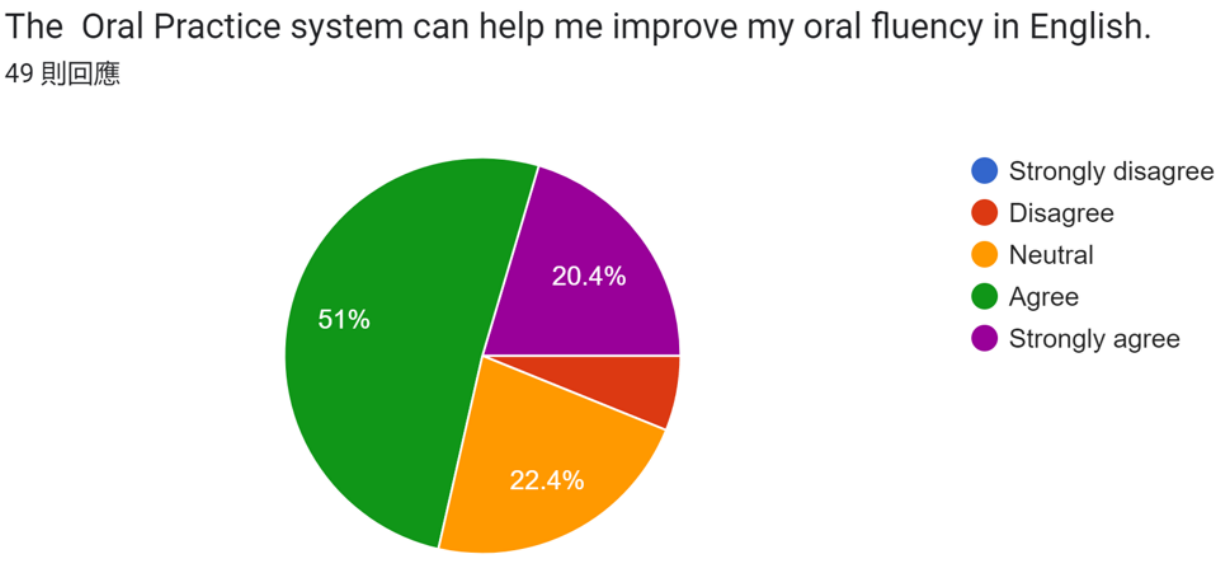


圖 3 的問卷結果顯示超過成八的學習者同意透過這個練習系統比較母語者和學習自己的發音後可以糾正自己的發音。

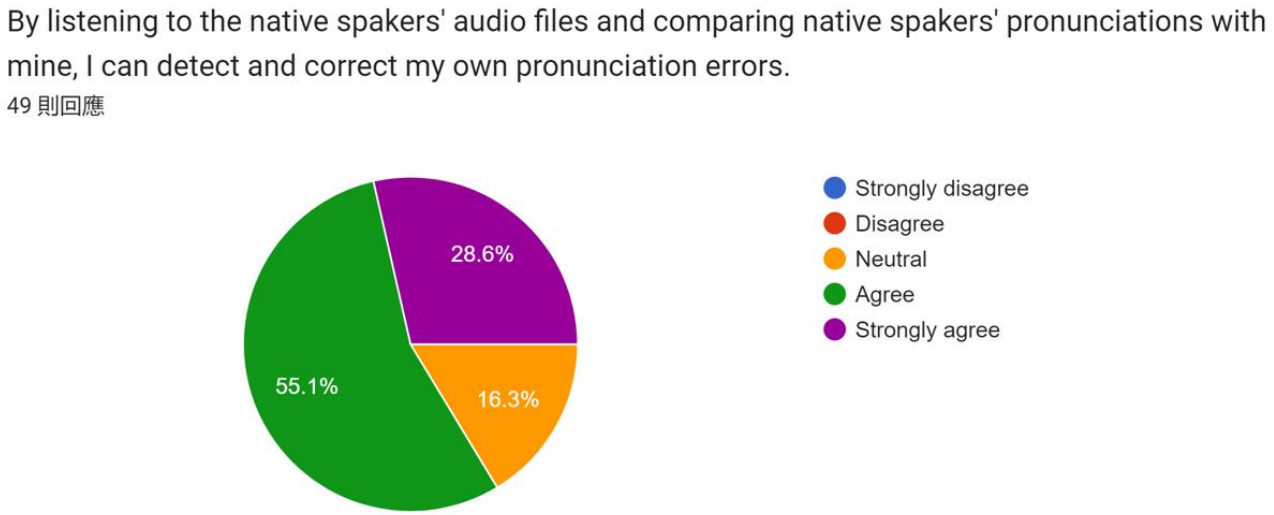


圖 2 及圖 3 的問卷結果進一步證實超過 7 成的受試者同意使用這套系統後，對學習策略和後設認知產生影響。圖 3 也顯示七成的受試者同意他們可以在自主學習的情形下利用本計畫所建立英語口說訓練系統與方法來提升英語視譯的能力。

研究問題 3：受試者是否認為可以在自主學習的情形下利用本計畫所建立英語口說訓練系統與方法來提升英語視譯的能力？對學習策略和後設認知的影響為何？

除了以上問卷調查的量化證據，受試者在問卷調查所提供的文字回饋意見也顯示一致的結果。也就是大部分受試者肯定本研究計畫所設計的口說系統和學習方法，可以有效提升英語口語能力，特別在詞彙和句型的學習，此外，大部分受試者也同意藉由這套口說系統和學習方法看可以在自主學習的情形下，自我糾正發音的錯誤。問卷調查結果反應我們設計的系統和提出的方法對者的學習學習策略和後設認知產生正面的影響，特別是在詞彙和句法以及發音錯誤的自我糾正。受試者回饋的意見在呈現在(3) 學生學習回饋這一節。

圖 4

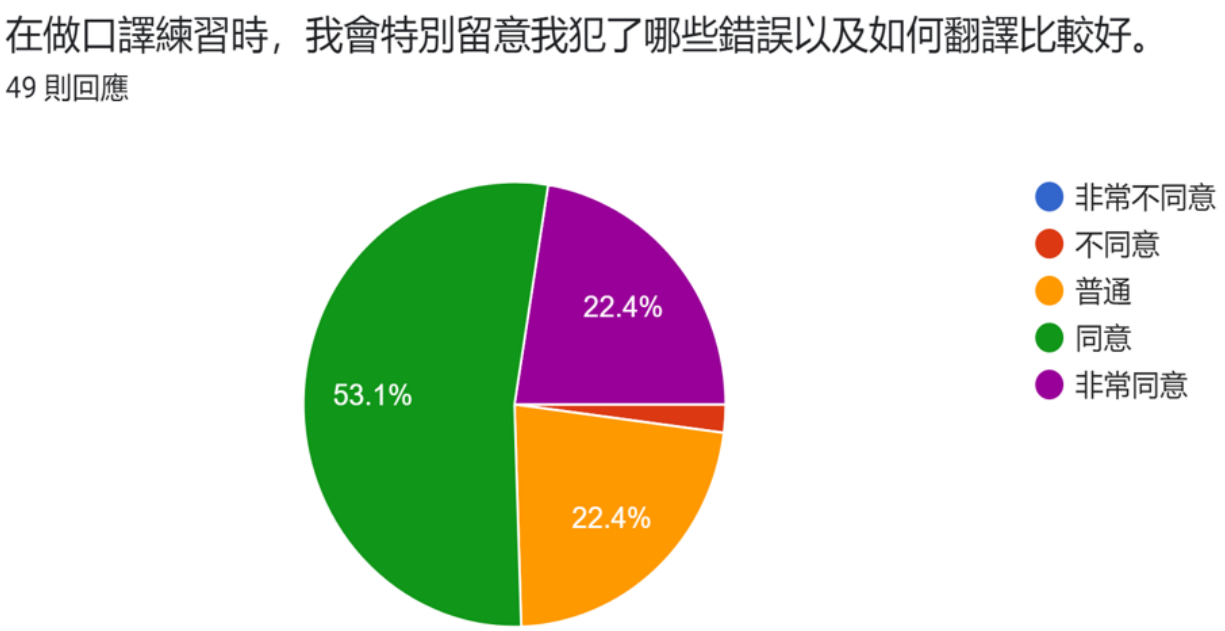
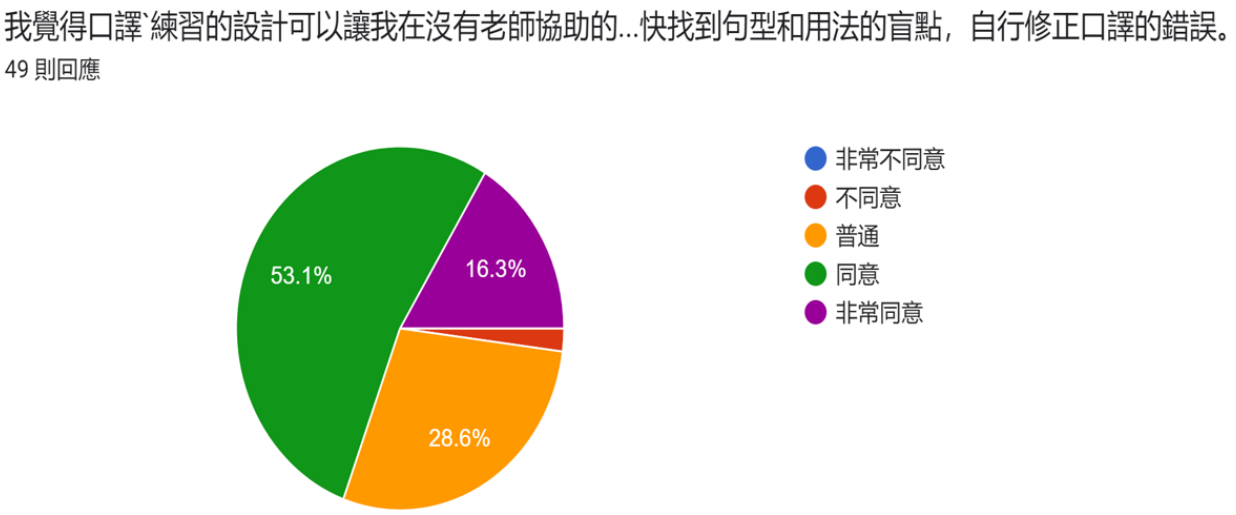


圖 5



研究問題 4 如何設計一套可以對視譯練習自動評分的 AI 系統？

由於自動評分系統相當複雜，需要收集受試者的練習音檔以及人工評分資料，我們在暑假完成自動評分功能。自動評分以及及時回饋的功能有助於增進學習者學習的動機並提升學習的效果。此一部分來不及在本年度的教學實驗中進行實驗，有待後續的研究證

實。圖 6 是程式跑評分模型的過程。目前自動評分模型正確率已經達到實用程度。實驗的結果顯示 Random Forest 演算法的評分模型最佳，0－5 分評分，分數完全預測正確有 64.5%，一分之內的有 96.1%，但是我們可以觀察到仍然有極少數的資料差距 2 分以上，甚至到 4 分。這部分的數據顯示演算法仍然有一些改良的空間，在引進學習者聲學特徵包括停頓，語速，speech repair 等可望進一步改良演算法的正確率。

	human_score	bleu_score	st_score	spacy_bleu_score	difficult_words	lexicon_count	dale_and_chall
0	4.0	4.797597e-78	0.890214	5.518020e-78	1.0	8.0	46.61
1	1.0	3.506226e-155	0.306182	3.413983e-78	0.0	3.0	61.93
2	4.0	5.341736e-01	0.951509	6.690484e-01	2.0	8.0	34.73
3	5.0	5.341736e-01	0.985151	6.690484e-01	2.0	8.0	34.73
4	4.0	2.979715e-01	0.863045	5.384952e-01	3.0	7.0	18.46
5	4.0	5.614022e-78	0.955887	6.954173e-78	1.0	7.0	45.60
6	0.0	0.000000e+00	0.071694	0.000000e+00	0.0	0.0	64.00
7	4.0	7.320648e-78	0.938827	5.154487e-01	1.0	6.0	44.03
8	4.0	5.728717e-78	0.894281	5.728717e-78	2.0	6.0	28.19
9	5.0	5.506953e-01	0.838758	7.365443e-01	2.0	7.0	32.03
10	5.0	1.000000e+00	1.000000	1.000000e+00	2.0	9.0	36.68

圖 6 程式跑評分模型的過程

表 2 Random Forest 演算法的評分模型的正確率

模型預測的分數與實際分數的差距	在差距的分數內所涵蓋樣本資料的百分比
0	64.5%
1	96.1%
2	97.6%
3	97.6%
4	98.1%
5	100%

圖 7　顯示影響分數最多的兩個特徵，分別是 sentence transformers 的語義相似度，學習者英文翻譯與參考答案兩者之間詞性標記順序的 Bleu 相似度。

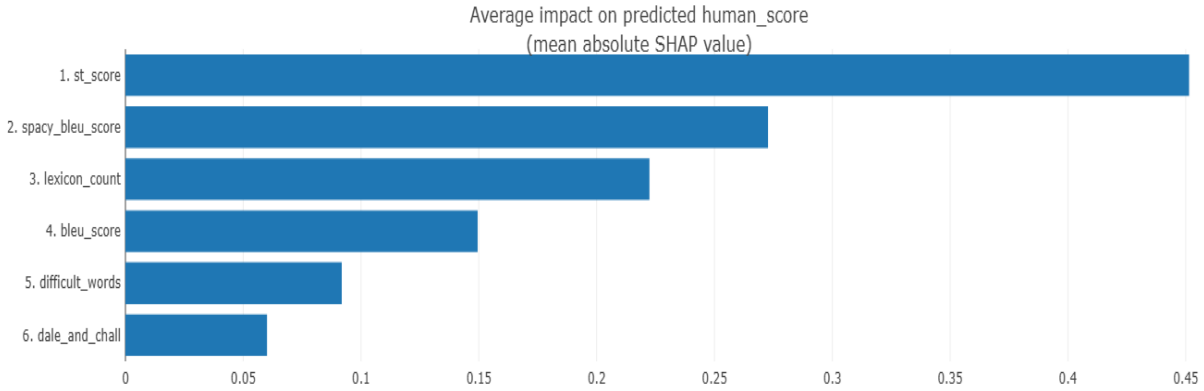


圖 7 評分模型中影響分數最多的幾個特徵。

(2). 教師教學反思

本計畫雖然大致達成預期目標，但是系統完成的時間太晚，導致學期最後兩週才開始進行教學實驗，如果訓練時間可以延長到一學期，甚至一學年，相信受試者在後測的表現很有可能顯著優於前測。另外，前後測句型和字彙複雜度的控制仍有改善的空間。這些變項如能控制得更好，實驗的成果就更能複製及推廣。

(3). 學生學習回饋

以下資料取自問卷調查中的文字回饋意見，可以觀察到受試者的意見大部分都是正面的。

- 介面簡單容易上手，能夠以錄音及語音正確唸法，在單人情況下**找到口說的問題**。
- 很有用，**可以馬上比對自己的發音**
- 我覺得這個系統還不錯用，雖然每一個單字或是句子都要唸五次會覺得有點疲乏，但是透過一次又一次仔細去比較自己的發音和母語者有何不同之處後，漸漸懂得如何改正自己的發音。平常自己練習英文時比較不會去注意這些發音的小細節，但透過這個系統讓我**可以發現自己的錯誤在哪，增進自己的口說能力**。且透過自己先翻過的英文句子後，再比對與正確答案的異同之處，**可以從中學習到像是一些單字常用的搭配**

詞、或是一些時態的小細節等等。

- 我覺得口說系統非常好用，不僅**可以幫助我改正我的發音，在練習句子時也能使我反思句子還能有哪些說法**。
- 我覺得用這種方式確實能讓我**學習到一些新的單字以及它們的發音**，但是有有些音檔的發音也沒有有很清晰，所以會讓我有點難反應過來，至於自己翻譯並嘗試唸出來的部分我覺得很好，有練習的效果
- 滿喜歡的，我的英翻能力比較薄弱，因此先試著自己翻，然後有答案可以參考，對我來說**既有挑戰性又不會有太大壓力**，我覺得這是很棒的練習
- 我覺得**可以學到更多不同的翻譯法**，很不錯
- 五次的練習有點花時間，但**我覺得中翻英的練習真的有對即時翻譯還有句型架構有幫助**，系統還是有一些容易當掉的時候，或是沒有正確計數練習的次數，但**整體我覺得是很有用且不會太枯燥**
- 我**從其中的音檔可以學習其語速及重音**，並且不斷反覆練習，錄音至第五次時幾乎能跟示範音檔的時間長度一致。慢慢習慣這種速度**能讓我的英語口說越來越自然及順暢**。
- **系統能有效幫助我們修正發音**。我認為口譯練習非常好，**因為可以先自己想一遍，之後再看答案，能有效率的學習英文句型**。不過網站系統蠻常當機的，重刷網站有時候會轉不出來原本該有的頁面。
- 口說的內容沒有很艱深卻是容易發音錯誤的內容，我**藉此糾正長期念錯的發音，利用重複練習記在腦海中**。
- 除了要一直重新整理有點麻煩之外，其他方面是個好用的程式
- 有些字的發音比較奇怪，且句子有些唸得快有些唸得慢，會比較難模仿，但**大致上都能讓我練習正確讀音與口譯的技巧**。
- 練習口說不確定怎麼說還是要先說出口，好處是**幫助我對組織語言的過程進行修正**。第一次不對，第二次知道怎麼說，之後漸漸熟練。只有少數錄音檔有問題，**幾乎所有錄音都對我修正發音和流暢度有幫助，因為我可以短時間內多次模仿。句子口譯的示範錄音則提供我英語的慣用法，我覺得很有幫助**。我覺得他如果漸漸完善，**最大優點**

是回饋很快速，可以吸住人的注意力。跟打電動一樣，一場又一場，成功的話可以縮短練習的時空。一旦愈完善，人也會愈喜歡用，因為人喜歡進展的感覺，虛擬世界使人上癮就是因為他回饋比現實世界快，打遊戲很快得到虛擬的成就感，追劇有虛擬的慰藉。如果學習也這麼快有回饋，我覺得他就是提高學習意願的和效果的工具。

- **在重複的練習下，有些我可能容易唸錯或唸不標準的發音，也逐漸進步且更加熟悉**。
- 我認為**重複的練習可以讓我自己糾正發音語調上面的錯誤並自行修正**。
- **口說系統可以即時找到自己和母語人士發音的不同之處**，我覺得很棒!
- 其實還不錯 那個翻譯得很酷
- 在使用英語口說系統時，題目的單字有部分是我平常比較不熟悉的，在讀的時候也就相對容易放錯重音，不過**透過比較音檔跟自己唸的版本就可以非常快速的被糾正出來，我也可以很快速地抓到錯誤點改正**。然而，當真的要把單字用在句子裡唸的時候，我發現我還是常常會卡住，證明我在口說的流暢程度上還有不足的地方需要加強。 另外，**因為錄音的第一遍要先自己試翻看看句子，我需要先去了解這個單字的意思跟用法才能造出句子，所以儘管是口說練習，我還是學習到了很多新的單字和其用法，完成作業後真的受益良多！**

7. 建議與省思 (Recommendations and Reflections)

傳統上認為一對一的口說能力的訓練效果最好，然而在課堂上受限於學生人數太多，師生比過高，一整節課一個學生或許只能輪到一次回答的機會。AI 的科技有可能改善這個長久以來的問題。本計畫應用最新的 AI 科技，結合傳統的機器學習和最新的深度學習技術，透過人工評分 1013 份受試者的音檔，發展自動評分和自動回饋功能。雖然這項新技術來不及在學期當中提供給受試者使用，但是根據問卷調查的結果，即使在教學實驗進行時的系統尚未具備自動評分功能，但是系統要求受試者先自行嘗試視譯句子並上傳音檔的 2 次過程中，受試者必須先產出視譯的思考如何翻譯再說出來，受試者如果

講太慢，導致音檔超過 10 秒，那麼必須重新練習再上傳，當練習達到兩次時，受試者一上傳音檔即可聽到參考答案的音檔，受試者比較兩者的差異即可發現自己發音和句型單字哪些地方需要修正。由於練習的句子刻意選擇語義即以及結構類似句子，受試者反覆練習即可掌握相關的單字和句型。本計畫限於執行期間太短，無法優化系統的功能，未來加入自動評分及回饋功能後，並將訓練時間拉長，加入一些遊戲化的元素，讓受試者有競賽的感覺，可望增強學習者的動機，進一步強化學習的效果。

8. 參考文獻(References)

胡家榮、廖柏森 （2009）臺灣大專中英口譯教學現況探討， 編譯論叢第二卷 第一期，pp. 151-178.

劉敏華、張嘉倩、吳紹銓（2008）口譯訓練學校之評估作法：臺灣與中英美十一校之比較。編譯論叢第一卷第一期，pp.1-42.

蔡小紅（2006）口譯評估. 中國對外翻譯出版公司.

Ai, Haiyang & Lu, Xiaofei (2013). A corpus-based comparison of syntactic complexity in NNS and NS university students' writing. In Ana Díaz-Negrillo, Nicolas Ballier, and Paul Thompson (eds.), *Automatic Treatment and Analysis of Learner Corpus Data*, pp. 249-264. Amsterdam/Philadelphia: John Benjamins.

Allen, D. (2009). Lexical Bundles in Learner Writing: An Analysis of Formulaic Language in the ALESS Learner Corpus. *Komaba Journal of English Education*, 1, 105-127.

Bashori, Muzakki, van Hout, Roeland, Strik, Helmer & Cucchiarini, Catia. (2022). Web-based language learning and speaking anxiety. Computer-assisted Language Learning, Volume 35, 2022 - Issue 5-6, pp. 1058-1089.

Chall, Jeanne Sternlicht, & Dale, Edgar (1995). Readability revisited.

Coxhead, Averil (2000). A New Academic Word List. TESOL Quarterly, 34(2): 213-238.

Crossley, S. A. & McNamara, D. S. (2009).Computationally assessing lexical differences in second language writing. *Journal of Second Language*

Writing, 17 (2), 119-135.

Crossley, S. A., & Salsbury, T. (2011). The development of lexical bundle accuracy and production in English second language speakers. *IRAL: International Review of Applied Linguistics in Language Teaching, 49* (1), 1-26.

Crossley, S. A., Salsbury, T., McNamara, D. S., & Jarvis, S. (2011). Predicting lexical proficiency in language learners using computational indices. *Language Testing, 28*(4), 561-580.

Crossley, S. A. & Crossley, S. A., & McNamara, D. S. (2013). Applications of Text Analysis Tools for Spoken Response Grading. *Language Learning & Technology, 17* (2), 171-192.

Cheung, Andrew K. F. & Li, Tianyun. (2018). Automatic speech recognition in simultaneous interpreting: A new approach to computer-aided interpreting. Ewha Research Institute for Translation Studies International Conference 2018At: Ewha Womans University.

Defranq, Bart , Desmet, Bart , Vandierendonck, Mieke. (2018). Simultaneous interpretation of numbers and the impact of technological support. In

Fantinuoli, Claudio (ed.). 2018. Interpreting and Technology (Translation and Multilingual Natural Language Processing 11). Berlin: Language Science

Press.

Dizon, G. (2017). Using intelligent personal assistants for second language learning: a case study of Alexa. TESOL Journal, 8(4), 811-830.

https://doi.org/10.1002/tesj.353

Dizon, G., & Tang, D. (2019). A pilot study of Alexa for autonomous second language learning. In

F. Meunier, J. Van de Vyver, L. Bradley & S. Thouësny (Eds), CALL and complexity – short papers from EUROCALL 2019 (pp. 107-112). Research-

publishing.net. https://doi.org/10.14705/rpnet.2019.38.994

Dizon, G. (2020). Evaluating intelligent personal assistants for L2 listening and speaking development. Language Learning & Technology, 24(1), 16－26.

https://doi.org/10125/44705

Evers, Katerina & Chen, Sufen. (2022). Effects of an automatic speech recognition system with peer feedback on pronunciation instruction for adults.

Computer-assisted Language Learning, Volume 35, 2022 - Issue 8, pp. 1869-1889.

Fantinuoli, Claudio (ed.). 2018. Interpreting and technology (Translation and Multilingual Natural Language Processing 11). Berlin: Language Science Press.

François, J. & Albakry, M. (2021). Effect of formulaic sequences on fluency of English learners in standardized speaking tests. Language Learning & Technology, 25(2), 26－41. http://hdl.handle.net/10125/73429

Fryer, L. K., Coniam, D., Carpenter, R., & Lăpuşneanu, D. (2020). Bots for language learning now: Current and future directions. Language Learning & Technology, 24(2), 8－22. Retrieved from http://hdl.handle.net/10125/44719

Jiang, Kai & Lu, Xi. (2021). The Influence of Speech Translation Technology on Interpreter's Career Prospects in the Era of Artificial Intelligence. Journal of Physics: Conference Series 1802(4):042074.

Kim, S. H. (2014). Developing autonomous learning for oral proficiency using digital storytelling. Language Learning & Technology 18(2), 20－35. Retrieved from http://llt.msu.edu/issues/june2014/action1.pdf

Lu, Xiaofei (2009). Automatic measurement of syntactic complexity in child language acquisition. International Journal of Corpus Linguistics, 14(1): 3-28.

Lu, Xiaofei (2012). The relationship of lexical richness to the quality of ESL learners' oral narratives. The Modern Language Journal, 96(2), 190-208.

Lu, Xiaofei (2013). The D-Level Analyzer. Version 2.1, released March 22, 2013. http://www.personal.psu.edu/xxl13/download.html

Lu, Xiaofei. (2014), Computational Methods for Corpus Annotation and Analysis. Springer.

Papineni, K.; Roukos, S.; Ward, T.; Zhu, W. J. (2002). BLEU: a method for automatic evaluation of machine translation (PDF). ACL-2002: 40th Annual meeting of the Association for Computational Linguistics. pp. 311－318.

Pastor Gloria Corpas. (2021). Interpreting and Technology: Is the Sky Really the Limit? Translation and Interpreting Technology Online, pages 15－24.

Peng, Gracie. (2022). Exploring Automatic Speech Recognition Technology for Undergraduate Sight Translation Training. Compilation and Translation Review, Vol. 15, No. 2, pp. 199-242.

West, Michael. (1953). A General Service List of English Words. London: Longman, Green and Co.

軟體

AntWordProfiler. https://www.laurenceanthony.net/software/antwordprofiler/

Berkeley Neural Parser. https://pypi.org/project/benepar/

COCA. https://www.english-corpora.org/coca/

DeepSpeech. https://github.com/mozilla/DeepSpeech

D-Level Analyzer. http://www.personal.psu.edu/xxl13/downloads/d-level.html

L2 Syntactic Complexity Analyzer.
http://www.personal.psu.edu/xxl13/downloads/l2sca.html

Lexical Complexity Analyzer. https://aihaiyang.com/software/lca/single/

Nltk. https://www.nltk.org/

Pycarret. https://pycaret.org/

Range. https://www.lextutor.ca/range/

readability 0.3.1. https://pypi.org/project/readability/

scikit-learn. https://scikit-learn.org/stable/

sentence-similarity. https://huggingface.co/tasks/sentence-similarity
spacy_sentence_bert. https://pypi.org/project/spacy-sentence-bert/
https://github.com/MartinoMensio/spacy-sentence-bert/

Tgrep. https://web.stanford.edu/dept/linguistics/corpora/cas-tut-tgrep.html

Vocabulary Profiler. https://www.lextutor.ca/vp/eng/

Wave2vec. https://ai.facebook.com/research/impact/wav2vec/

Whisper. https://openai.com/blog/whisper/

附件 (Appendix)

授課計畫書

*開課時段	☐上學期 ☐下學期 ☐寒假 ☐暑假 ☐其他(請說明_____)
*授課教師	☐主授 ☐合授 (若為多人合授課程請詳列教師姓名與單位，並於課程進度表註記各教師負責部分)
*開課系(所)	外國語文學系
*中文課程名稱	英文(附二小時英聽)二
*英文課程名稱	Freshman English (with 2-hour Aural Training) (II)
*課程屬性	☐系所/學程/學院必修(請填寫系所/學程/學院名稱_____) ☐系所/學程/學院選修(請填寫系所/學程/學院名稱_____) ☐共同科目 ☐通識課程 ☐其他(需為學校正式學制採計畢業學分之課程)_____
*學分數	____3____學分(如無學分數，請填「0」)
*授課時數	總計___48_____小時(___3___小時/週) (實習時數不計入)
實習時數	總計____0____小時(___0___小時/週)
*授課對象	☐專科生(_____年級) ☐大學部學生(____1____年級) ☐碩士生 ☐博士生
*過去開課經驗	☐曾開授本門課程 ☐曾開授類似課程 ☐第一次開授本門課程
*預估修課人數	**36**
*主要授課語言	☐國語 ☐臺語 ☐客語 ☐原住民族語 ☐英語 ☐其他(_____語)
*教學目標	本課程訓練學生整合英文聽，說，讀，寫四種語言能力。目標是 2 學期的訓練後能達到下列能力指標。聽：可以聽懂 TED 90%以上。說：能夠在有準備的情形下不看草稿做某一主題的簡報。同時具備基本中翻英句

<table>
<tr><td></td><td>子口譯的能力。讀：能夠不查字典看懂紐約時報 90%的內容，並能掌握文章主旨脈絡。寫：能夠利用各種辭典（英英辭典，搭配語辭典）網路資源及 Word 文書處理器寫出通順而沒有嚴重文法錯誤的句子，並能夠注意搭配語的用法。字彙：掌握最常用的學術詞彙 Academic Word List。</td></tr>
<tr><td>*教學方法</td><td>(1) 做線上練習（自主學習）（包括每週 AWL 字彙中翻英線上口譯練習 20 個句子）（2）課文講解 （3）分組討論 （4）每人輪流報告及回答問題 （5）即席寫作與報告。以上全部以英文進行。除了課堂上的練習，學生必須於課後自行連上本課程教學網頁及 e-Learning 系統訓練英文聽力與字彙與翻譯。每 4 週有小考。(5). 學習使用各種科技工具以及電腦輔助語言教學系統提升聽說讀寫能力。</td></tr>
<tr><td>*成績考核方式</td><td>(1). 作業及網路練習 20%，包括每週 AWL 字彙中翻英線上口譯練習 20 個句子 (2). 階段測驗共 3 次 占 20% (3) 摘要寫作及口頭報告(每 2 週一次) 60%。 本學期口頭報告約 6 次。報告前需先寫草稿於報告完時繳交。報告時不可照稿子唸。(4). 修課學生參加 3 項雙語教育中心主辦或補助之活動、修讀外教中心開設之混成式英語達人課程或英語短期課程，學期成績可最多額外加 6 分。(5).如通過全民英檢中高級初試，加總分 1 分記得報名，學校提供一次免費考試，本學期考試日期 5 月 14 日。(6). 如通過 Coursera，edX 等英語授課之線上課程並有證書加總分 2 分。</td></tr>
<tr><td>*課程進度</td><td>請簡述每週(或每次)課程主題與內容，自行依照所需增減表格
(若課程進行方式非以週進行，請以堂次進行填寫)
<table>
<tr><th>週次(堂次)</th><th>課程主題</th><th>內容【說明】</th><th>備註</th></tr>
<tr><td>1</td><td>Finding Success Starts with Finding Your Purpose https://hbr.org/2022/01/finding-success-starts-with-finding-your-purpose Online Vocabulary Exercises http://nlp.csie.org/~sound/English/Fr</td><td></td><td></td></tr>
</table></td></tr>
</table>

		eshmanEnglish2021/Finding%20Success%20Starts%20with%20Finding%20Your%20Purpose%20(Definitions).htm http://nlp.csie.org/~sound/English/FreshmanEnglish2021/Finding%20Success%20Starts%20with%20Finding%20Your%20Purpose%20(Sentences).htm Listening Sisters of the Valley - The Pot Selling Nuns https://www.youtube.com/watch?v=QqBPIEyHEnk&t=192s		
	2	Reading Plants Fight for Their Lives - Issue 112: Inspiration https://nautil.us/plants-fight-for-their-lives-13506/?utm_source=RSS_Feed&utm_medium=RSS&utm_campaign=RSS_Syndication Vocabulary Exercise http://nlp.csie.org/~sound/English/FreshmanEnglish202		

		1/Plants%20Fight%20for%20Their%20Lives.htm The Myth of Jason and the Argonauts https://www.ted.com/talks/iseult_gillespie_the_myth_of_jason_and_the_argonauts		
	3	Reading Darling Letters: On How to Keep a Hopeful Heart https://blog.darlingmagazine.org/darling-letters-on-how-to-keep-a-hopeful-heart/ http://nlp.csie.org/~sound/English/FreshmanEnglish2021/On%20How%20to%20Keep%20a%20Hopeful%20Heart.htm Powerpoint file SUPERBLOCKS: HOW BARCELONA IS TAKING CITY STREETS BACK FROM CARS https://drive.google.com/file/d/1OibhiarMkDtEyznPVToVNBPOp1D1SLeS/view		
	4	Reading The Bumblebee's		

		Decline Shows How We Get Conservation Wrong https://undark.org/2022/02/03/the-bumblebees-decline-shows-how-we-get-conservation-wrong/ Online Vocabulary Exericse http://nlp.csie.org/~sound/English/FreshmanEnglish2021/The%20Bumblebee's%20Decline%20Shows%20How%20We%20Get%20Conservation%20Wrong.htm		
	5	Quizzes based on what is taught in the first four weeks		
	6	Reading 15 Inspiring Audre Lorde Quotes https://www.biography.com/news/audre-lorde-quotes Online Vocabulary Exercise http://nlp.csie.org/~sound/English/FreshmanEnglish2021/15%20Inspiring%20Audre%20Lorde%20Quotes.htm		

	7	Reading Monsters in the Bible - Is it a Credible Source? https://www.phantomsandmonsters.com/2022/02/monsters-in-bible-is-it-credible-source.html		
	8	Reading AI to help find UAP evidence in satellite images of Earth, Harvard astronomer says https://www.mysterywire.com/ufo/ai-to-help-find-uap-evidence-in-satellite-images-of-earth-harvard-astronomer-says/ Vocabulary Exercise http://nlp.csie.org/~sound/English/FreshmanEnglish2021/AI%20to%20help%20to%20find%20UAP%20evidence.htm http://nlp.csie.org/~sound/English/FreshmanEnglish2021/AI%20to%20help%20UAP%20evidence%20PART%20II.htm		
	9	Reading The Responsibility of Intellectuals by		

		Noam Chomsky The New York Review of Books, February 23, 1967 https://chomsky.info/19670223/ Online Cloze Exercise http://nlp.csie.org/~sound/English/FreshmanEnglish2021/The%20Responsibility%20of%20the%20Intellectuals%20by%20Noam%20Chomsky.htm Sci-Fi Short Film “Plurality" https://www.youtube.com/watch?v=pocEN5HprsM		
	10	Quizzes based what is taught from the sixth to the ninth week		
	11	Reading What Does It Mean to Be A Grown-Up? Embracing playfulness is important at any age. https://www.psychologytoday.com/intl/articles/202201/what-does-it-mean-be-grown		
	12	Reading How robots and bubbles		

		could soon help clean up underwater litter https://horizon.scienceblog.com/1935/how-robots-and-bubbles-could-soon-help-clean-up-underwater-litter/		
	13	Presentations		
	14	Reading Heroism in the Face of Hopelessness: The Story of the SS Dorchester https://www.saturdayeveningpost.com/2022/02/heroism-in-the-face-of-hopelessness-the-story-of-the-ss-dorchester/		
	15	Quizzes based on what is taught from the eleventh to the fourteenth week		
	16	Final Presentations		
*學生學習成效	預期學生在聽說讀寫四項技能都可以獲得顯著的進步，並學會如何運用科技工具改善聽說讀寫四項能力。			
*預期個人教學成果	預期可以提升學生的學習成就感和對課程的滿意度。			
*學習成效評量工具(如前後測、學生訪談、問卷調查等)	前後測、學生訪談、問卷調查、學生練習資料量化指標			
其他補充說明				

(如課程參考網址)	

備註：*為必填項目；本計畫如擬搭配其他課程，請依課程分別羅列。