

行政院國家科學委員會專題研究計畫成果報告

應用語音辨識與合成技術於語言障礙復健之可攜式裝置的研究與設計

A Portable Device Using Speech Recognition and Synthesis Applied to
Rehabilitation of Subjects with Speech Disorder

計畫編號：NSC 89-2614-E-038-002

執行期限：89 年 08 月 01 日至 90 年 07 月 31 日

執行單位：台北醫學大學 醫學資訊研究所

主持人：徐建業 台北醫學大學 醫學資訊研究所

共同主持人：潘力誠 台北醫學大學 物理學科

一、中文摘要

語言及聽力是人類最主要的溝通工具。具備正常的語言及聽覺能力可以正確的表達自己的思想並瞭解對方傳遞的訊息。語言障礙普遍存在於多類的身心障礙中，他們的說話清晰度不佳，導致生活、學習與工作上的適應困難。

根據多位學者專家研究結果顯示，臺灣 7 地區成人語言障礙發生率約為 10—12%，對於輕度及中度語言障礙病患，大多可經由語言治療有效改善其語言異常；對於中重度或重度者，則需藉助語言治療及溝通輔具的協助，才能有效改善言語清晰度、語言表達能力以及語言理解能力。根據美國研究報告顯示，需要溝通輔具者比例應遠大於總人口數 1%。

輔助溝通系統 (Augmentative and Alternative Communication; AAC) 拜電子科技高度發展之賜，逐漸受到重視。使用患者口語輸入，經過系統的辨識並加以修正的想法是有，但研究發展可說是相當稀少。由於高科技溝通輔具具有多面相之溝通功能，適合不同類型及嚴重度之語言障礙病患，且能配合肢

體控制能力，家人支持度，生活環境以及實際溝通需求，能有效提高語言治療療效及溝通效果。本計劃已完成一硬體系統，可供使用者進行語音指令之辨識並輸出正確發音的語音，以提供語障者進行溝通之用。本系統使用數位信號處理晶片與單晶片控制器之多晶片架構，並且配合語音樣本比對與語音合成技術，將語障者的非正確語音經由語音樣本比對而轉譯為正常之發音。如由語音重複性較高的語障患者使用，有不錯的辨識率，如能再配合產業界之技術將成本降低，本裝置將會有更大的實用性。

關鍵詞

語音辨識、語音合成、輔助溝通系統、輔具科技

Abstract

Individuals with mental and/or physical disabilities are often also impaired in speech production. Their speech intelligibility is so poor that adjustment in learning, work and other life situations have become very difficult for them. Three approaches to the rehabilitation for speech disorders are to

enhance the speech production skills, to apply augmentative and alternative communication systems, and to train their conversation partner perception for the impaired speech. According to many researches, about 10 to 20 percent of people have speech disorder problems in Taiwan and many of them need AAC to help their communication task.

Speech transformations via automatic speech recognition systems is a promising technique. However, there is a major problem. That is, the cross-and within-speaker variability is huge and the phonetic classification could be far too distorted in the impaired speech.

The main purpose of this research project are: First, we will try to investigate the problems of the perceiving of impaired speech by machine. Second, we will apply the present speech techniques including speech recognition and speech synthesis to develop an electronic device with intelligent and efficient user interface for speech impaired subjects. And also, we will evaluate the device according to the improvement of the communication task for subjects.

A PC based software simulation programs that can recognize and synthesize speech commands have been developed. The performance of the system has being evaluated. The software system is being transferred to a DSP emulation and development system. A hardware prototype device based on TMS 320C31 and a control chip AT8515 has been integrated and tested. The final version of the prototype will be evaluated and make it portable.

There are very few projects which are presently researching for the topic of adopting speech recognition technology in AAC system. Almost all the AAC systems using speech technology are developed oversea, and are difficult to be used directly in therapy. We believe that the study proposed by this project

will obtain a valuable result for the development of speech related AAC systems in the future.

Keyword

Speech recognition, Speech synthesis, Augmentative and Alternative Communication, assistive technology

二、緣由與目的

語言及聽力是人類主要的溝通工具之一。因此，具備正常的語言及聽覺能力方可正確的表達自己的思想並瞭解對方傳遞的訊息，其生活的範例之一包括，透過語言的交流取得別人的合作，建立合宜的人際關係，與他人分享經驗等。若語言表達的構造或功能因腦中風、運動神經障礙、唇顎裂、發音方式錯誤、用聲不當、腦傷、智能不足、情緒障礙、聽力障礙等受到損害，將造成語言或聽覺能力異常，進而引起構音異常、運動言語異常、失語症、嗓音異常、語暢異常、語言發展遲緩以及音啞等症狀。而這些人口音常無法有效的表達或與他人進行有效溝通。這種溝通障礙極易導致學業成績低落、人際關係疏離、工作適應不良、心理及行為偏差等現象而引起學校、家庭與社會問題。甚至對於成長中的幼兒，語言的障礙將會影響將來的社交活動而造成學習遲緩或障礙，因此為身心障礙者開發一電腦輔助裝置是一件非常重要的事。

輔助溝通系統 (Augmentative and Alternative Communication; AAC) 拜電子科技高度發展之賜，已逐漸受到重視，不過仍以合成或數位化語音的系統為多。一般其方法是由使用者口語輸入，經過系統的辨識並加以修正，但在文獻上這類的研究可說鳳毛麟角相當稀少。不

僅如此台灣目前亦十分缺少此類型本土化之溝通輔具工具，大大影響臨床上語言治療的效果，並限制病患溝通能力。所以，本計劃之主要目的即在於研究如何應用現行的語音科技於語言障礙復健上，其實施方法除語音引擎的改良與搜尋演算法的改良外並將包括自行設計與開發一電子裝置並評估其效果。進而可應用高科技溝通輔具達到多面相(multiple-phases)之溝通功能，發展適合不同類型及嚴重度之語言障礙病患，且能配合肢體控制能力，家人支持度，生活環境以及實際溝通需求，能有效提高語言治療療效及溝通效果之本土化高科技溝通輔具。

三、結果與討論

本系統裝置採用德儀公司(Texas Instruments, USA)生產之 TMS 320C3x DSK 發展評估板為主要之硬體核心，並加上智控科技(台北, 台灣)之功能延伸板，組合成一數位信號處理實驗板；本計劃利用此一發展系統為裝置雛形，進行語音辨識裝置的開發，為了方便發展的緣故，本系統必須配合個人電腦作一開發介面，也就是利用個人電腦進行程式的撰寫與編譯，在模擬驗證無誤後，下載至上述之實驗板，經實驗板將程式載入後即可進行程式運作，並開始進行語音辨識的動作。

本系統分為兩個運作週期，一個是訓練週期，一個是辨識週期；在訓練週期時，使用者依照畫面指示說出語音，系統此時將語音錄下，並進行特徵值抽取，當作參考樣本，此為訓練週期。在訓練完所有命令後，即可進入辨識週期，在辨識週期中，系統不斷錄下聲音，並即時從中抽取特徵值，與系統在訓練週期時所內存的基準語音特徵進行比對，並將相似度大

於內定之臨界值的相符命令加以挑出並對相似度進行排序，在第二次相符命令出現後即進行前後文關聯判斷，並將符合關聯且相似度最大之辨識結果求出，此一辨識結果將作為輸出語音函式的參數；語音輸出函式則根據辨識結果參數選擇其相對應的音檔，也就是語障者本應發出的正常語音，將其撥放，如此便完成一道命令的辨識。

本系統採用大量的積體電路(IC)，期望藉此將成本降至最低。系統中負責處理語音信號編解碼的 IC 是實驗板上的語音編解碼晶片 CS4215(Cirrus Logic, USA)，其功能為對語音進行數位化及輸出語音的動作，以便讓數位信號處理晶片可以對語音進行處理。

數位信號處理核心為美國德儀公司所生產的 TMS 320C31，其匯流排寬度 32 位元，運算時脈 33MHz，其主要工作為負責聲音信號的數位信號處理及特徵值萃取包括：分框、預強調、取窗、求線性預測編碼參數、倒頻譜運算和特徵值的比對工作。

決策判斷核心為一顆單晶片控制器 AT8515(Atmel, USA)負責，其主要工作為對語音信號的比對結果進行排序，與最後的決策動作，也包含了整體系統人機介面操作的部分。系統方塊圖如圖一所示。

總結的說，在硬體方面是以最普及的 IC 來設計的電路系統，所以並無很多的資源可供應用，譬如負責決策與人機介面的單晶片控制器的可定址空間太小，數位信號處理晶片的處理時脈不夠快，整體功率消耗太大等等；所以硬體方面仍有改善空間，具體建議是如能結合業界技術將可改善上述缺點。

在軟體方面，採用命令控制與語音合成技術結合，而語音辨識的核心則是採直接動態時間扭曲(Dynamic Time Warping, DTW)比對技術，採取有限數目命令的策略，並且加大命令之間的

特徵距離以期望提高辨識率。這樣的做法並不是最理想的，但就成本與複雜度考量也不失為一權宜之計；如果能再進一步對語障者進行標音系統的建立，並且引入以統計為基礎隱藏式馬可夫模型，將隱含在語音信號中的統計特性找出，想必會對辨識率與字彙數目的提昇有很大的幫助。

四、計劃成果自評

本系統應用語音辨識與合成技術與電子裝置技術完成一語障者溝通輔具，此一裝置的完成，將可以帶給無法以直接治療予以矯正的語障者一個有效與他人溝通的途徑，本計劃目前正進行臨床的錄音及辨識測試。因為裝置與開發環境已經完備，所以如果能藉由足夠的臨床測試數據加以修正，則本系統會更趨實用。

此外，對語障者建立一標音系統將是要持續戮力的目標，當然因為重度語障者的變異性太大或是發音器官的損傷，使得這條路困難重重，但針對一部份具有發音重複性高的語障者來說，本系統仍有高度的價值，所以本系統將是值得持續發展的。

最後，整個系統的成本仍嫌太高，這將使得原本的好意變成身心障礙者遙不可及的夢想，所以如果能結合業界的力量，將此一好的構想加以降低成本，變成人人買得起的商品，幫助到真正需要幫助的人，也就不枉費本計劃所投入的心血與初衷的美意了。

五、參考文獻

- Alm N. and Parnes P., Augmentative and alternative communication: past, present and future, *Folia Phoniatrica et Logopedica*. 47(3):165-92, 1995.
- American Speech-Language-Hearing

Association. The prevalence of speech and language disorders in preschool, school age, teenager and adults in the United States in 1977. From www.asha.org, 2000.

- B. R. Bai, Lee-Feng Chien, L. S. Lee, "Syllable-based Relevance Feedback Techniques for Mandarin Voice Record Retrieval Using Speech Queries," Proceedings of the 1997 IEEE International Conference on Acoustics, Speech and Signal Processing, German, 903-906 (ICASSP '97).

Peng, Yu-Hao, Hsu, Chien-Yeh, Pan, Li-Chern, 應用語音辨識與合成技術於語言障礙復健之裝置模擬, in Proceeding, 中華民國八十九年醫學工程年會暨醫工科技研討會, December, 2000.

Yu-Hung Kao, P.K. Rajasekaran, A Low Cost Dynamic Vocabulary Speech Recognizer on a GPP-DSP System, Texas Instruments,

Yifan Gong, Yu-Hung Kao, Implementing a High Accuracy Speaker-Independent Continuous Speech Recognizer on a Fixed-Point DSP,

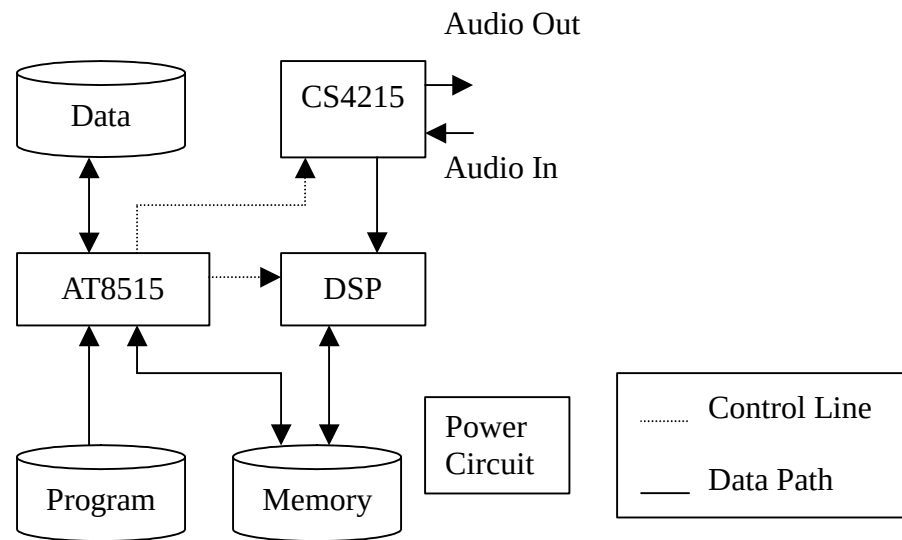
王小川：國家科學委員會專題研究計劃成果報告。總計畫：適合視障者使用之電腦介面技術與系統設計（三）。子計畫一：國語關鍵詞語音之強健性辨認方法及其在視障者電腦之應用（三）。Y2K生醫科技工程論文研討會(2000)。

曾進興、王文容、徐靜音、劉惠美與黃國祐(1996)說話清晰度的概念及測量。中華民國特殊教育學會八十五年年會專輯：特殊學生的學習與轉銜。中華民國特殊教育學會。

林寶貴。我國四至十五歲兒童語言障礙出現率調查研究。特殊教育, 12, 10-13, 民73。

李淑娥。非口語溝通法：重度語障之最佳管道。聽語會刊, 3, 28-34, 民75。

張傳濱、陳友倫：供身心障礙者使用之頭控電腦滑鼠系統之研發。Y2K生醫科技工程論文研討會(2000)。



圖一、系統方塊圖