

# 國語文句翻台語語音系統之研究

## A Study for Mandarin Text to Taiwanese speech System

†王駿發 †黃保章 \*林順傑

{wangjf, hwangbj, linsj}@server2.iie.ncku.edu.tw

†國立成功大學電機工程學系

\*國立成功大學資訊工程學系

### 摘要

本論文建立了一套能說台語的中文文句翻台語語音系統。使用者只要輸入中文詞句，便可輕易地得到對應的台語發音。文中主要針對台語語音性質在語音系統實作時所產生的：(1) 漢字的一字多音、(2) 文句的斷詞、(3) 聲調運作的問題、(4) 合成單元處理等四個問題進行探討。

### 1. 簡介

就文字轉語音合成系統而言，由於國語為普遍性的溝通語言，研究單位也多偏向此領域來進行研究，並且有不錯的研究成果發表。但是在方言的語音合成系統方面，卻鮮少論及。由於母語在現今傳播媒體的使用頻率與日俱增，母語的語音合成系統，在運用上也變得有其需要性。有鑑於此，乃嚐試開發會說台語的文字轉語音合成系統。

國語文句斷詞處理完在漢字對應於台語發音所產生的一字多音問題，本文是先對各音做優先音排序，若仍有取音混淆情形則將某音以詞庫窮舉的方式來解決。

台語聲調的變化是與國語差異最大的部份。本論文先決定了一套語音合成系統上所適用的台語聲調種類與變調規則；然後再提出一套規則來處理句子中，何字該變調，何字不該變調的問題。最後，對於本系統所用的合成單元是直接取自本論文提出的「固定音高旋律性錄音法」所錄得的語音資料。

本論文的章節架構如下：第二節是研究現況與研究背景；第三節則介紹我們資料庫的建立；第四節說明研究方法-斷詞、台語的取音及聲調變化的處理情形[21]；實驗的結果則在第五節描述；最後，第六節則是結論與討論。

## 2. 研究現況與研究背景

### 2-1 研究現況

台灣語音溝通科技之研究肇始於 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)系統的開發。現階段的台語研究，較專注於台語語言學特性的探討，而台語計算語言學研究仍屬萌芽階段。近年來台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[15][16][17]，台灣大學陳信希教授[18]、成功大學吳宗憲、鍾高基教授的文字轉語音系統[19][20]等。

### 2-2 拼音符號介紹

台語和國語皆屬漢語語系分支，是單音節語言，由所謂聲母、韻母和聲調三部份所組成。本文修改部分【國台雙語辭典】的音標方案，以利於論文中的台語語音標示，其音標形式詳見表 2-1。聲母有 17 種、韻母 48 種，各欄名稱解釋如下：「國」是指注音符號，「台」、「楊」是指國台雙語詞典所採行的音標符號，「黃」是指本論文的音標方案，「教羅」台灣教會所採行的音標方案，「例字」是發此音的漢字。

### 2-3 一字多音的處理

由於時間悠久、空間遼闊及其他因素造成台語的一字多音的現象。本系統的漢字對應台語音個數，一個漢字對應台語音節的數量以九個音為最多，計有「撓那落」三個字。傳統台語的漢字發音教學方式，在由字發音時，多半只教文白音的分別念法，並未指定規則來擇取。今若以系統實做考量，這種分類方式是很籠統的，況且單一漢字的台語音節數目也不止兩種而已。要決定取該字的哪一個音節來發音，便成為猶疑難決的問題。

例如：香港的香很香。(hiong2-kang2 ye7 vhiol ziong8 pang1)三個『香』皆發不一樣的音，而字典中「香」有五種讀音(vhiu1 白話音,vhiol 白話音,hiong1 文言音,hiang1 文言音,pang1 俗音)。

國台聲母 (CONSONANT)

國	ㄅ	ㄆ	ㄇ	ㄏ	ㄉ	ㄊ	ㄋ	ㄌ	ㄍ	ㄍ	ㄎ	ㄏ	ㄗ	ㄗ	ㄘ	ㄙ					ㄗ	ㄗ		
台	ㄅ	ㄆ	ㄇ		ㄉ	ㄊ	ㄋ	ㄌ	ㄍ	ㄍ	ㄎ	ㄏ	ㄗ	ㄗ	ㄘ	ㄙ					ㄗ	ㄗ		
黃	b	bh	p	m	d	t	n	l	g	gh	k	h	z(i)	zh	ch(i)	s(i)					z	zh(u)	ch	s
楊	b	v	p	m	d	t	n	l	g	q	k	h	z(i)	j	c(i)	s(i)					z	j(u)	c	s
教	p	b	ph	m	t	th	n	l	k	g	kh	h	ts	j	tsh	s					ts	j	tsh	s
羅													c		ch	chh					ts	c	ch	chh
例字	褒	帽	波	冒	刀	桃	奴	囉	哥	鵝	科	號	之	字	癡	施					資	裕	此	思

國台韻母 (VOWEL)

國	ㄩ	ㄩ	ㄩ	ㄩ	ㄩ	ㄩ	ㄩ	ㄩ	ㄩ	ㄩ	ㄩ	ㄩ	ㄩ	ㄩ	ㄩ	ㄩ	ㄩ	ㄩ	ㄩ	ㄩ	ㄩ	ㄩ	ㄩ	ㄩ
台					ㄩ	ㄩ	ㄩ	ㄩ	ㄩ	ㄩ	ㄩ	ㄩ	ㄩ	ㄩ	ㄩ	ㄩ	ㄩ	ㄩ	ㄩ	ㄩ	ㄩ	ㄩ	ㄩ	ㄩ
黃					a	va	er	e	ve	ai	vai		au	vau		am	an	(n)	ang	*	m	ng		
楊					a	a <sup>n</sup>	ə	e	e <sup>n</sup>	ai	ai <sup>n</sup>		au	au <sup>n</sup>		am	an	(n)	ang	*	m	ng		
教羅					a	a <sup>n</sup>	o	e	e <sup>n</sup>	ai	ai <sup>n</sup>		au	au <sup>n</sup>		am	an	(n)	ang	*	m	ng		
例字					膠	監	阿	高	家	更	皆	閒	交	艾	甘	干					江		姆	鋼

國	ㄩ	ㄩ	ㄩ	ㄩ	ㄩ	ㄩ	ㄩ	ㄩ	ㄩ	ㄩ	ㄩ	ㄩ	ㄩ	ㄩ	ㄩ	ㄩ	ㄩ	ㄩ	ㄩ	ㄩ	ㄩ	ㄩ	ㄩ	ㄩ
台	ㄩ	ㄩ	ㄩ	ㄩ	ㄩ	ㄩ	ㄩ	ㄩ	ㄩ	ㄩ	ㄩ	ㄩ	ㄩ	ㄩ	ㄩ	ㄩ	ㄩ	ㄩ	ㄩ	ㄩ	ㄩ	ㄩ	ㄩ	ㄩ
黃	i	vi	ia	via	iam	ian	iang	iong	im	in	ing	io	vio	ir					iau	viau	iu	viu		
楊	i	i <sup>n</sup>	ia	ia <sup>n</sup>	iam	ian	iang	iong	im	in	i/eng	io	io <sup>n</sup>	iə					iau	iau <sup>n</sup>	iu	iu <sup>n</sup>		
教羅	i	i <sup>n</sup>	ia	ia <sup>n</sup>	iam	ian	iang	iong	im	in	eng	無	io <sup>n</sup>	io					iau	iau <sup>n</sup>	iu	iu <sup>n</sup>		
例字	居	梔	迦	驚	兼	堅	姜	恭	金	巾	經		驚	茄					嬌	梔	梔	梔	梔	梔

國	ㄩ	ㄩ	ㄩ	ㄩ	ㄩ	ㄩ	ㄩ	ㄩ	ㄩ	ㄩ	ㄩ	ㄩ	ㄩ	ㄩ	ㄩ	ㄩ	ㄩ	ㄩ	ㄩ	ㄩ	ㄩ	ㄩ	ㄩ	ㄩ
台	ㄩ	ㄩ	ㄩ	ㄩ	ㄩ	ㄩ	ㄩ	ㄩ	ㄩ	ㄩ	ㄩ	ㄩ	ㄩ	ㄩ	ㄩ	ㄩ	ㄩ	ㄩ	ㄩ	ㄩ	ㄩ	ㄩ	ㄩ	ㄩ
黃	u	(vu)	ua	vua	uai	vuai		uan	ue	ue <sup>n</sup>	ui	ui <sup>n</sup>	un	uang	o	vo	om	ong						
楊	u	u <sup>n</sup>	ua	ua <sup>n</sup>	uai	uai <sup>n</sup>		uan	ue	ue <sup>n</sup>	ui	ui <sup>n</sup>	un	uang	o	o <sup>n</sup>	om	ong						
教羅	u	u <sup>n</sup>	oa	oa <sup>n</sup>	oai	oai <sup>n</sup>		oan	oe	oe <sup>n</sup>	ui	ui <sup>n</sup>	un	oang	o <sup>o</sup>	o <sup>on</sup>	om	ong						
例字	龜		瓜	官	乖	糜		觀	檜		規	關	君	光	沽	姑	箴	公						

表 2-1 台語音標符號相關表

## 2-4 聲調特性介紹：

眾所周知，國語聲調計有陰平(1)，陽平(2)，上(3)，去(4)，輕聲，五種，本文姑且稱為本調聲調。當上聲音節接上聲音節時，前音節會變成陽平聲以便接上聲；另外還有「一」、「不」的聲調變化情形。我們稱此現象為國語聲調的變調。

國語聲調產生變調的原因乃是基於發音困難而不得不變。但台語則不然，台語聲調產生變調的情形不僅較國語多，而且台語的變調運用，並不是單單基於發音困難與否而產生，變與不變二者往往有辨義或語態上的不同作用。

譬如：

“後(yau7)日(zin8)”

台語念(yau3-zin8)是”改天”的意思，若念成(yau7-zin3\*)則是”後天”的意思。

“驚(vgia1)死(si2)”

台語念(vgia7-si2)是”怕死”的意思，台語念(vgia1-si3)則是”嚇死”的意思。

### 2-4-1 聲調種類：本調與變調

自古以來台語聲調種類的歸納分類並未脫離「八音」之說，意即台語有八種不同聲調的意思，名稱如表 2-2，括弧內為其對應的聲調代號。但是現在的臺灣除了鹿港發音外，已無陽上(6)，只剩七種聲調。

	平	上	去	入
陰	陰平(1)	陰上(2)	陰去(3)	陰入(4)
陽	陽平(5)	陽上(6)	陽去(7)	陽入(8)

表 2-2 台語聲調的傳統分類

觀察台語音調本調的頻率走勢，取現在的台南口音來描繪實際的本調變調頻率走勢(如圖 2-1(b))，比較王育德博士(台南人)在 1954 年夏天於東京大學理工研究所曾做實驗所畫出的台語音調走勢[3]如圖 2-1(a)，可知台南的聲調的變動情形很小。

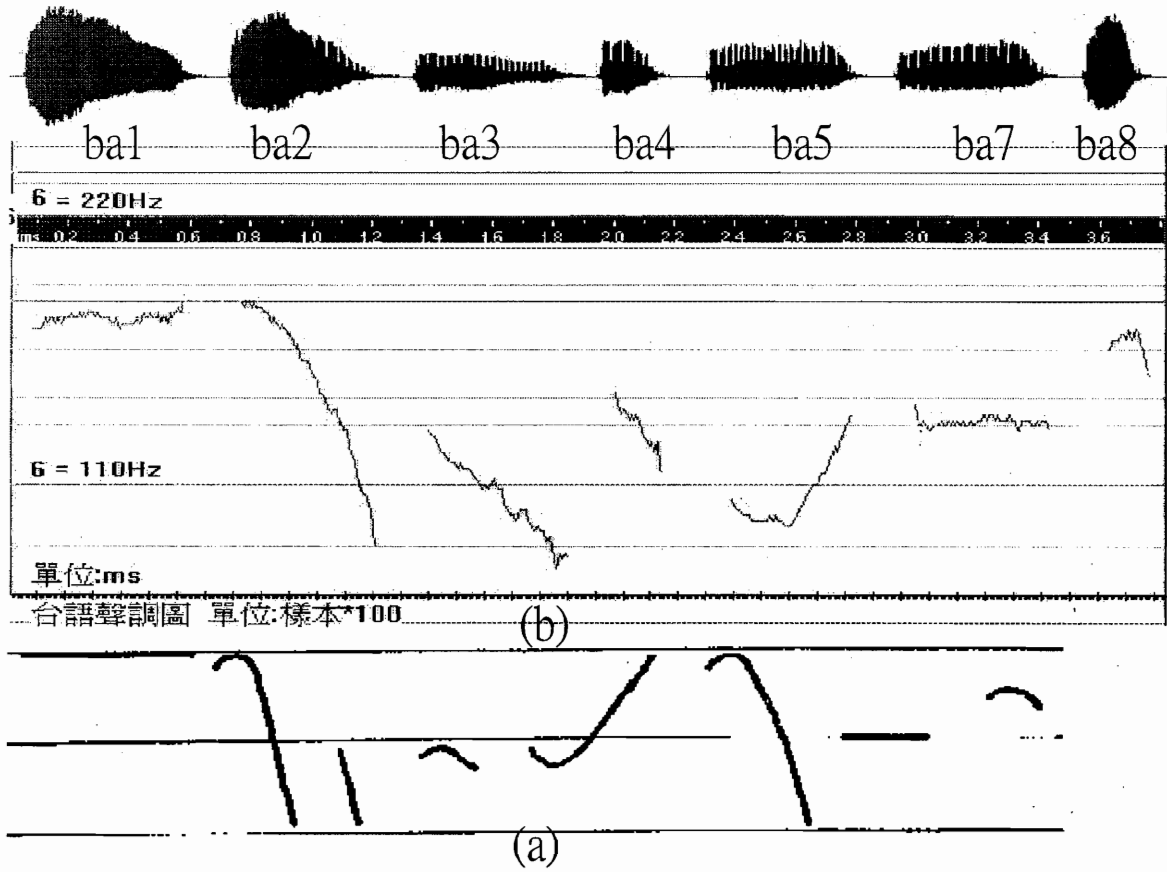


圖 2-1 台南口音本調頻率走勢圖

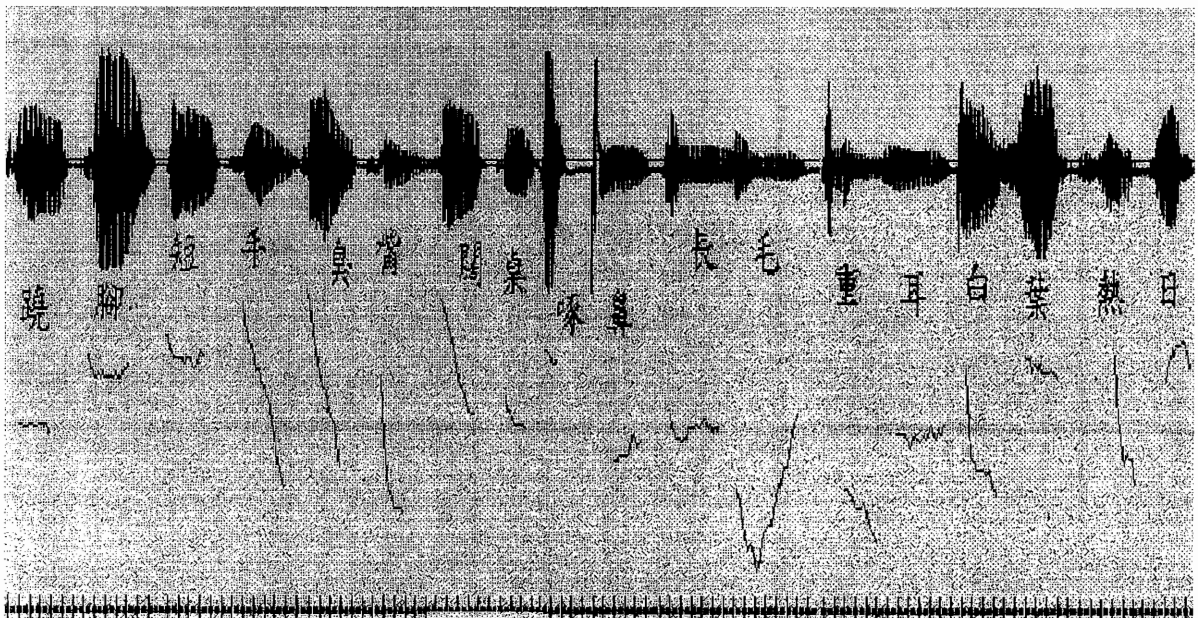


圖 2-2 台南口音變調頻率走勢圖

### 2-4-2 一般變調規則

再觀察二字詞的發音情況，它的頻率走勢如圖 2-2 所示。參考現今台語教學著作所描述台語變調規則計有三種版本[1][12][13][14]（見圖 2-3(a)、(b)、(c)）。分析其中差異出現在入聲調變化的問題上。以台南音，台北音而言，經實驗證實可以用圖 4-2 的規則來處理台語發聲系統的聲調轉換。此規則與圖 2-2 的頻率走勢相驗，是無問題的，並且經本系統證實是可用的。

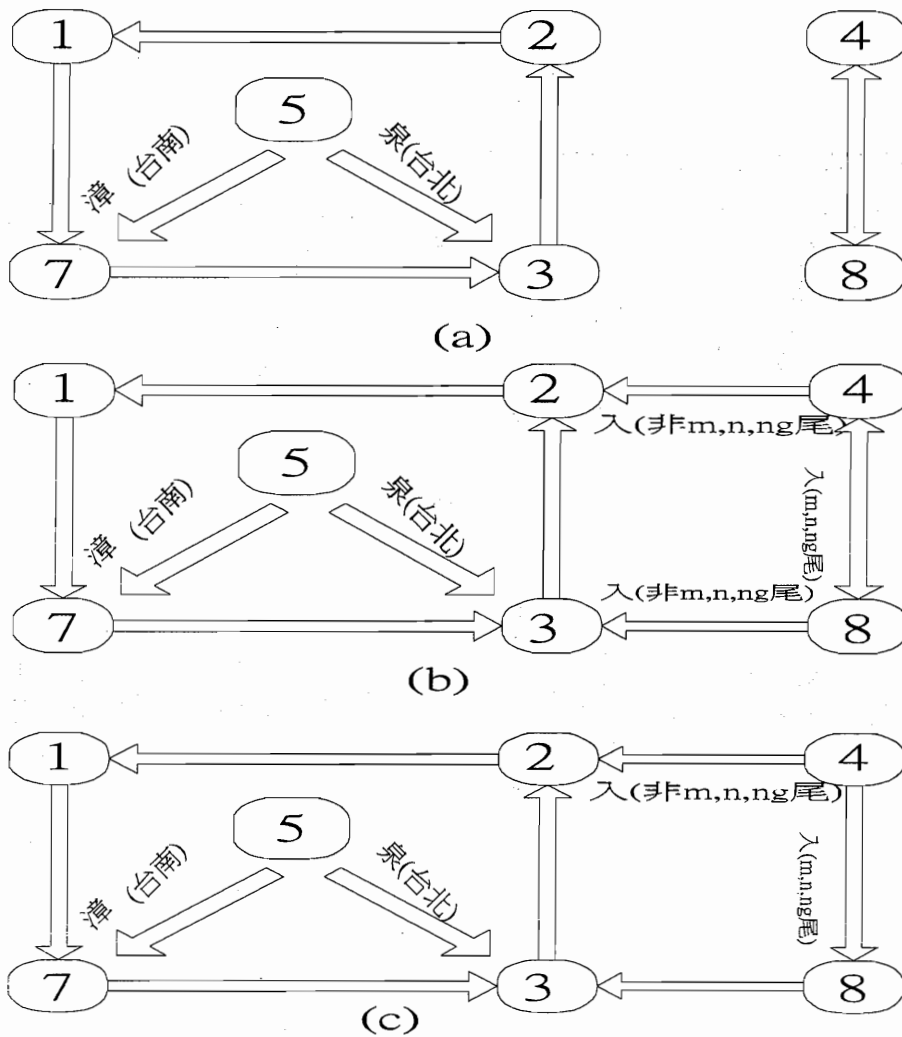


圖 2-3 台語一般變調規則(a), (b), (c)

### 3. 資料庫的建立

#### 3-1 系統架構

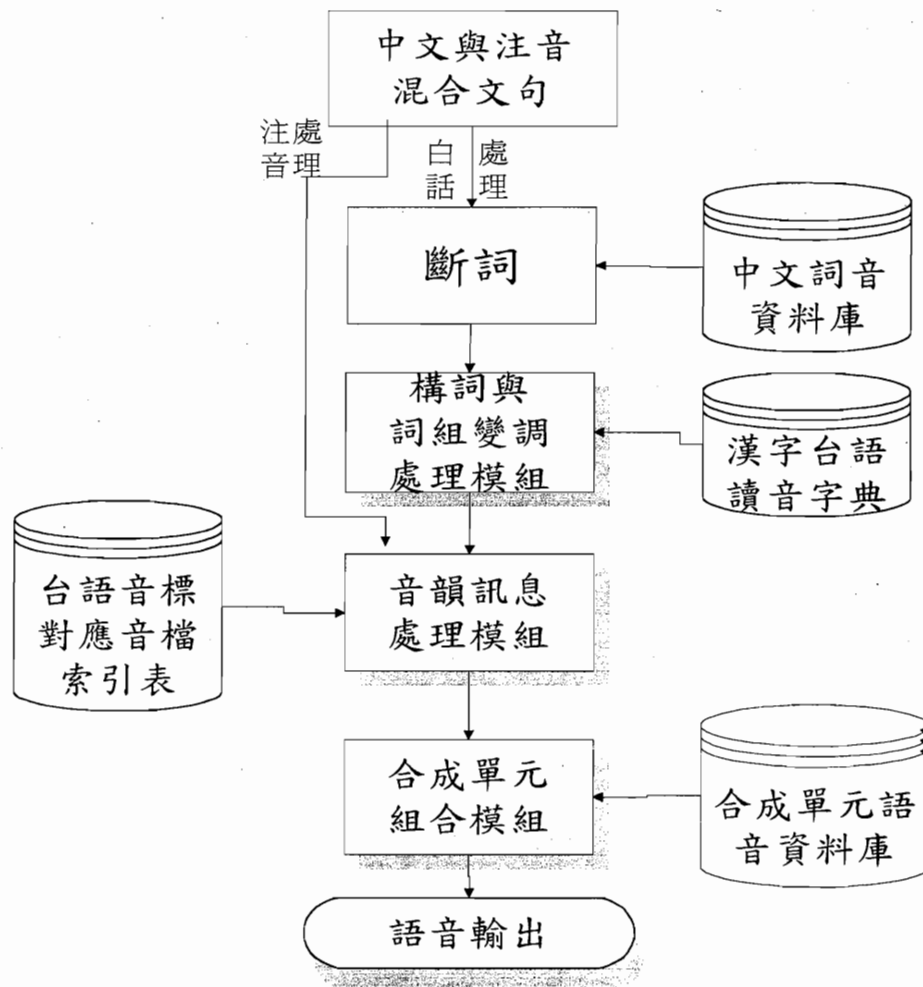


圖 3-1. 系統架構圖

整個系統架構如圖 3-1 所示，由四個資料庫及四個處理模組構成。本節先介紹資料庫，下一節再介紹四個處理模組方塊運作情形

#### 3-2 漢字台語讀音字典與中文詞音資料庫

漢字台語讀音字典決定斷詞後漢字對應的台語注音字，以及作為單字發音查詢使用。本系統漢字對台語音字典的建立，主要是以楊青矗先生所編寫的【國台雙語辭典】為範本而建立[2]。本系統收錄漢字對應台語音的字計有 8168 字，其中對應注音數目如下表 3-1：

總音數	1	2	3	4	5	6	7	8	9
總字數	4418	2504	866	247	93	27	7	6	3

表 4-1 漢字對應注音總數表

而中文詞音資料庫共建立二字詞有 18730 條，三字詞有 4207 條，四字詞有 3983 條，共有 26920 條的常用詞彙，已包含用窮舉方式所解決的單一漢字取台語音混淆情形的詞條，作為斷詞結果的主要依據。

### 3-3 台語音標對應音檔索引表

此表由 552 個音節各含 7 個聲調的組合，共 3864 個注音音符組成，包含了本系統所使用的所有音節符號，用以對應到該音節的聲音檔索引名稱，對於各種形式的注音(拼音)音符間的對應關係，可以對此表做擴充，以增加系統的運用彈性。

當輸入編輯器接受所輸入注音字或漢字後，皆可透過此表轉換成聲音檔的索引名稱，以便供給變調規則運用和語音樣本擷取時所需的資訊。

### 3-4 合成單元語音資料庫

觀察人講話的語音特性，在不同時空之下，同樣一句話，其音高特性會有差異，也就是說語音音高特性是相對的，不是如樂音旋律音高值是絕對的。

對系統的語音合成所需取得音節聲音樣本而言，各同調音節語音樣本卻極需具有相同的音高特性，並且不同聲調音節的語音樣本間的音高特性更需控制在一定範圍內，以便產生良好的合成語音音質輸出。合成單元語音資料是以本論文所提出的「固定音高旋律性錄音法」錄音方法錄製而成，共 3864 個音節，並且紀錄各音節的起點，結束點，以及詞中、詞尾的端點。

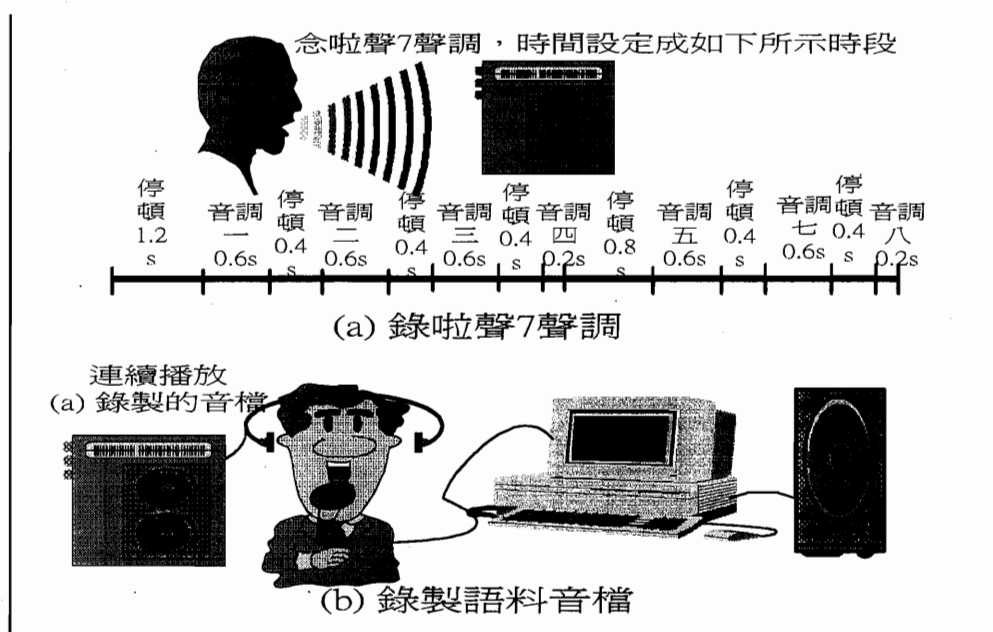


圖 4-2 固定音高旋律性錄音法步驟一與步驟二



#### 4. 研究方法-斷詞、聲調變化及台語的取音

##### 4-1 斷詞

本論文採用蒐集大量的詞彙利用統計分析建立一個詞庫加以斷詞[4]，然而有限的詞庫亦無法包含所有的詞彙，因此有可能發生在詞庫找不到匹配的詞而造成錯誤，但是若要獲致滿意的結果，最終雖仍需要靠語法及語意的分析才能解決。

一般語言模型的機率可以下式表示：

$$P(W) = \prod_{i=1}^n P(W_i | W_1 W_2 \dots W_{i-1}) \quad (\text{式 4-1})$$

其中  $P(W_i | W_1, W_2, \dots, W_{i-1})$  代表詞  $W_i$  出現在字串  $W_1, W_2, \dots, W_{i-1}$  之後的機率，但是此機率在實際的情況下是無法估算出來的，因為假如我們的語言含有  $L$  個詞，就可能須要記錄  $L^{(i-1)}$  筆資料，這是相當大的資料量，所以我們簡化這機率公式，假設  $W_i$  出現的機率只和它的前幾個詞有關而已，和在之前的詞無關。例如  $W_i$  只和它之前的一個詞有關時，即為 bigram 關係，以 4-2 式來表示：

$$P(W) = \prod_{i=1}^n P(W_i | W_{i-1}) \quad (\text{式 4-2})$$

其中前一詞和下一詞之間相連的機率  $P(W_i | W_{i-1})$  的機率估算由下式來計算：

$$P(W_i | W_{i-1}) = f(W_i | W_{i-1}) = \frac{C(W_{i-1} W_i)}{C(W_{i-1})} \quad (\text{式 4-3})$$

上式中的分子分母分別是 bigram 函數和 unigram 函數。結果如圖：

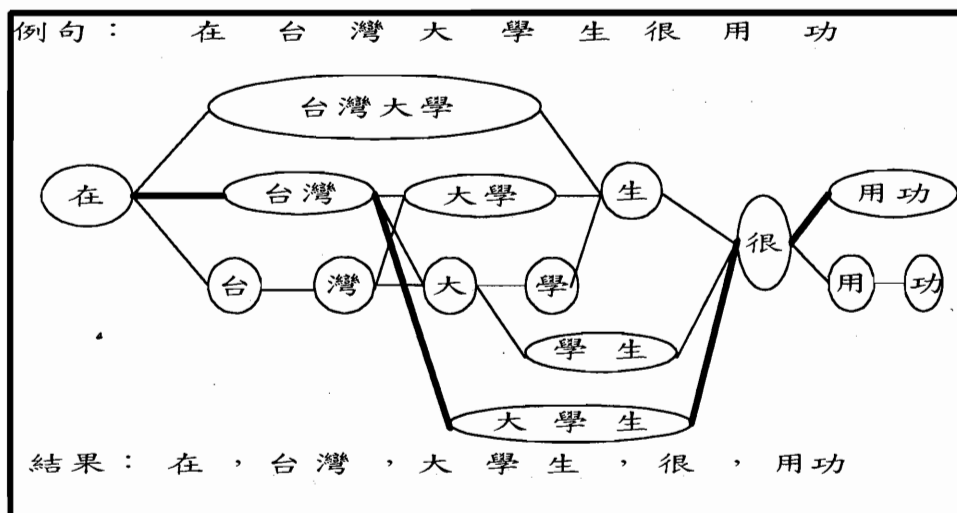


圖 4-1 斷詞狀態轉移圖

#### 4-2 構詞、選音與詞組邊界變調處理

台語文句的音調變化的準確與否，決定了聽覺的流利度感覺與辨義功能。就詞彙本身而言，一般詞本身的基本變調規則為：最後一字不變調，前面全變調。

譬如：

『安』念 yan1   『安平』念 yan7-beng5   『安平港』念 yan7-beng7-gang2  
『碼』念 beng2   『碼頭』念 beng1-tau5

但是有少數詞彙及大量成語皆有特例性質，不適用上述規則，需以詞庫建立處理。

如：心\*悶(sim1-bhun7)，縱虎\*歸山(ziong2-ho2 gui7-san1)

上述的變調規則為系統斷詞結果與自然人說話所認定的斷詞點相同時所採用的規則，然而有些詞並無法在系統詞庫中列舉出來，如數字，人名，地名等，所以我們需以構詞方式處理來補強斷詞結果。下一小節將介紹系統處理數字的方法。

另外還得解決詞組邊界的聲調處理。意義以下列例子說明：

看右面句子：在台灣大學生很用功

斷句成：在,台灣,大學生,很,用功

台語注音音節表示為：

zai7,dai7-wuan5,dai3-hang3\*-seng1,ziong4,yiong3-gong1.

由上說明即知：所謂詞組邊界的聲調處理，意即是要決定『在，灣，生，很，功』等字是否要變調。圖 4-2 為本系統所採用的變調規則，其中 3\*是當音節尾音有 m, n, ng 音，且聲調值是 8 的變調後聲調值。

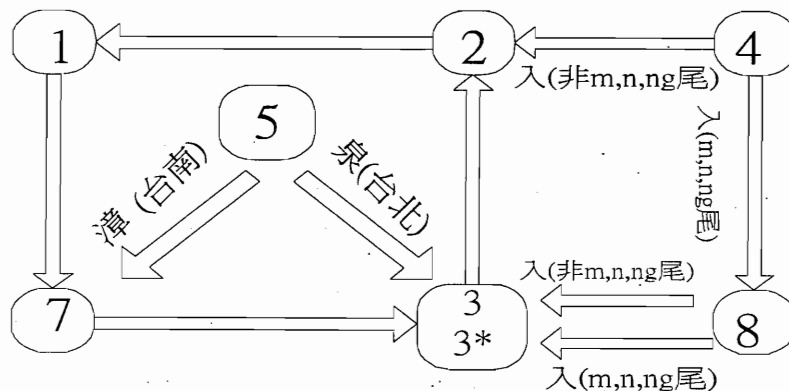


圖 4-2 本系統採用的台語一般變調規則

#### 4-2-1 數詞的構詞、選音與變調規則處理

本文雖以「詞庫窮舉法」的方式解決台語系統一字多音選音混淆的問題，但是仍有難以決定誰是第一順位的問題存在，這在數字念法上特別明顯。處理說明於下。

漢字(1)	漢字(2)	文言(1)	白話(2)
一	壹	yin4	zinj
二	貳	zhi7	lng
三	參	sam1	vsac
四	肆	su3	siq
五	伍	vgho2	gho
六	陸	liong8	langj
七	柒	chin4	
八	捌	ban4	be4
九	玖	giu2	gau2
十	拾	sim8	zam8
	零		leng5
百	佰	beng4	ba4
千	仟	chian1	cheng1
	萬	bhan7	
	億	yeng4	
	兆	diau7	

表 4-1 數詞注音表

表 4-1 是數詞注音表，其文言音與白話音在口語中常並列使用，前者常用在成語或術語中，後者則用在數量詞用語中。文言音因為詞庫畢竟無法全部包含，故選音排名為優先地位。當我們以構詞方式定出數量詞用語，需再改變注音符號以達成正確選音的需求。再來還需探討數量詞連結時，變調處理的問題。

基本上，我們將數詞分成 A，B 二類，如下表 4-2：

A 類	一，二，三，四，五，六，七，八，九，十
B 類	百，千，萬，億，兆，零

表 4-2 數詞分類表

則詞或字結合點的變調處理規則為：

只有 A 的組合則念文言音。如：一九九九年。

- A+B 的組合則念白話音。如：一百五十年，但是二十、十二，的”二”須念文言音。
- $[N^*(A+B) \text{ or } N^*(A+B+B), N>1]$  的組合時，雖是念白話音，但是千，萬，億，兆，固定不變調。如：二千一百三十五億人，「千」念本調音。
- 組合詞是 (數詞+量詞) 或 (數詞+名詞)，則數詞尾字需變調，該組合詞視為獨立詞。
- 組合詞是 (數詞+量詞+名詞) 則數詞、量詞尾字均需變調，該組合詞視為獨立詞。

#### 4-2-2 詞組邊界的聲調處理

- 中文詞庫的詞屬性分成變動詞、獨立詞、連接詞、的、數詞、量詞、可形容詞化名詞、感嘆詞(語助詞)。
- 所有詞最後只剩 變動詞 與 獨立詞。  
變動詞接任何詞則變動詞尾音節需變調。獨立詞則否。
- 「連接詞」、「的」視為句子段落，之前若有斷詞斷點，該分界不需變調。
- 「的」本身成為變動詞。
- 有「連接詞」、「的」的句子，該詞即視為句子段落。也就是說，之前若有斷詞斷點，該分界不需變調。而「的」本身之後若有斷詞斷點，則本身需做變調處理。「連接詞」分成二類，一類需要變調處理，如「既然」等，一類不需要變調處理，如「但是」等。
- 將初步斷詞為一字詞或二字詞或三字詞或四字詞者，分成變動詞與獨立詞，變動詞與下一詞相連時，尾音節需變調，獨立詞則否。
- 句子尾字與獨立詞尾字念本調。

- 句子尾字若是感嘆詞或語助詞如「啊，呀，而已」等語助詞則其前一字需念本調，本身毋需做變調處理。

以例子說明：

「安平港碼頭」念 yan7-beng7-gang2 beng1-tau5.

因為安平港，碼頭皆為獨立詞，所以接合處音節 gang2 無須變調。

再看下面句子：

「在台灣大學生很用功」，斷句成：在,台灣,大學生,很,用功。

「在，很，用功」等字是變動詞，所以接合處音節須變調，但「功」在句尾，所以不變調。而「台灣，大學生」皆為獨立詞，所以接合處音節無須變調。故結果變成 zai3,dai7-wuan5,dai3-hang3\*-seng1,ziong8,yiong3-gong1.

經過上述詞組式變調處理後，文句大致可以得到還不錯的結果。

#### 4-3 聲音輸出處理

語音合成器基本上可分為兩大類：第一類是利用全極點(all-pole)的語音發聲模型，將發聲的機制分為聲帶激發訊號(glottal excitation)及一個口腔模型(acoustic tube model)。應用此一模式所發展出來的語音合成技術多源自語音壓縮技術，如線性預估編碼(LPC vocoder)[5][6]，單、多脈衝激發(multi-pulse excitation)[7][8][9]等。

第二類的合成器則是所謂波型合成器，此類合成器在時域上直接調整語音基週頻譜特性而不需藉由簡化的發聲數學模型來處理合成語音的音韻。雖然儲存的空間需求較大，但所合成出來的語音品質較第一類合成器的效果要好，而且此類合成器的運算量較少。目前的波型合成器所使用的方法以時域基週同步疊加法(Time Domain Pitch synchronous Overlap-Add, TD-PSOLA)[10][11]最廣為採用。本系統並未運用上述技術，而是以音節直接加以合成，整個作法說明如下：

經由斷詞及構詞以及詞組變調旗號處理之後，我們得到該文句各字的注音字、變調旗號、音長旗號、停頓間隔旗號，送入音韻訊息處理模組。

音韻訊息處理模組便根據文句分析之後所得到的上述參數，將注音字轉換成發音代碼，並以一般變調規則進一步轉換成實際的發音代碼，之後再藉由音長旗號與停頓間隔旗號決定各音節在合成時所需的語音樣本長度及所需的停頓間隔。

對於台語的入聲音節，由於本身是短促音，口語化時與相鄰音節相接時必有一停頓現象，可以本身音節長度做為靜音長度，以增加語氣的流利度。

經由上述處理，合成單元組合模組便可藉由適當的音節斷點與靜音長度的組合，自合成單元語音資料庫取得語音樣本，便將語音結果輸出。

## 5. 實驗

### 5-1 實驗環境

本系統是以 Pentium 133 個人電腦加上 16 位元聲霸卡在 Win95 作業系統下以 Microsoft Visual C++ 5.0 開發完成，而實驗環境也是在如上述的環境下進行。各項目簡述如下：

- 測試人員：分一般人士組及實驗室同仁組兩組，各組 5 人。
- 測試樣本：選取二、三及四字詞包含七聲調變化調與本調各 20 條。報紙短句 10 句。
- 測試方式：

可辨度：

在可辨度的評估方面，在系統合成測試樣本後，由受測者將所聽到的語音以聽寫方式寫出。最後統計正確文句與受測者所寫文句之間的差異，以作為評估可辨度的方式。

自然度：

我們採用平均鑑定分數(Mean Opinion Scores, MOS)做為評估的標準。這種評估方式將合成語音的自然度分為優良(excellent)、良好(good)、尚可(fair)、差(poor)及極差(unsatisfactory)五個等級，分別給予 5 至 1 不等的分數。測試人員在聽過合成的語音後，以所感覺到的自然度評分。

### 5-2 實驗結果

實驗結果如下所示：

## 一、可辨度

測試種類	數量	一般人士組可辨度	實驗室同仁組可辨度
二字詞	20	95.0% (38/40)	95.0% (38/40)
三字詞	20	90.0% (54/60)	93.3% (56/60)
四字詞	20	87.5% (70/80)	93.8% (75/80)
短句	10	60.7%	60.2%
平均		83.3%	85.58%

表 5-1. 可辨度實驗結果

由實驗結果的數據顯示，本系統合成語音的平均可辨度為 84.4%，而且測試樣本字數愈多可辨度愈低。推究其原因，是因為字數長度愈長，變調處理錯誤機會也越高，可見變調問題的處理對台語系統的發音的重要性。

## 二、自然度

測試種類	數量	一般人士組自然度	實驗室同仁組自然度
二字詞	20	4.0	4.0
三字詞	20	4	3.5
四字詞	20	3.5	3.7
短句	10	3	3.1
平均		3.625	3.575

表 5-2. 自然度實驗結果

自然度評估的結果，平均鑑定分數為 3.6125，介於尚可與良好之間，表示本系統所輸出的語音雖未在合成技術上多所著墨，但是採用的錄音方法，亦可達成某一程度的水準。另外，變調在台語流利度的感受占著極大的因素，也由實驗數據可以得知。

## 6. 結論與討論

在本文中，我們實作了一套能說台語的國語文句翻台語語音的系統。在面對漢字一字對台語多音的取音抉擇問題上，試以詞庫窮舉的方式來解決。

本論文亦建立了一套語音合成上可用的台語一般變調規則，並對詞與詞連接時，所衍生前一詞的詞尾是否變調的問題處理提出作法。而以「固定音高旋律性錄音法」所錄

得語音資料，可以有效降低聲音樣本因時間差而產生的語音頻譜差異過大的現象。這對於連音部分自然度的摩擬，很有幫助。此種錄音方式亦有利於單音節多聲調的系統做初步的開發。譬如國語或客語發音系統。

由於變調處理在台語發音上，是為必要的成分，所以應該變調之處就必須變調，不該變調處就不要變調，所以斷詞正確與否對台語發音有關鍵性的影響。本文所用的斷詞演算法可增加斷詞的正確性。同時，也建立構詞組詞規則來增加變調處理的正確性。系統雖完成初步成果，但離自然流利的口語仍很遠。以下提出了未來應該改進的幾個方向：

- 語法的分析對提昇斷詞、變調組位置及字轉音選哪音的正確性有很大的幫助。尤其是斷詞結果必與轉調相關聯，也就是對於合成語音的自然度有很大的影響。因此，一個包含語法分析器的文句分析模組是大家該努力的方向。
- 漢語的特色是單音節多聲調，當可嚐試建立一套共同技術系統，來直接開發各種漢語方言的發音系統。

## 7. 參考文獻

- [1] 鄭良偉，鄭謝淑娟合編，「台灣福建話的語音結構及標音法」，學生書局台北，民76年8月四刷
- [2] 楊青矗，「國台雙語辭典」，敦理出版社，台北市，1996.2五版
- [3] 王育德，「台灣話講座」，台北，pp. 34，自立出版社，1993.5
- [4] 陳世達，「應用音中仙中文聽寫機之全球資訊網語音輸入查詢系統」，成大電機研究所碩士論文，民國87年6月
- [5] 歐陽明、李琳山，「一套中文的文句翻語音系統」，台大電機研究所碩士論文，1985年6月
- [6] L. R. Rabiner and R. W. Schafer, "Digital Processing of Speech Signals", p398-p403, 1987.
- [7] 劉繼謚等，「以線性預測編碼為合成器的中文文句翻語音系統」，電信研究季刊，第19卷第3期，民國78年9月
- [8] S. Singhal and S. Atal, "Amplitude Optimization and Pitch Prediction in Multipulse Coders", IEEE Trans. Acoust, Speech, Signal Processing,



ASSP-37, pp.317-327, March 1989

- [9] 朱國華等，「一套以多脈衝激發語音編碼器為架構之即時中文文句翻語音系統」，電信研究季刊，第21卷第4期，民國80年12月
- [10] J. Charpentier and M.G. Stella, "Diphone Synthesis Using an Overlap-Add Technique For Speech Waveforms Concatenation", Intern. conf. on ASSP, ICASSP-86, pp. 2015-2019, 1986
- [11] Christian Hamon, Eric Moulines, Francis Charpentie, "A diphone synthesis based on time-domain prosodic modifications of speech", ICASSP, pp.238-241,1989
- [11] 施炳華，「台語教學法」，啟人書局，台南市，民82
- [12] 林繼雄，「台語教學法」，大夏出版社，台南市，民79
- [13] 方南強，「大家來說台灣母語（閩南語篇）」，時報出版公司，1996.5初版七刷
- [15] Chiang Yuang-Chin, et. al., "A New Hybrid Duration Hidden Markov Model with Application to Large Vocabulary Taiwanese(Min-nan) Word Recognition," 1<sup>st</sup> International Symposium on Chinese Spoken Language Processing(ISCSP98),Dec, 1998, Singapore .
- [16] Ren-yuan Lyu, Yuang-jin Chiang, Ren-zhou Fang, Wen-ping 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, p. 139-149
- [17] Ren-yuan Lyu, Yuang-jin Chiang, Wen-ping 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", ICSP1998 .
- [18] 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。
- [19] Y.J.Sher, K.C. Chung, C. H. Wu, "Taiwanese Syllable-based Synthesis Units Database", Accepted by Chinese Journal of Medical and Biological Engineering (CJMBE-233),1999,volume 19,no 1, page47-58.
- [20] Y.J.Sher, K.C. Chung, C. H. Wu, "Establish Taiwanese 7-Tones Syllable-based Synthesis Units Database for the Prototype Development of Text-To-Speech System", Accepted by ROCLING XII .
- [21] 鄭良偉，「台語的語音和詞法」，遠流出版社，台北，民86年出版。