

第1章 緒論

1.1 研究背景

許多語言教學研究證明：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新語言的記憶和理解。而新語言的發音最好在以該語言為母語的環境下來培養，或是透過與該語言為母語的人士經由長時間一對一的練習訓練。但是，對於身處非英語系國家的英語學習者，這種條件是很難達到的。於是近年來便有許多電腦輔助語言學習（Computer Assisted Language Learning, CALL）系統設計用來提供英語學習者在虛擬環境下練習的機會。讓 CALL 系統結合語音辨識技術的系統，系統於是能夠理解學習者的輸入語音，並依據內容進行分析和評估，然後進行互動反應或是提供矯正回饋，這種系統稱為電腦輔助發音訓練（Computer Assisted Pronunciation Training, CAPT）系統。一般語言學習者皆對用語音進行互動的系統感興趣，但許多語言專家學者仍對 CAPT 系統提升學習者發音能力的有效程度相當懷疑，原因在於語音辨識技術的能力仍是有限。

對此，於是有許多關於建立有效 CAPT 系統的研究和成果。研究後發現系統問題大致可以歸為兩類，一是系統對於使用者的口音或是錯誤發音的辨識效能差，二是系統無法提供有意義的發音品質評估。而在接下來的章節，我們提出一種偵測語者錯誤發音的方法，並設計實驗來驗證方法是否有效。

1.2 文獻回顧

近十年來，已有許多使用 CAPT 系統學習矯正發音的相關實驗研究。

1997 年，Ronen 等人提出 Mispronunciation Networks 的概念[1]。

他們設計一個讓美國人學習法文發音的訓練系統。他們建立了法國人和美國人的聲學模型，讓系統在辨識的時候同時比對兩套聲學模型，之後再以統計美式拼音在辨識結果中出現的比率作為評估發音的好壞。1999 年，語言學家證明 L1 (Mother Tongue) 和 L2 (Learning Language) 兩種語言存在一種取代對應關係[8]，於是 Silke Witt 提出三種組合調適語者獨立聲學模型的方法[2]，讓系統根據語言對應關係來結合兩種語言聲學模型，能讓語者獨立的語音辨識系統對 L2 學習者帶有母語口音的語句有較高的辨識率，改善 CAPT 系統對 L2 學習者語音辨識效果差的問題。而他的實驗結果證明，用 MM (Model Merging) 和 MLLR (Maximum Likelihood Linear Regression) 對語者獨立聲學模型進行調適能有效提升系統對 L2 學習者語音的辨識率。2000 年，Menzel W. 等人設計一個讓義大利人和德國人學習英文發音的訓練系統[3]，他們的系統能針對使用者的發音錯誤提供矯正性回饋。系統首先對使用者輸入的語音進行辨識，並對辨識結果做可信度分析。檢查辨識結果找出使用者的錯誤發音，然後根據使用者的母語歸納出使用者的錯誤類型，最後提供使用者適當的矯正指引。但可惜他們的系統實驗結果，由於聲學模型訓練不佳導致辨識率低，系統偵測錯誤發音的正確率因而降低，效果並不好。2002 年，Yasushi 等人則提出以 Pronunciation Error Networks 來偵測日本學生發音錯誤的類型[4]，而系統會根據偵測出的錯誤類型挑選同類型的練習語句讓使用者加強練習。實驗結果顯示利用這個方法能正確有效地偵測出學生的發音錯誤。2005 年，分別有湯士民和徐偉棠兩篇碩士論文進行改良 CAPT 系統的研究[5][6]。他們整理了 L2 學習者可能有的錯誤發音類型，建立發音混淆網路 (Pronunciation Confusion Networks, PCN)，然後透過 PCN 來預測使用者的錯誤發音。湯士民設計的 CAPT 系統以 PCN 偵測出使用者的錯誤發音後，利用 3D 動畫合成正確發音應有的唇型和舌頭動作，以動畫呈現方式提供使用者修正回饋。徐偉棠的 CAPT 系統則是在 PCN 偵測出錯誤發音後，利用

共振峰這項特徵計算辨識結果的可信度。而他的實驗證明了共振峰能提升系統偵錯準確率。

1.3 研究動機與目的

總結多位前人的研究成果，我們發現目前 CAPT 系統普遍有兩項問題。第一，因為第二語言學習者（L2 Learner）容易發生以母語發音取代新語言發音的現象，而造成一般語者獨立語音辨識系統辨識效能不佳。第二，目前的 CAPT 系統雖利用 PCN 來偵測使用者的錯誤發音型態，但一般系統中 PCN 所套用的錯誤規則皆是事先整理定義的，能預測出的錯誤情形受到限制，無法偵測出定義之外的錯誤。每個人的發音習慣和方式必不相同，使用者有可能因為具有個人獨特口音，導致特殊的錯誤類型，而這些錯誤系統無法偵測出來。

為改善以上問題，我們提出以下方法。首先，我們將學習者的母語資訊增加到辨識系統中，使系統能偵測語句中是否發生以母語方式發音的情形，改善系統對參雜母語口音語句的辨識率。接著，讓系統能自動偵測使用者個人的發音錯誤型態，動態新增至錯誤規則定義中，將 PCN 進行使用者個人化調適，建立個人化發音混淆網路，讓系統能更有效地抓出使用者的錯誤發音。

1.4 研究方法簡介

在系統開始運作前，先有兩項準備工作。一是找出中英文兩種不同子母音之間可能的取代對應關係，二是歸納整理語言學者提出的以中文為母語的學習者在學習英文發音時容易有的錯誤情形，建立成發音錯誤規則。利用這發音錯誤規則和尋找到的中英文對應關係，將所有可能發音建立在辨識網路中（包括正確發音與所有可能的錯誤發音）。在系統運作時，會對使用者輸入的語音進行兩次辨識，第一次辨識是利用建立好的辨識網路來偵測使用者是否有符合錯誤規則中定義的錯誤情形；

第二次辨識則是進行偵測使用者個人特殊錯誤型態，若有則將之定義成使用者的個人錯誤規則，並建立包含個人錯誤型態的辨識網路。讓辨識網路能依使用者進行調適，並透過調適後的辨識網路來偵測使用者的發音錯誤型態。

1.5 章節概要

本論文共分為下列幾個主要章節：

- 一、 序論：說明研究背景、文獻探討、研究動機及目的，以及所提出的系統架構與研究方法。
- 二、 系統架構：系統架構圖的說明，概括性的介紹整個系統運作的流程，及本系統各個模組的作用。
- 三、 發音錯誤類型之偵測：介紹本論文提出的方法，說明系統個人化辨識網路的內容和建立方法，和偵測使用者錯誤發音的流程。
- 四、 實驗結果與討論：針對論文提出方法的各步驟進行實驗，介紹各實驗內容和結果並進行分析。
- 五、 結論與未來研究方向