

運用英語單字語音辨識系統降低英文聽力學習之認知負荷

Using English speech recognition system to reduce the cognitive load of English Listening Comprehension

葉芯妤¹ 張智凱²

YEH HSIN YU¹ CHANG CHIH KAI²

¹國立臺南大學 數位學習科技學系 學生

¹National University of Tainan Department of Information and Learning Technology Student

E-mail : quill4749@gmail.com

²國立臺南大學 數位學習科技學系 教授

²National University of Tainan Department of Information and Learning Technology Professor

E-mail : chihkai@mail.nutn.edu.tw

摘要

英語溝通是現代社會中必備的技能，傳統正規化的英語教學多以緩慢清楚地發音為主，導致學習者不熟悉英語口說的特性，出現英語聽力理解的困難。許多研究肯定多媒體素材為學習英語較有效率的方式，因為能同時提供豐富的視覺與聽覺訊息，成為許多研究推薦的學習素材(黃寶仙，2007)。然而，在英語學習影片中以字幕輔助時，學習者易著重在字幕的閱讀，而忽略聽力技巧且增加外在認知負荷(extraneous cognitive load)。有鑑於此，本研究提出一英語單字語音辨識系統，先讓學習者選擇影片，以較困難的字詞為基礎，利用手機應用程式聆聽其發音，並透過 Google 語音辨識出學習者所說出的單字，提供作為發音對照參考進而降低英語聽力時的認知負荷。本研究共 40 人進行實測，分析認知負荷問卷結果，實驗結果顯示利用此方式相較於傳統方式，能使學習者更有效率地提升英語聽力，同時增進英語口說的技巧。

關鍵字：英語口說、英語聽力理解、認知負荷、學習輔助系統、語音辨識

Abstract

As globalization becomes a trend, the ability to communicate in English became one of the most important skills. Comparing to speaking, reading, and listening, most of the English learners considers "Listening" to be the most difficult skill to acquire. The traditional way of teaching English as a second language mostly involves in slow and clear pronunciations. However, learners have difficulties learning through the traditional method. Many second language learning researchers pointed out that the reduced forms of English pronunciation has caused the difficulty in learning English

for students (Fan, 1993; Rubin, 1994). The best way to learn the reduced form of English pronunciation is to listen to an extensive amount of English conversations in real world situations (Rosa, 2002; Norris, 1993, 1995). Movies and TV series can reflect real life situations; therefore, they are recommended for learning the reduced form of English according to many pieces of research (黃寶仙, 2007; Ito, 2001; Norris, 1993, 1995). Subtitles are often used to assist the understanding of the English learning videos.

The learners may focus on the reading of the subtitles and tend to neglect the listening section. This research aims to select videos suitable for language learning, and also to develop a system that uses the Inventor 2, an application developed by the MIT Center for Mobile Learning, to train the pronunciation of the students and enhance the listening ability of the students.

Keywords : English Speaking, English Listening Comprehension, Cognitive load, learning support system, Speech Recognition

壹、前言

現代社會中，英語能力是人人必備的第二語言，更是決定未來吸取知識廣度及深度的關鍵，許多研究者都希望能找出更有效的學習方式，以提升學習者的英語溝通能力。研究顯示，一般人與他人溝通時，聽力花費的時間最多，約占 40% 至 50%。英語聽力在溝通上扮演了重要的角色，為必須增進的能力(Leloup & Ponterio, 2003; Verdugo & Belmonte, 2007)；臺灣屬於 EFL(English as a Foreign Language)國家，缺乏口說英語的真實環境，因此英語聽力理解對於外語學習者來說，難度較高，而口語發音縮減特徵為影響英語聽力的一個重要因素(Brown & Yule, 1983)。

隨著科技進步，不同於傳統教學中透過老師反覆唸句子進行不斷地練習，現今有許多學習工具與媒介提供學習者選擇，例如：觀看英語影片、收聽英語廣播與聆聽英語歌曲等，影片能讓學習者產生興趣與學習動機，創造成功的學習氣氛，學習者在觀看影片時，擁有聽覺與視覺上的探索，有利於聽力的瞭解(Allan, 1985; Lonergan, 1989)；此外，學習者也能從影片中的情境更真實地接觸外國文化，藉由沉浸在真實語言環境的脈絡下，學習者能更有效率地增進自己的英語能力。學習者觀看影片時，若以英語字幕作為輔助，字幕的呈現能提升聽力的了解並增進閱讀學習的能力(Markham, 1999)，但若以全文字幕呈現，學習者會將注意力放在字幕的閱讀上。本研究之研究動機為希望學習者能在觀看影片的同時，也能兼顧聽力和口說的學習，利用更高效率的方式增進英語能力，因此，本研究希望設計一個語口說輔助學習系統，以口語發音縮減特徵、多音節以及片語的英語字詞作為口說練習的字詞，讓學習者聆聽三次該發音之後並將其錄下，以 Google 語音

辨識的方式，輔助學習者的口說練習以及聽力理解。

貳、 文獻探討

一、 口語發音縮減特徵

口語發音縮減特徵(Reduced Forms)是以英語為母語者在非正式場合中，或日常生活對話中，為了使說話更為流暢，字詞的發音會隨之改變，產生特有的說話特徵。

早期研究中，Norris(1993)歸納出在非正式口語英語中會出現的特徵：縮式(Contraction)、省略音(Elision，省去該音節不發音)、同化音(Assimilation)及弱化音。Fan(1993)則發現連音(Liaison，把兩個單字連接)也是常存在於口語英語中的一個特徵，並指出特徵中的連音及省略音這兩項會影響聽力。之後，有些研究從Norris(1993)和 Fan(1993)的研究中整理出口語發音縮減特徵的種類，刪除其中的弱化音，形成縮式、省略音、同化音與連音四種特徵。趙御筌(2005)則將縮式與省略音合併為「消音」一詞，並將同化音以「變音」一詞代稱，最後加上連音與弱化音，形成消音、變音、連音以及弱化音四種口語發音縮減特徵，其中「連音」被認為是四種特徵中最為困難的一種(鐘靜怡，2008)。

二、 電腦輔助發音訓練

隨著科技的進步、網路的普及，在外語學習發展的過程，許多語文學家從只注重讀與寫的文法翻譯法(Grammar Translation Method)中注意到，外語學習不應該只偏重於訓練閱讀、解析文法或修辭的書面言語，也應注意口語表達的能力；此外，字彙記憶能力和閱讀理解能力與學習者的發音準確度有正向相關性，正確的發音能有效協助對語言的記憶和理解(Stanovich, 1986)。而外語的發音學習需要在以該語言為母語的環境下培養為佳，對於身處非英語系國家的學習者是很難達成的條件。由於個人電腦的普及和運算速度的提高，「電腦輔助語言學習」(CALL, Computer Assisted Language Learning)的研發已經到達可以商業化的階段(Chapelle, 2008)，其中又以「電腦輔助發音訓練」(CAPT, Computer Assisted Pronunciation Training)最受矚目，因為其實用性高且互動性強，其系統部份結合了音訊處理與語音辨識技術，提供英語學習者在虛擬環境下訓練發音的機會，大量彌補了口說語言師資的不足(DAGAR, 2013)。利用語音辨識技術的電腦輔助發音訓練系統(CAPT)，不僅提供學習者無壓力的環境，同時也能針對發音問題給予回饋，例如：語言學習軟體，其功能可在一般個人電腦上進行完整的語音評分，並包含各種語言，以達到更客觀的評分與分析。

三、 MyET 軟體

MyET 是由艾爾科技所研發出來的軟體，艾爾科技成立於西元 2002 年，為專注於先進語言教學系統研發的教育軟體公司。其課程內容來自於世界各地知名

出版社、大學、補習班等內容提供者(Content Provider)，包含新聞英語、商用英語、校園英語、兒童英語、實用生活會話、專業英語等等，使用者可以選擇較適合自己的課程來學習。根據「聽說教學法」(Audio-Lingual Method)及「溝通式學習法」(Communicative Approach)所設計，搭配「自動語音分析系統」(ASAS, Automatic Speech Analysis System)檢定功能，能在分析使用者的發音後，精確指出問題出在哪一個字、哪一個音節，針對發音、語調、流利度、音量等項目做出評分，並告訴使用者如何有效地改進發音(Chen, 2011)。

下圖 1 為 MyET 的單字發音練習聲紋比對系統，運用語音辨識技術，系統將學習者錄下的語句進行聲紋比對，提供每一句子中的單字發音功能。學習者聽老師唸完一個句子後，用滑鼠的右鍵點選句子中的單字聆聽老師的發音，並進行口說練習，系統將學習者與老師的聲紋呈現出來，提供學習者做為參考調整，進一步模擬老師發音的方式及技巧。學習者進行口說練習時，朗讀影片中的關鍵字幕，系統會記錄其語音訊號，與影片中字詞的語音訊號相比對，通過一定的準確率則可繼續影片撥放，若未達標準則重複口語練習，如此反覆練習可增加使用者口說經驗。

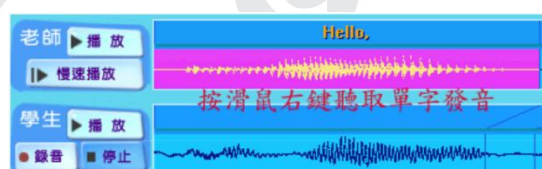


圖 1 MyET 聲紋比對系統(圖片來源：myET 使用手冊)

四、 SpeakingPal 應用程式

SpeakingPal 是一間開發教育行動裝置應用程式的公司，其開發軟體所專注的重點在於英語口說能力。透過高端多媒體豐富的內容結合自動語音識別，提供使用者一個極具吸引力和高互動的方式進行英語口說練習。使用者在使用「SpeakingPal 英語家教」應用程式時，需先選擇一段視頻，視頻分為初級、中級、高級三個等級，每部影片中配有關於如何發音的小測驗以及聽力練習題，視頻選擇完成後，如下圖 2 所示，使用者藉由和影片中的角色對談練習英語口說，程式將使用者說的話錄下，利用自動語音識別功能，與一位英語母語者進行比對，針對使用者所說的單詞或句子，立即給予回饋。回饋的方式是在獲得使用者的發音後，和原始影片中的聲紋比對進行標示以及評分的即時反饋，並以易於辨識的“紅綠燈”顏色標示，提供使用者作為參考。



圖 2 SpeakingPal 口說練習使用介面

參、 研究實施與設計

一、 研究方法

(一)英語口說學習輔助系統介紹

影片學習為一種能夠建構出完整情境脈絡的學習模式，當純正的外語發音環境和影像呈現在學習者面前，使其直接建立語言與外界經驗的聯繫，在潛移默化之下累積經驗並習得外語(廖曉青，2002)。學習者可以透過觀看影片從情境中演繹出的語言來瞭解語境(context)的使用，內化語言並遷移至實際應用中(吳旻純，2011)。因此針對學習者在觀看影片時，時常忽略聽力上理解的問題，本研究設計一英語口說輔助學習系統，幫助學習者在觀看影片的同時增進聽力理解與口說能力。

本研究利用 MIT 行動學習中心(MIT Center for Mobile Learning)所開發的 App Inventor 2 設計一手機應用程式，即為英語口說學習輔助系統，提供學習者便利的英語學習環境，使用方式如下圖 3 所示。首先，學習者開啟應用程式後，可選擇兩種方式進行英語聽力與口說的練習，第一種方式為個人學習方式，只要輸入想學習的字詞在上方文字格中，即可直接進行自主練習。而第二種為教師提供字詞給學習者練習的方法，學習者需先點選「Import Word」後進入第二個介面，再點選「SCAN」，將會跳轉至條碼掃描器。學習者掃描教師給予的 QRcode 後，畫面的文字格將顯示條碼中的字詞，接著點選「Start Practicing」會回到口說練習的介面，QRcode 中的第一個字詞將會出現在上方文字格中。學習者可開始進行口說練習，點擊「Tap To Listen」，系統將會重複三次該字詞的發音，接著點擊「Tap To Speech」，便會跳出 Google 語音辨識的畫面，偵測使用者所說出的單字，並顯示在下方文字格中，若與上方文字格中的字詞一致，即可點選「Next Word」進行下一個字詞的練習。



圖 3 應用程式使用示意圖

(二) 學習活動流程

首先，由教師選擇欲教學的影片以及影片中的字詞，以口語發音縮減特徵為基礎，延伸出多音節以及片語這類較難發音的字詞，將欲練習的重要字詞轉換為 QRcode 提供給學習者。學習者利用應用程式掃描 QRcode 後，點擊開始按鈕後則進入聽力與口說練習的介面，同時，教師播放影片後，當遇到所選擇的字詞時，影片先暫停播放，學習者便可開始練習該字詞的發音，應用程式會將該字詞重複播放三次。最後利用 Google 語音辨識將學習者的聲音錄下，並顯示偵測後的單字，提供作為對照參考，學習者可藉此比對所說出的字詞是否與教師提供的字詞相符，若相符方能繼續觀看影片並進行下階段學習。若不相符，則學習者需繼續練習直到與教師所選擇之字詞一致為止，在練習的過程中，不僅能提升聽力的理解能力，也能增進英語口說的技巧，學習活動流程如下圖所示。

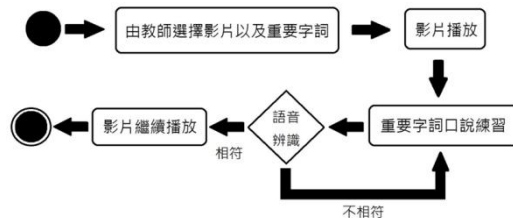


圖 4 英語口說練習學習活動示意圖

(三) 實驗設計

本計畫之實驗對象為國立臺南大學數位學習科技學系一年級學生，受測人數約 40 位，採取常態編班方式。實驗地點為電腦教室，實驗開始前學生須先安裝本實驗設計之手機應用程式，應用程式包含 QRcode 掃描以及口說練習的功能。首先，要求學生依照一般過往英語聽力課程學習方式填寫認知負荷問卷(前測)，之後則進行上述 4-2 學習活動流程。當學習活動結束後，學生在未使用系統的情況下，將影片重新再觀看一次，最後則要求學生再次填寫認知負荷問卷(後測)。

本研究選擇認知負荷問卷了解受測者使用傳統教學學習方法與本系統後的兩者認知負荷是否有差異(詳細問卷內容請見附錄二與附錄三)，受測者須根據對於使用本系統學習英語的方式，與在平時課堂中的學習方式，例如：朗讀、對話等，相互比較之後的感受來填寫問卷。其中，問卷前測為瞭解受測者對於傳統學

習方法的認知負荷，而後測為瞭解受測者對於本系統學習方法的認知負荷，前測與後測皆有七題，每題只能圈選一個數字，數字範圍為一到九，數字愈小表示感受力愈弱，數字愈大表示感受力愈強。此外，問卷題目除了認知負荷外，在後測問卷中另外加入三題開放式問題，依據受測者所提供的使用本系統後的感想，對於系統進行修正，讓系統更加貼近使用者，進而提升學習的效率並降低學生的認知負荷。最後，本研究根據受測者所填寫之問卷結果，分析學生對於兩種學習方式的認知負荷差異，整個實驗流程如下圖 5 所示。

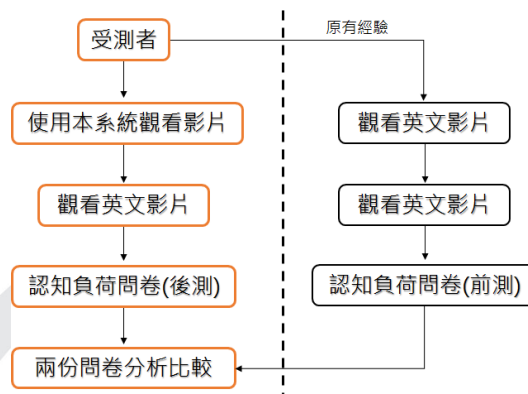


圖 5 實驗流程圖

此外，本研究期望接受口說練習輔助之學習者對於系統具有高滿意度，依照使用過程中的介面、系統、功能等觀感進行主觀滿意度調查與有用性，後測問卷中三題開放式問答題調查，受測者填寫對於系統與提供教材設計等相關建議。希望藉由此方式獲得受試者對於系統的主觀意見，因此，問題方向必須讓受試者不受影響自由作答，例如：使用後的心得或是對系統整體的意見等。透過受試者問卷中的答案瞭解學習者需求，並針對其需求改善系統設計，以達到促進學習者對影片內容的理解以及英語口說能力的進步。

肆、 結果與討論

本研究採用認知負荷問卷以及三題開放式問題，收集共 40 位受測者的回饋，並對其進行數據分析；認知負荷問卷分為前測與後測，前測為詢問受測者對於傳統上課方式的認知負荷，而後測為詢問受測者對於本系統學習方式的認知負荷，前測與後測的實驗結果，如下表 1 所示，在前測的部份，其感受力的平均值為 41.13、標準差為 14.333。而在後測的部份，感受力的平均值為 26.13、標準差為 12.635。在獨立樣本 t 檢定結果中，p 值小於 0.001 達到顯著差異 ($p = 0.000 < 0.001$)，由統計結果得知，大部分受測者認為使用此系統之學習方式，相較於傳統課堂的學習方式，所感受的認知負荷較低。

表 1 前測與後測之實驗結果

| 組別 | 人數 | 平均數 | 標準差 | T 值 |
|----|----|-------|--------|--------|
| 前測 | 40 | 41.13 | 14.333 | 9.965* |
| 後測 | 40 | 26.13 | 12.635 | |

註:*表示 $p < 0.05$

在本次實驗中，每一位受測者使用本系統總共做了五次口說體驗，且體驗後填寫認知負荷問卷。有些受測者表示 iPhone 手機無法安裝此應用程式，在受測過程中只能借用他人手機體驗，有些則認為若能加入整個句子的朗讀功能，能使應用程式更方便使用，也能讓整體功能更加完整。此外，受測者們除了對於系統的功能提出改善的建議之外，也肯定本系統對於學習英語聽力與口說的幫助，許多受測者認為，本系統不僅能提供教師在課堂上使用，也相當適合課後個人的自主學習，可以透過本系統自行解決並提升聽力與口說的能力，只要利用系統輸入想學習的單字，即可馬上知道在哪一個部份有發音上的不足，不用每當遇到問題就必須請教老師。以下為彙整受測者所提供的建議之問卷調查結果：

受測者在問卷上主要提出三個建議與改善的方向：第一，此次系統無法支援 iOS 的受測者，使得受測範圍受到一定的局限，希望本系統能改善支援限制，使更多學習者使用本系統；第二，在口說訓練方面，期待本系統能夠一次辨識更多單字，使學習者在較長的句子或文章中，透過系統的語音播放與跟讀學習正確的發音方式，並矯正平常口說對話的語調；第三，整體的系統介面非常地簡潔清楚，使受測者容易上手，如果能讓系統和學習的影片能在同一個介面上進行英文學習，將大幅提升系統的方便性與使用效率。

本次的受測者中將近有四分之三的人從未接觸過英語口說學習系統，因此對於本研究設計之應用程式感到相當新奇。而許多受測者認為，透過手機應用程式練習英文聽力與口說，使得學習更加方便，只要輸入想學習的英文單字，不用等到課堂時間詢問教師，便能即時聆聽正確的發音方式並立刻進行練習，在一個無壓力的環境下進行練習，解決許多學習者因害怕而不敢面對他人說英文的情形。

伍、 未來展望

近年來網際網路日漸普及，網路上有大量豐富的影片可以取得，觀看影片學習英文的趨勢亦漸提升。希望未來能繼續擴增本系統的功能並改善原有的問題，使之成為一個能與 MyET 和 SpeakingPal 功能相近的英語口說學習輔助系統，提供學習者藉由英語影片進行口說練習，以提升英語單字量、口語流暢度、英語發音等能力，並且讓學習者在欣賞影片的過程中，藉此瞭解口語發音縮減特徵字彙

以及發音較困難的字詞，同時進行英語口說練習以及聽力的訓練，有助於提升使用者的英文能力。

致謝

本研究承蒙臺灣行政院科技部專題研究計畫贊助，計畫編號：MOST 105-2511-S-024 -007 -MY3，謹此致謝。

參考文獻

一、中文部分

- Hua, P.(2008)。臺灣大學生透過電腦輔助軟體學習英語發音的研究。政治大學英國語文學研究所學位論文，頁 1-191。
- 吳旻純(2011)。自主性線上語音觀看模式對語言學習之影響。生活科技教育月刊。44(4)，22-32。
- 黃寶仙(2007)。電腦輔助語音縮減教學對提升整體英語聽力理解之研究。未出版之碩士論文，國立臺北教育大學教育傳播與科技研究所，臺北市。
- 廖曉青(2002)。英語教學法。臺灣：五南。
- 趙御筌(2005)。說出好聽力。臺灣：笛藤。
- 鍾靜怡(2008)。建置影片式英語口語發音縮減特徵學習系統輔助英文聽力學習。未出版之碩士論文，國立中央大學網路學習科技研究所，桃園市。

二、英文部分

- Allan, M. (1985). *Teaching English with video*. New York: Longman.
- Brown, G., & Yule, G. (1983). *Teaching the spoken language*. Cambridge: Cambridge University Press.
- Brown, J. D., & Hilferty, A. (1986). The effective of teaching reduced forms of listening comprehension. *RELC Journal*, 17(2), 59-70.
- Chapelle, C. A. (2008). Computer assisted language learning (pp. 585-595). Blackwell Publishing Ltd.
- Chen, Y. C. (2011). A Study on the User Experience of MyET Elearning System.
- DAGAR, M. (2013). Computer Assisted Pronunciation Training.
- Fan, Y. (1993). Listening: problems and solutions. *English Teaching Forum*, 31(2), 16-21.
- Ito, Y. (2001). Effect of reduced forms on ESL learners' input-intake process. *Second Language Studies*, 20(1), 99-124.
- Leloup, J. W., & Ponterio, R. (2003). Foreign language study and the brain. *Language*,

Learning & Technology, 7(1), 2-3.

Lonergan, J. (1989). *Video in language teaching*(4th edition). NY:Cambridge University Press.

Markham, P. (1999). Captioned videotapes and second-language listening wordrecognition. *Foreign Language Annals*, 32(3), 321-328

Norris, R. W. (1993). Teaching reduced forms: an aid for improving lower-level students' listening skills. *Fukuoka Women's Junior College Studies*, 46, 49-56.

Norris, R. W. (1995). Teaching reduced forms: putting the horse before the cart. *English Teaching Forum*, 33, 47-50.

Rosa, M. (2002). Don'cha know? A survey of ESL teachers' perspectives on reduced forms instruction. *Second Language Studies*, 21(1), 49-78.

Rubin, J. (1994). A review of second language listening comprehension research. *The Modern Language Journal*, 78, 199-221.

Stanovich, K. E. (1986). Matthew effects in reading: Some consequences of individual differences in the acquisition of literacy. *Reading research quarterly*, 360-407.

Verdugo, D. R., & Belmonte, I. A. (2007). Using digital stories to improve listeningcomprehension with spanish young learners of english. *Language Learning and Technology*, 11(1), 87-101.

