

國立交通大學

電信工程學系碩士班

碩士論文

客語文句翻語音系統之實作

An Implementation of Hakka Text-to-Speech System



研究生：林東毅

指導教授：王逸如 博士

中華民國九十五年八月

客語文句翻語音系統之實作

An Implementation of Hakka Text-to-Speech System

研究生：林東毅

Student : Dong-Yi Lin

指導教授：王逸如 博士

Advisor : Dr. Yih-Ru Wang

國立交通大學

電信工程學系碩士班



Submitted to Institute of Communication Engineering
College of Electrical Engineering and Computer Science

National Chiao Tung University

in partial Fulfillment of the Requirements

for the Degree of

Master of Science

in

Electrical Engineering

August 2006

Hsinchu, Taiwan, Republic of China

中華民國九十五年八月

客語文句翻語音系統之製作

研究生：林東毅

指導教授：王逸如 博士

國立交通大學電信工程學系碩士班

中文摘要



本論文完成一套客語文句翻語音系統。它由四個主要部分組成：文句分析器、RNN 韻律訊息產生器、語音波形樣本資料庫和 PSOLA 語音合成器。輸入文句經由文句分析後產生適當的語言參數，RNN 韻律訊息產生器根據這些參數產生相對應的韻律參數。PSOLA 語音合成器則根據合成音節碼從語音波形樣本資料庫取出適當的語音波形樣本，將其依照韻律參數調整後，得到合成語音波形輸出。在此研究中，為了能使合成語音的韻律參數更接近實際的情形，我們嘗試用人工調整切割位置與修正音高軌跡。最後，我們使用一個單一文件界面的文字編輯器配合語音合成核心製作了一套在 Windows 平台上的展示系統。

An Implementation of Hakka Text-to-Speech System

Student : Dong-Yi Lin

Advisor : Dr. Yih-Ru Wang

Institute of Communication Engineering
National Chiao Tung University

Abstract

In this thesis, a Hakka Text-to-Speech (TTS) system is implemented. It consists of four main parts: Text Analyzer, RNN prosody generator, waveform inventory of synthesis units and PSOLA synthesizer. The input text is first tagged in the text analyzer into word sequence. Then, the RNN prosody generator is used to generate the prosodic information by using linguistic feature extracted from the word sequence. The Waveform corresponding to the word sequence is then extracted from the waveform inventory and prosodically-adjusted to generate the output speech. The basic implementation of the system follows the Mandarin TTS system developed previously in NCTU. A demo system operating on the Windows platform by using a SDI(Single Document Interface)text editor with the synthesis kernel was last realized. Informal listening tests show that most synthesized speeches sound fair.

致謝

一個 TTS 系統，需要許多資料庫，如：音節單音語料庫、詞典、訓練語料庫、構詞表格…等，在碩一時，幫忙博班學長的統計詞典資料時，就踏入客語語音合成系統的領域了。這些資料整理，包含：挑選文稿、開始錄音、語料切割、調整切割位置、校正音高軌跡、標記 Pitch Mark、人工斷詞、人工標記詞典詞性、詞典網路蒐集、拼音系統轉換…等，這些都是從無到有、要處理的量多、費時、加上幾乎都是需要人工處理。研究所兩年就在整理客語資料中度過。一個人是絕對做不成的，這裡特別感謝余秀敏老師，還有老師帶領的得意門生，秋香、姿伶、婉宜、雅婷、怡玲、玉婷等，有了你們的耐心與細心，論文才可順利進行。還有龔萬灶老師為了客語語言文化的推廣，幫忙標記文章音節與提供許多客語資料，讓文字資訊更加完整。

在 TTS 系統架構上，感謝阿德、希群學長，有學長帶領學到的東西真的很多，還有從你們身上也學習到做人處事的觀念，謝謝你們。碩一學弟們，啟風、存智、友駿、小傅、宏宇、獻文、銘彥、伯蒼、胤賢，在面試國防役還好有你們在幫忙，不然會真的忙不過來，跟你們打球、玩賽車遊戲真的很過癮。碩二的同伴們，強者國興，實驗室有了你真的不一樣，謝謝你不耐煩的教我程式，讓我進步許多，老人家振豐，最後幾個月一起去健身房、到發條桔子買飲料，從你身上學到很多，陽光男孩世哲，打球最佳伙伴，相互配合下，幾乎跟你同一隊都是穩贏的，熱血男孩見惶，實驗室有你在真的歡笑很多，電玩高手鴻彥，相處快七年了，榮登哈拉天王非你莫屬。特別感謝陳信宏教授與王逸如教授，在百忙之中，不斷指導學生的研究方法與方向，有教授的指點才可以找出自己要改進的地方，謝謝教授兩年來的教誨。

最後，感謝家人在背後的支持，讓我無後顧之憂，還有女友舒茹的貼心照顧，有了你們，我才可以順利完成論文。

目錄

中文摘要	I
英文摘要	II
致謝	III
目錄	IV
表目錄	VI
圖目錄	VII
第一章 諸論	1
1.1 研究動機	1
1.2 客語文句翻語音系統概述	1
1.3 研究方向	3
1.4 章節概要	3
第二章 客家話語文特性	4
2.1 苗栗四縣客家話	4
2.2 四縣客家話之聲母	5
2.3 四縣客家話之韻母	6
2.4 四縣客家話之聲調	7
2.5 四縣客家話之音節	8
第三章 T T S系統原理	10
3.1 韻律訊息產生器	10
3.1.1 產生韻律訊息	11
3.2 基頻同步疊加合成方法	13
3.2.1 基頻同步分析	14
3.2.2 基頻同步變換	14
3.2.3 基頻同步疊加合成	14
3.2.4 時長變化	15
3.2.1 音高變化	16
第四章 訓練語料庫的前處理過程	19
4.1 語料庫說明	19
4.2 切割資訊	20

4.2.1 切割資訊的求取	-----	20
4.2.2 切割資訊的修正	-----	20
4.3 基頻軌跡參數的求取	-----	23
4.3.1 基頻軌跡檢查前處理	-----	23
4.3.2 基頻軌跡檢查流程	-----	24
4.3.3 基頻軌跡檢查結果與討論	-----	25
4.4 基頻軌跡參數	-----	29
第五章 文字轉語音系統之製作	-----	30
5.1 文句分析器	-----	30
5.1.1 文句輸入輸出格式	-----	32
5.1.2 詞庫的收集狀況	-----	34
5.1.3 構詞規則的整理狀況	-----	34
5.1.4 客語文字分析結果範例	-----	34
5.2 韻律訊息產生器	-----	37
5.3 聲音波形資料庫	-----	41
第六章 結論與未來展望	-----	42
參考文獻	-----	43
附錄一	-----	44
附錄二	-----	50
附錄三	-----	51



表目錄

表 2-1：四縣客家話聲母表	5
表 2-2：四縣客家話韻母表	6
表 2-3：四縣客家話音節類型表	8
表 3-1：標點符號分類表	13
表 4-1：錄製音檔軟硬體設備與設定	19
表 5-1：文字分析器輸出結果範例	32
表 5-2：客語詞典的詞長統計	34
表 5-3：由處理後語音訊號中抽取出來的韻律訊息參數範例	38
表 5-4：韻律參數之均方根誤差	39



圖目錄

圖 1-1：客語文句翻語音離形系統之基本架構圖 -----	2
圖 2-1：漢語方言的音節結構 -----	4
圖 2-1：四縣客家話六個聲調音高軌跡圖 -----	7
圖 3-1：人腦產生韻律資訊的概念模型 -----	11
圖 3-2：細分之韻律產生的概念模型 -----	12
圖 3-3：遞迴式類神經網路方塊圖 -----	12
圖 3-4：以時域基頻同步疊加改變時長 -----	16
圖 3-5：以基頻同步疊加改變音高 ($\beta = \gamma$) -----	17
圖 3-5：以基頻同步疊加改變音高 ($\beta \neq \gamma$) -----	18
圖 4-1：非濁音聲母的音節切割 -----	21
圖 4-2：連音中舒聲韻輔音韻尾與濁音切割處 -----	21
圖 4-3：當聲母受前一個音節影響，而有週期特性時的切割位置 -----	22
圖 4-4：人工調整與原始切割位置位移統計圖 -----	22
圖 4-5：音節起頭與結尾部分不規則的跳動 -----	23
圖 4-6：利用內插法修補不平滑的音高軌跡 -----	24
圖 4-7：音節 F0 示意圖 -----	24
圖 4-8：無法求取音高軌跡之範例 -----	26
圖 4-9：音尾發生 double pitch -----	27
圖 4-10：音節結尾是入聲韻輔音韻尾(p.t.k)時的音高軌跡 -----	27
圖 4-11：基頻軌跡出現 Half Pitch 現象 -----	28
圖 4-12：音高軌跡少於四點 -----	28
圖 5-1：客語斷詞器架構圖 -----	30
圖 5-2：音高軌跡圖 -----	36
圖 5-3：語音訊號處理範例 -----	37
圖 5-4：RNN 產聲韻律參數之範例 -----	40
圖 5-5：圖 5-5 基本音節“man”切割及 pitch marks 標示結果 -----	41
圖 5-5：圖 5-5 基本音節“tiet”切割及 pitch marks 標示結果 -----	41

第一章 緒論

語音，是人與人傳達訊息的橋樑。廣義來說，語音是一種可以傳達思想和感情的有意義的聲音，其中包含語義和語法的概念。科技不斷追求的目標是帶給人類生活上的便利，資訊設備，如：電腦、手機、PDA、智慧型家電等，已經成為人類生活的一部份，如果能讓資訊設備聽得懂人說的話和讓資訊設備也能夠像人一樣的說話，那麼人們使用資訊設備會如同和人溝通一樣便利。文字轉語音系統(Text-to-Speech, TTS)，是語音技術中重要發展之一，其系統的功能在於經由文字輸入，讓電腦發出近似人類流暢說話聲音。未來可應用於汽車導航語音輔助系統、語言學習、盲人輔助閱讀器、PDA/Phone 之簡訊閱讀、機器人人機溝通介面等。

1.1 研究動機



目前台灣客家族群約佔總人口七分之一，但根據《九十二年客家民眾使用客語使用狀況調查報告》中指出阻礙客語傳承的主要因素是“本身不太會講”。四縣腔是台灣客家話中最通行的語言，全台灣只要有客家人的地方，大部分都能用四縣話溝通。為了響應客語語言文化的傳承，本論文的研究動機是發展出第一套四縣客家話文字轉語音系統(Hakka Text-to-Speech System, HTTS)，輸入客語文句，讓電腦能以流暢的客語語音輸出。

1.2 客語文句翻語音系統概述

語音合成包括聲學、語言學、數位訊號處理、多媒體技術等多項學科技術，是語音訊號處理領域的一項前瞻技術。解決的主要問題就是如何將文字轉化為流暢的語音訊號。使以往只能用眼睛看的文字資訊，也可以用耳朵來聽。通俗地說，就是讓電腦開口說話的技術。近幾年來，國立交通大學電信工程學系語音實驗室致力於

國語與閩南語文句翻語音系統的發展，並且已經得到不錯的結果可以合成出相當流利的語音，本論文延伸過去的研究，發展 HTTS。本論文所使用 TTS 系統的基本架構如圖 1 所示：

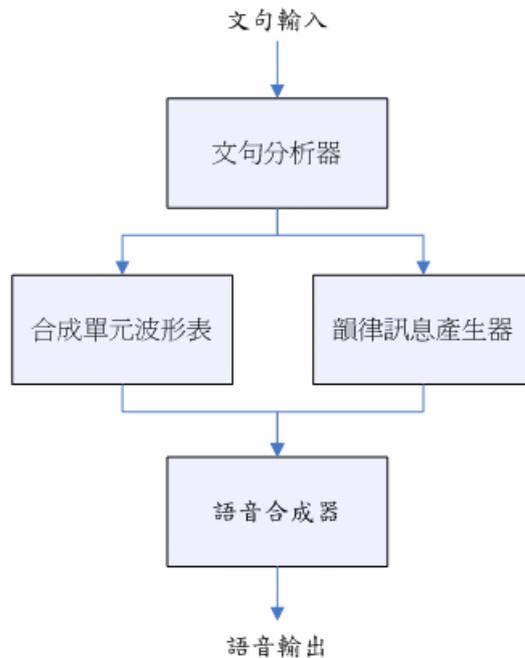


圖 1 客語文句翻語音雛形系統之基本架構圖

這是一套純軟體的系統，包含四個模組：文句分析 (text analysis)、韻律信息產生器 (prosody generation)、合成單元波形表 (acoustic inventory)、語音合成器 (speech synthesis)，文句分析是將輸入文句做斷詞分析，以產生詞串、詞類串、音節串；韻律信息產生器是由文句分析的結果抽取一些語言參數，來產生適當的韻律信息，包含基頻軌跡、音量和音長；合成單元波形表即是聲音波形資料庫儲存所有客語的基本音節波形訊號，它使用文句分析的結果，來組合出對應音節串的基本音節波形串；語音合成是使用韻律信息來調整基本音節波形串的韻律，使其成為自然流利的語音。

1.3 研究方向

本論文的研究重點，分成四個階段，資料彙整、訊息調整、模組訓練和最後系

統整合。資料彙整，包含：建立客語音節清單、蒐集客語詞典、客語文稿、錄製四縣客語話語料庫等；訊息調整，包含：校正語音切割位置、標記 671 單音波形音高、修正音高軌跡、人工初步斷詞、標記詞性等；模組訓練，主要是將前述所整理的韻律訊息，拿到韻律訊息產生器訓練韻律模型，最後將與文字分析器與合成器作系統整合，而完成一套 Windows 平台上的客語文句翻語音系統。

1.4 章節概要

本論文共分六章，各章節編排如下：

第一章 **緒論**：介紹文句翻語音系統的應用、客語文句翻語音系統的基本架構和說明本篇論文的研究動機與研究方向。

第二章 **客語語文特性**：介紹漢語方言的音節結構、四縣腔客家話的音韻系統。

第三章 **TTS 系統原理**：介紹語音合成系統各模組的所使用的演算法，包含：遞迴式類神經網路、基頻同步疊加法…等。

第四章 **訓練語料庫的前處理過程**：介紹處理大量語音資料庫的流程，以求得正確的語音韻律參數；過程中包含語音自動切割與修正、基頻軌跡的求取和調整。經過這樣的步驟後，方可掌握整個語料庫，語音的韻律變化。

第五章 **文字翻語音系統之製作**：介紹TTS系統中文字分析器、韻律訊息產生器，產生結果分析。

第六章 **結論與未來展望**：對本篇論文所提出的方法下結論，並說明未來改進的方法。

第二章 客家話語文特性

客家話的音節結構和其他漢語方言一樣，可以劃分兩大部分：聲母和韻母。聲母是指音節的第一個輔音，而韻母則包含介音及韻腳，韻腳又包括主要元音及韻尾。如圖 2-1 即是漢語方言的音節結構[1]。結構元素中，只有聲調和主要元音是不可或缺的要素，其他則可有可無。有了音節結構的初步概念後，便可從聲母、韻母和聲調等三個層面來分析客家話的語音。

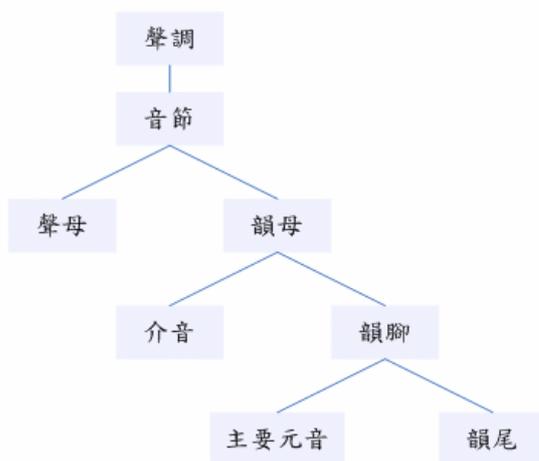


圖 2-1 漢語方言的音節結構

2.1 苗栗四縣客家話

台灣客家話有許多次方言，其中以四縣腔為最通行的腔調，地區以苗栗縣市為最大本營，其他腔調如：海陸腔、饒平腔、紹安腔等，這些次方言之間在聲母、韻母、聲調、詞彙上有不同程度的差異[1]。本論文考量到客語的代表性和實用性，因此選用客語人口最多的苗栗四縣客家話為發展客語語音合成技術的語料來源，其中苗栗四縣客家話的音韻系統由中華大學余秀敏老師所統整、客語文章的撰寫主要由苗栗的兩位退休教師所主筆，一位龔萬炆老師、另一位為陳碧娥老師，而陳老師也

是語料的主要發音人，為方便起見以下僅以「四縣客家話」指稱本論文的客語語料腔調。

2.2 四縣客家話之聲母

四縣客家話的聲母是處於音節開頭的輔音，如果音節開頭沒有輔音稱為空聲母。四縣客家話聲母中除了 m、n、ng、l、v 具有週期特性，其餘聲母皆為清音，若加上空聲母，四縣客家話共有 17 個聲母，表 2-1 所列即為四縣客家話的聲母。以教育部通用拼音的符號為代表。

表 2-1 四縣客家話聲母表(括號內為例字)

發音 方式	塞音		塞擦音		鼻音	邊音	擦音	
	清音		清音		濁音	濁音	清音	濁音
	不 送 氣	送 氣	不 送 氣	送 氣				
發音 部位								
雙唇	b(飛)	p(片)			m(慢)			
唇齒							f(飯)	v(位)
舌尖	d(知)	t(梯)			n(泥)	l(來)		
舌尖前			z(正)	c(鄭)			s(師)	
舌根	g(高)	k(扣)			ng(咬)			
喉							h(後)	

2.3 四縣客家話之韻母

音節中聲母後面的部分叫做韻母，韻母又可以分韻頭、韻腹和韻尾三部分，其中韻頭就是介音，可以是 i 或 u。韻腹即是主要元音，是每個音節必須有的，四縣客家話的主要元音有六個，分別是 a、e、i、o、u、ii。韻尾有兩種，一種是元音韻尾，如：i、u 等；另一種是輔音韻尾，輔音韻尾又因聲調分成兩種，以鼻音（m、n、ng）結尾的舒聲韻，另一種是由塞音（p、t、k）結尾的入聲韻。

四縣客家話的韻母共有 71 個，如表 2-2。

表 2-2 四縣客家話韻母表

代碼	韻母	代碼	韻母
1	ii	37	ion
2	i	38	un
3	e	39	iun
4	ie	40	ang
5	ue	41	iang
6	a	42	ong
7	ia	43	iong
8	ua	44	ung
9	o	45	iung
10	io	46	uang
11	u	47	iip
12	iu	48	ip
13	iui	49	iep
14	ai	50	ep
15	uai	51	ap
16	oi	52	iap
17	ioi	53	iit
18	ui	54	it
19	eu	55	et
20	au	56	iet
21	ieü	57	uet
22	iau	58	at
23	im	59	uat
24	im	60	ot
25	em	61	ut
26	iem	62	iut
27	am	63	ak
28	iam	64	iak
29	in	65	ok
30	in	66	iok
31	en	67	uk
32	ien	68	iuk
33	uen	69	FNUL
34	an	70	ing
35	uan	71	uak
36	on		

2.4 四縣客家話之聲調

聲調方面國語有四個，閩南語七個，四縣客家話聲調系統是幾個次方言中較為簡單的一個，共有六個聲調，即陰平(調號1)、陰上(調號2)、陰去(調號3)、陰入(調號4)、陽平(調號5)、陽入(調號8)。如圖 2-2 是四縣腔六個聲調的音高軌跡圖。

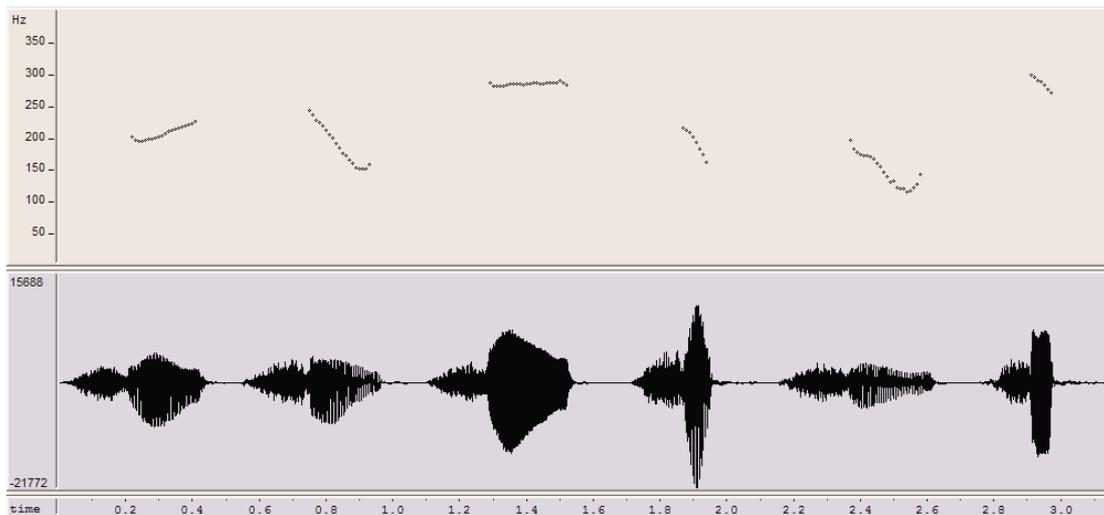


圖 2-2 四縣客家話六個聲調的音高軌跡圖，依左至又分別是 si1 (西)、si2 (死)、si3 (序)、sit4 (卸)、si5 ()、sit8 (砂)

四縣客家話的變調規則相當簡單，只有三個[2]，

1. 由兩個陰平(調號1)字構成的字彙，讀時前字變調讀陽平(調號5)

陰平(調號1) + 陰平(調號1) → 陽平(調號5) + 陰平(調號1)

例：「新衫」 sin1 sam1 → sin5 sam1

「買新衫」 mail sin1 sam1 → mai5 sin5 sam1

2. 陰平(調號1)字與陰去(調號3)字構成的詞彙，讀時前字變調讀陽平(調號5)

陰平(調號1) + 陰去(調號3) → 陽平(調號5) + 陰平(調號1)

例：「針線」 ziim1 sien3 → ziim5 sien3

「拿針線」 nal ziim1 sien3 → na5 ziim5 sien3

3. 陰平(調號 1)字與陽入(調號 8)字構成的詞彙，讀時前字變調讀陽平(調號 5)

陰平(調號 1) + 陽入(調號 8) → 陽入(調號 8) + 陰平(調號 1)

例：「音樂」 im1 ngok8 → im5 ngok8

「聽音樂」 tang1 im1 ngok8 → tang5 im5 ngok8

2.5 四縣客家話之音節

紀錄客語的文字大多是以音節為單位，一個國字代表語言裡一個音節。依據聲母、韻母、聲調的音韻結構，可能的音節組合形式有：17x71x6=7242種，但實際上如同國語，四縣客家話有嚴格的聲韻組合規則，依據以上原則，可以衍生出 13 種不同的音節類型，如表 2-3 所示，四縣客家話音韻系統共有 671 個音節，其詳細列表如附錄一。

表 2-3 四縣客家話音節類型表

音節類型	例字
元音	o2(襖)、i5(于)
元音化的鼻音輔音	m5(毋)、ng5(魚)
元音 + 輔音韻尾	it4(一)、ap4(鴨)
元音 + 元音韻尾	ai2(矮)、oi3(要)
介音 + 元音	iu3(柚)、ia5(爺)
介音 + 元音 + 輔音韻尾	iok8(浴)、iuk4(育)
介音 + 元音 + 元音韻尾	ieu5(謠)、iaul(飮)
輔音 + 元音	ho5(何)、pa1(爸)
輔音 + 元音 + 輔音韻尾	nap8(納)、ziip4(汁)
輔音 + 元音 + 元音韻尾	cai2(採)、tui3(退)
輔音 + 介音 + 元音	mial(摸)、dui3(對)
輔音 + 介音 + 元音 + 輔音韻尾	giuk(趨)、liok8(略)

輔音 + 介音 + 元音 + 元音韻尾

tiau5(跳)、diao3(釣)



第三章 TTS 系統原理

在一個 TTS 系統中可分為四個部分，如第一章圖一所示，在本章節中我們將會簡單描述 TTS 中韻律訊息產生器與語音合成器的兩個部分。

3.1 韻律訊息產生器

在人類的發音模式中，是從肺部的氣流流經發音器官，藉由氣流的變化，可控制聲音的強弱、快慢，另一方面藉由聲帶的震動可控制聲音的頻率高低。韻律訊息產生器在文字翻語音系統中佔極重要的角色，也是合成流暢語音的主要關鍵。韻律訊息的變化是指發音的長短（包含音節和停頓的長度）、音量的大小和聲調的抑揚頓挫，目前對韻律合成的方法大致可分為：

1. 規則法[4,5]：運用語言學的分析方法，歸納出人類發音規則，利用這些規則來產生合成語音的韻律。但是人類說話的方式並非單一無變化，要能有效掌握並不是一件容易的事，必須經由長期的觀察研究才能得到良好的結果。
2. 統計法[6]：使用統計演算法，從大量的語音資料中統計出韻律變化的規則，運用這些規則來控制合成語音的韻律變化。但是此方法需要大量的語音資料，而且很難得到各種韻律特性的語音資料與不容易掌握文句的整體韻律變化。
3. 類神經網路[7,8]：利用一組複雜的網路來模擬人腦的記憶與學習能，其學習方法是採用漸進式的修正錯誤與更新記憶的方式，在經由長時間的學習訓練後，可以得到不錯的效果。

類神經網路的方法不但可以避免前兩者方法的不同，更可以掌握文句的整體性的韻律變化。以下就介紹類神經網路韻律訊息產生器的運作原理。

3.1.1 產生韻律訊息

語音的韻律訊息包含音調的高低、音量的大小、音長的快慢，人類產生韻律訊息的概念圖，如圖 3-1 所示，首先分析輸入文句，獲得語言參數，語言參數可分為詞層次語言參數與音節層次語言參數，詞層次語言參數包含詞、詞類、標點符號等；音節層次語言參數包含音節的子音、母音、聲調等，最後將語言參數送入韻律訊息產生器，產生韻律參數。

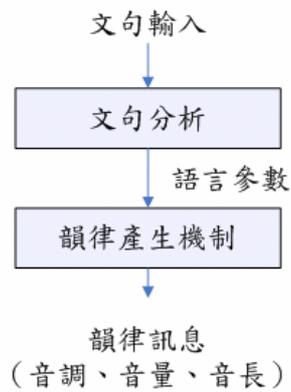


圖 3-1 人腦產生韻律資訊的概念模型

由於詞層次語言參數在文句分析而言是屬於上層語言參數，對於整體韻律訊息的影響較大，所以將韻律產生機制再細分成兩個部分，一為韻律模型，另一為韻律產生器，如圖 3-2 所示。韻律模型的輸入是詞層次語言參數，由輸出的韻律狀態來控制文句韻律的整體趨勢。韻律產生器則是在韻律模型的輔助下，依據音節層次語言參數而產生適當的韻律參數。根據上述模型，我們以遞迴式類神經網路

(Recurrent-Neural-Network, RNN) 來完成圖 3-1 所示的韻律產生機制，其方塊圖如圖 3-3 所示。

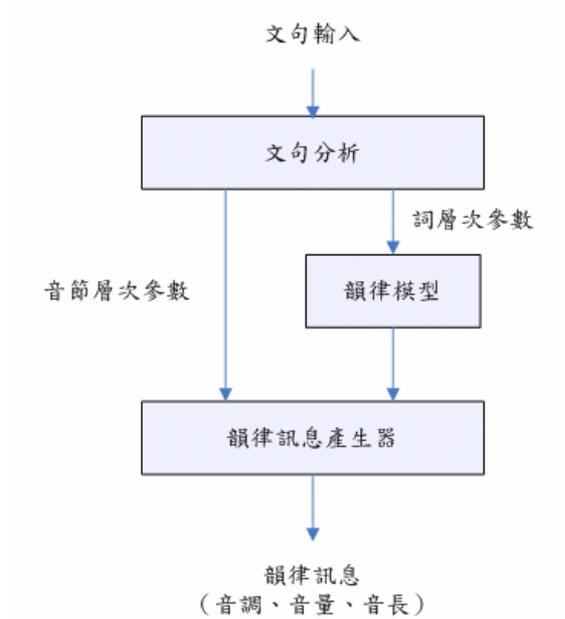


圖 3-2 細分之韻律產生的概念模型

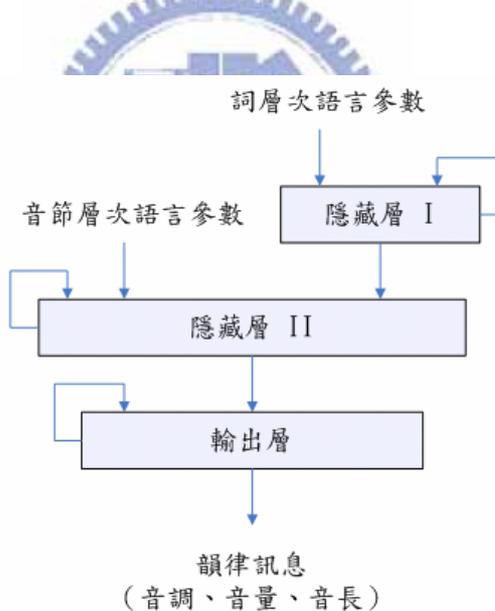


圖 3-3 遞迴式類神經網路方塊圖

如圖 3-3 所示的遞迴式類神經網路為一具有兩個隱藏層的四層網路。它的功能可以分成兩個部分。第一個部分是韻律模型由輸入層和第一個隱藏層所組成，其操作作為詞層次。它是利用詞層次語言參數來探索文句與對應之語音的韻律結構，以詞為單位輸出每個詞之韻律狀態參數。韻律模型輸入的詞層次語言參數包含：目前欲

處理的詞之詞類、詞長、下一個詞的詞類、詞長以及前後的標點符號。其中詞長從一字詞到八字詞共八類；標點符號分成 4 類，如表 3-1 所示；詞類分成 43 類，如

表 3-1 標點符號分類表

類別	標點符號
1	句邊界
2	，。；！
3	「」：、
4	？

第二部分是由第二隱藏層和輸出層所組成，其操作為音節層次，它接受第一部份韻律模型所產生之韻律狀態與音節層次語言參數，並以音節為單位產生每個音節的韻律訊息。為了降低網路的複雜度，將第二隱藏層和輸出層分成三個群組，分別輸出四個基頻軌跡參數、四個音長參數和一個能量參數。音節層次語言參數包括：目前欲處理的音節部分：聲母類別、韻母類別、聲調、詞頭、中、尾、單字詞的分類。

下一個音節：聲母類別、聲調。



3.2 基頻同步疊加合成方法

基頻同步疊加法 (Pitch Synchronous Overlap Add, PSOLA) 分成三個步驟：

Step1、基頻同步分析：

將原來語音訊號與一序列基頻同步視窗函數相乘，可以得到一序列的短時 (short-term) 訊號。

Step2、基頻同步變換：

將這些短時訊號進行適當的時域或頻域轉換，得到與合成基頻軌跡同步的一序列合成短時訊號。

Step3、基頻同步疊加合成：

將合成短時訊號重疊相加而合成語音。

3.2.1 基頻同步分析

將原始訊號與一序列基頻同步的視窗函數相乘，得到一序列的短時訊號：

$$x_m(n) = h_m(t_m - n) \cdot x(n) \quad (3.1)$$

其中：

$x(n)$: 原始語音訊號。

t_m : 基頻同步標記的位置。

$h_m(n)$: 中心位於 t_m 的視窗函數。

視窗一般是使用 Hamming window，視窗長度大於一個基頻 (pitch)，這樣使得相鄰的短時信號總有一部份重疊。通常視窗長度取相對位置處的基頻 p 的 u 倍， $2 \leq u \leq 4$ 。所以，我們可以得到視窗函數長度的變換公式：

$$h_m = h\left(\frac{n}{up}\right) \quad (3.2)$$

其中 $h(t)$: 具正規化長度的視窗函數。



3.2.2 基頻同步變換

將分析短時訊號序列 $x_m(n)$ 轉換成與合成語音基頻標記 \tilde{t}_q 同步的合成短時信號序列 $\tilde{x}_q(n)$ ，此轉換包含三個步驟：改變短時信號個數、改變短時信號間的時間延遲和改變每個獨立的短時信號的波形。合成基頻標記 \tilde{t}_q 的個數和基頻調節因子 β 和時長調節因子 γ 有關，透過適當的演算法找出 $\tilde{t}_q \rightarrow \tilde{t}_m$ 的映射關係，進而找出由哪些分析信號產生哪些合成信號。

3.2.3 基頻同步疊加合成

合成短時信號經過重疊相加的步驟後可以得到最後的合成語音輸出。重疊相加的方法有數種：

1. 最小平均方重疊相加：

$$\tilde{x}(n) = \frac{\sum_q \alpha_q \tilde{x}_q(n) \tilde{h}_q(\tilde{t}_q - n)}{\sum_q \tilde{h}_q^2(\tilde{t}_q - n)} \quad (3.3)$$

其中：

$\tilde{h}_q(n)$: 合成視窗序列。

α_q : 正規化因子，用來補償由於基頻變換所造成的能量變化。

從頻譜上解釋這種合成方法是使合成短時信號 $\tilde{x}_q(n)$ 的頻譜與相對應的合成信號 $\tilde{x}(n)$ 的短時頻譜的平方誤差最小。

2. 簡單重疊相加：

$$\tilde{x}(n) = \frac{\sum_q \alpha_q \tilde{x}_q(n)}{\sum_q \tilde{h}_q(\tilde{t}_q - n)} \quad (3.4)$$

在 (3.4) 式中的分母是個時變的正規化因子，用來補償相鄰視窗的重疊不相同所造成能量變化。在窄頻帶條件下，特別是當合成視窗的長度取合成信號基頻週期的二倍時，這個因子幾乎是常數。在這種情況下，如果假設 $\alpha_q = 1$ ，則 (3.4) 式可以簡化成：

$$\tilde{x}(n) = \sum_q \tilde{x}_q \quad (3.5)$$

此時，合成信號只是合成短時信號的線性和。

3.2.4 時長變化

時長變化可以與基頻變換同時進行，也可以單獨進行。如果作時長變換，就不需要作頻域變換，只需要使用時域基頻同步疊加法就可以。最簡單的情形是當時長變換因子 γ 為一常數時。在這種情況下，如果要降低語音速度，我們只需要重複部分分析短時信號。反之，如果要增快語音速度，我們只需刪除部分分析短時信號。在圖 3-4 中，帶箭頭直線分別為分析時間軸和合成時間軸上的基頻標記，虛線代表之間的對應關係。圖 3-4 (a) 表示未為降低語音速度，有兩個分析短時信號 P1 與

P2 被重複。圖 3-4(b) 表示為增加語音速度，有兩個分析短時信號 P3 與 P4 被刪除。

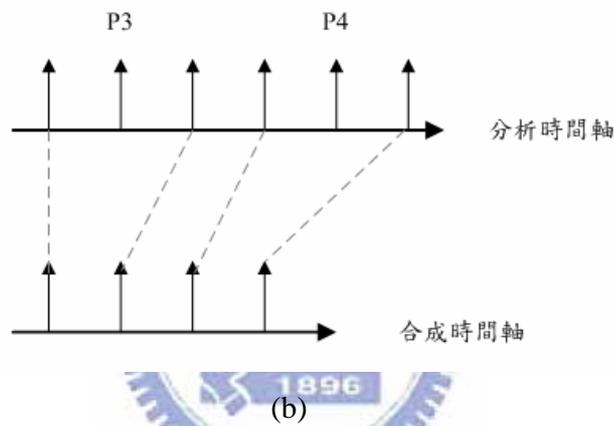
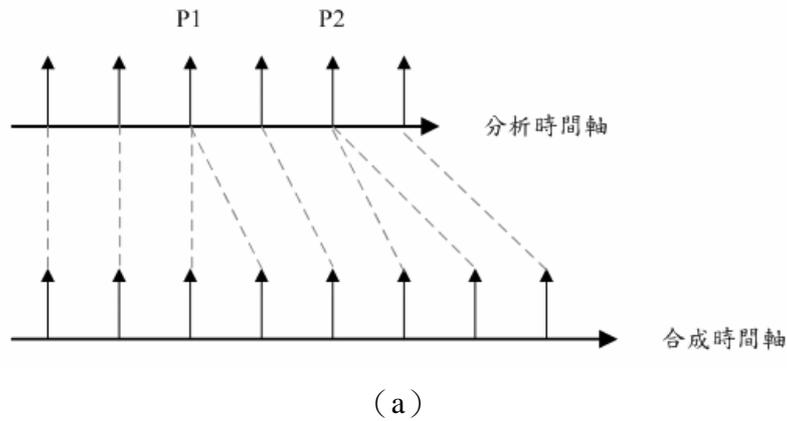
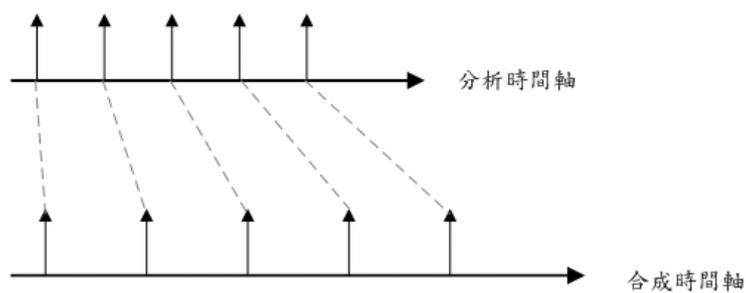


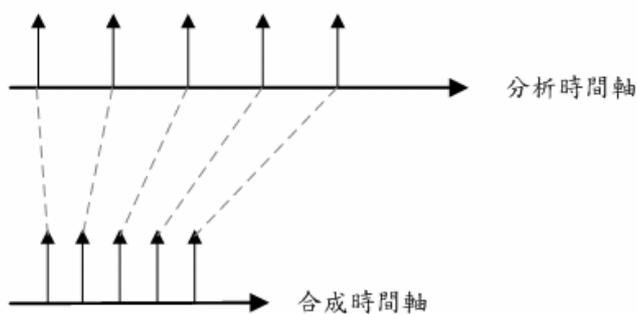
圖 3-4 以時域基頻同步疊加法改變時長

3.2.5 音高變化

音高變化就是按照基頻變化因子 β 改變相鄰合成短時信號的時間延遲。這時音高變化隱含著時長的變化。最簡單的一種情形是當基頻變化因子等於時長變換因子的時候，此時合成基頻標記與分析基頻標記是一一映射的，如圖 3-5 所示，圖 3-5 (a) 表示為降低語音音高，圖 3-5 (b) 表示為升高語音音高。



(a)



(b)

圖 3-5 以基頻同步疊加改變音高 ($\beta = \gamma$)

然而在一一般的情形下，這兩個因子是不相等的。但是，我們可以將視其為先作一次時長變換因子與基頻變換因子都為 β 的變換，再做一次時長變換因子為 $\frac{\gamma}{\beta}$ 的變換。如圖 3-6 所示，分析短時信號先經過一次音高變換，此時音高與音長同時改變，再經由一次時長變換得到適當的音長。事實上，我們可以根據基頻變化因子和時長變化因子找出分析時間軸與和誠實間軸的映射關係，進而找出分析短時信號與合成新號的映射關係，如此指需要進行一次變換即可完成所需的基頻和時長變換。

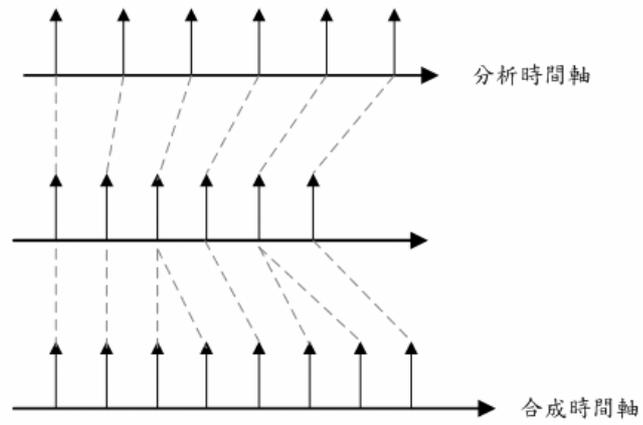


圖 3-6 以基頻同步疊加改變音高 ($\beta = \gamma$)



第四章 訓練語料庫的前處理過程

類神經網路是運用複雜網路來模擬人腦的記憶與學習功能，以漸進式的學習方式修正錯誤與更新記憶，經由大量的學習訓練後，可以得到不錯的效果。在類神經網路學習前，我們必須建立出一套訓練樣本（training pattern）即語料庫，使類神經網路在學習的過程中有一個參考。用來進行網路訓練的輸出層資料稱為目標值（Target），本系統中，目標值包含子音長度、母音長度、音節間停頓、能量、基頻軌跡的正交化係數。而整理訓練語料庫的品質，會影響整個 TTS 系統合成聲音的優劣，所以本章就介紹訓練語料庫的處理過程，包含語料的切割資訊的求取與調整、基頻軌跡參數的求取與修正。

4.1 語料庫說明

四縣客家話語料庫文字稿，已收錄 57 篇文章，處理過程依序為蒐集文章、錄製語料、音節標記，主要由苗栗的兩位退休教師一位是龔萬灶老師、另一位為陳碧娥老師；與中華大學余秀敏教授所主筆，而陳老師也是語料的主要發音人。文章中有 57 篇已完成錄音，共有 42,337 個音節，語音檔共 325 個，407Mbytes。錄音軟硬體設備及格式如下表所示：

表 4-1 錄製音檔軟硬體設備與設定

錄音軟體	Adobe Audition 1.0
麥克風	單一指向性(Uni-directional)
錄音場所	普通房間
錄音情境	依照文稿唸出
取樣頻率 (Sampling rate)	20 kHz
發音速度	4.3 syllable/second
取樣大小	16 bits
聲道	單聲道(mono)
檔案格式	pcm

4.2 切割資訊

在此小節中，將介紹四縣客家話語料庫如何使用工具軟體，求得基本的切割位置，並為了精準化，進一步介紹人工調整切割位置的方法。

4.2.1 切割資訊的求取

雖然目前四縣客家話語料庫中已有 42,337 個音節，但在 671 基本音節中仍有約 99 個音節尚未出現，而只出現 5 次以內的音節有 215 個。本論文是利用 HTK(Hidden Markov Toolkit) 軟體訓練音素(Phone)的 HMM Model，再對所有的語料進行強制切割 (Forced Alignment)。其中參數的設定為 38 維的參數，包含 12 階的梅爾倒頻譜參數(Mel-frequency Cepstral Coefficients, MFCCs)與能量對數值(log energy)，及其一階微分與二階微分，扣除原本的能量對數值後共 38 維，音框大小(frame size)設為 32ms，音框位移(frame rate)設為 5ms。

4.2.1 切割資訊的修正

為了提供韻律訊息產生器訓練良好的韻律模型，將得到的初步切割位置，用人工修正。以下介紹人工調整切割位置的判斷方式。

方法一：除了聲母是濁音 (m、n、ng、l、v) 之外的音節，其子母交界處可以運用時域與聲譜圖 (spectrogram) 的週期特性或非濁音之聲母在聲譜圖上無低頻成分的特性來分辨之，如圖 4-1。

方法二：音節的起始與結束位置，即聲母的開始與韻母的結束，也可以運用聲譜圖的能量分佈判斷之，如圖 4-1。

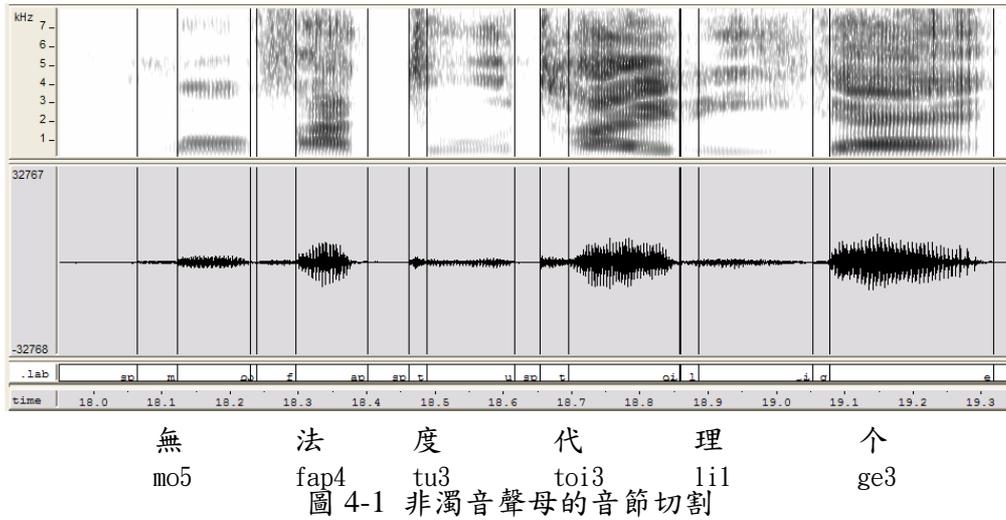


圖 4-1 非濁音聲母的音節切割

方法三：當連音字串中，第一個音節是舒聲韻輔音韻尾即鼻音韻尾（m、n、ng），第二個音節的聲母是濁音（m、n、ng、v、l）時，運用因口腔形狀的改變造成波形的變化，來分辨之，如圖 4-2。

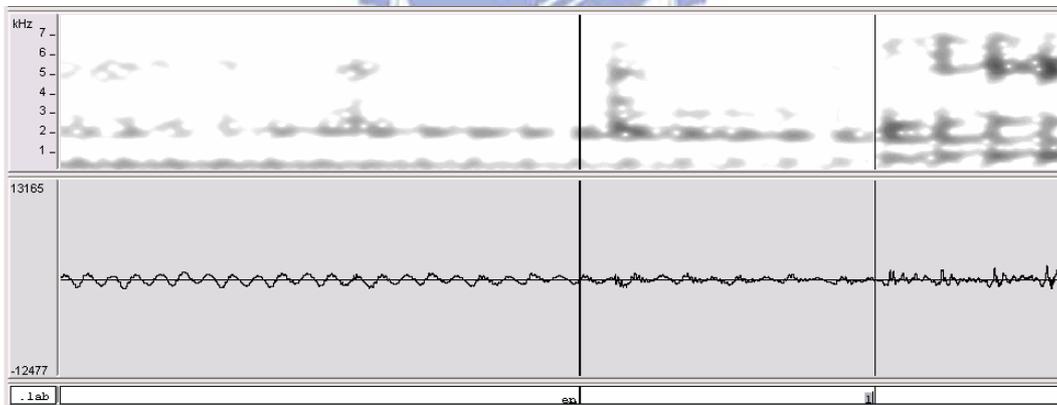
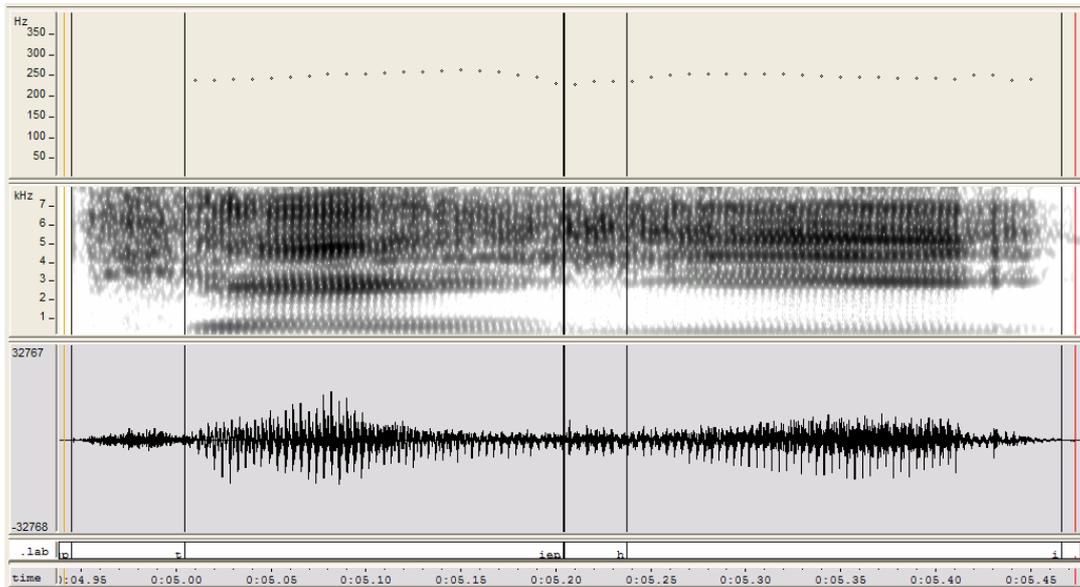


圖 4-2 連音中舒聲韻輔音韻尾與濁音切割處 1

方法四：部分聲母因受前後韻母影響，而有週期特性，可以利用滑順音高軌跡的轉折處判斷兩個音節的交界。此例常發生在連音中第二個音節的聲母為 h 時，如圖 4-3。



電 tien3 器 hi3
圖 4-3 當聲母受前一個音節影響，而有週期特性時的切割位置

經由以上利用語音特性來人工調整切割位置方法後，再比對原始切割語料，如圖 4-4 為比對過後的統計圖，從圖中可得知人工調整過後的音素開始與結束位置皆比原始切割位置右移 3 至 4 個 frame (10m sec/frame)。

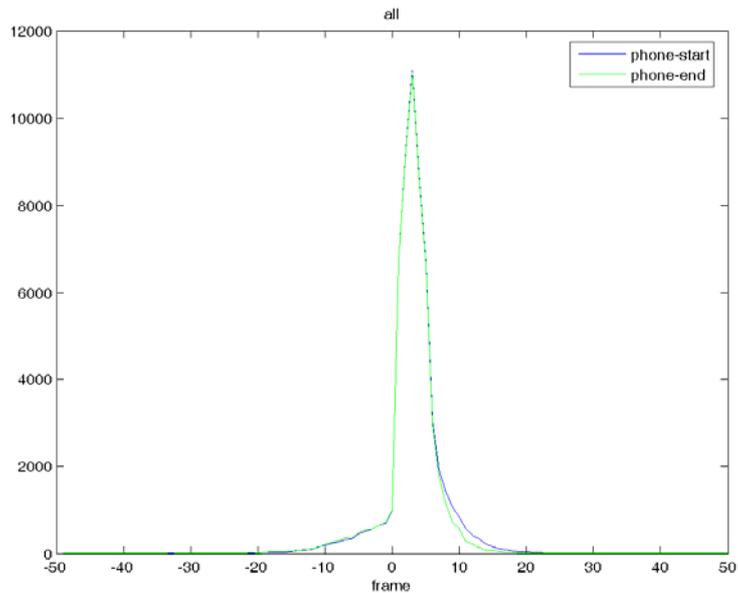


圖 4-4 人工調整與原始切割位置位移統計圖

4.3 基頻軌跡參數的求取

本論文是透過WaveSurfer軟體，使用ESPS演算法將整段音檔的基頻軌跡求出；如同其它的求取基頻軌跡的演算法，二倍基頻（Double pitch）、半倍頻（Half pitch）等的情況會出現在所求得的基頻軌跡裡，為了能正確取得基頻資訊，我們試著建立一套檢查方式，利用在音節切割位置正確的情況下，將錯誤的基頻位置找出，以人工修正的方式更正之。

4.3.1 基頻軌跡檢查前處理

因為WaveSurfer所求得之基頻軌跡，並沒有自動修正的功能，所以常在音節的起頭或結尾之基頻軌跡有偏高或偏低等不連續現象，主要原因是音節起頭與結尾容易有週期特性不穩定的現象，而本論文的修改方法是將此情況的不連續點捨去，如圖4-5。另一個情形是發生在若一段音節基頻軌跡中，出現單點或一小段的跳躍等不連續現象，則利用內插法將此點補回，如圖4-6。

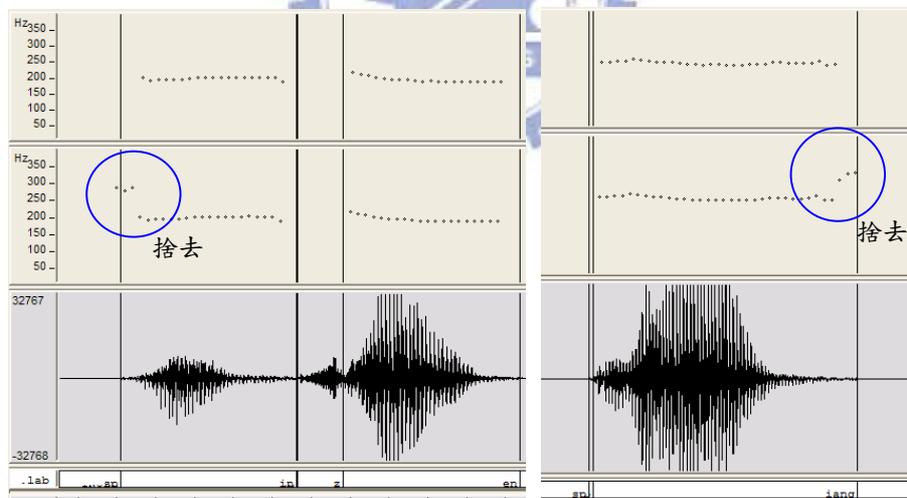


圖4-5 音節起頭與結尾部不規則的跳動

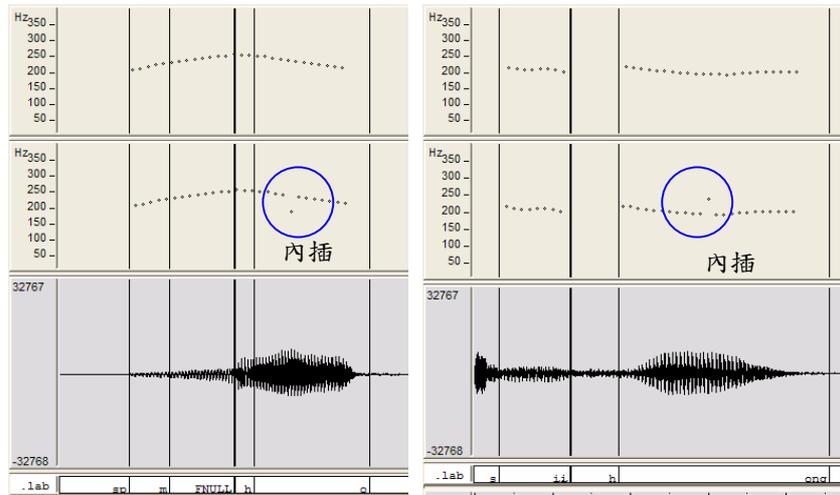


圖4-6 利用內插法修補不平滑的音高軌跡

4.3.2 基頻軌跡檢查流程

下圖4-7為一音節區間內，基頻（F0）數值示意圖，虛線表syllable boundary位置。利用我們所提出的檢驗方式，可將音節區間內基頻有問題的地方標示出，其檢查的流程如下：

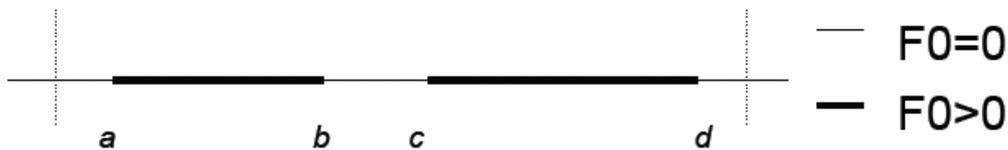


圖4-7 音節F0示意圖

1. 如圖4-7，先找出區間內第一個F0非零的位置a，由此位置開始記錄連續非零的F0值，直到下一個F0為零的位置b為止。
2. 接著再找下一個非零值位置c，若b、c間只有一個值為0的F0，則用此位置左右各一點自動內插出此點之F0（此動作可在前處理中發現），若有二個以上為0之F0，則利用上述1之方法，從c點開始將此區間所有F0非零的區段找出。
3. 在正常的情況下，一個音節區間內必有F0，且F0會連續，因此只會找出一段F0非零的區段。若區段數為零，表示此音節求不出F0，若區段數大於等於二個，則表

示切割位置有問題。

- 4.若求得區段只有一個，為了求得基頻軌跡參數，F0的個數要大於等於四，因此判斷區段內F0個數，若小於四個，則輸出錯誤訊息。
- 5.若個數大於四個，求取音節F0的平均值M，並依序檢查每個F0的變動是否超過一個範圍，依據觀察的結果，我們將此範圍設為 $\pm 30\text{Hz}$ ，若超過此範圍，則有Double pitch或Half pitch的現象發生。
- 6.依步驟1~5，求出每個音節的 M_j （並非所有音節皆可求出M值），比較 M_j 是否在一個範圍內（人說話時的F0約在60~400Hz），若不在範圍內則視為錯誤發生。
- 7.依序檢查 M_j 變動的範圍，若變化過大，則發生pitch jump的現象，但在prosodic phrase boundary時，M值會突然變化，此現象屬於正常的變動，因此為了避免誤判，我們將連續觀察三個音節的M是否變動過大，才決定是否有pitch jump發生。

4.3.3 基頻軌跡檢查結果與討論

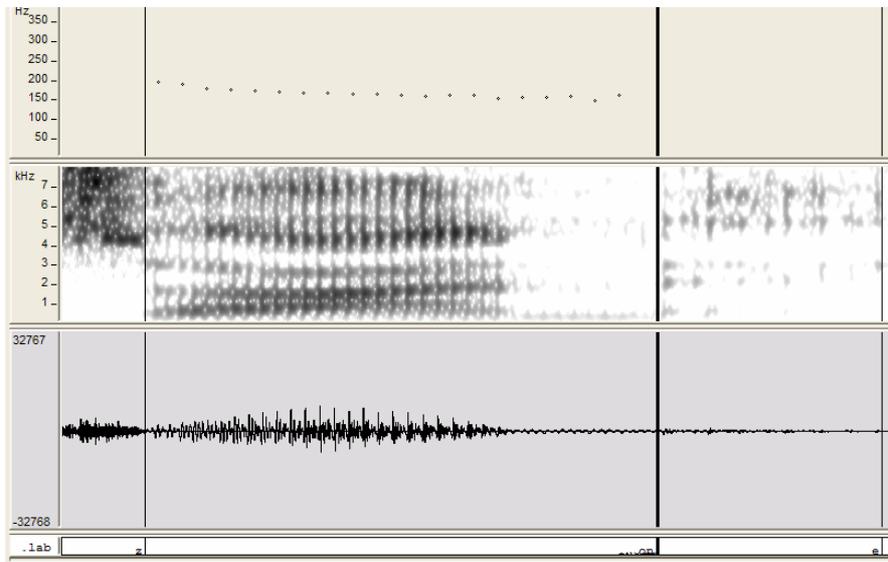
經由上面的檢查，我們可將錯誤分為：

- 1.沒有求得F0；
- 2.音節內基頻軌跡不連續；
- 3.基頻軌跡個數少於四個；
- 4.Half pitch；

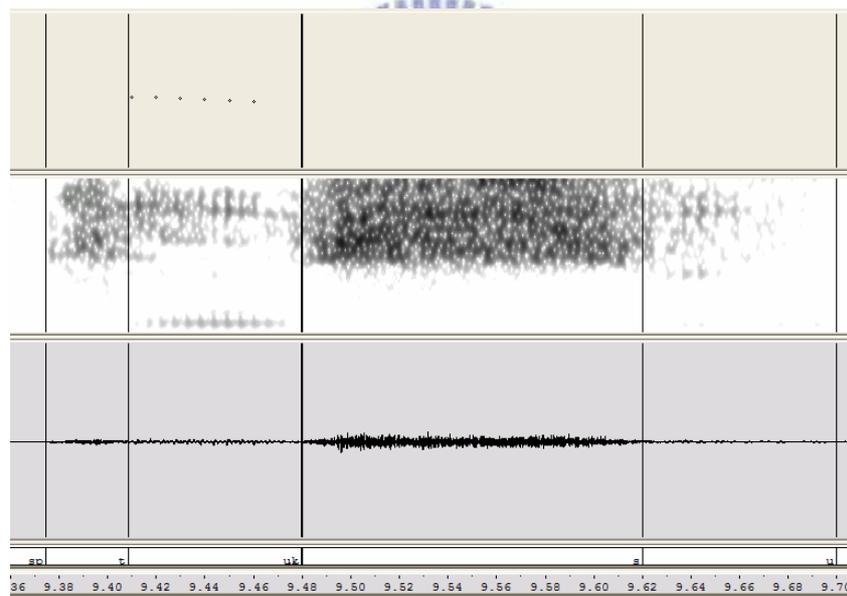
除了音節內基頻軌跡不連續為切割位置錯誤造成外，其它錯誤發生原因可歸為下面幾類：

- 1.音節發音時間短，或是在音節中，unvoiced部分比voiced部分時間長很多，會造成求不出F0及F0個數小於四個的錯誤，尤其當子音為摩擦音類時，容易發生第二種狀況，如圖4-8，（a）為音節發音過短，（b）為子音長度過長所造成求不出F0的現象。





(a) 為音節發音過短一轉仔 it4 zong2 e5



(b) 為子音長度過長，讀書 tuk8 su1

圖4-8

2. 音節尾音拉得太長時，或音節與下個音節發生耦合現象時，會有double pitch錯誤發生。此問題作法目前是将影響整段基頻軌跡的不連續的尾部軌跡刪除，如圖4-9所示，為音尾發生double pitch的現象。

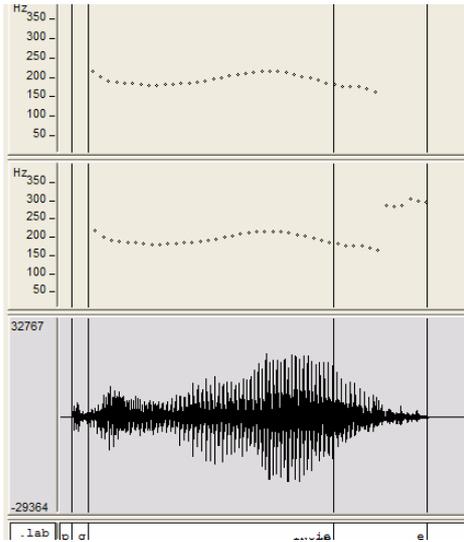


圖4-9 為音尾發生double pitch

3. 音節結尾是入聲韻輔音韻尾(p.t.k)時，常與整段音節基頻軌跡不連續，為了不影響整段基頻軌跡的平滑性，只要是入聲韻輔音韻尾造成的不平滑的部分刪除，如圖4-10。

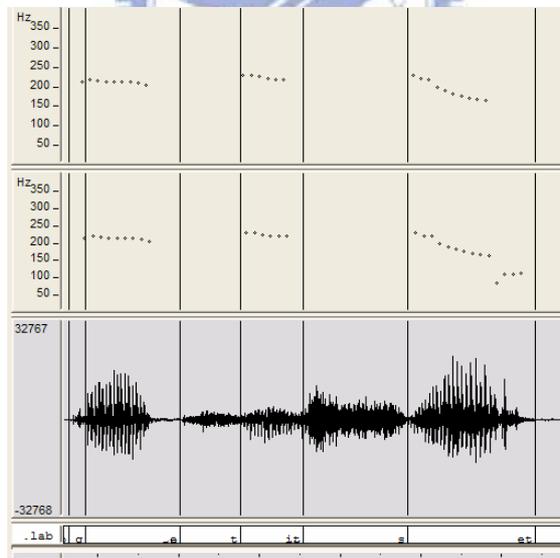


圖4-10

5. 因為錄製語料為女性錄音者，ESPS求取的音高軌跡，Half Pitch的情況多於Double Pitch，當遇到此情況，將Half Pitch部分恢復原來的音高值，使整個音節其基頻軌跡有平滑性，如圖4-11。

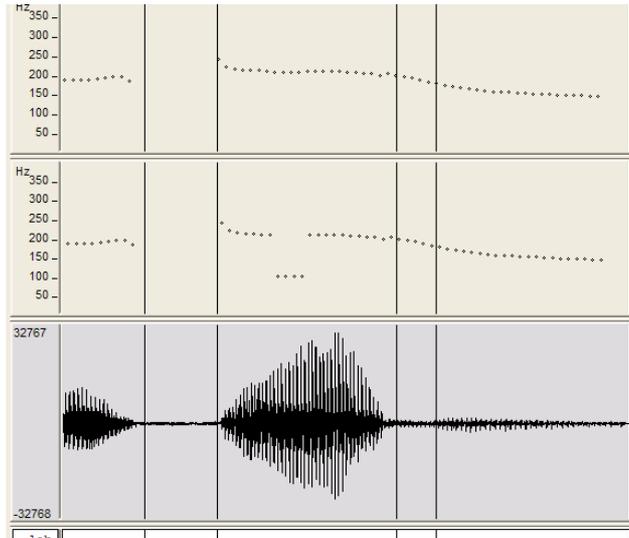


圖4-11 基頻軌跡出現Half Pitch現象

6. 可能是因能量過低或週期特性不明顯，使求出的基頻軌跡少於四點，本論文作法是將此音高軌跡依照其平滑方向順補至四點，如圖4-12。

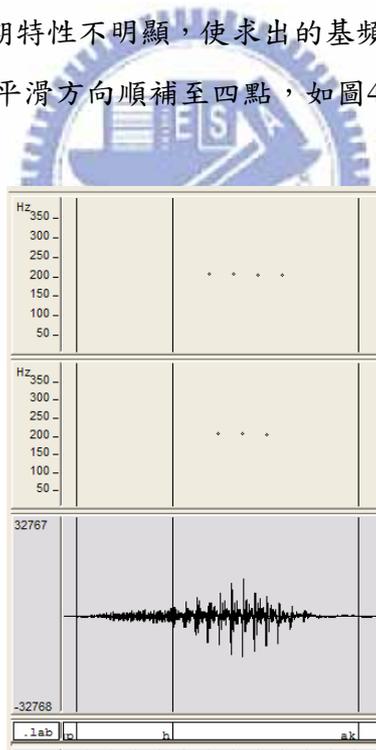


圖4-12 音高軌跡少於四點

經過以上的步驟，將所有錯誤找出後，我們以人工修正的方式，修訂基頻軌跡，當修正完成後，即可利用下一節介紹的方式，求得基頻軌跡參數。

4.3.4 基頻軌跡參數

語音合成的實作上，為了降低儲存所有基頻軌跡所花費的大量記憶體及運算時間，我們以基頻軌跡參數[3]代替原有基頻軌跡；基頻軌跡參數為所有基頻軌跡正交化展開之前四階係數，其詳細數學式如下：

$$a_j = \frac{1}{N+1} \sum_{i=0}^N \text{Pitch}(i) \cdot \Phi_j\left(\frac{i}{N}\right) \quad (4.1)$$

其中 (a_0, a_1, a_2, a_3) 代表基頻軌跡的四個參數， $\text{Pitch}(i)$ 為原始基頻軌跡， $0 \leq i \leq N$ ， $N+1$ 為基頻軌跡的長度。而 $\Phi_j\left(\frac{i}{N}\right), 0 \leq j \leq 3$ 為正交化函數，其中定義如下：

$$\Phi_0\left(\frac{i}{N}\right) = 1 \quad (4.2)$$

$$\Phi_1\left(\frac{i}{N}\right) = \left[\frac{12 \cdot N}{(N+2)} \right]^{\frac{1}{2}} \cdot \left[\left(\frac{i}{N}\right) - \frac{1}{2} \right] \quad (4.3)$$

$$\Phi_2\left(\frac{i}{N}\right) = \left[\frac{180 \cdot N^3}{(N-2)(N+2)(N+3)} \right]^{\frac{1}{2}} \cdot \left[\left(\frac{i}{N}\right)^2 - \left(\frac{i}{N}\right) + \frac{N-1}{6 \cdot N} \right] \quad (4.4)$$

$$\Phi_3\left(\frac{i}{N}\right) = \left[\frac{2800N^5}{(N-1)(N-2)(N+2)(N+3)(N+4)} \right]^{\frac{1}{2}} \cdot \left[\left(\frac{i}{N}\right)^3 - \frac{3}{2} \left(\frac{i}{N}\right)^2 + \frac{6N^2 - 3N + 2}{10N^2} \left(\frac{i}{N}\right) - \frac{(N-1)(N-2)}{20N^2} \right] \quad (4.5)$$

若知基頻軌跡參數，可以利用下式重建基頻軌跡。

$$\text{Pitch}'(i) = \sum_{j=0}^3 p_j(k) \cdot \Phi_j\left(\frac{i}{N}\right), 0 \leq i \leq N \quad (4.6)$$

第五章 文字轉語音系統之製作

四縣客家話文字轉語音系統，包含四個模組：文句分析器、韻律信息產生器、合成單元波形表、語音合成，系統發展過程之工作說明如下：

5.1 文句分析器

文字分析器為 TTS 系統中第一級處理的部分，也是最重要的單元；文字分析結果的優劣會直接影響合成聲音的品質和正確性，所以在做文字分析時，必須使用正確的資訊和方法做分析，以提供後級韻律訊息產生器好的各項資訊。其系統架構圖如下：

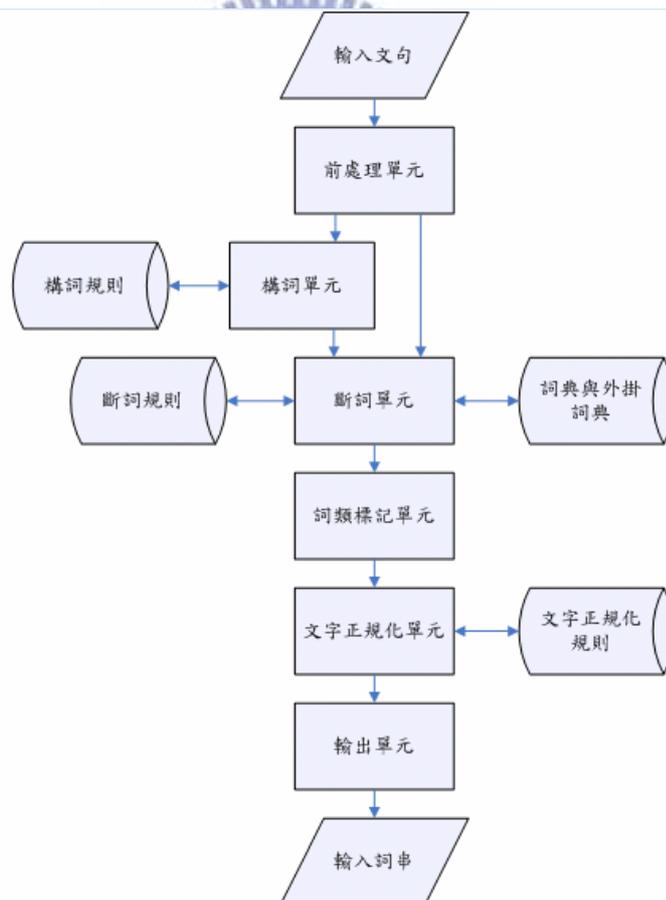


圖 5-1 客語斷詞器架構圖

(1) 前處理單元：

為了使系統處理格式統一，將輸入文句中含有 ASCII code 與 Big-5 code 的字串，全部統一轉成 Big-5 code 格式，例：「西元 1945 年,對福建廣東心舅仔人數介調查,」轉為「西元 1 9 4 5 年，對福建廣東心舅仔人數介調查，」。另一方面，客語文字如同台語書寫，也有漢字/拼音夾雜情況，目前解決方式是利用 Big5-code 中保留區域對應其客語 671 音節拼音，此作法也可以處理連續拼音現象。

(2) 斷詞單元：

初步將全形字串利用 Matching Algorithm 與客語詞典進行比對，再將延用中文斷詞器所使用的中研院詞庫小組所提出的斷詞規則，依序為長詞優先、詞長標準差最小者優先、附著語素最小者優先、定量複合詞字數合最小者優先、一字詞詞頻最高者優先，與總詞頻最高者優先等，選取最適當的斷詞結果。

(3) 構詞單元：

如常見「西元 1980 年」、「三十公斤」等數量複合詞或其他同字疊字型態出現的詞，如：「結結」（結實的樣子）、「極極」（很擁擠的樣子）、「洗洗湯湯」（泛指洗滌的工作）等，都是無法全部收錄之詞典中，所以利用有規律的詞串組合，提供良好的斷詞結果。

(4) 詞類標記單元：

此單元的目的是將斷詞後的詞串給予對應的詞類標記（Part of Speech, POS）

(5) 文字正規化單元：

書寫與口讀會有差異，如：書寫是寫「90%」，但口讀是「百分之九十」，所以此單元的目的是將部分阿拉伯數字、詞、符號等由寫法轉為口讀寫法。

(6) 輸出單元：

輸出韻律訊息產生器所需要的詞、音節碼、詞性、斷詞位置等語言參數。

以下將針對文句輸入輸出格式、詞典的收集、文字分析程序、文字分析結果範例等四個部分作討論。

5.1.1 文句輸入輸出格式

客語文章如同閩南語文章都有漢羅夾雜與書寫系統不一致的特性。尤其是當文章用字遣詞方式越接近口語、文章內容越屬於鄉土文學，漢羅混用的情形就越明顯，例如『當時 ge2 竹筴，就像這下介 hiap4 錢箱樣仔』就是一個混用漢字和羅馬拼音的句子。另外，書寫不統一，例如「我」字會用「崖」、「吾」代替，但其拼音皆是 ngai5。目前共收錄 57 篇短文，其中包含故事、俚語、諺語等文章內容。

以客語文字分析器分析句子，可以得到以下輸出資訊：

- (A) 斷詞結果：分析輸入文句，轉換成以詞為單位的詞串。
- (B) 詞長：分析文句後，詞串中每個詞的長度。
- (C) 音節在詞中的位置：文字在詞中的位置，是詞首、詞中或詞尾。
- (D) 音節碼：各詞所擁有的客語音節碼，標音原則則以辭典內所收集的詞／拼音對照表中的拼音為標音結果。
- (E) 詞類標記：以辭典內詞所擁有的詞類中出現頻率最高的詞類作為詞類標記的標記結果。

表 5-1，為輸入文句『它內容大約是講客家是中國按多民族中最進步介民族，』所得到的斷詞結果輸出形式：

表 5-1 文字分析器輸出結果範例

詞串	音節碼	詞長與 音節在詞的位置	詞性
它	1178	101	24
內	3219	201	12
容	5651	202	12
大	3182	101	1
約	4665	101	6
是	3385	101	11

講	2467	101	12
客	4620	201	1
家	1439	202	1
是	3385	101	11
中	1316	201	12
國	4478	202	12
按	3645	101	12
多	1139	101	1
民	5074	201	12
族	8383	202	12
中	1316	101	1
最	3296	101	7
進	3306	201	37
步	3036	202	37
介	3438	101	7
民	5074	201	12
族	8383	202	12
，	9001	101	49

在輸出詞串中，各詞保有本身的詞類標記、音節碼，以及由詞本身可以計算出來的詞長、音節在詞中的位置等資訊，以提供下一級韻律訊息產生單元所需的韻律參數。

需要得到上述的各項資訊，我們必須建立一套擁有這些資訊的詞典，以及適當的文字分析處理程序。以下將針對詞庫的狀況以及處理程序作進一步說明。

5.1.2 詞庫的收集狀況

詞庫是文字分析器最重要的資料之一，以詞庫的各項資訊可決定斷詞結果的輸出。客語詞庫的收集來源目前主要從「台大客家社詞典」、「台北市客委會 - 現代客語詞彙彙編」、「醫院地區標示語」、「華、閩、客互譯詞典」、「客語實用手冊」、「客家語言能力認證基本詞彙-初級題庫」...等中收錄客語詞，另外也從所蒐集文章中人工斷詞的部分，挑出詞典尚未收錄的客語詞，加入詞典。由中華大學余秀敏教授提供拼音系統轉換格式，再用程式轉換成實驗室的四縣客家話拼音系統。表 5-2 為針

對目前已收錄且標記詞性的客語詞典的各詞長長度統計的各詞詞數。

表 5-2：客語詞典的詞長統計

字詞	一字詞	二字詞	三字詞	四字詞	五字詞	六字詞	七字詞	八字詞	總計
數量	7326	18549	5927	3776	413	139	99	30	36259

5.1.3 構詞規則的整理狀況

構詞規則在中文文字分析架構下已經被證明效果非常好且正確，所以加入構詞規則對於目前的客語文字分析器來說是非常重要的。目前客語構詞規則除了延用中文構詞規則外，也從《客語實用手冊》第十三章客語中的量詞章節中，收錄客話特有的量詞，如：尋 (kim5)，兩手平張，從左手指尖到右手指尖的長度為「一尋」。例句：「這枝竹篙有兩尋長」。詳細如附錄三。另外，將部分國語詞轉成客語書寫念法，如：國語的『今日』在客語的書寫念法是『今晡日』。

5.1.4 客語文字分析結果範例

以下為數個基於目前的客語文字分析器流程所分析出來的斷詞結果範例，為方便比較，以下將以自動／人工斷詞結果並排的方式列出並做簡單的討論。表格中第一行為客語文字，第二行為四位數音碼，最左邊第一位數代表聲調，後面三位數為音節帶碼，範圍由 001 至 0671，其代碼對應的音節如附錄一，第三行為音節在詞中的資訊，如：成功，「成」在此詞是詞首所以在第三行欄為填 201，「功」在詞中第二個字，所以填 202。

例句一：

自動斷詞：

但係(Cbb)無半樣(VH)成功(Nb)，總係(D)半(Da)途(Na)而(Caa)廢(A)。
dan3he3 mo5ban3iong3 siin5gung1 ， zung2he3 ban3 tu5 i5 fi3 。

人工斷詞：

但係(Cbb)無半樣(VH)成功(Nb)，總係(D)半途而廢(VH)。

dan3he3 mo5ban3iong3 siin5gung1 , zung2he3 ban3tu5i5fi3 。

此例句斷詞結果是好的，只有在句尾的「半途而廢」，因在斷詞器的詞典尚未收錄，所以無法斷成一個詞，解決方法是把自動斷詞與人工斷詞結果進行比對，從人工斷詞的文句中收錄尚未收錄的客語詞。目前從 57 篇文章（共 42186 個音節）中，不包含一字詞，人工斷詞部分有 12,583 個客語詞，自動斷詞有 12029 個，所以約有 500 多個客語詞可加進總詞典中。

例句二：

自動斷詞：

第二個(DM)賴仔(Na)也(D)已(D)煞猛(A)，學著(VC)一手(DM)殺豬(Na)的(DE)技術(Na)。

ti3ngi3ge3 lai3e2 ia1 i2 sat4mang1 , hok8do2 it4su2 cii5zu1 ge3

gi1sut8 。

人工斷詞：

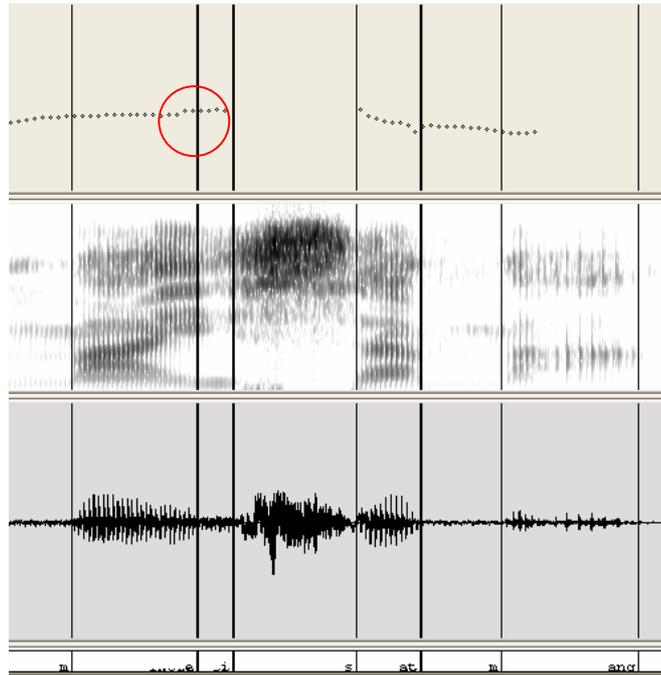
第二個(DM)賴仔(Na)也(D)已(D)煞猛(A)，學著(VC)一(A)手(Na)殺豬(Na)的(DE)技術(Na)。

ti3ngi3e3 lai3e2 me3 i3 sat4mang1 , hok8do2 it4 su2 cii5zu1 ge3

gi5sut8 。

此句斷詞結果，可從以下幾方面來探討：

1. 人工斷詞部分，其音碼欄位由龔老師聽語料，再進行標音，「個」字，因口語自然的表達，此字受前一字影響原本單字應念 ge3 而唸成 e3。
2. 「也」其人工、自動斷詞拼音不一致，原因為一字多音，即破音字。在語料中，此句「也」發 me3，但在詞典一字詞中「也」是以 ia3 收錄，這也是目前 TTS 系統待解決的問題之一。
3. 「已」為人工與自動斷詞結果，聲調不一致，從此句的音高軌跡圖 5-2 來看，為聲調 3，所以此字也是一字多音。



也 已 熬 猛

圖 5-2 音高軌跡圖

4. 自動斷詞已加入構詞規則，所以會將「一手」構成一個詞，而人工斷詞錯誤係人為疏忽。
5. 因自動斷詞目前未將變調規則加入，所以在「技」字，其聲調正確為 5。

總結，人工與自動斷詞的比較結果，主要有：

1. 音節不一致：

如果是單字詞，最有可能的原因是一字多音，如果是二字詞以上，可能是音節在此字受前後音影響，有念法有所不同。

2. 聲調不一致：

客語聲調主要有三個，如第二章客語語文特性中提到，如果將構詞規則加入斷詞器單元，最後合出的聲音，才可接近口語化。

3. 斷詞資訊不一致：

人工斷詞是需要的，可找到自動斷詞中詞典所未收錄的客語詞。

5.2 韻律訊息產生器

韻律訊息，包含音節基頻軌跡、音節能量、聲(韻)母長度、及音節間停頓長度等，本系統使用的方法是類神經網路 (neural networks)，它由一個大的語料庫來學習語言參數到韻律參數的對應關係，訓練時藉調整類神經網路的神經元間連結的加權 (weights)，在合成時，將語言參數輸入來產生韻律參數。

首先我們請陳碧娥老師錄製了一個朗讀的語料庫，共有文章 57 篇，其中包含：短文、諺語、俚語等，共 42186 字，文章是陳碧娥老師及余秀敏所蒐集，我們對文章進行處理，作了斷詞、標示詞類、標示音節碼的工作。

語音的錄製是使用 CoolEditor 軟體，由 PC 上的聲霸卡直接錄製，使用 16k Hz 的取樣率，量化採用 16 bits。錄製完之語音先訓練一個 HMM models，用來做 forced alignment 切割音節及聲韻母邊界，再以人工調整切割位置；另外我們使用 ESPS 軟體做 pitch detection，再以人工更正音高軌跡，詳細內容第四章所描述，如圖 5-3 為一段處理完後的語音範例，表 5-3 為由處理後的語音訊號中抽取出來的韻律訊息參數，這些參數將作為 RNN 類神經網路學習時之目標值。音節或聲韻母啟始與結束位置其單位是 sample，正交化軌跡參數也是以 sample 為單位。

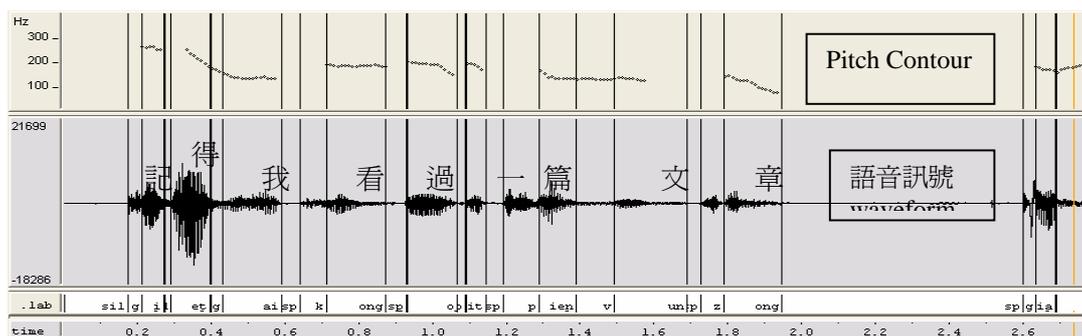


圖 5-3 語音訊號處理範例

表 5-3 由處理後語音訊號中抽取出來的韻律訊息參數範例

文字	音碼	斷詞	詞性	音節起始位置	聲母、韻母交界處	音節結束位置	能量參數	基頻軌跡參數 ¹	基頻軌跡參數 ²	基頻軌跡參數 ³	基頻軌跡參數 ⁴
記	3436	201	12	2802	3405	4363	58.17	61.20	1.05	0.56	0.00
得	4166	202	9	4398	4680	6360	64.45	72.06	5.86	-0.00	0.22
我	5551	101	24	6402	6920	9470	52.18	105.80	-3.13	-16.06	-10.15
看	3523	201	9	10251	11400	13960	49.40	85.41	0.285	-0.51	0.51
過	3442	202	9	14844	14915	17061	55.66	79.963	15.98	-6.34	11.87
一	4656	401	12	17434	17452	18338	52.90	72.59	25.89	-14.29	13.79
篇	1044	402	12	19083	20609	22223	54.89	101.73	22.63	-19.65	13.26
文	5125	403	12	22240	23880	27029	44.94	120.39	0.67	0.61	1.44
章	1314	404	12	27632	28625	31144	50.30	121.07	7.52	-0.69	-0.29
，	9001	101	49	0	0	0	0	0	0	0	0
它	5440	101	24	41608	42175	43000	64.19	91.56	2.40	-0.69	-0.19
內	3219	201	12	43080	44818	47317	54.72	72.61	-13.01	3.10	1.67
容	5651	202	12	47337	47376	49127	57.10	82.95	17.47	5.06	0.04
大	3182	201	1	52143	53318	55089	55.74	76.35	11.51	-13.66	11.99
約	4665	202	6	55128	55187	56997	57.85	82.66	-5.09	-13.95	-15.70

本系統之輸入語言參數包含兩類，一類是音節層次 (syllable-level) 的參數包括前後音節的 tones、現在音節的 17 類聲母及 47 類韻母、現在音節在詞中的四種位置資訊(單音節詞、多音節詞的詞首、詞中及詞尾)等，另一類是詞層次 (word-level) 的參數包括前後詞的詞類 (part of speech, POS)、詞長、詞前後是否有標點符號等。

在完成訓練後，此韻律信息產生器之效果不錯，產生之韻律信息可用以合成流利自然的語音，圖 5-4 為一個韻律參數產生之範例，由圖中可看出 RNN 產生的各種韻律參數和人語音的韻律變化整體而言有一致的之變化，表 5-4 列出各種韻律參數之均方根誤差，表 5-4(a)為將 57 篇短文，共 42,186 音節，表 5-4 (b) 為拿掉文章為諺語、俚語內容，共有 40,390 個音節，從這結果可以看出，因為念俚語、諺語時，人類說話的速度與韻律變化不是以平常自然的方式說出，尤其在說話停頓的地方。所以未來如欲增加語料，需以短文為主，但仍有少許的較大誤差，可能是 tone sandhi 效應所引起，需要在未來做進一步的改進。

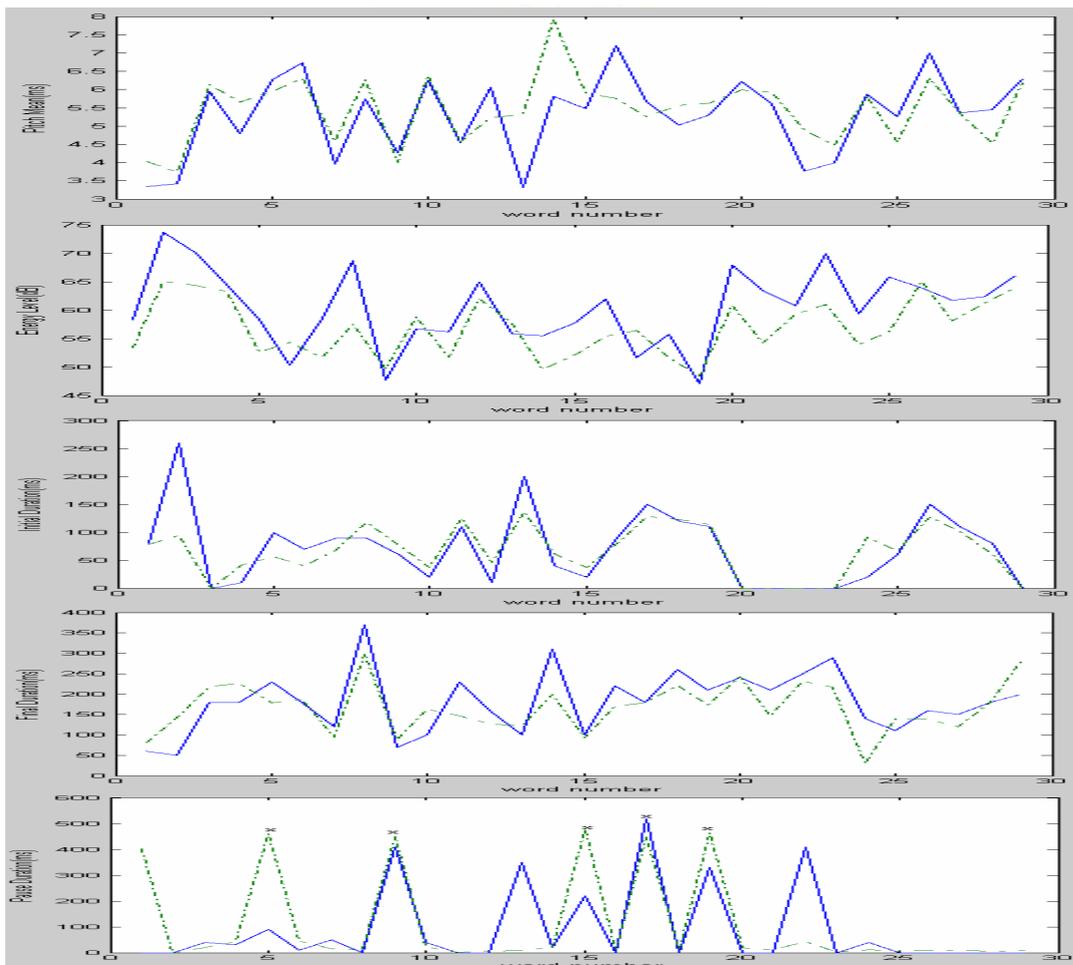
表 5-4 韻律參數之均方根誤差

(a)

	Inside Test
F0 Contour	30.397(sample)
Pause Duration	918.336(sample)
Initial Duration	324.910(sample)
Final Duration	689.168(sample)
Energy Level	3.794(db)

(b)

	Inside Test
F0 Contour	29.545(sample)
Pause Duration	886.671(sample)
Initial Duration	322.782(sample)
Final Duration	677.068(sample)
Energy Level	3.721(db)



註一：虛線表示合成結果，實線為人工標示求出；

註二：x 表示此字之前為標點符號

圖 5-4 RNN 產生韻律參數之範例：(a) pitch mean, (b) energy level, (c) initial duration, and (d) final duration of syllables as well as (e) inter-syllable pause duration. 文章為“第二叫著「田頭地尾」：客家婦女蒔田、割禾、除草、剷秧仔樣樣不會輸細賴仔。”

5.3 聲音波型資料庫

時域基頻同步疊加是一種利用波形表 (Wave Table) 來合成語音的合成方法，因此我們必須整理出適用的語音波形。由於客語和漢語一樣是以音節為基本的發音單元，音節的數目並不多，我們因此使用不帶聲調的 671 個基本音節做為 TTS 系統的基本合成單元，本系統共使用 17 種聲母及 71 種韻母，其中包含空聲母、空韻母以及帶入聲之韻母。原本音節單音語料庫是設計了一個載字句將每一音節放在句中特定的位置，由陳碧娥老師以正常的速度朗讀，再以人工處理先將基本音節波形切出，但許多包含鼻音的音節會與前後音節產生耦合現象，使欲擷取的音節波形在啟始或結束位置不明確，之後也從語料庫中擷取音節波形，但仍有大部分的音節無法找到合適的。

最後，決定以單音錄製音節波形，以下是如何挑選出適合語音合成的音節波形判斷方法：

1. 初步由人耳判斷：
 - (1) 音節聲音清楚，無雜訊
 - (2) 音節波形，無受前後音節影響
 - (3) 說話速度一致
2. 從韻律訊息判斷：
 - (1) Duration
 - a. 音節波形長度不宜過長或過短，通常約在 0.15 - 0.4 秒以內。
 - (2) Energy
 - a. 音量不宜過大
 - b. 子音能量不宜過大
 - c. Energy contour 在音節波形起始段平滑上身；在結束段平滑下降
 - d. 子音不被雜訊影響
 - (3) Pitch

- a. 週期特性明顯
 - b. Pitch Mean 值
3. 從音節所在位置判斷
- a. 句首、句中、句尾
 - b. 單音

統計 671 音節平均長度約在 0.36 秒，如附圖 5-5，音節為 man，虛線為 pitch mark，綠色為子母音交界處。

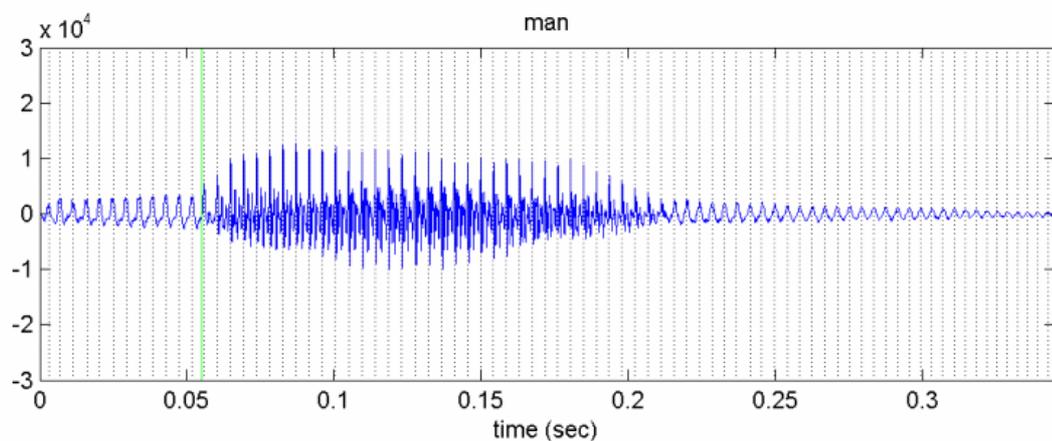


圖 5-5 基本音節“man”切割及 pitch marks 標示結果

另外以 p.t.k 為音節音節韻尾，因其週期特性有轉換，所以不加以標記，如圖 5-6，音節為 tiet。

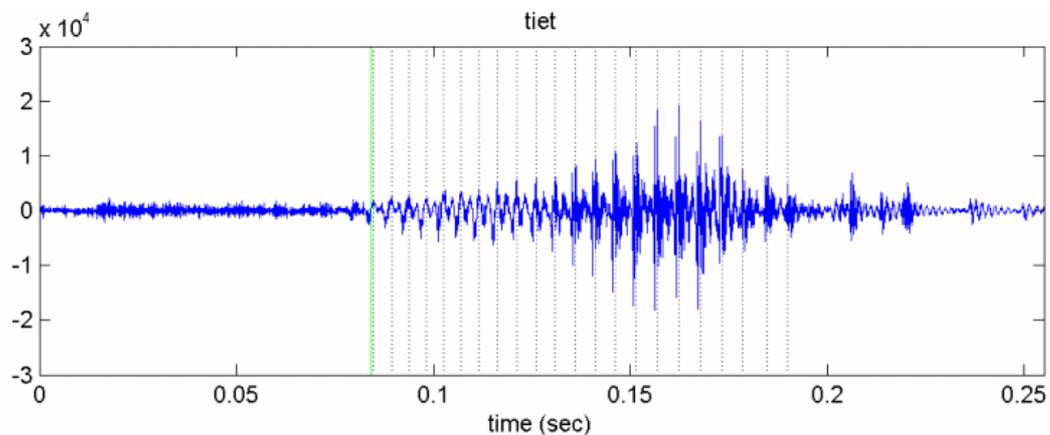


圖 5-6 基本音節“tiet”切割及 pitch marks 標示結果

第六章 結論與未來展望

本論文中，我們實現了視窗版的四縣客語話文字轉語音系統。整個實作過程中，我們得到以下的結論：

1. 合成系統中韻律訊息產生器所需的訓練語料庫，其切割位置經由人工調整會比原始從 HTK 求得的位置更加準確，加上校正基頻軌跡不連續情況，可以得到較好的 RMSE。
2. 合成語料庫，其語料文稿應以短文為主，從實驗結果可得知，訓練韻律訊息時，不包含諺語、俚語的語料，會得到較好的 RMSE。從這邊也可以得知語料音節數目到了一定的量時，其與改善 RMSE 並沒有直接對應的關係。
3. 合成單元波形表，音節的波形挑選，其波形的啟始與結尾需平滑的上升與下降，使合成出來的聲音比較悅耳。

同時，我們覺得有些地方仍需要的改進：

1. 擴充四縣客家話語料庫，目前在語料庫中，671 音節裡仍有 99 個音節尚未出現，而只出現 5 次以內的音節有 215 個，充足的語料才將能訓練良好的韻律模型。
2. 客語詞典在 TTS 系統佔非常重要的角色，目前客語詞典共有 36259 個詞，與國語詞典 121341 個詞仍明顯不足，尤其在一字詞，客語有 7326 個一字詞，國語有 13110 個一字詞。。
3. 合成單元波形表，目前是以單音錄音的方式，往前國語與台語將之從語料庫擷取可得到合適的音節波形，未來可從這方面著手。
4. 目前客語合成語音，並未對入聲韻偉(p. t. k)處理，未來合成音節波形，可需之獨立處立。

參考文獻

- [1] 鍾榮富，2004，〈〈 台灣客家語音導論 〉〉，五南圖書出版股份有限公司。
- [2] 龔萬灶，2003，〈〈 客話實用手冊 〉〉，國家圖書館出版。
- [3] S. H. Hwang, S. H. Chen, and Y. R. Wang, "A Mandarin Text-to-Speech system," in Proc. ICSLP-96, pp.1421-1424, Oct. 1996.
- [4] L. S. Lee, C. Y. Tseng, and M. Ouh-Young, "The synthesis rules in Chinese text-to-speech system" IEEE Trans. Acoust, Speech, Signal Processing, vol. 37, n0. 9, p1309-1319, Sep. 1989
- [5] L. S. Lee, C. Y. Tseng, and C. J. Hesih, " Improved tone concatenation rules in a formant-based Chinese text-to-speech system," IEEE Trans. Speech and Audio Processing, Vol. 1, No. 3, pp. 287-294, July 1993.
- [6] S. H. Chen, S. G. Chang, and S. M. Lee, " A statistical model based fundamental frequency synthesizer for Mandarin speech," J. Acoust. Soc. Am. , 92(1), pp. 114-120, July 1992
- [7] S. H. Chen, and S. H. Hwang, and Y. R. Wang, " An RNN-based prosodic information synthesizer for Mandrain text-to-speech," IEEE Trans. Speech and Audio Processing, vol. 6, no. 3, pp. 226-239, May 1998.
- [8] 黃紹華，" 中文文句翻語音系統中韻律訊息產生器之研究"，國立交通大學博士論文，民國八十五年六月。

附錄一

四縣腔客家話 671 個音節、17個聲母、71個韻母對照表

音碼	音節	聲母	韻母
1	bi	b	i
2	be	b	e
3	ba	b	a
4	bia	b	ia
5	bo	b	o
6	bu	b	u
7	biu	b	iu
8	bai	b	ai
9	boi	b	oi
10	beu	b	eu
11	bau	b	au
12	bin	b	in
13	ben	b	en
14	bien	b	ien
15	ban	b	an
16	bun	b	un
17	bang	b	ang
18	biang	b	iang
19	bong	b	ong
20	biong	b	iong
21	bung	b	ung
22	bit	b	it
23	bet	b	et
24	biet	b	iet
25	bat	b	at
26	bot	b	ot
27	but	b	ut
28	bak	b	ak
29	biak	b	iak
30	bok	b	ok
31	buk	b	uk
32	pi	p	i
33	pa	p	a
34	pia	p	ia
35	po	p	o
36	pu	p	u
37	piu	p	iu
38	pai	p	ai
39	poi	p	oi
40	peu	p	eu
41	pau	p	au

音碼	音節	聲母	韻母
42	pin	p	in
43	pen	p	en
44	pien	p	ien
45	pan	p	an
46	pon	p	on
47	pun	p	un
48	pang	p	ang
49	piang	p	iang
50	pong	p	ong
51	piong	p	iong
52	pung	p	ung
53	pit	p	it
54	pet	p	et
55	piet	p	iet
56	pat	p	at
57	put	p	ut
58	pak	p	ak
59	piak	p	iak
60	pok	p	ok
61	piok	p	iok
62	puk	p	uk
63	mi	m	i
64	me	m	e
65	ma	m	a
66	mia	m	ia
67	mo	m	o
68	mu	m	u
69	mai	m	ai
70	moi	m	oi
71	meu	m	eu
72	mau	m	au
73	miau	m	iau
74	min	m	in
75	men	m	en
76	mien	m	ien
77	man	m	an
78	mun	m	un
79	mang	m	ang
80	miang	m	iang
81	mong	m	ong
82	miong	m	iong

音碼	音節	聲母	韻母
83	mung	m	ung
84	mit	m	it
85	met	m	et
86	miet	m	iet
87	mat	m	at
88	mut	m	ut
89	mak	m	ak
90	mok	m	ok
91	muk	m	uk
92	fi	f	i
93	fe	f	e
94	fa	f	a
95	fo	f	o
96	fu	f	u
97	fai	f	ai
98	foi	f	oi
99	feu	f	eu
100	fam	f	am
101	fin	f	in
102	fen	f	en
103	fan	f	an
104	fon	f	on
105	fun	f	un
106	fang	f	ang
107	fong	f	ong
108	fung	f	ung
109	fap	f	ap
110	fit	f	it
111	fet	f	et
112	fat	f	at
113	fok	f	ok
114	fut	f	ut
115	fuk	f	uk
116	vi	v	i
117	ve	v	e
118	va	v	a
119	vo	v	o
120	vu	v	u
121	vai	v	ai
122	voi	v	oi
123	van	v	an

音碼	音節	聲母	韻母
124	von	v	on
125	vun	v	un
126	vang	v	ang
127	vong	v	ong
128	vung	v	ung
129	vet	v	et
130	vat	v	at
131	vut	v	ut
132	vak	v	ak
133	vok	v	ok
134	vuk	v	uk
135	di	d	i
136	de	d	e
137	da	d	a
138	dia	d	ia
139	do	d	o
140	dio	d	io
141	du	d	u
142	diu	d	iu
143	dai	d	ai
144	doi	d	oi
145	dui	d	ui
146	deu	d	eu
147	dau	d	au
148	diau	d	iau
149	dem	d	em
150	dam	d	am
151	diam	d	iam
152	din	d	in
153	den	d	en
154	dien	d	ien
155	dan	d	an
156	don	d	on
157	dun	d	un
158	dang	d	ang
159	dong	d	ong
160	diong	d	iong
161	dung	d	ung
162	dep	d	ep
163	dap	d	ap
164	diap	d	iap

音碼	音節	聲母	韻母
165	dit	d	it
166	det	d	et
167	diet	d	iet
168	dat	d	at
169	dot	d	ot
170	dut	d	ut
171	dak	d	ak
172	diak	d	iak
173	dok	d	ok
174	diok	d	iok
175	duk	d	uk
176	ti	t	i
177	te	t	e
178	ta	t	a
179	to	t	o
180	tu	t	u
181	tiu	t	iu
182	tai	t	ai
183	toi	t	oi
184	tui	t	ui
185	teu	t	eu
186	tiau	t	iau
187	tiem	t	iem
188	tam	t	am
189	tiam	t	iam
190	tin	t	in
191	ten	t	en
192	tien	t	ien
193	tan	t	an
194	ton	t	on
195	tun	t	un
196	tang	t	ang
197	tong	t	ong
198	tiong	t	iong
199	tung	t	ung
200	tep	t	ep
201	tap	t	ap
202	tiap	t	iap
203	tit	t	it
204	tet	t	et
205	tiet	t	iet

音碼	音節	聲母	韻母
206	tat	t	at
207	tot	t	ot
208	tut	t	ut
209	tak	t	ak
210	tok	t	ok
211	tuk	t	uk
212	ni	n	i
213	ne	n	e
214	na	n	a
215	no	n	o
216	nu	n	u
217	niu	n	iu
218	nai	n	ai
219	nui	n	ui
220	neu	n	eu
221	nau	n	au
222	niau	n	iau
223	nem	n	em
224	nam	n	am
225	nin	n	in
226	nen	n	en
227	nan	n	an
228	non	n	on
229	nun	n	un
230	nang	n	ang
231	nong	n	ong
232	nung	n	ung
233	nap	n	ap
234	nit	n	it
235	net	n	et
236	nat	n	at
237	not	n	ot
238	nak	n	ak
239	nuk	n	uk
240	li	l	i
241	le	l	e
242	la	l	a
243	lo	l	o
244	lio	l	io
245	lu	l	u
246	liu	l	iu

音碼	音節	聲母	韻母
247	lai	l	ai
248	loi	l	oi
249	lui	l	ui
250	leu	l	eu
251	lau	l	au
252	liau	l	iau
253	lim	l	im
254	lem	l	em
255	lam	l	am
256	liam	l	iam
257	lin	l	in
258	lien	l	ien
259	lan	l	an
260	lon	l	on
261	lion	l	ion
262	lun	l	un
263	lang	l	ang
264	liang	l	iang
265	long	l	ong
266	liong	l	iong
267	lung	l	ung
268	liung	l	iung
269	lip	l	ip
270	lep	l	ep
271	lap	l	ap
272	liap	l	iap
273	lit	l	it
274	let	l	et
275	liet	l	iet
276	lat	l	at
277	lot	l	ot
278	lut	l	ut
279	lak	l	ak
280	liak	l	iak
281	lok	l	ok
282	liok	l	iok
283	luk	l	uk
284	liuk	l	iuk
285	zii	z	ii
286	zi	z	i
287	ze	z	e

音碼	音節	聲母	韻母
288	za	z	a
289	zia	z	ia
290	zo	z	o
291	zio	z	io
292	zu	z	u
293	ziu	z	iu
294	zai	z	ai
295	zoi	z	oi
296	zui	z	ui
297	zeu	z	eu
298	zau	z	au
299	ziau	z	iau
300	ziim	z	iim
301	zim	z	im
302	zem	z	em
303	zam	z	am
304	ziam	z	iam
305	ziin	z	iin
306	zin	z	in
307	zen	z	en
308	zien	z	ien
309	zan	z	an
310	zon	z	on
311	zun	z	un
312	zang	z	ang
313	ziang	z	iang
314	zong	z	ong
315	ziong	z	iong
316	zung	z	ung
317	ziung	z	iung
318	ziip	z	iip
319	zip	z	ip
320	zep	z	ep
321	zap	z	ap
322	ziap	z	iap
323	ziit	z	iit
324	zit	z	it
325	zet	z	et
326	ziet	z	iet
327	zat	z	at
328	zot	z	ot

音碼	音節	聲母	韻母
329	zut	z	ut
330	zak	z	ak
331	ziak	z	iak
332	zok	z	ok
333	ziok	z	iok
334	zuk	z	uk
335	ziuk	z	iuk
336	cii	c	ii
337	ci	c	i
338	ce	c	e
339	ca	c	a
340	cia	c	ia
341	co	c	o
342	cio	c	io
343	cu	c	u
344	ciu	c	iu
345	cai	c	ai
346	coi	c	oi
347	cui	c	ui
348	ceu	c	eu
349	cau	c	au
350	ciau	c	iau
351	ciim	c	iim
352	cim	c	im
353	cam	c	am
354	ciam	c	iam
355	ciin	c	iin
356	cin	c	in
357	cen	c	en
358	cien	c	ien
359	can	c	an
360	con	c	on
361	cion	c	iong
362	cun	c	un
363	cang	c	ang
364	ciang	c	iang
365	cong	c	ong
366	ciong	c	iong
367	cung	c	ung
368	ciung	c	iung
369	cip	c	ip

音碼	音節	聲母	韻母
370	cap	c	ap
371	ciap	c	iap
372	ciit	c	iit
373	cit	c	it
374	cet	c	et
375	ciet	c	iet
376	cat	c	at
377	cot	c	ot
378	cut	c	ut
379	cak	c	ak
380	ciak	c	iak
381	cok	c	ok
382	ciok	c	iok
383	cuk	c	uk
384	ciuk	c	iuk
385	sii	s	ii
386	si	s	i
387	se	s	e
388	sa	s	a
389	sia	s	ia
390	so	s	o
391	sio	s	io
392	su	s	u
393	siu	s	iu
394	sai	s	ai
395	soi	s	oi
396	sui	s	ui
397	seu	s	eu
398	sau	s	au
399	siau	s	iau
400	siim	s	iim
401	sim	s	im
402	sem	s	em
403	sam	s	am
404	siam	s	iam
405	siin	s	iin
406	sin	s	in
407	sen	s	en
408	sien	s	ien
409	san	s	an
410	son	s	on

音碼	音節	聲母	韻母
411	sun	s	un
412	sang	s	ang
413	siang	s	iang
414	song	s	ong
415	siong	s	iong
416	sung	s	ung
417	siung	s	iung
418	siip	s	iip
419	sip	s	ip
420	sep	s	ep
421	sap	s	ap
422	siap	s	iap
423	siit	s	iit
424	sit	s	it
425	set	s	et
426	siet	s	iet
427	sat	s	at
428	sot	s	ot
429	sut	s	ut
430	sak	s	ak
431	siak	s	iak
432	sok	s	ok
433	siok	s	iok
434	suk	s	uk
435	siuk	s	iuk
436	gi	g	i
437	ge	g	e
438	gie	g	ie
439	ga	g	a
440	gia	g	ia
441	gua	g	ua
442	go	g	o
443	gio	g	io
444	gu	g	u
445	giu	g	iu
446	guai	g	uai
447	goi	g	oi
448	gui	g	ui
449	gau	g	au
450	gieu	g	ieiu
451	giu	g	iu

音碼	音節	聲母	韻母
452	gim	g	im
453	giem	g	iem
454	gam	g	am
455	giam	g	iam
456	gin	g	in
457	gen	g	en
458	gien	g	ien
459	guen	g	uen
460	gan	g	an
461	guan	g	uan
462	gon	g	on
463	gun	g	un
464	giun	g	iun
465	gang	g	ang
466	giang	g	iang
467	gong	g	ong
468	giong	g	iong
469	gung	g	ung
470	giung	g	iung
471	guang	g	uang
472	gip	g	ip
473	giep	g	iep
474	gap	g	ap
475	giap	g	iap
476	git	g	it
477	giet	g	iet
478	guet	g	uet
479	gat	g	at
480	guat	g	uat
481	got	g	ot
482	gut	g	ut
483	gak	g	ak
484	giak	g	iak
485	gok	g	ok
486	giok	g	iok
487	guk	g	uk
488	giuk	g	iuk
489	guak	g	uak
490	ki	k	i
491	kie	k	ie
492	kue	k	ue

音碼	音節	聲母	韻母
493	ka	k	a
494	kia	k	ia
495	kua	k	ua
496	ko	k	o
497	kio	k	io
498	ku	k	u
499	kiu	k	iu
500	kai	k	ai
501	kuai	k	uai
502	koi	k	oi
503	kioi	k	ioi
504	kui	k	ui
505	keu	k	eu
506	kau	k	au
507	kieu	k	ieu
508	kiau	k	iau
509	kim	k	im
510	kam	k	am
511	kiam	k	iam
512	kiem	k	iem
513	kin	k	in
514	ken	k	en
515	kien	k	ien
516	kan	k	an
517	kuan	k	uan
518	kon	k	on
519	kun	k	un
520	kiun	k	iun
521	kang	k	ang
522	kiang	k	iang
523	kong	k	ong
524	kiong	k	iong
525	kung	k	ung
526	kiung	k	iung
527	kip	k	ip
528	kep	k	ep
529	kap	k	ap
530	kiap	k	iap
531	kit	k	it
532	kiet	k	iet
533	kat	k	at

音碼	音節	聲母	韻母
534	kut	k	ut
535	kiut	k	iut
536	kak	k	ak
537	kiak	k	iak
538	kok	k	ok
539	kiok	k	iok
540	kuk	k	uk
541	kiuk	k	iuk
542	kuak	k	uak
543	ngi	ng	i
544	ngie	ng	ie
545	nga	ng	a
546	ngia	ng	ia
547	ngo	ng	o
548	ngio	ng	io
549	ngu	ng	u
550	ngiu	ng	iu
551	ngai	ng	ai
552	nguai	ng	uai
553	ngoi	ng	oi
554	ngui	ng	ui
555	ngau	ng	au
556	ngieu	ng	ieu
557	ngiau	ng	iau
558	ngim	ng	im
559	ngiem	ng	iem
560	ngam	ng	am
561	ngiam	ng	iam
562	ngin	ng	in
563	ngien	ng	ien
564	ngan	ng	an
565	nguan	ng	uan
566	ngion	ng	ion
567	ngiun	ng	iun
568	ngang	ng	ang
569	ngiang	ng	iang
570	ngong	ng	ong
571	ngiong	ng	iong
572	ngiung	ng	iung
573	ngip	ng	ip
574	ngap	ng	ap

音碼	音節	聲母	韻母
575	ngiap	ng	iap
576	ngit	ng	it
577	ngiet	ng	iet
578	ngat	ng	at
579	ngut	ng	ut
580	ngiak	ng	iak
581	ngok	ng	ok
582	ngiok	ng	iok
583	ngiuk	ng	iuk
584	hi	h	i
585	he	h	e
586	ha	h	a
587	hia	h	ia
588	ho	h	o
589	hio	h	io
590	hiu	h	iu
591	hai	h	ai
592	hoi	h	oi
593	heu	h	eu
594	hau	h	au
595	hieu	h	ieu
596	hiau	h	iau
597	him	h	im
598	hem	h	em
599	ham	h	am
600	hiam	h	iam
601	hin	h	in
602	hen	h	en
603	hien	h	ien
604	han	h	an
605	hon	h	on
606	hiun	h	iun
607	hang	h	ang
608	hong	h	ong
609	hiong	h	iong
610	hiung	h	iung
611	hiing	h	iing
612	hip	h	ip
613	hap	h	ap
614	hiap	h	iap
615	hit	h	it

音碼	音節	聲母	韻母
616	het	h	et
617	hiet	h	iet
618	hat	h	at
619	hot	h	ot
620	hak	h	ak
621	hok	h	ok
622	hiuk	h	iuk
623	i	INULL	i
624	e	INULL	e
625	ie	INULL	ie
626	a	INULL	a
627	ia	INULL	ia
628	o	INULL	o
629	io	INULL	io
630	iu	INULL	iu
631	ai	INULL	ai
632	oi	INULL	oi
633	eu	INULL	eu
634	au	INULL	au
635	ieu	INULL	ieu
636	iau	INULL	iau
637	im	INULL	im
638	em	INULL	em
639	am	INULL	am
640	iam	INULL	iam
641	in	INULL	in
642	en	INULL	en
643	ien	INULL	ien
644	an	INULL	an
645	on	INULL	on
646	iun	INULL	iun
647	ang	INULL	ang
648	iang	INULL	iang
649	ong	INULL	ong
650	iong	INULL	iong
651	iung	INULL	iung
652	ip	INULL	ip
653	ep	INULL	ep
654	ap	INULL	ap
655	iap	INULL	iap
656	it	INULL	it

音碼	音節	聲母	韻母
657	et	INULL	et
658	iet	INULL	iet
659	at	INULL	at
660	ot	INULL	ot
661	iut	INULL	iut
662	ak	INULL	ak
663	iak	INULL	iak
664	ok	INULL	ok
665	iok	INULL	iok
666	iuk	INULL	iuk
667	m	m	FNULL
668	ng	ng	FNULL
669	iui	INULL	iui
670	ngung	ng	ung
671	nguk	ng	uk



附錄二：詞類表

編號	標記	詞類	編號	標記	詞類
1	A	非謂形容詞	24	Nh	代名詞
2	Caa	對等連接詞	25	I	感嘆詞
3	Cab	連接詞，如：等等	26	P	介詞
4	Cba	連接詞，如：的話	27	T	語助詞
5	Cbb	關聯連接詞	28	VA	動作不及物動詞
6	Da	數量副詞	29	VAC	動作使動動詞
7	Dfa	動詞前程度副詞	30	VB	動作類及物動詞
8	Dfb	動詞後程度副詞	31	VC	動作及物動詞
9	Di	時態標記	32	VCL	動作接地方賓語動詞
10	Dk	句副詞	33	VD	雙賓動詞
11	D	副詞	34	VE	動作句賓動詞
12	Na	普通名詞	35	VF	動作謂賓動詞
13	Nb	專有名詞	36	VG	分類動詞
14	Nc	地方詞	37	VH	狀態不及物動詞
15	Ncd	位置詞	38	VHC	狀態使動動詞
16	Nd	時間詞	39	VI	狀態類及物動詞
17	Neu	數詞定詞	40	VJ	狀態及物動詞
18	Nes	特指定詞	41	VK	狀態句賓動詞
19	Nep	指代定詞	42	VL	狀態謂賓動詞
20	Neqa	數量定詞	43	V_2	有
21	Neqb	後置數量定詞	44	DE	的，之，得，地
22	Nf	量詞	45	SHI	是
23	Ng	後置詞	46	FW	外文標記
59	VP	動詞片語			

附錄三：新增量詞

量詞	拼音	例子
盞	ang1	「一盞覆菜」、「一盞酸梅汁」。
臂	bi2	「一臂羊肉」、「一臂豬羊份」。
舫	biong1	「一舫火車」、「畜一舫鴨仔」、「頭舫車幾多點？」。
擺	bai2	「愛講幾多擺，你正聽得識？」、「頭擺人毋識飛行機」。
炊	coi1	「一炊香茅」、「一炊樟腦」。
材	cai5	「這筒樹仔有幾多材？」、「一百材 枋仔」等。
車	ca1	「一車香茅草」、「一車觀光客」等。
碓	doi3	早期把穀子去殼有用石碓的，每碾成一次的量稱為「一碓」。
埕	de3	「一埕豬肉」、「一埕菜頭版」、「一埕枋仔」
斷	don3	「這枝梯牌有十一斷」、「該條石崎仔有三百零斷」。
到	do3	「一擺毋會，再做加一到」、「一儕做一到」。
檣	fong5	「水檣」、「菁檣」、「一檣鹹菜」、「一檣香茅」、「一檣糯穀」等。
窟	fut4	「一窟水」、「一窟塘虱」、「歸窟仔湖鰱仔」
笕	gong5	「一笕糯米兩笕米」。
肛	gong1	「該隻豬 養過三肛子」、「狗仔養一肛子，大體仔有四、五隻」。
槓	gong3	「一槓落峨」、「一槓牛欄」、「一槓店面」
弓	giung1	「這頭弓蕉樹打一大弓弓蕉」。
莢	giap4	「一莢番豆」、「一莢四季豆」、「一莢荷蘭豆」。
挾	giap4	「一挾薤菜」、「一挾豆腐皮」。
梗	guang1	稻田裡的秧苗，每列五叢為「一梗」。菜圃的菜苗一列也稱「一梗」。算盤珠子，每一行也叫「一梗」。
穴	hiet8	「一穴風水」、「一穴家塚」。
批	ia2	「一批沙仔」、「一批粃仔」、「一批糙米」、「一批塵灰」。
尋	kim5	「這枝竹篙有兩尋長」。
擲	kuan3	「一擲粽仔」、「一擲豬肉」。
袴	kua2	「一袴樹袴」、「一翕牛眼」、「一袴牛眼」、「一袴人」。
袴	pa2	「一袴樹袴」、「一翕牛眼」、「一袴牛眼」、「一袴人」。
綦	kien3	「一綦魚丸」、「一綦綦孔錢」。
坵	kiu1	「一坵足水田」、「幾坵仔旱田」、「一坵菜園」等。
孔	kung2	「一孔泥胚」、「若衫項有五、六孔」、「腳底爛兩、三孔」。
坎	kam2	「一坎店仔」、「一坎布行」、「一坎雜貨店」、「一坎超級市場」。
料	liau3	「一料紙條仔」、「一料豬肉」。
領	liang1	「一領襯衫」、「一領被單」、「一領水衣」。
蕊	lui1	「一蕊牡丹花」、「一蕊大紅花」。
卵	lon2	「一卵薑」、「一卵嫩薑」。
隸	li3	「豆隸」、「你種幾多隸四季豆？」。

廉	liam3	「一廉西瓜」。
齧	ngat4	一個刻度稱之，如用「秤」時，秤桿上的刻度，用齧來算。
棚	pang5	「一棚苦瓜」、「一棚菜瓜」、「雙棚絞」、「稗棚」、「一棚禾稗」。
坏	pu5	「一坏牛屎」、「一坏蟻公泥」。
胚	poi1	「一胚泥胚」。
禽	poi1	「一禽牛眼」、「一禽荔枝」。
皮	pi5	「一皮樹葉仔」、「一皮菜蓋」、「一皮參仔」。
餅	piang1	「一餅稗餅」、「一餅豬油」、「一餅雨棚」。
儕	sa5	「一儕人」、「請五、六儕仔來 手」、「去三儕就罉哩」。
析	sak4	「一粒西瓜破四析」、「一析竹析」、「一析黃梨」、「一析柑仔」。
托	tok4	「姐婆買一托弓蕉來」。
渡	tu3	「一渡橋」、「這條索仔有幾多渡結頭？」
甕	vung3	「一甕高粱酒」、「一甕蘿蔔乾」、「一甕米醬」。
鑊	vok8	「一鑊糖」、「一鑊長年菜」、「一鑊藥草」。
廂	siong1	「一廂菜地」、「一廂秧地」。
相	siong2	「一枝竹仔有幾多相？」、「其腳相已長」
仙	sien1	「請一仙神明來服祀」、「這座廟項有幾多仙神明」。
抓	zal	「一抓薤菜」、「一抓豆仔」、「一抓香心骨」、「一抓筷仔」。
轉	zon2	「行仔五、六轉，就還無見著佢」、「一日就走三、四轉」等。

資料來源：龔萬灶，2003，〈〈客話實用手冊〉〉

