

第5章 結論與展望

在本論文中，我們針對母語為華語的使用者所使用的英文發音訓練系統，提出新方法來改良系統偵測錯誤發音所使用的發音混淆網路，讓發音混淆網路能動態彈性地依不同使用者進行個人化調適。

我們提出一種方法用於尋找中文音素和英文音素發音上的對應關係，並以此種對應關係建立 BPCN，期望能藉此改善系統辨識使用者口音效力不佳的問題。但實驗結果顯示並不符合預期，包含母語發音對應的 BPCN 其辨識率比固定式 PCN 平均降低了 3.81%。

而我們提出另一種方法則是用於偵測並建立使用者個人化發音錯誤規則，並以個人化規則建立 PPCN。讓一般 CAPT 中使用的固定式 PCN 能針對不同語者的錯誤習慣進行調適，突破固定式 PCN 僅能偵測出定義內含的發音錯誤種類的限制。實驗結果顯示，藉由此方式建立的 PPCN 進行錯誤發音偵測，相較於固定式 PCN 提升約 3.97% 的正確率，並且對於非英語系語者的正確率提升幅度較大。我們推測其原因為非英語系語者的發音通常較不標準，較容易有一般常見錯誤發音定義之外的個人特殊錯誤情形，於是以我們提出的方法進行偵測個人化錯誤發音，能有較高的提升辨識率效果。英語系語者由於本身發音較標準，特殊錯誤情況較少，於是提升辨識率的效果相對較低。

未來可延續的研究方向，可以從以下幾點著手：

- 一、進行反向尋找母語發音對應：本論文中設計找出母語發音對應的方式是單向的，是以中文聲學模型對英文語料進行強迫對應。若能反向實驗以英文聲學模型對中文語料進行強迫對應，再將兩種對應結果做比對，並再以實驗進行母語對應關係的可信度分析，應能找出更可信的對應關係。
- 二、增加判斷對應的條件，並將兩套語言聲學模型進行調整後再建立 BPCN：由於實驗中建立 BPCN 時，並未考慮兩套語言聲學模型之間的差異，直接以機率排序決定兩種語言的對應音素，並直接將對應到的兩種音素聲學模型合併在一起

進行辨識，導致機率值偏移而喪失模型之間的互補性。因此，若在尋找音素對應時增加判別條件，多考慮兩種語言之間的細節差異，例如：讀音時間的長短、讀音時的語調高低或聲調起伏等。而在確定發音對應後，先將兩套語言聲學模型進行正規化調整，再組合建立成 BPCN，以避免機率值偏移。

三、將母語發音對應進行個人化調適：由於實驗系統中使用的母語發音對應並未進行個人化調適，推測因此造成徒增加不適當的路徑到發音混淆網路中，使複雜度增加導致辨識正確率下降。使用者會有以母語發音方式取代英語發音是可確定的，但依習慣口音不同，不同使用者的母語發音對應也應不相同。若能將母語發音對應進行調適後再加入發音混淆網路，應能有較好的輔助辨識效果。

四、刪減不正確或不常發生的錯誤規則：本論文提出能自動建立使用者個人錯誤規則的方法，是藉由分析 Continuous Phone Recognition 的辨識結果來產生的。但辨識結果並未能達到百分百正確，於是系統有可能產生不正確的錯誤規則。因此，若能評估系統產生的錯誤規則是否真確，刪減多餘無用的規則，如此能減少系統錯誤辨認的情形。

五、設定錯誤規則發生機率：一種音素可能有兩種以上的錯誤型態發生，但未必各種錯誤發生的機率會是相同的，有可能只有其中一種或少數的發生機率較高，其餘則是非常偶然或是規則並不正確。若能建立一套機制去統計每種錯誤發生的次數，評估錯誤的發生頻率並能設定辨識網路中各種路徑的發生機率，如此能加強系統正確辨識錯誤發音的能力。