

臺語多聲調音節合成單元資料庫暨文字轉語音雛形系統之發展

余永吉、鍾高基、*吳宗憲

國立成功大學醫學工程研究所、*資訊工程研究所

Establish Taiwanese 7-Tones Syllable-based Synthesis Units Database for the
Prototype Development of Text-To-Speech System

Yung-Ji Sher, Kao-Chi Chung, *Chung-Hsien Wu

*Institute of Biomedical Engineering, *Department of Computer Science and
Information Engineering, National Cheng Kung University, Tainan, Taiwan*

摘 要

臺語為具有入聲與鼻音等特性的多聲調語言，故建立聲調組合的音節資料庫，為計算語言學研究的基礎。本研究之文字轉語音雛形系統，係由漢字或臺語現代文輸入、單音字詞資料庫、字轉音處理模組、合成單元音檔資料庫及臺語語音輸出處理模組等子系統所構成。2557 個音節合成單元資料庫的建立，係參考漢字與羅馬拼音文獻及字典。合成單元音檔採用載字句錄音，合成器使用基週同步疊加法。系統效能評估則以臺語現代文語詞輸入，聽寫測驗十位受測者，紀錄可辨度與自然度。結果顯示可辨度平均為 83.5%，MOS 自然度量表平均為 3.188。

關鍵詞：臺語、合成單元資料庫、音韻學與數位訊號處理、語音合成

緒 論

閩南語或稱廈門話，源起於西元三世紀以前，隋唐時代中原的口語，隨著族群陸續遷徙到閩南；十六世紀起，再隨著移民逐漸散佈到臺灣暨東南亞各國，如泰國、新加坡、菲律賓、馬來西亞、印尼等[1]。東南亞地區約有四千九百萬人的母語是閩南語，人口數在世界語族排名第二十一位，比某些歐洲語言如義大利語及荷蘭語還多[2]。

臺灣是多語言的社會，現有語言源自於原住民語、閩南語、客語及北京語。依據學者之研究，語言依語系及起源地區形成不同特性，東方語系屬聲調語言(tone languages)，原住民語屬南島語系，閩南語與客語屬漢藏(Sinitic)語系，北京語屬阿爾泰(Altaic)語系[1, 3]。在日常生活中，各族群經常使用自己的母語來交談，週邊國家的語言如英語及日語也常出現於某些特定的場合。本研究所稱的臺語，係指占母語人口數四分之三強的閩南語。閩南語(Amoy Chinese)源起於福建省南方之漳、泉州、及廈門，早期隨移民遷徙至臺灣後分別形成不同的腔調，如府城(今臺南)、鹿港(偏泉州)、艋舺(今臺北萬華)及宜蘭(偏漳州)腔等。因口音交流、歷史經驗、地理背景、族群融合與外來文化，臺語各不同腔調的口音逐漸融合，且形成有別於原鄉閩南之獨特語音(phonology)、語彙(morphology)、語意(semantics)及語法(syntax)[4]。臺語具有多聲調、入聲、詞前轉調、多鼻音(nasals)而缺乏捲舌音等音韻學上之特色，直接影響計算語言學之語音研究方法與實驗設計，例如基本合成單元(basic synthesis units)之選取、字轉音規則庫之建立、語音特徵參數(parametric extraction)之選擇與量化分析[5]、語音合成法則(synthesis algorithm)之決定等。目前臺語相關之計算語言學研究相當缺乏，此領域的語音資訊實屬研究開發的處女地。

臺灣語音溝通科技之研發肇始於 1980 年代的中文語音合成與辨識，現有研究成果應用於人性化溝通之電腦人機介面、電話語音系統等，均專注於北京語(Mandarin)系統的研發。現階段的臺語研究，較專注於臺語語言學特性的探討[6, 7, 8, 9]，而臺語計算語言學研究仍屬萌芽階段[10]。近年來臺語計算語言學論文發表的研究群，包括長庚大學呂仁園與清華大學江永進教授的大量詞彙辨識系統[2, 11, 12]、臺灣大學陳信希教授[13]、成功大學王駿發、吳宗憲、鍾高基教授的文字轉語音系統[14, 15]等。臺語計算語言學研究所遭遇之困難，主要是缺乏高效率的文字系統及參考文獻。臺語為典型的口語式語言(oral language)，尚無統一之文字，對於人文科技之發展與本土文化之傳承，深具負面影響。目前研究所採用的文字系統，多沿用以北京語發音、語彙及語法為主的漢字。漢字字集龐大複雜且總字數無法確定，不僅文字音、形分離，發音及轉調對應規則複雜且特殊情況甚多；至於採用罕見漢字之臺語漢字系統，在目前普遍使用的中文電腦內碼系統中並不完整，導致甚難提供科學化語音合成、辨識與理解系統之研發[16, 17]。

臺灣目前大力提倡母語教育，但首先需要針對書寫的文字作科學化的分析與研究，以發展有效率的母語教材。目前使用的臺語文字根據互異之基本音素(phonemes)單元，可分成三大類，包括：漢字(Chinese-substituted)、拼音字(Spelling)與音漢混合字。漢字系統之使用，最早追溯於 1566 年的”荔鏡記”文獻，教會羅馬字(Romanization，以下簡稱教羅)，源於 1832 年西方傳教士所建的”廈門白話字”，是臺語拼音字的雛形。以“東”字為例，漢字”東”的臺語標音方式為”德紅切”，即採用前字的聲母與後字的韻母和聲調，來表示該字的發音。而教羅僅需標示”tang”，音形意合一[1, 18]。因此，臺語漢字音、形分離，缺乏教育及訓練的效率性。閩南語十五個子音最早記載於西元 1800 年黃謙的”彙音妙悟”字典，教羅則以六個最基本的單母音 a, e, i, o[•], o, u 及其組合來標示臺語的母音。傳統閩南語字典使用漢字標音，難以準確標示讀音，無法符合資訊化處理的原則。教羅針對臺語語言的特色，應用特殊符號或數字來標示聲調，以字尾變化配合特殊符號來區分入聲。其缺點乃未善用拼音字母，且屬於表音式音標而非拼音文字。

臺語為多聲調的語言，一般將臺語分成七或八個聲調[17]，不同聲調的音節即具有不同的語意，聲調記號的標示是臺語拼音文字化的重要課題。臺語的特色是具有許多的鼻音，以子音”k”與母音”oa”所組合的音節”koa”為例：前鼻音與後鼻音分別代表”汗”及”縣”，亦即相同子音與母音組合的音節，會因鼻音在前與鼻音在後之不同，而代表不同的意念。使用漢字無法直接表達讀音，而教羅則以”n”的上標與否表達鼻音位置之不同，例如”koaⁿ”及”koan”分別代表”汗”及”縣”，相同的字母組合無法明確的表達鼻音的差異[18]。臺語最大特色在於複雜的轉調規則，例如單音節位於單字詞與詞末時，發本調音；若位於多字詞的前幾個字時，則發轉調音，稱為詞前轉調的特性。若以漢字紀錄臺語，必須正確斷詞後才能轉調，斷詞是漢字文章的一大問題，因此漢字完全無法表達轉調的特色。且漢字的臺語讀音常為一字多音，難以取捨讀音，例如”香港的香很香”。教羅為音節(syllable)單元的記音式音標(phonemic transcription)，一律紀錄本調音，無法明確表達詞前轉調的特性，難以記憶背誦與教育訓練。

林繼雄教授於 1943 年提出以語詞(words)為單元之音、形、意三合一文字概念，稱為臺語現代文(Modern Literal Taiwanese, MLT)[19]。臺語現代文承襲教羅，

逐一改進缺點，共使用 27 個字母及其組合來記錄臺語，包括英文的 26 個字母，及特殊字母'o'。首先將易混淆的母音'o'及'o'，標示成'o'及'o'(或'o')；子音'c'依所接母音之不同分成'c'及'z'；前鼻音化音節則以'v'為標記的字母，改接於子音字母之後；選取教羅不常用或不使用之字母為聲調記號，透過不同的字母代表臺語的八個聲調，分別是高調音(raised tone)、上突音(pushes out tone)、下突音(depressed tone)、低促音(low stop tone)、迴旋音(bend tone)、迴升音(bend-up tone)、基調音(fundamental tone)及高促音(high stop tone)；再以字尾之不同字母變化標示入聲；以語詞為單位，語詞拼字乃直接拼寫轉調後的拼字組合等。臺語現代文符合臺語現代化與國際化的需求，期許經由減少方言腔調爭議之妥協拼字，改善現有臺文表達困難與不確定性(ambiguous)等缺點，進而達到臺語文字化、資訊化發展之終極目的。

根據語言學及音韻學原理[20]，所有語言的語音，均可由有限之單元(units)所組成。語音合成系統的合成單元包括：音素(phonemes)、單音節(syllables)、雙音(diphones)、三音(triphones)、多音(polyphones)，次音節(semi-syllables or demi-syllables)與不定長度(non-uniform)等，應依需求選取適當的合成單元。西方語系屬句調語言(intonation language)，音節僅指子音與母音的組合，音節內僅區分重音與次重音等。

近年來文字轉語音系統，多使用連續語音資料庫來建立合成單元，其流程包括基週偵測及平滑化、語音單元的平滑化、頻譜參數的求取、合成單元挑選、以及人工檢驗。基週偵測及平滑化指對於語音資料庫中的每一個語音單元，使用聲母/韻母分割的演算法以及自相關法來自動地求取基週。基週軌跡的量化是以離散的雷建德多項式[5]，轉換成四維的基週向量 (a_0, a_1, a_2, a_3) 來表示基週軌跡，並將基週軌跡作平滑化的處理。由於某些語音單元的發音太短促、音高太高或太低等，故必須過濾這種語音單元，稱為語音單元的平滑化。相鄰合成單元間頻譜軌跡的連接越平滑，產生的語音會比較平順悅耳，因此在語音合成時，以較接近共振頻率(Formant Frequency)的線頻譜對頻率(Line Spectrum Pair Frequency，簡稱LSF)來作為頻譜參數[21]，選取最小的音節間頻譜失真的單元。合成單元挑選則使用兩種失真度的量測來決定每一種發音的基本合成單元：第一種為「音節失真

度」，乃對於同一種發音的語音單元，計算分析其間的距離；第二種是「音節間失真度」，用來計算分析不同發音單元之間的頻譜失真度。目前合成單元之篩選，係先由電腦自動化輔助過程，再經由人工檢查，剔除音質不好的合成單元，最後挑選一個適當的合成單元。

研究目的及重要性

本研究之目的為應用語音溝通的數位訊號處理科技以及臺語現代文，探討與建立臺語語音合成單元基本資料庫並發展文字轉語音合成系統雛形。本論文的特定目標(Specific Aims)為：

1. 探討分析臺語音韻學中的音素、音節、聲調、相關拼音法則與轉調規則的特性，建立資訊化處理時最佳紀錄及分類方式，以發展臺語多聲調單音節(syllable-based) 合成單元基礎資料庫；
2. 依據語言學的原則與方法：選擇載字句，錄製多聲調載字句電子音檔，並利用數位訊號處理方法切割、評估與篩選，以建立合成單元語音資料庫；
3. 依據臺語聲調與轉調特性，發展字轉音處理模組，以建立漢字與臺語現代文轉語音合成系統之雛形。

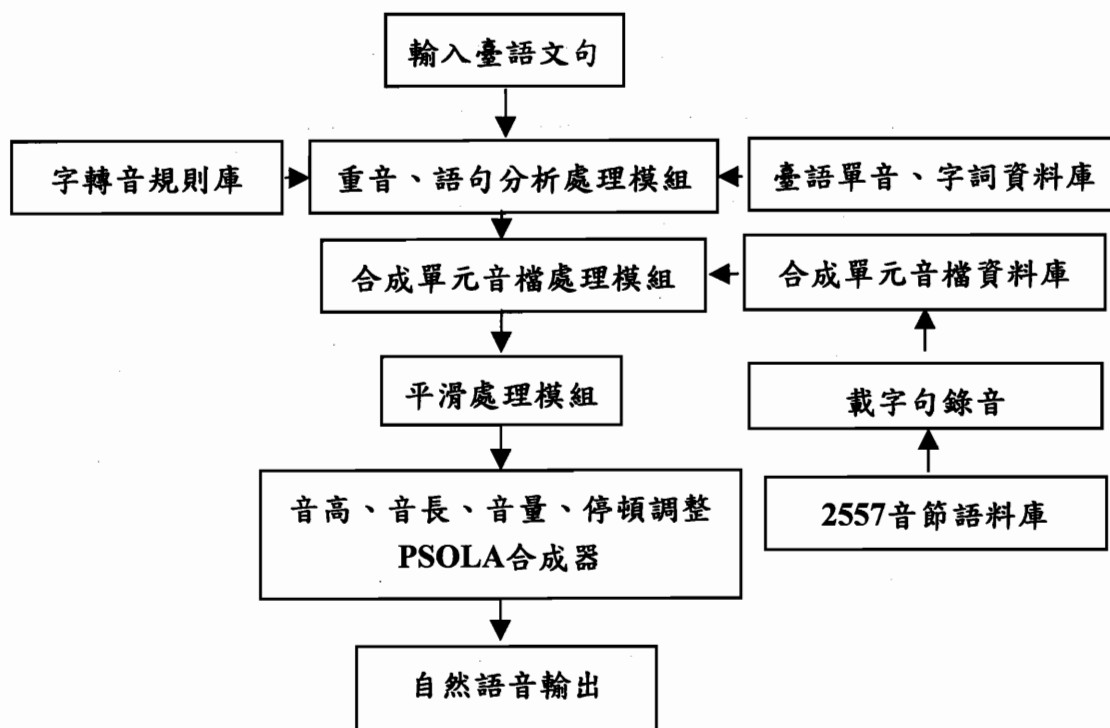
本研究的重要性乃針對臺語文字化、現代化、資訊化及國際化的前瞻性發展需求，提供臺語資訊化發展及計算語言學相關研發的基礎，同時廣泛的應用到一般成人與兒童早期的母語教育、電腦輔助語言教育與訓練、聽語障礙者特殊教育及語音溝通輔助復健科技之發展。

方 法

本研究考量臺語的聲調特性與內含的音韻訊息，因此採用于音、母音與聲調組成的音節作為基本合成單元，並根據林繼雄教授的臺語現代文，將鼻音化的子音與不同聲調的母音，均歸類為相異之基本音素單元，以明確區分不同聲調的特性。故本系統的臺語基本音素包括 32 個子音、33 個單聲調母音與其八個聲調的變化組合[15]。

實驗設備/特徵參數萃取：錄音系統為具方向性且抗背景雜音之電容式麥克風，架設在 Pentium PC 上，取樣頻率為 22.05 kHz，解析度為 16 bits。以短時間視窗(short-time windowing)之時域及頻域分析及萃取特徵參數，每一個視窗長度約為 40 ms，且至少包含七個基週。並依據能量及過零率曲線之分析，區分音節與語音段。

實驗方法與步驟：本研究之整體架構如圖一所示，系統發展分成三個階段：第一階段為文獻探討及考量不同的音節單元組合，發展單音、字詞資料庫、字轉音規則庫與分析處理模組，並建立以單音節為基礎之臺語七聲調基本合成單元資料庫；第二階段進行載字句之錄製、切割、評估與篩選，建立七聲調音節合成單元電子音檔資料庫；第三階段為七聲調臺語語音合成系統之雛形研發及系統效能評估。具體的步驟與方法詳述如下：



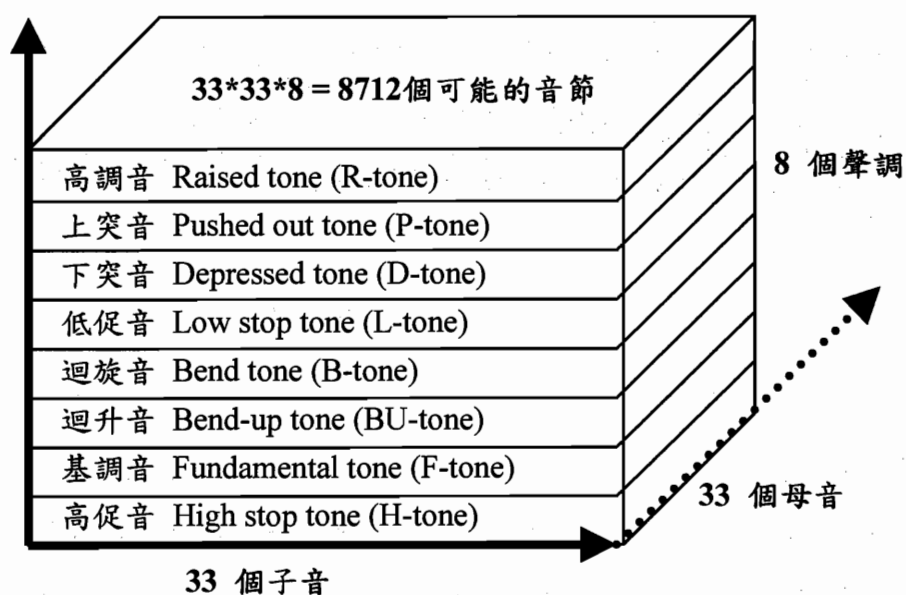
圖一：本研究之架構圖

1. 發展及建立多聲調音節合成單元資料庫：

1-1 首先建立臺語字典與辭典電子檔，並針對臺語的特性、轉調規則、文白讀音與不同腔調，發展與建立字轉音規則庫及分析處理模組。經由文獻探討歸納出子音音素共計 33 個，包括：1 個音節內只存在母音而無子音之情況、18 個一般子音及 14 個後接‘v’的前鼻音化子音。其中 b, j, l, m, n 五個子音的鼻音不存在或本身即為鼻音，鼻音韻母則在母音字母前加上字母‘v’。

1-2 本研究中的臺語母音音素，不特意區別元音、介音與韻尾，只將其組合成的單母音、雙母音及複母音，均視為互異之母音。故母音的總數為 264 個，包含 33 個單聲調母音及其八聲調之組合。雖然母音總數較多，然其主要的優點為不同聲調的音節可經由不同的拼字方式予以明確區別。

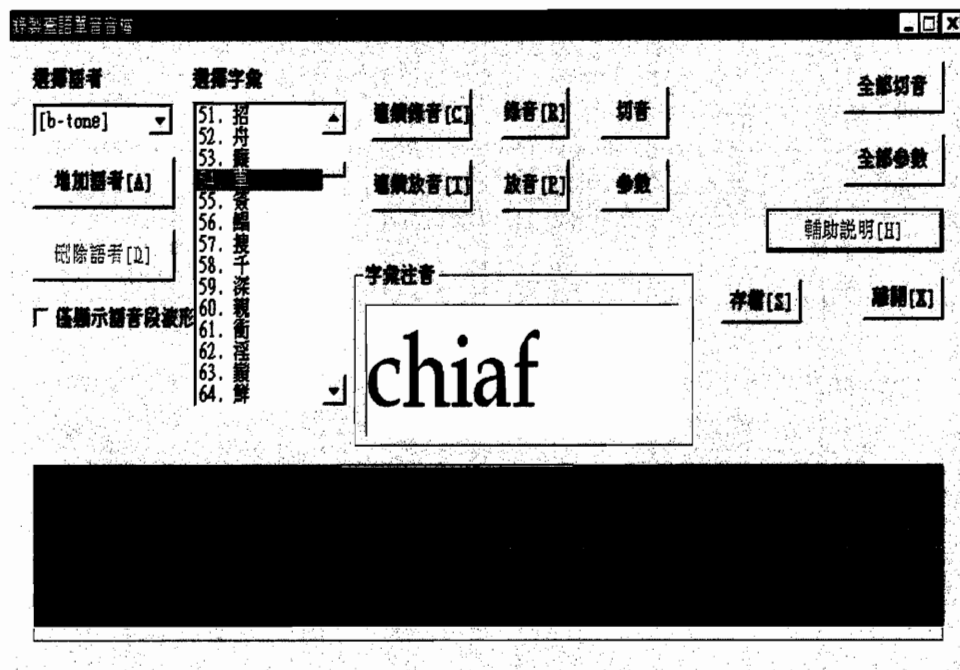
1-3 依據上述基本音素的訂定，共形成 8712 ($33*33*8$)個音節(如圖二)，組合成的音節不一定具有語意。根據拼音臺文漢字輸入法、臺語電子字典、辭典等相關文獻的探討，將 8712 個音節中，無對應漢字的音節剔除，以建立單音節基本合成單元資料庫[15, 22, 23]。



圖二：臺語單音節的組合

2. 建立多聲調合成單元電子音檔語音資料庫

首先錄製上述多聲調音節的載字句。載字句係指一段連續的多音節語句，涵蓋欲錄製的音節，卻未必具有特定語意。載字句選擇的標準，是以該載字句前後單字與所欲錄製語音音節的界線是否清晰為原則。本研究所選用的載字句為”Goafthak ‘S’ ciuxhor” (我讀’S’就好)，’S’表示欲錄製的語音音節，置於載字句之間，以控制音高、音長與音量之變異，並保留較佳的音韻訊息。錄音程式可以同時顯示所欲錄製的單音節漢字與所對應之臺語現代文(如圖三)。錄製載字句連續語音過程中，每秒速度約為三個單音節，音量大於 65dB，音高、音量、速度、及音長儘量保持一致。載字句的錄音錄製過程非一次完成，因此每次錄音時均播放先前錄音之音檔，以維持音高的一致性。



圖三：錄音程式之視窗介面

載字句電子音檔錄製完成後，重新播放以評估其音高、音長、音質、音量與可辨度；根據波形的時域觀察，並經由能量曲線(energy contour)及過零率曲線(zero-crossing-rate contour)分析之輔助篩選，判斷音節界線並切割所需之音段，切割後每個單音節語音音段約為 300 毫秒。若可辨度不佳，或音高、音量、速度及音長相差太大，則重新錄製與分析。

3. 臺語多聲調語音合成系統之雛形研發及系統效能評估

本系統包括文字輸入、單音字詞資料庫、字轉音規則庫與分析處理模組、合成單元音檔處理模組、及語音合成輸出等子系統。系統提供臺語現代文或漢字輸入之功能：若以臺文拼字輸入，則由拼字檢查(spelling check)程式檢查拼字；若為漢字輸入時，則經由漢字轉臺文之詞翻譯系統轉換成臺文拼字，再以字轉音程式規則庫、語句分析處理模組等，調整詞調及句調，對應到合適的合成單元電子音檔，經由平滑化處理及 PSOLA 合成器，最後輸出合成單元資料庫中經過調整處理的自然音節語音。

3-1 漢字轉臺文之詞翻譯系統：現有電腦大五碼(big-5)中的所有漢字，均已鍵入於本系統所建立的臺語單音資料庫中，亦即所有現存於電腦系統的通行漢字，均可找到對應之臺語拼字。故當本系統以漢字輸入時，首先經由電子辭典進行斷詞(segmentation)，再根據臺語語詞的漢字及拼音文字對照資料庫，利用語詞為單位轉換到對應之臺語現代文拼字；無法在辭典中找到的語詞均視為單字詞，並以臺文漢字字典逐字轉換成臺文拼字。

3-2 字轉音規則庫：直接輸入或經詞翻譯後的臺語現代文拼音字，均具有八聲調變化，但是當臺語連續語音整句合成輸出時，有豐富且複雜的轉調規則。基於臺語有文、白讀音、與各地腔調等不同的歧異，例如臺南(偏漳州)與臺北(偏泉州)的迴旋音與入聲，分別適用不同的轉調規則，且漢字的臺語讀音常有一字多音的現象，本研究嘗試根據學者的歸納與整理，發展字轉音的規則程式，以配合字、辭典的音標選取，減少錯誤讀音的發生率。

3-3 重音、語句分析處理模組：在整句臺語的連續讀音時，必須考量加重音、輕讀或綴詞等之讀音轉調，故應分析處理以提高合成語音的品質。經過語句的分析處理，可以作為詞調與句調調整的依據，以提供合成單元音檔處理模組、音韻調整模組、與平滑處理模組調整參數的基本資料。

3-4 其他處理模組及 PSOLA 合成器：本研究目前仍持續收集臺語語料之電子檔，以提供 PSOLA 合成器參數調整規則的歸納或參數調整訓練的參考。

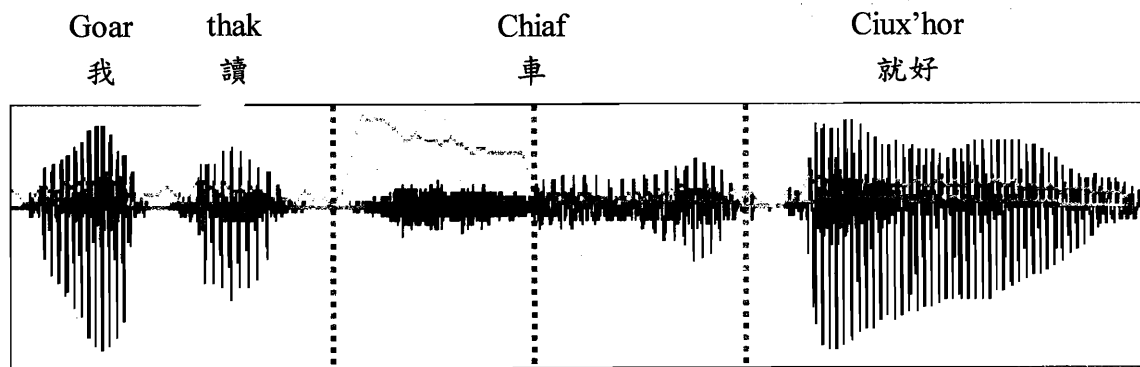
3-5 本系統之表現評估：包括可辨度(intelligibility)與自然度(naturalness)，受測者共計十人，分成一般人士組五人，實驗室研究生五人。可辨度之評估，係選取日常生活最常使用的 20 個二字詞、20 個三字詞、20 個四字詞成語、與臺語教材文句 10 句。其中二字詞與三字詞多為白話音，四字詞則為文言音，文句則屬生活對話(如附錄一)，不同子音開頭之語詞均加以選取，由系統合成臺語語音，受測者以聽寫方式寫下所聽到的語詞，計算受測者書寫語詞之正確率，作為可辨度的評估結果。

自然度則採用平均鑑定分數(mean opinion scores, MOS)，區分為五級，即優良(excellent)、良好(good)、尚可(fair)、差(poor)、極差(unsatisfactory)，由受測者在聆聽系統合成語音，分別依所感覺的自然度給予 5 至 1 分的評分。

結 果

建立 2557 個七聲調的音節資料庫 本研究建立的 2557 個七聲調音節資料庫，包括：394 個基調音的音節(如附錄二)、418 個高調音的音節、426 個上突音的音節、394 個下突音的音節、292 個低促音的音節、387 個迴旋音的音節單元、及 246 個高促音的音節單元。傳統臺文僅有七個聲調，臺語現代文將迴升音視為一個獨立的聲調，然文獻中甚難確定迴升音的音節總數，此情況頗類似國語的輕聲。本研究僅建立傳統七個聲調的音節資料庫，至於臺語的聲調區分，有待進一步的探討與量化。

建立七聲調合成單元語音資料庫 對應上述之音節資料庫，共錄製 2557 個載字句的電子音檔，錄音時的速度、音量、音高及音長，以及訊號處理時擷取的長度(length of window)，均採用中文系統目前研發的經驗參數值。此未必符合臺語之語音特性，故將來更需要進一步的探討。以臺語的”chiaf”(車)為例，載字句”Goarthak chiaf ciuxhor”(我讀車就好)的時域波形及過零率、能量、基週曲線如圖四所示。自左端算起第三個較大的波形為本實驗所切割的音節合成單元”chiaf”(車)，其前後音節”讀”(~thak)和”就”(ciux~)的母、子音和”chiaf”形成可區分的連音波形，音節界線清晰而易於切割。圖五為”chiaf”(車)切割後的時域波形，前段振幅較小為子音’ch’，後段類週期訊號為母音’ia’，’f’為高調音標示記號。



基週曲線
 過零率曲線
 能量曲線

取樣頻率：22.05 kHz
 解析度：16 bits
 時域波形

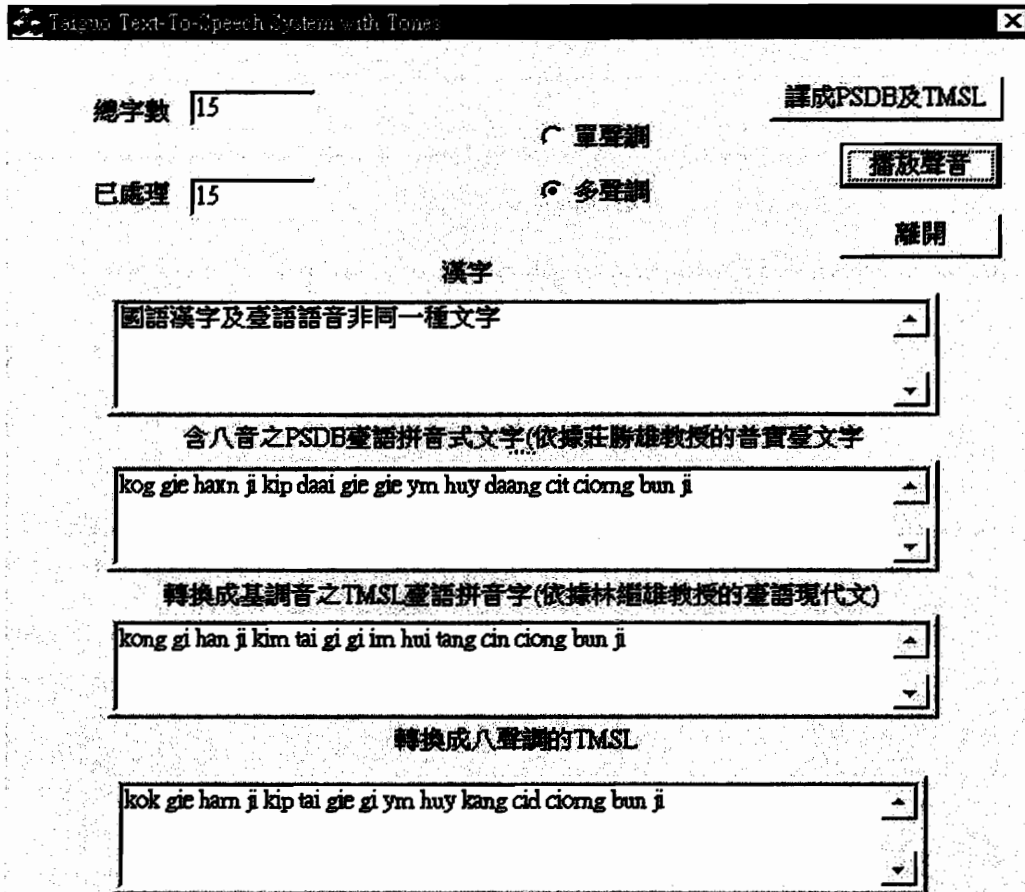
圖四：載字句 “Goafthak chiah ciuxhor.” (我讀車就好) 語音訊號之特徵



圖五：合成單元“chiah”(車)的時域波形

七聲調語音合成之雛形系統

雛形系統之程式係在中英文電腦上編譯，為程式編寫方便，必須英文字母化，故本研究將臺語母音'ò'以'Q'標示。臺語並無統一之文字，文獻中記載的不同文字系統均具參考價值，因此本研究針對最為廣泛使用之漢字、教會羅馬字、臺語現代文與普實臺文四種文字系統，建立文字及音標方案之音素對照表，提供不同臺語文字系統的轉換。教會羅馬拼音及普實臺文文字系統皆經由上述音素對照表自動轉換譯成臺語現代文，而漢字系統則必須經由詞翻譯系統轉換成臺語現代文，以合成語音輸出，如圖六所示。普實臺文文字系統目前具有電子字辭典、拼字檢查程式、及臺文輸入轉漢字之功能。



圖六：臺語文字轉語音系統雛形

本研究目前所發展的臺語現代文單字詞資料庫包括 20,000 個漢字與拼字之對照表，及多字詞資料庫包括 40,000 個語詞漢字與拼字之對照表(表一、二)；且針對上述單字詞及多字詞資料庫所對應的臺語文言音、白話音、漳州音、泉州音等逐步增訂音標及建立字轉音規則。拼音檢查程式乃根據上述發展之資料庫，檢查比對輸入之臺文拼字是否正確。為配合臺灣官方漢字使用之現況，發展漢字轉臺文的詞翻譯系統，以提供使用漢字作為文字輸入的使用者，仍可透過系統輸入漢字以合成臺語語音(如圖六)。本雛形系統包括：漢字轉臺文的詞翻譯系統、字轉音規則庫與分析處理模組、及語音處理模組等。

本系統可辨度測試結果如表三、表四所示；自然度測試結果如表五、表六所示。

表一：臺語單字詞資料庫

編號	漢字	臺文
6194	棠	toong
6200	棟	toxng
6231	棉	mii
.....

表二：臺語多字詞資料庫

編號	漢字	臺文	音節
2355	今日	kinafjit	kin'af'jit
3967	螢火蟲	Huefkimkof	huef'kim'kof
.....

表三：可辨度實驗結果

一般人士組(5人)	測試詞句	數量	可辨度
	二字詞	20	88.1%
	三字詞	20	83.3%
	四字詞成語	20	79.6%
	文句	10	72.2%
	平均		80.8%

表四：可辨度實驗結果

實驗室研究生(5人)	測試詞句	數量	可辨度
	二字詞	20	84.5%
	三字詞	20	88.3%
	四字詞成語	20	92.0%
	文句	10	81.9%
	平均		86.68%

表五：自然度實驗結果

一般人士組(5人)	測試詞句	數量	自然度(MOS)
	二字詞	20	3.6
	三字詞	20	3.5
	四字詞成語	20	3
	文句	10	2.5
	平均		3.150

表六：自然度實驗結果

實驗室研究生(5人)	測試詞句	數量	可辨度
	二字詞	20	3.3
	三字詞	20	3.5
	四字詞成語	20	3.3
	文句	10	2.8
	平均		3.225

此2557個音節資料庫係經由剔除不具語意之合成單元及篩選出具有對應漢字之音節，有效減少合成單元資料庫的數量，而不影響整體的完整性。

本系統目前輸出的臺語語音，係經由載字句切割出音節單元所合成。由於載字句未必具有語意，也不一定符合文法規則及語用情境，故所錄製的合成連續語音，其音韻品質較差。二字詞、三字詞與四字詞之可辨度均優於整句文句，係因臺語除詞末音節發本調音外，其餘音節皆必須依規則轉調，稱為詞前轉調。本系統採用之臺語現代文以語詞單位拼寫，其語詞拼字乃直接拼寫詞前音節的轉調音，及詞末音節的本調音，故系統不必處理詞前轉調的問題。但是整句文句除詞調之轉調規則外，仍必須考慮句調之轉調，目前之系統無法有效調整句調，必須進一步錄製平衡句以萃取句調參數。臺灣目前相當缺乏臺語連續語音之音檔處理資料，先天性的限制有關臺語合成法則及合成器參數之研究。因此建立平衡詞句資料庫，實為發展高品質臺語語音合成系統之當前最重要的課題。

臺灣教育部頒訂常用的繁體標準漢字約4,800餘字，均可對應到本系統的電子音檔，並透過語音合成系統輸出臺語語音。臺文中某些常使用的罕見漢字，在現行電腦的通行內碼中並不完整。長期受到百越族方言、日語、英語等之影響，現代臺語中有些常使用的語音沒有對應的漢字[19]，因此必須經由日常生活中之對話、臺語廣播與電視新聞的語音資料、及拼音式文字的文獻中持續的研發探討，以增進本資料庫的完整性。

國語與臺語實質上具有不同之音韻、語意、文法、語用等特徵[16]，例如國語之”颱風”與臺語之”風颱”為語詞的倒裝，故有些國語漢字與臺語語音轉換時，必須考慮機器翻譯(machine translation)的使用。臺語語音合成系統之研發初期，首先應以臺文文字轉臺語語音為基礎，待雛形系統完成後，再考量國語漢字轉臺文之機器翻譯系統等，以減少在初步系統研發過程中的複雜度，且可符合目前臺灣官方使用國語漢字之現況。

本研究發現許多國語漢字無法直接翻譯成符合臺語口語的語詞，主要原因乃目前的系統係以語詞為基礎的直接對照翻譯。本系統若以國語漢字輸入，經詞

庫斷詞後，直接對照漢字之臺語文言音拼字，由於尚未建立臺語口語之拼音文字對照，故系統輸出為臺語之文言音，而非普遍使用之臺語口語詞；若以臺語漢字輸入，由於尚未建立臺語漢字詞庫，故無法以臺語詞庫斷詞，所輸入的漢字將被系統判斷為單字詞，直接對照單字詞庫中之臺語拼音，因此系統輸出為漢字的單字詞拼音，而非以語詞為單位之臺語口語詞。解決之道為：第一種方法乃針對國語漢字詞庫建立對照的臺文口語詞拼音辭典，並依據臺語之口語語詞加入綴字，例如漢字語詞”今日”或”今天”，依臺語口語習慣，最好對應成”kin'afjit”而非”kim'jit/kim'tiefn”，國語漢字語詞對照臺文口語拼字時，必須依據許多文獻考證之結果，以減少對應之爭議；第二種方法為訓練使用者直接以臺語拼音文字作為系統之文字輸入。

臺語漳州與泉州腔之轉調規則不同，若欲產生不同腔調之臺語連續讀音，仍須建立適當之機讀轉調規則。臺語連續語音中，有加重音、輕讀音、介係詞轉調等特殊規則，必須仰賴系統所建立的規則庫，以轉換成正確之音調。語句分析處理模組負責剖析輕讀字、介係詞、重音等之位置，並依規則轉調後，再由合成單元音檔處理模組對應正確之合成單元。

音韻調整處理模組負責句調之調整，本系統目前的研發進度侷限於肯定句、否定句與疑問句三種句型之音高、音長、音量及停頓參數調整。為提高系統語音合成之品質，仍需持續的針對更多之句型探討與選擇適當的調整參數。

臺語現代文係以多音節語詞為拼音化的基本單元，其優點是符合使用者的語言學習經驗，可以有效減少發音與斷詞的不確定性。本研究應用臺語現代文的語言學理論及音韻規則，發展與建立語音合成系統，期待透過教育與訓練，並經由視覺與聽覺的回饋，發展具親和性及效益性的電腦人機介面應用。

結論與未來展望

本研究應用臺語現代文、音韻學學理及數位訊號處理科技，發展並建立 2557 個單音節合成單元語料庫，經由載字詞音檔錄製 2557 個音節合成單元電子音檔資料庫，配合單音、字詞資料庫、字轉音規則庫、文句分析與音韻調整模組之程式發展，研發完成臺語文字轉語音(text-to-speech)系統的雛形。

本研究未來的發展方向：(1) 首先蒐集文獻資料並運用統計方法，發展臺語平衡句語料庫；(2) 利用自相關法(auto-correlation) 或離散的雷建德多項式(discrete Legendre polynomial)求出基週 (pitch)、能量(energy)與音段(duration)曲線(contour)的特徵參數；(3) 由頻域(frequency domain)計算共振波(formant)參數、倒頻譜參數(MFCC: mel frequency cepstrum coefficient)及差值倒頻譜參數(Delta MFCC)等特徵值，共同組成每一個單音節語音之特徵參數矩陣；(4) 利用統計之多變量分析(multi-variable)區分不同之音素、音高、聲調與音韻訊息，以選取最佳化合成單元及調整音韻訊息參數，改善目前臺語語音合成系統的音韻品質。

使用母語具有許多無法取代之優點，而發展一套合乎資訊化、國際化的臺語文字系統，為當前臺灣步入資訊科技時代最重要的課題。發展音、形合一的拼音式現代臺灣語言，必能提高教育、訓練與學習的效率，提供科學研究的有利條件，提昇臺灣人文與科技的品質，對人類文明提供更具體之貢獻。

誌 謝

承蒙林繼雄教授悉心的指導與提供臺語現代文詳盡的文獻資料，本論文得以完成。感謝莊勝雄教授研究群提供的文獻與字、辭典檔案、臺文輸入法與拼音檢查工具程式。感謝成大資訊所學弟林順傑、郭啟祥、醫工所邱毓賢、劉俊明先生等協助系統之架設。

參 考 文 獻

1. 施炳華，”附錄：臺語文常識—臺語尋根之旅，”府城國中雙語教材第一冊，pp. 35-41。
2. Y. C. Chiang, et. al., “A New Hybrid Duration Hidden Markov Model with Application to Large Vocabulary Taiwanese (Min-nan) Word Recognition,” 1st International Symposium on Chinese Spoken Language Processing (ISCSLP98), 1998.
3. 許極敦，臺灣語概論，臺灣語文研究發展基金會，臺北市，初版，1990，pp. 33-54。

4. 許極敦，臺語文字化的方向，自立晚報，臺北市，1992，pp. 3-55。
5. S. H. Chen and Y. R. Wang, "Vector Quantization of Pitch Information in Mandarin Speech," *IEEE Transactions on Communications*, 38(9), 1990, pp. 1317-1320.
6. F. H. L. Jian, "Boundaries of Perception of Long Tones in Taiwanese Speech," *ICSLP1998*.
7. S. H. Peng, "Production and Perception of Taiwanese Tones in Different Tonal and Prosodic Contexts," *Journal of Phonetics*, 25, 1997, pp. 371-400.
8. F. H. Jian, "Perception of Long and Short Tones in Taiwanese Speech," *J. Acoust. Soc. Am.* (1997), 102(5), pp. 3095.
9. S. H. Chen, "Phonetograms of Normal Taiwanese Young Adults," A Thesis of Doctor of Philosophy in University of Wisconsin-Madison, 1996.
10. F. H. L. Jian, "Classification of Taiwanese Tones Based on Pitch and Energy Movements," *ICSLP1998*.
11. R. Y. Lyu, Y. J. Chiang, R. Z. Fang, W. P. Hsien, "A large-vocabulary Taiwanese (Min-nan) speech recognition system based on inter-syllabic initial-final modeling and lexicon-tree search," *ROCLING XI Conference 1998*, pp. 139-149.
12. R. Y. Lyu, Y. J. Chiang, W. P. Hsieh, "A Large-Vocabulary Taiwanese (Min-Nan) Multi-Syllabic Word Recognition System based upon Right-Context-Dependent Phones with State Clustering by Acoustic Decision Tree," *ICSLP1998*.
13. 林川傑 & 陳信希，"中文到閩南語之線上翻譯及閩南語之語音合成，"1999 語文處理技術研討會集刊。
14. 黃保章，"國語文句翻臺語語音系統之研究，"成功大學電機工程研究所碩士論文，1999。
15. Y. J. Sher, K. C. Chung, C. H. Wu, "Taiwanese Syllable-based Synthesis Units Database," *Chinese Journal of Medical and Biological Engineering*, 19(1), 1999,

pp. 47-58.

16. 鄭良偉，走向標準化的臺灣話文，自立晚報文化出版部，1989，pp. 69-100。
17. 洪惟仁，臺灣河佬話語聲調研究，自立晚報，臺北市，1985，pp. 1-47。
18. 林繼雄，林華英，陳煜楠，由漢字注音到臺語拼字文，育德文教基金會，初版，1997，pp. 7-15, 72-133。
19. 林繼雄，臺語現代文，大夏出版社，初版，1990，pp. 1-19。
20. J. P. Gee, An Introduction to Human Language, New Jersey, Prentice-Hall Inc., 1993, pp. 65-134.
21. S. Furui, "Digital Speech Processing, Synthesis, and Recognition," Marcel Dekker, 1989, pp. 5-43.
22. 林繼雄，Taiwanese Dictionary of Words with Modern Spelling，1988。
23. Maryknoll, English Amoy Dictionary, Taichung, 1995.

附錄一：可辨度及自然度評估詞彙集

二字詞部份 (共 20 個)

編號	臺文縮寫	漢字	臺語現代文
1	AK	愛睡(睏)	Aekhuxn
2	EK	浴室	Egkefng
3	HH	高興(歡喜)	Hvoahie
4	KI	喜歡	Kah'ix
5	KT	客廳	Khehthviaf
6	SZ	刷牙	S0efzhuix
7	PJ	小便	Parngji0
8	PK	房間	Pangkefng
9	PaS	大便	Parngsae
10	PiS	廁所	Piexnsor
11	PV	醫院	Pvixvi
12	SiB	失望	Sitbong

13	SqB	洗臉	S0efbin
14	SL	想(思)念	Suliam
15	SeK	逛街	Seqkef
16	SiK	生氣	Siuxkhix
17	SP	散步	Sarnpo
18	SS	傷心	Siongsym
19	ZS	穿衣	Zhexngsvaf
20	TI	討厭	Th0'iax

三字詞部份 (共 20 個)

編號	臺文縮寫	漢字	臺語現代文
1	SSK	洗澡(身軀)	S0efsefngkhw
2	BSK	不舒服	B0sofngkhoaix
3	GBH	我不會	Goar buexhiao
4	GKK	我感覺	Goar kafmkag
5	GMA	我不要	Goar m aix
6	KPZ	看報紙	Khvoax p0rzoar
7	KTS	看電視	Khvoax tiexnsi
8	ZIS	找醫生	Zh0exisefng
9	TOL	臺灣人	Taioaan-laang
10	HZT	蕃薯湯	Hanzuu-thngf
11	TJP	糖尿病	Thngji0xpvi
12	CSL	一世人(一輩子)	Cidsielaang
13	TTK	土地公	Thoftixkofng
14	HZS	現此時	Hiexnzhwsii
15	HTP	虎頭蜂	Hofthauphafng
16	IET	游泳池	Iu'efngtii
17	PTK	怠惰骨	Pintvoaxkud
18	PTH	布袋戲	Port0exhix
19	SLK	少年家	Siaolienkef
20	ZHT	臭火乾	Zhaoh0eftaf

四字詞部份 (共 20 個)

編號	漢字	臺語現代文
1	人山人海	Jinsafn-jinhae
2	三頭六臂	Samthi0o-liogpix
3	千變萬化	Chienpiexn-baxnhoax
4	山珍海味	Santyn-haybi
5	不三不四	Putsafm-putsux

6	不義之財	Putgi-cy-zaai
7	心甘情願	Sym-kafm-zeeng-goan
8	甘拜下風	Kampaix-haxhofng
9	安分守己	Anhun-siwkie
10	作惡多端	Zok'og-t0toafn
11	門當戶對	Buntofng-hoxtuix
12	莫名其妙	Bogbeeng-kibiau
13	耀武揚威	Iauxbuo-iong'uy
14	繡花枕頭	Siuhoaf-cymthi0o
15	談天說地	Tamthiefn-soatte
16	滿面春風	Boafnbien-zhunhofng
17	寬宏大量	Khoanhoong-taixliong
18	福如東海	Hog-juu-tonghae
19	飲水思源	Ymsuie-sugooan
20	頂天立地	Tefngthiefn-libte

文句部份 (共 10 句)

編號	漢字	臺語現代文
1	放棄自己所有的一切	Parngsag kaki sofu ee itzhex.
2	他一時糊塗偷拿錢	Y cidsii hotoo thautheq cvii.
3	你會曉講英語沒	Lie exhiao korng Enggie b0e?
4	廢除壞的風俗習慣	Hoeatii phvae ee hongsiok sibkoaxn.
5	媽媽教我如何掃地板	Mamaf kax goar afnnar saux thokhaf.
6	厝裡有一點仔事情要辦	Zhuxnih u cidtiafmar taixcix boeq pan.
7	大家都是台灣人	Taixkef lorng Si Taioaan-laang.
8	那內面就是冷凍的物件	Hit laixbin ciux si lefngtoxng ee miqkvia.
9	太陽從東邊升起	Jidthaaui tuy tangpeeng khielaai.
10	我要養一隻小隻的鳥仔	Goar b0eq chi cid ciao s0eaciaq ee ciawar.

附錄二：臺語 2557 個七聲調的音節資料庫—僅節錄基調音 394 個

