

國立交通大學

電信工程研究所

碩士論文



基於隱藏式馬可夫模型之客語文句轉語音系統

An HMM-based Hakka Text-to-Speech System

研究生：蔡依玲

指導教授：王逸如 博士

中華民國九十九年七月

基於隱藏式馬可夫模型之客語文句轉語音系統

An HMM-based Hakka Text-to-Speech System

研究生：蔡依玲

Student: Yi-Ling Tsai

指導教授：王逸如 博士

Advisor: Dr. Yih-Ru Wang

國立交通大學

電信工程研究所

碩士論文



Submitted to Institute of Communication Engineering

College of Electrical and Computer Engineering

National Chiao Tung University

in Partial Fulfillment of the Requirements

for the Degree of

Master

In

Communication Engineering

July 2010

Hsinchu, Taiwan, Republic of China

中華民國九十九年七月

基於隱藏式馬可夫模型之客語文句轉語音系統

An HMM-based Hakka Text-to-Speech System

研 究 生：蔡依玲

指 導 教 授：王逸如 博士

國立交通大學電信工程研究所



本論文完成一套客語文句轉語音系統。它由斷詞系統、停頓預估系統、文脈分析器以及基於隱藏式馬可夫模型之語音合成器所組成。文句輸入經由斷詞系統做分詞和詞性標記，由於客語語料的不足，本論文在此結合中文斷詞條件隨機域模型、客語外掛詞典、中文內部詞典及客語構詞規則形成客語斷詞系統。斷詞系統輸出斷詞資訊後，由停頓預估系統預估詞後是否停頓，再由文脈分析器輸出和語音合成器相對應之合成單元及語言參數，依據文字標記檔及合成器訓練端所得到的模型進行音長、音高、頻譜參數的預測，而得到合成語音輸出。最後我們分別設計實驗，對斷詞系統、停頓預估系統、合成器以及客語 TTS 系統做出效能評估，而客語 TTS 系統在主觀評分上得到了不錯的分數，顯示出它是一套不錯的系統。

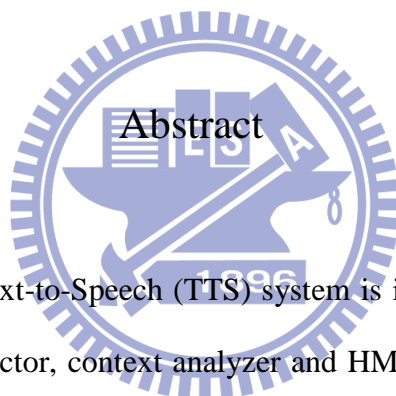
An HMM-based Hakka Text-to-Speech System

Student : Yi-Ling Tsai

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: parser, pause predictor, context analyzer and HMM-based synthesizer. The input text is first tagged in the text analyzer into word sequence. Due to the lack of a large text corpus to train a robust Hakka parser, we adopt a new approach to constructing a Hakka parser via extending an existing CRF-based Chinese parser to add a Hakka dictionary and incorporate some Hakka word construction rules. Then, the pause predictor estimates the inter-syllable locations to insert pauses. The context analyzer then generates the synthesis unit and some language parameters. Lastly, the HMM-based synthesizer produces duration, pitch, and spectral parameters to generate the output synthesized speech. Some experiments are also designed to evaluate the performances of the parser and the pause predictor, as well as the quality of the synthesized speech. A good MOS score obtained in the subjective quality test confirms that the Hakka TTS system is a promising one.

致謝

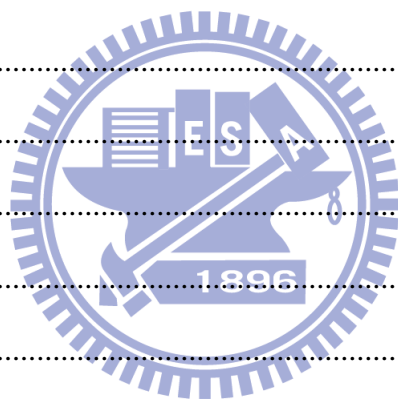
耶耶耶，我畢業了!!!!感謝陳信宏老師和王逸如老師，感謝陳信宏老師明明很忙卻還是會常常來實驗室巡邏督促我們的進度，明明很忙每個禮拜都還是會抽空咪聽；感謝王逸如老師每次都讓我跑進去辦公室問一大堆問題，而且總是很溫和的告訴我答案解救我；感謝你也太厲害了的小小老闆性獸，我只能說，沒有你哪有我們，真的!!!感謝吃素很不嚴謹的楊智合，每次有很幼稚的問題也都讓我會去一直煩你；感謝也太不把我當女生看的阿德，我以後一定隨手帶個 20 妹讓你挑；感謝常常跟我分享自己風流故事的輝哥，以後聽不到故事了怎麼辦，而且我好像還欠你好幾頓大餐；感謝愛護地球的希群，應該有好幾棵樹是因為你而沒被砍掉的吧；感謝很帥氣漂泊的巴金叔叔，工作真的很難找啦哪有你說得那麼簡單；感謝有一大堆人妻和學妹 MSN 的嘍馬，我記得 C 到底怎麼寫得好像是你教的，可是 20 妹你應該不需要所以我沒有可以回報你的了；感謝我不管多晚來都還能遇到的承燁，你再撐一下就要出頭天了不用擔心；感謝耳背很嚴重傻傻的舒舒，我會先去瘋狂面試再告訴你哪間比較好的；感謝同是客語人的財祿，我們終於擺脫客語了耶!!!感謝修課時常常借我抄作業的浩翔，你口試一定順啦；感謝比 google 還好用的宥余，我好像真的轉過去問問題的次數比估的次數多耶，以後去 htc 要幫我留隻手機；感謝學長們普屋、小宋、QQ、小帥哥，我好想念你們喔；感謝碩一學弟人妻殺手胖胖、豆腐喵、銘傑、大胖、智障、小蝦、啟全；感謝這兩年給我工作的廠商們，才能讓不愁吃穿；感謝我的好朋友們，我朋友真的很少但我還是懶的列舉了；感謝我的爸爸媽媽弟弟，沒有家人的支持就沒有今天的我呀；最後要感謝我的阿媽，阿媽，對不起我不能常常回家看你，可是我畢業了耶!!!

目錄

中文摘要	I
致謝	III
目錄	IV
圖目錄	VII
表目錄	VIII
第一章 緒論	1
1.1 研究動機	1
1.2 客語文字轉語音系統概述	1
1.3 章節概要說明	2
第二章 四縣客語語料庫介紹	4
2.1 四縣客家話之簡介	4
2.1.1 四縣客家話之聲母	5
2.1.2 四縣客家話之韻母	5
2.2 四縣客家話之聲調	6
2.3 四縣客家話之音節	7
2.4 語料庫簡介	8
2.4.1 客語詞典	9
第三章 斷詞系統	10
3.1 斷詞系統概述	11
3.2 中文分詞與詞性標記	12
3.2.1 中文條件隨機域模型	13
3.2.2 中文分詞	14

3.2.3 中文詞性標記	16
3.3 客語分詞與詞性標記	17
3.3.1 客語詞典	17
3.3.2 構詞規則	18
3.3.3 客語借字問題	19
第四章 HMM-based 語音合成器	21
4.1 HMM-based 語音合成系統	21
4.2 考慮動態特性的參數生成演算法	23
4.2.1 動態參數	23
4.2.2 參數生成演算法	25
4.3 切割位置	27
4.4 文本相關資訊	28
4.5 狀態合併決策樹問題集	29
4.6 停頓預估	30
4.6.1 停頓預估模型	30
第五章 實驗結果	32
5.1 斷詞系統之效能分析	32
5.1.1 中文分詞和詞性標記系統之效能分析	32
5.1.1.1 訓練、測試語料及中文詞典	32
5.1.1.2 評估方法及實驗結果	33
5.1.2 客語分詞和詞性標記	36
5.1.2.1 測試語料	37
5.1.2.2 實驗方法	37
5.1.2.3 實驗結果分析	38
5.2 停頓預估模型之效能分析	41

5.2.1 訓練及測試語料	42
5.2.2 實驗設計	42
5.2.3 實驗結果分析	42
5.3 客語語音合成實驗結果與分析	44
5.3.1 實驗與測試語料	44
5.3.2 客觀評估	44
5.4 文字轉語音系統效能分析	45
5.4.1 實驗方法	45
5.4.2 主觀評分	46
第六章 結論與未來展望	48
6.1 結論	48
6.2 未來展望	48
文獻回顧	49
附錄一	50
附錄二	58
附錄三	59



圖目錄

圖 1-1 客語文字轉語音系統.....	2
圖 2-1 客語音節結構.....	4
圖 3-1 斷詞系統.....	11
圖 4-1 基於隱藏式馬可夫模型之語音合成系統架構.....	22
圖 4-2 一階及二階動態參數關係圖.....	24
圖 4-3 轉換矩陣表示示意圖.....	25
圖 4-4 參數生成示意圖.....	26
圖 5-1 實驗之 MOS 比較.....	47



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 四縣客家話聲調表.....	7
表 2-4 四縣客家話音節類型表.....	8
表 2-5 客語詞典統計表.....	9
表 3-1 中文分詞之條件隨機域訓練特徵.....	15
表 3-2 中文分詞之條件隨機域訓練特徵樣版.....	15
表 3-3 中文詞性標記之條件隨機域訓練特徵樣版.....	16
表 3-4 中文詞性標記之條件隨機域訓練特徵樣版.....	16
表 3-5 客語外掛詞典統計表.....	17
表 3-6 國客語共用構詞規則.....	19
表 3-7 國客語平行構詞規則.....	19
表 4-1 文脈相關資訊格式.....	28
表 4-2 停頓預估模型之特徵.....	31
表 4-3 停頓預估模型之特徵樣版.....	31
表 5-1 中文內部詞典統計表.....	33
表 5-2 中文分詞實驗結果.....	34
表 5-3 中文詞性標記實驗結果.....	36
表 5-4 客語分詞實驗結果.....	38
表 5-5 客語詞性標記實驗結果.....	39
表 5-6 客語詞性標記實驗結果.....	40
表 5-7 實驗正確率.....	43

表 5-8 實驗之四種機率值.....	43
表 5-9 客觀評估.....	44
表 5-10 音長之 RMSE 值.....	45
表 5-11 音高之 RMSE 值.....	45



第一章 緒論

1.1 研究動機

語音合成的技術主要有兩大類，一為基於語料庫之合成法(corpus-based)，近期的單元選取合成系統(unit selection synthesis system)所合成出的語音品質已相當穩定，但在輸出的語者特性或情緒上難以調適，並且須保存大量語料庫；第二種是基於數據處理之合成法(data-driven)，不同於直接從語料庫挑選語音片段單元合成，在訓練完模型參數後，只需保留訓練得到的模型，在可攜性(Potable)及適應性(Flexibility)上有很大的進步。近年來，基於隱藏式馬可夫模型的語音合成系統(HMM-based Speech Synthesis System, HTS)是基於數據處理之合成法中最廣為應用的合成技術，其應用上的彈性大，能合成出較自然的語音。

客家話在台灣是屬於次語言，隨著世代交替，不會說客家話的「隱性客家人」也變得越來越多。國語和閩南語的文字轉語音系統(Text-to-Speech System, TTS System)已經有一定程度的發展，但在客語方面的相關研究卻是少之又少，本研究將使用基於隱藏式馬可夫模型語音合成系統完成客語文字轉語音之系統實作。

1.2 客語文字轉語音系統概述

語音合成涵蓋的領域非常廣，是語音訊號處理領域的一項前瞻技術，包括聲學、語言學、數位訊號處理、多媒體訊號處理等多項技術。TTS 系統的輸入是文句，而輸出是語音訊號，也就是將只能用眼睛看的文字資訊，也可以讓耳朵來聽。本論文的客語 TTS 系統基本架構如圖 1-1 所示：

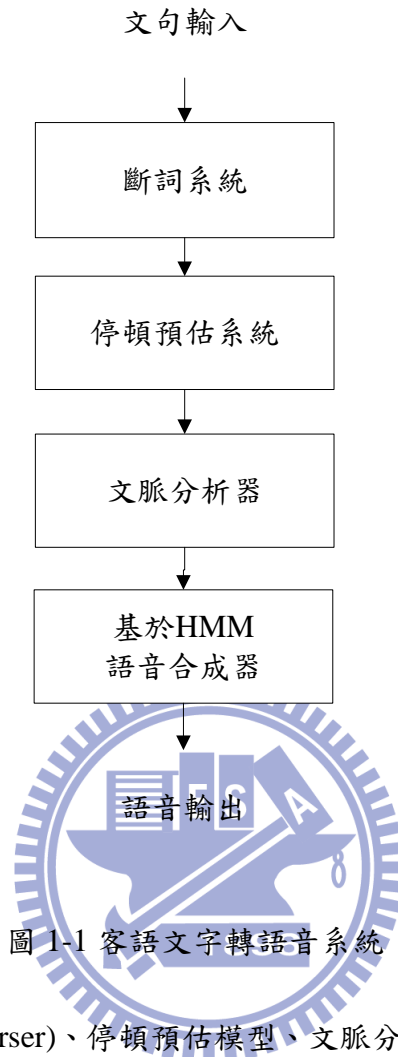


圖 1-1 客語文字轉語音系統

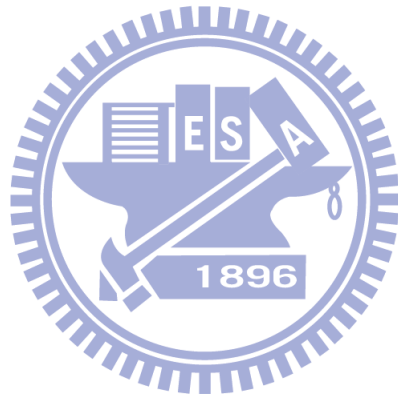
本系統包含斷詞系統(Parser)、停頓預估模型、文脈分析器以及基於 HMM 之語音合成器；文句輸入經由斷詞系統做分詞(Segmentation)及詞性標記(Part of Speech Tagging)，接著由停頓預估模型預估詞後是否有停頓，文脈分析器將前一階段的輸出轉成合成單元序列及其對應語言參數，接著 HTS 找到相對應之頻譜、音長和音高的 HMM 模型，再還原成語音訊號輸出。

1.3 章節概要說明

本論文的內容共分為五章：

第一章 緒論：介紹本論文之研究動機與研究方向。

- 第二章 四縣客語語料庫介紹：介紹客語音節結構、音韻以及本論文之語料庫
- 第三章 斷詞系統：介紹本論文的客語斷詞系統
- 第四章 基於 HMM 之語音合成系統：介紹基於 HMM 之語音合成系統原理、文脈相關資訊以及決策數問題集
- 第五章 停頓預估：介紹停頓預估模型之訓練及實驗結果
- 第六章 實驗結果：評估本論文之斷詞系統效能以及基於 HMM 之語音合成系統效能
- 第七章 結論與未來展望：對本篇論文所提出的方法下結論，並說明未來的改進方法



第二章 四縣客語語料庫介紹

在進行客語語音合成 HMM 模型的訓練之前，我們需要準備一套客語語料庫，而合成單元的選取更直接影響了合成語音的品質，本章將介紹四縣腔客家話的音節結構以及聲調，並且說明本論文語料庫來源以及其錄製方法。

2.1 四縣客家話之簡介

台灣的客家話有四縣客家、海陸客家、饒平客家、紹安客家等，其中四縣客家是最通行的腔調，不同腔調在聲調及詞彙上有不同程度的差異。本論文採用四縣腔調做為研究之語料來源，為了方便起見以下以「四縣客家話」指稱本論文的客語語料腔調。

客語有 671 個音節(syllable)，客家話的音節結構和國語一樣，分為聲母和韻母，韻母又再分為介音及韻腳，韻腳內是主要原音和韻尾，圖 2-1 為客家話音節結構。客語的結構因素（聲調、聲母、介音、主要元音、韻尾）中，聲調和主要元音是不可缺少的。

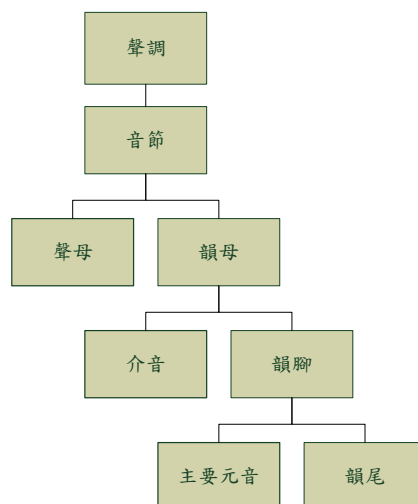


圖 2-1 客語音節結構

2.1.1 四縣客家話之聲母

四縣客家話的聲母是位於音節開頭的輔音，空聲母是指音節開頭沒有輔音，四縣客家話共有 17 個聲母，其中空聲母也包含在內；四縣客家話的聲母只有 m、n、ng、l、v 為發音時聲帶有振動具有週期性的濁音，其餘聲母都屬於清音。我們依照發音部位以及發音方式用表 2-1 將四縣客家話的聲母分類，其中括號內為例字。

表 2-1 四縣客家話聲母表

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)	

2.1.2 四縣客家話之韻母

韻母是音節中位於聲母後面的部分，韻母可以再細分三部分，韻頭、韻腹以及韻尾，其中韻腹是必須要有的，韻頭和韻尾則可有可無。韻頭就是介音，有 i 或 u；韻腹為主要元音，四縣客家話的主要元音有 a、e、i、o、u、ii 共 6 個；韻尾分為元音韻尾和輔音韻尾，元音韻尾包含 i、u 等，輔音韻尾因聲調的不同又分為兩種，一種是以鼻音節尾的舒聲韻，有 m、n、ng，入聲韻則是由塞音節尾，共有 p、t、k。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	代碼	韻母
1	ii	25	em	49	iep
2	i	26	iem	50	ep
3	e	27	am	51	ap
4	ie	28	iam	52	iap
5	ue	29	iin	53	iit
6	a	30	in	54	it
7	ia	31	en	55	et
8	ua	32	ien	56	iet
9	o	33	uen	57	uet
10	io	34	an	58	at
11	u	35	uan	59	uat
12	iu	36	on	60	ot
13	iui	37	ion	61	ut
14	ai	38	un	62	iut
15	uai	39	iun	63	ak
16	oi	40	ang	64	iak
17	ioi	41	iang	65	ok
18	ui	42	ong	66	iok
19	eu	43	iong	67	uk
20	au	44	ung	68	iuk
21	ieu	45	iung	69	FNULL
22	iau	46	uang	70	iing
23	iim	47	iip	71	uak
24	im	48	ip		

2.2 四縣客家話之聲調

漢語當中國語的聲調有四個，閩南語有七個，四縣客家話的聲調和其他次方言相較之下較為簡單，共有六個，其調類、調號和例字如表 2-3 所示：

表 2-3 四縣客家話聲調表

調類	陰平	陰上	陰去	陰入	陽平	陽入
調號	1	2	3	4	5	8
例字	千	錢	淺	賤	絕	切

四縣客家話和中文相同，也有變調的情形，變調規則會跨詞出現，但不會出現在句子與句子之間。其中變調規則共有三種：

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」 mai1 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) + 陰去(調號 3)

例：「針線」 ziim1 sien3 → ziim5 sien3

「拿針線」 na1 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) → 陽平(調號 5) + 陽入(調號 8)

例：「音樂」 im1 ngok8 → im5 ngok8

「聽音樂」 tang1 im1 ngok8 → tang5 im5 ngok8

2.3 四縣客家話之音節

客語文章中紀錄的單元大多是音節，一個中文字代表一個音節。依據聲韻組合規則，四縣客家話可以衍生出 13 種不同的類型，這也是四縣客家以及其它漢語方言所有音節的縮影，表 2-4 列出這 13 種音節結構，由於聲調是每個音節所必備，在下表音節類型裡，聲调用數字表示。而實際上四縣客家話的音韻系統共有 671 個音節，詳細列

表如附錄一。

表 2-4 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) 、iau1 (飪)
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) 、diau3 (釣)

2.4 語料庫簡介

本論文所採用的四縣客家話語料庫有男性語料庫及女性語料庫，語料發音人分別為龔萬灶老師和陳碧娥老師，文章來源主要由龔萬灶老師所主筆的「阿啾箭个故鄉」一書。男性語料庫包含 42 篇文章，語音檔分為 639 個，共有 63158 個音節，音檔為 20kHz 的取樣頻率及 16-bit 之單聲道 pcm 格式。語音檔是由發音人在錄音室依照文稿唸出，平均發音速度為 0.24 second/syllable，錄音軟體為 Adobe Audition 1.0，並使用單一指向性麥克風。女性語料庫中包含 57 篇文章，325 個語音檔，總音節數為 42337 個，由發音人在普通房間依照文稿唸出，音檔取樣頻率 16kHz，16-bit 之單聲道 pcm 格式，平均發音速度為 0.21 second/syllable。

2.4.1 客語詞典

詞典在 TTS 系統中有非常重要的功能，在分詞和標記詞性階段能提供詞和其詞性資訊，在合成語音前的文字分析器當中必須查詢詞典得到對應的音標和聲調。我們主要的收集來源為「完整版中級詞彙彙整檔案」、「台北市客委會-現代客語詞彙彙編」、「教育部客家語常用詞辭典」等；詞典內容包含詞、四縣音標、音節碼、詞性、中文翻譯以及詞來源。目前所收錄的客語詞之數量如表 2-5 所示：

表 2-5 客語詞典統計表

	詞數
一字詞	6747
二字詞	18078
三字詞	5095
四字詞	4217
五字詞	250
六字詞	18980
七字詞	60
八字詞	14
總計	34541

第三章 斷詞系統

斷詞系統包含分詞以及詞性標記兩階段，傳統的中文斷詞系統使用長詞優先規則，如中央研究院的中文斷詞系統；近年來，條件隨機域（Conditional Random Field）方法被提出，並且可達到很好的斷詞結果，因此被有效的使用在文字處理中。

客家話的文字系統大多使用中文字表示而成，而且客語詞的詞性分類與國語相同，共 46 種詞性標籤。因此在客語分詞和詞性標記中，使用和中文相同的方法，由條件隨機域訓練模型是可行的。但是我們已有標記正確分詞和詞性結果的客語語料只有 63158 個單元，這樣的資料量對於訓練條件隨機模型來說是不足的。

我們分析客語文章可以發現大量和國語相同的詞彙，例如一客語句「阿爸歸生人就為著三餐在該打拚」，其對應的中文為「父親一輩子都為了三餐在那打拚」，我們將此例在下表列出，其中以空格表示詞間，本例中共有 8 個詞，而其中國客語完全相同的詞就有 3 個，其餘詞雖然國客無法共用，但卻可以一一對應，而且每個相對應的國客語之詞性標籤也完全相同。

客語：阿爸 歸生人 就 為著 三餐 在該 打拚
國語：父親 一輩子 都 為了 三餐 在那 打拚

由於國客語的詞性結構及用詞相似的特性，我們使用中文條件隨機域模型來對客語進行斷詞；另一方面，我們在客語詞典的收錄已有不錯的結果，而基於條件隨機域的斷詞系統中，可利用增加外掛詞典提高其系統效能，因此本論文結合客語外掛詞典、國語詞典、客語構詞規則以及中文條件隨機域模型組合而成客語斷詞系統。本章將先介紹中文斷詞系統，以及如何加入客語詞典及構詞規則使其成為客語斷詞系統。

3.1 斷詞系統概述

斷詞是文字轉語音系統中第一步驟，此步驟的輸入是一個完整的句子，輸出是已標記好的詞彙序列以及其相對應的詞性標籤。斷詞的好壞會直接影響合成聲音的品質與正確性，是文字轉語音當中很重要的單元。其系統架構如圖 3-1：

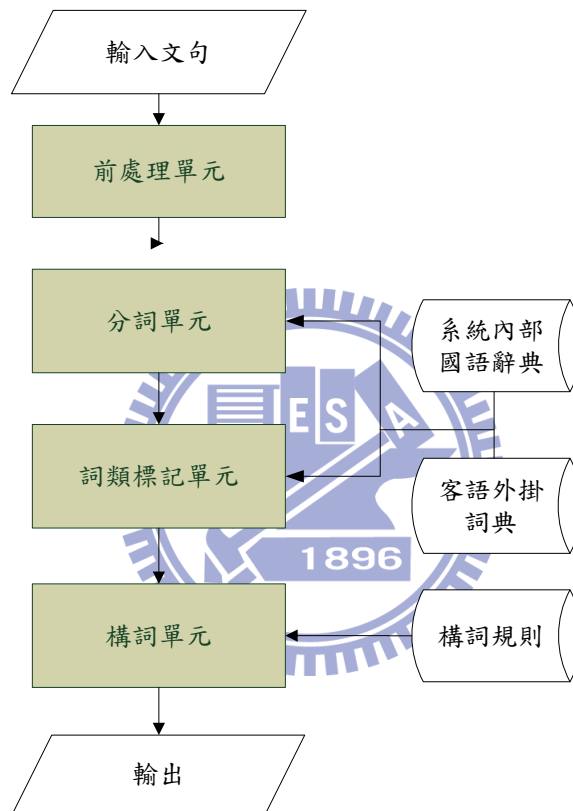


圖 3-1 斷詞系統

其中各個單元之功能在以下做簡單介紹：

(1) 文字正規化單元

輸入的文句當中可能包含 ASCII code 或 Big-5 code，在進入斷詞單元之前，我們將所有的 ASCII code 轉成 BIG5 code 格式，使字串格式統一。例如「西元 1945 年」轉為「西元 1 9 4 5 年」；而中文的用字歧異和標點符號的統一也都在此一單元中解決，

例如「裏」和「晒」轉為「裡」和「曬」；另外，標點符號中如逗號「，」皆統一轉成「,」；此外，客語文章和台語文章一樣，有許多造字情況，目前解決方式是利用 Big-5 code 中的保留區域對應其造字。

(2) 分詞單元

此單元目的是將輸入的文句做出適當分詞，也是整個系統當中最核心的部分。分詞單元是最前面的分詞階段，後來的「詞性標記單元」會利用此單元所產生的候選句標記詞性。

(3) 詞類標記單元

詞類標記單元利用機率統計模型，標記所有可能的詞性組合，並從候選句當中選出最合乎語法的句子。

(4) 構詞單元

四縣客家話中如「結結」(結實的樣子)同字疊字型態，或是名詞後加詞綴的「細人仔」(小孩子)等，都是無法全部收錄至詞典當中，因此利用構詞規則合併詞組以提供良好的斷詞結果。



3.2 中文分詞與詞性標記

中文分詞與詞性標記第一步是透過詞典及中文分詞條件隨機域模型得到可能的分詞候選句；接著經由查詢詞典得到候選句中每個詞彙的詞性標籤組合，最後利用詞性標記條件隨機域模型，從候選句中選出最佳結果，而此最佳結果為已標上詞性標籤的詞彙序列；本節將對中文分詞以及詞性標記做介紹。

當輸入某句文字序列序列 C 時，我們要找到其最佳對應的詞彙序列 W ，以及詞性

標記序列 T ，可用以下公式表示：

$$P(W, T|C) = P(T|W, C) \times P(W|C) \quad (3-1)$$

可分為 $P(T|W, C)$ 和 $P(W|C)$ 兩機率值的相乘， $P(W|C)$ 表示出現某個文字序列時，出現某個詞彙組合的機率，即出現某個分詞組合之機率，我們由條件隨機域模型得到最高機率前幾名作為分詞單元輸出候選句； $P(T|W(S, B))$ 是在某個詞彙序列的情況下，出現某個詞性標記序列的機率，候選句配合詞典中每個詞彙可能出現的詞性標籤輸入至條件隨機域模型，模型會依照學習到的句法結構，給予候選句（從第一個到最後一個）每一個詞彙一個最高機率數值，即最可能出現詞性標籤，以及其對應的機率； $P(T|W(S, B))$ 和 $P(B|S)$ 兩機率相乘得到候選句的分數，最後選取分數最高組合的做為最佳分詞與詞性標記結果。

3.2.1 中文條件隨機域模型

條件隨機域（Conditional Random Field）常用在標註和分析序列資料，例如自然語言文字序列，是一種鑑別式的機率模型，在 2001 年由 Lafferty[8] 等人提出。條件隨機域是一種無向性的圖模型，在給定需要標記的觀察序列條件下，計算整個標記序列的聯合機率分佈；而不是在給定當前狀態條件下，定義下一個狀態的狀態分佈；也就是給定觀察序列 X ，使 $P(Y|X)$ 值為最大時的 Y 即為最佳答案，如以下式子所表示：

$$Y^* = \arg \max_Y P(Y|X) \quad (3-2)$$

X 是本研究中代表中文字元序列 C ，例如序列 X 「感動的事」， C_1 是「感」、 C_2 是「動」、 C_3 是「的」而 C_4 為「事」； Y 則代表詞位標籤序列，詞位標籤將於下節做詳細

介紹，本例所對應的正確詞位標籤為「B₁ESS」。而 $P(Y|X)$ 如下式所表示：

$$P(Y|X) = \frac{1}{N(X)} \exp\left(\sum_{t=1}^T \sum_{i=1}^I \lambda_i f_i(Y_t = y, Y_{t-1}, X)\right) \quad (3-3)$$

$N(X)$ 為正規化數值，為了保證所有標記總和值為一， $\sum_y P(Y|X) = 1$ ，

$f_i(Y_t = y, Y_{t-1}, X)$ 表示序列 X 在位置 t 時的第 I 個特徵函數，此函數對應到 0 或 1。

$$f_i(Y, Y_{t-1}, X) = \begin{cases} 1, & \text{if } y = y_t \text{ is satisfied and } x = x_t \\ 0, & \text{otherwise} \end{cases} \quad (3-4)$$

例如，序列 X ：「感動的事」， Y 序列(標準答案)為「B₁ESS」($y_1 = B_1$ 、 $y_2 = B_2$ 、 $y_3 = S$ 、 $y_4 = S$)，當 $t=2$ 時 $f_i(Y_2 = E, Y_1 = B_1, X_2 = \text{動}) = 1$ 。而 $\lambda_1, \lambda_2, \dots, \lambda_k$ 為條件隨機域從訓練語料中學習的權重向量。



3.2.2 中文分詞

由於每個字在組成一個特定詞語時都佔據一個構詞位置，即字位，因此可以將分詞過程當作學習此字之字位訊息的機器學習過程。分詞是對字串中每一個字相對應做出一個在該處切分與否的決策過程；本研究採用 Zhao[4] 在 2006 年所提出的六字位標注集。六字位標注集之集合為 $T = \{B_1, B_2, B_3, M, E, S\}$ ，其中 B_1 、 B_2 和 B_3 分別表示該字位於一個詞的前三字位置， M 則代表該字位於詞中第四個字元之後但非詞尾的位置， E 表示該字位於一個詞的詞尾， S 代表單字詞。

我們使用線上開放的套裝軟體 CRF++ 來訓練中文斷詞模型，此軟體被廣泛應用在許多自然語言處理領域的研究。在訓練階段時，CRF++ 必須透過特徵樣版產生特徵參數，特徵樣版表示我們擷取特徵的規則，本研究在訓練中文斷詞條件隨機域模型所選用的特徵如表 3-1 所表示。其中字元 C_n 、標點符號 PM_n 、數字字元 D_n 、英文字元 E_n 、

和詞典詞相對應的 WL_n 及 WT_n 幾項特徵為 Zhao 所提出；而本研究加入每個字元的部首資訊 B_n 以及 DP_n 和 DT_n 兩項未收錄詞典詞之特徵， DP_n 及 DT_n 是關於在同一個句子中是否找到重複出現的詞，可以補足詞典中未收錄之詞；例如在文章中常出現但詞典並未收錄的專有名詞或是長詞，這兩項特徵可以提高此部分的分詞效能。

在訂定適當的特徵後，我們將每個特徵依據其相關性及其前後資訊分組，稱為特徵樣版，我們所採用的特徵樣版詳細列在表 3-2。

表 3-1 中文分詞之條件隨機域訓練特徵

特徵	含意
C_n	當前字元
B_n	當前字元的部首
PM_n	當前字元是否為標點符號
WL_n	在詞典可找到包含當前字元的最長詞之詞長
WT_n	在詞典可找到包含當前字元的最長詞中，當前字元的詞位標注(B_1, B_2, B_3, M, E, S)
DP_n	在同一句子當中，可以找到包含當前字元的最長詞之詞長，而此詞在包含當前字元的詞出現前曾經出現過
DT_n	若當前字元可找到其包含當前字元的最長詞，而此詞在包含當前字元的詞出現前曾經出現過，當前字元在此詞當中的詞位標注
D_n	當前字元是否為數字
E_n	當前字元是否為英文字母

表 3-2 中文分詞之條件隨機域訓練特徵樣版

特徵樣版	含意
$C_{n-1}, C_n, C_{n+1}, (C_{n-1} C_n), (C_n C_{n+1}), (C_{n-1} C_{n+1}), PM_n$	字元與其臨近字元資訊
$(B_n B_{n-1}), (B_{n+1} B_n), (B_{n+2} B_{n+1})$	部首資訊
$(WL_n WT_n), (WT_{n-1} WT_n WT_{n+1}), (WL_n WT_n DL_n DT_n)$	詞典詞及重複詞資訊
$(C_n D_{n-1} D_{n+1})$	數字字元資訊
$(E_{n-1} E_n), (E_n E_{n+1}), (E_{n-1} E_n E_{n+1})$	英文字元資訊

3.2.3 中文詞性標記

詞性(Part of Speech, POS)可以代表一個詞在句子中的語法角色，而詞性標記單元的功能，就是在斷詞單元產生 N 個最佳候選句後(本系統中， $N=3$)，給予每個詞彙相對應的所有詞性組合，再由條件隨機域模型算出有最高機率的詞彙組合及其詞性標籤。本研究在訓練中文詞性條件隨機域模型所使用的特徵及特徵樣版如表 3-3 及表 3-4 所表示：

表 3-3 中文詞性標記之條件隨機域訓練特徵樣版

特徵	含意
PM_n	標點符號類別 0：沒有標點符號 1：「、」 3：「。」、「；」、「：」、「？」、「！」、「—」 2：其他標點符號
WL_n	當前詞長
$LPOS_n$	若詞典有該詞，則此詞收錄在詞典中的所有詞性組合 例如：「一」的 $LPOS_n$ 為「Cbb_Di_D_Neu」
FC_n	若該詞未收錄於詞典，則該詞得第一個字元 例如：「約克夏」的 FC_n 為「約」
LC_n	若該詞未收錄於詞典，則該詞得最後一個字元 例如：「約克夏」的 LC_n 為「夏」

表 3-4 中文詞性標記之條件隨機域訓練特徵樣版

特徵樣版	含意
$PM_{n-1}PM_nPM_{n+1}$	標點符號資訊
$LPOS_{n-1}, LPOS_n, LPOS_{n+1}, (LPOS_{n-2}LPOS_{n-1}), (LPOS_{n-1}LPOS_n), (LPOS_nLPOS_{n+1}), (LPOS_{n+1}LPOS_{n+2})$	POS 資訊
WL_n, FC_n, LC_n	未收錄詞典詞資訊

3.3 客語分詞與詞性標記

由於我們在客語語料上的不足，無法以條件隨機域方法獨自訓練出客語專用的斷詞系統，因此本論文採用客語外掛詞典、國語詞典、客語構詞規則以及中文條件隨機域模型組合成客語分詞與詞性標記系統。

在分詞階段，文句輸入經由優先搜尋客語外掛詞典，若是在詞典搜尋到該詞則將此詞切分開，其餘文字經由中文條件隨機域模型及國語內部詞典挑選出分詞候選句，接著查詢外掛詞典及內部詞典得到每個詞的詞性標記組合，最後由中文條件隨機域模型得到最佳斷詞與詞性標記序列。

3.3.1 客語詞典

客語外掛詞典是客語分詞與詞性標記系統中很關鍵的資訊，我們先將客語詞典中與中文內部詞典相同的共用詞彙刪除，其所有外掛詞典之資料量在表 3-6 詳細列出。

表 3-5 客語外掛詞典統計表

	詞數
一字詞	91
二字詞	5296
三字詞	5521
四字詞	3401
五字詞	293
六字詞	75
七字詞	80
八字詞	12
總計	14769

3.3.2 構詞規則

分詞會將輸入的文句和詞典做比對，我們無法將所有可能的詞條列於詞典當中，這些無法收錄於詞典之詞，有些是有規律的，可由「詞類標記單元」得到詞性標籤後再經由「構詞單元」結合出來。

客家話的構詞規則和國語非常相似，大致上分為四種，分別是重複，附加，附合及合併。所謂重複是指單一詞彙之重複，如「洗湯」重複成「洗洗湯湯」(表廚房內擦擦洗洗之事)。附加則是在詞的前後附加綴詞，如在「婆」之前加前綴詞「阿」形成「阿婆」一詞，後綴詞如在「刀」之後加「仔」形成「刀仔」。附合詞是指兩個獨立的詞結合而成另一個新詞，例如「青菜」是由「青」和「菜」組成，但「青菜」不一定是指「青色的菜」，而是所有蔬菜的泛稱，附合詞兩個詞之間不能插加任何詞語，由此可見附合詞並不單純是兩個詞的併列，而是有其密不可分的構詞特性。第四種構詞方式稱為合併，這是中文構詞規則所沒有的，如 cin1 (親) gal (家) 念快時會變成 cia1 (親家)，這種合併結果是單音節多詞素的構詞。我們解決附合詞的方式為新增詞條至詞典，而合併詞在四縣客家話當中並不常見，因此本系統所採用的構詞規則主要以重複和附加為主。

我們採用中央研究院詞庫小組所制訂的「中文分詞處理原則」，由其處理原則中挑選出符合客語文章之構詞規則加入本系統，詳細構詞規則如表 3-6、表 3-7 所示。表 3-6 列出了客語和國語完全相同的構詞規則，表 3-8 則為客語和國語平行的構詞規則，其構詞結構一樣，但關鍵字由中文詞替換為客語詞。

表 3-6 國客語共用構詞規則

國語構詞規則	國語範例	客語範例
XXY(Z)	說說看	日日大
X-XY-Y	乾乾淨淨	洗洗湯湯
V+了/著/過/看/看看+(V)	想過	歇過
名詞後詞綴：氏/度/性/家/長/師/員/們/子/倆	女性朋友們	細妹伴仔
排行:數字+姐/妹/哥/弟/伯/叔/姑	二伯	二伯
VG+為/成/作/不了/得了	身為	身為

表 3-7 國客語平行構詞規則

國語構詞規則	國語範例	客語構詞規則	客語範例
V+不+到/起/出	看不起	V+毋+到/起/出	看毋起
越+來+越	越來越少	緊+來+緊	緊來緊少
越+ X+越+Y	越穿越多	緊+ X+緊+Y	緊著緊多
VB/VH+不休/兮兮/不絕/不下/不得/開來	蜿蜒不絕	V+毋得	捨毋得
老/小/阿+Na	小妹妹	細+Na	細妹仔

3.3.3 客語借字問題

在客語的語料紀錄當中，遇到有音無字的情形時，解決方法除了造字以外，經常使用借中文字表示客語字，例如「每擺」（每次）中，「擺」音同「次」的客語發音，因此客語文章中使用「擺」代表中文字「次」；同時，「擺脫」（擺脫）是客語文章中的國客共用詞，「擺」在此的意思又與中文原意相同。

借字情形出現在二字詞以上時，我們新增此類情況的詞彙至客語外掛詞典，大部分可正確進行分詞與標記詞性；但在一字詞的情況下，如前例中客語文章中的「擺」代表了中文字「擺」和「次」，中文字「擺」有可能出現的詞性組合為「VAC_VC」，中文字「次」的詞性組合則為「Nes_Nf」，因此「擺」在客語文章中所有可能的詞性組合為「Nes_Nf_VAC_VC」，但此種組合並未在訓練時加入中文詞性條件隨機域模型當中，因此由本研究的斷詞器無法標記出正確的斷詞及詞性結果，並且容易造成搶

詞情形。

為了處理客語字之詞性組合和中文模型無法相對應的問題，我們從中文模型可允許的詞性組合中找出最相近的，上例中為「Na_Nf_VC」，並將其加入外掛詞典，如下例所表示：

客語字	語料庫中出現對應中文字/POS	對應中文字/POS	加入客語外部詞典 POS
擺	次/Nf	次/Nes_Nf	Na_Nf_VC
	擺/VC	擺/VAC_VC	

本實驗從客語語料中統計出現次數最多的 40 個一字詞，將有此問題之單元整理其對應詞性，並加入外掛詞典中，所加入之一字詞如附錄二所表示。



第四章 HMM-based 語音合成器

基於隱藏式馬可夫模型的語音合成系統(HMM-based Speech Synthesis System, HTS)是目前最廣為應用的合成法，與傳統的串接式合成系統比較，基於模型的語音合成方法具有較高的可攜性以及適應性，儲存模型所需要的空間，遠小於串接式合成系統所需儲存大量的音檔；另一方面，模型的可訓練性，能快速使模型參數進行調式或轉換。本章將詳細介紹 HTS 工作原理、我們所採用的決策樹問題集、文本相關資訊，以及對本論文在語音合成前所做的停頓預估系統加以討論。

4.1 HMM-based 語音合成系統

基於 HMM 之語音合成系統架構如圖 4-1 所示，主要分為訓練以及合成兩個階段：

(一) 訓練階段：

我們將得到切割資訊後的聲音語料，估算其特徵向量作為參數，包括頻譜部分(spectrum, 梅爾倒頻譜)以及其動態特徵向量、激發訊號部分(excitation, $\log F_0$) 以及其動態特徵向量；聲音語料所對應的文字則由斷詞系統和文本分析器分析出對應的文脈相關序列，使系統在訓練階段時可以透過合併分裂樹學習合成單元間的相連狀況；包含切割資訊的語音訊號以及文本相關序列先經由 HMM 訓練，產生各合成單元的音高特徵模型、頻譜特徵模型和音長特徵模型，再透過狀態合併決策樹問題集對模型進行分群，得到和文本相關的 HMM 模型。

(二) 合成階段：

使用者輸入文字，由斷詞和詞性標記系統得到斷詞資訊後，再經由停頓預估器預估出停頓位置（停頓預估系統將在 4.5 節詳細介紹），送入和訓練部分相同的文本分析器得到合成單元及其所對應的語言參數序列；透過分類及回歸樹挑選出相對應之

HMM 模型序列，根據參數產生演算法產生特徵參數，最後通過梅爾對數頻譜近似濾波器(Mel Log Spectrum Approximation filter, MLSA filter)合成語音訊號輸出。

基於 HMM 之語音合成器主要核心技術有三個；(1) 包含動態特性的參數生成演算法，此演算法將在下節做詳細介紹。(2) 多空間機率分佈隱藏式馬可夫模型 (Multi-space Probability Distribution HMM, MSD-HMM)，基頻在濁音處是有值的，但在清音處為零，為了解決基頻在濁音處和清音處轉換上的不連續，MSD-HMM 在此定義濁音處的參數維度是一，而清音處參數維度是零。(3)基於梅爾倒頻譜的聲音合成技術，其中包含訓練階段梅爾倒頻譜的抽取，以及合成階段 MLSA filter 將梅爾倒頻譜系數還原為語音訊號。

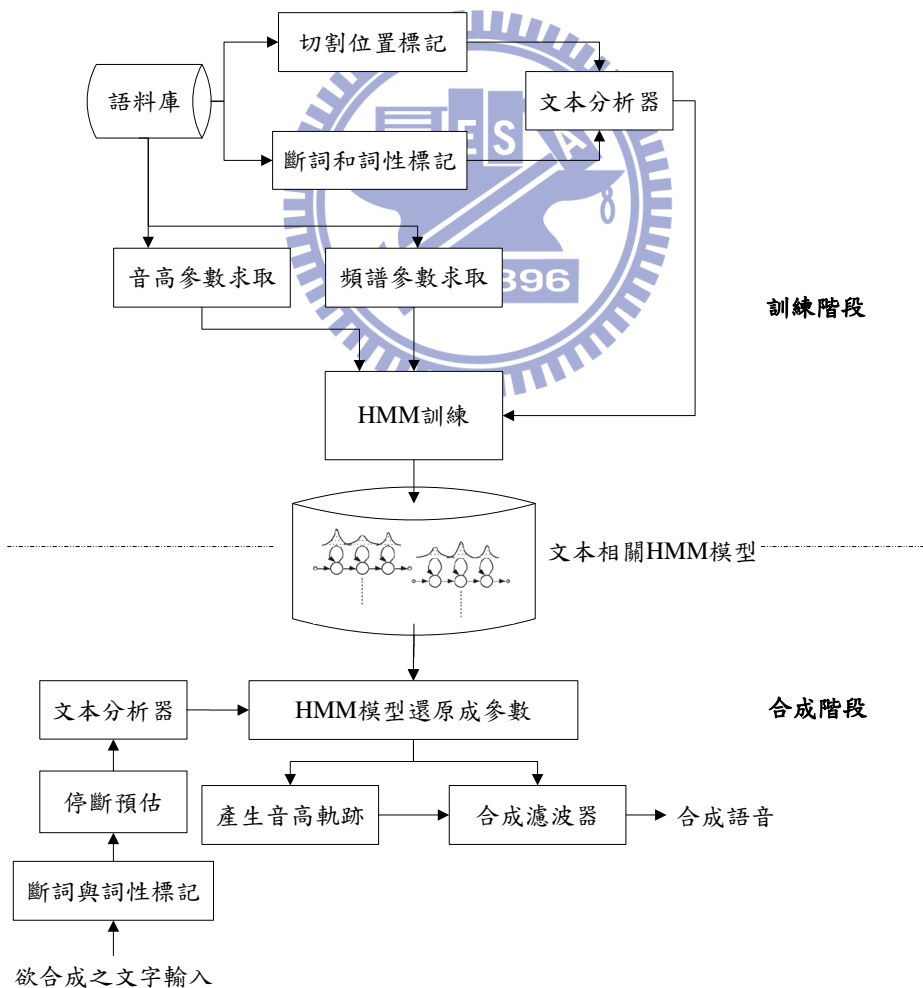


圖 4-1 基於隱藏式馬可夫模型之語音合成系統架構

4.2 考慮動態特性的參數生成演算法

考慮動態特性的參數生成演算法是 HTS 能合成出流暢自然語音的關鍵，本節將對動態參數以及其參數生成演算法加以詳細介紹。

4.2.1 動態參數

系統所抽取出的梅爾倒頻譜數值以及 logF0 數值稱為靜態參數(Static feature)，代表了音高走勢以及頻譜特徵的變化，而此種變化會受到前後合成單元的影響，也就是特徵參數的改變與前後時間的關聯性非常大，因此有學者提出動態參數(Dynamic feature)的觀念在語音辨識的 HMM 模型參數上，後來也在基於 HMM 的語音合成器中被使用。動態參數的考慮能讓訓練出的模型單元和單元間的連接更加流暢。基於 HMM 的語音合成器中的動態參數為計算靜態參數在時間軸上的前後關連性而得到，最後的觀察參數為原本的靜態參數以及其動態參數，不僅擴大模型的參數維度，也讓模型在前後時間關係有更好的描述。

令 $c_t = [c(1), c(2), \dots, c(M)]^T$ 為在時間點 t 的靜態特徵向量(例如頻譜參數)，而相對於 c_t 的一階動態參數 Δc_t ，以及二階動態參數 $\Delta^2 c_t$ 計算式如下：

$$\Delta^n c_t = \sum_{i=-L^{(n)}}^{L^{(n)}} \omega^{(n)}(i) c_{t+i}, \quad n=0, 1, 2 \quad (4-1)$$

$\omega^{(n)}(i)$ 表示第 n 階動態參數對於時間點 $t+i$ 的靜態參數所附予的權重係數， i 表示作用範圍，也就是時間點 t 往前後各 $L^{(n)}$ 個時間單位。

$$\omega^n = \left[\begin{array}{ccccccc} 0 & \dots & \omega_{-L^{(n)}}^n & \dots & \omega_t^n & \dots & \omega_{L^{(n)}}^n & \dots & 0 \\ & & & & & & & & \end{array} \right] \quad (4-2)$$

其中 $\omega^{(0)}(0)=1$ 以及 $L^{(0)}=0$ 。

動態參數表現在時間上的物理意義會因為各階動態參數所對應設計的 ω^n 而有所不同，本研究的動態參數計算如下：

$$\Delta c_t = \frac{1}{2}c_{t+1} - \frac{1}{2}c_{t-1} \quad (4-3)$$

$$\Delta^2 c_t = (c_{t-1} - c_t) + (c_{t+1} - c_t) \quad (4-4)$$

我們在圖 4-2 表示出一階及二階動態參數的關係：

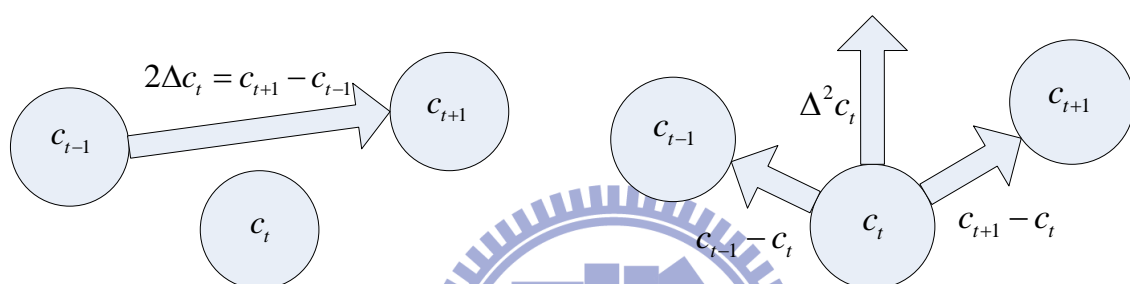


圖 4-2 一階及二階動態參數關係圖

由圖可以看出，一階動態參數和自己本身的靜態參數沒有關係，是下一個單元以及前一單元的差距；而二階動態參數則是由前一單元與自己的差距和後一單元與自己的差距相加。我們將靜態參數以及一階、二階動態參數作為模型的參數向量，則在時間點 t 的參數向量如下表示： $o_t = [c_t, \Delta c_t, \Delta^2 c_t]$ ，其中 $c_t = [c_t(1), c_t(2), \dots, c_t(M)]$ 。

而整個句子的參數向量是以 $O = [o_1, o_2, \dots, o_T]^T$ 表示， O 和 c 之間的對應關係矩陣 W 如圖 4-4 所示：

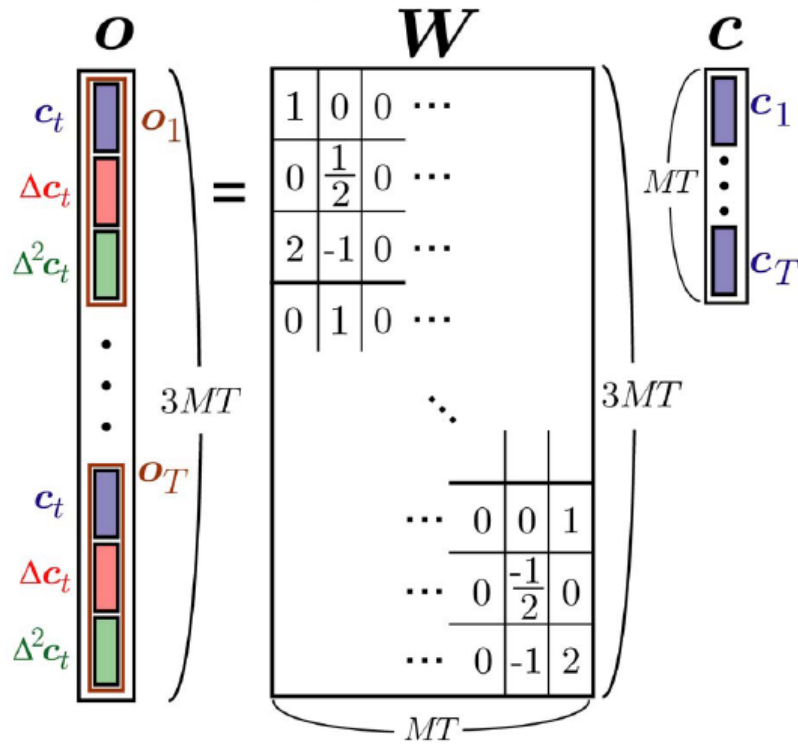


圖 4-3 轉換矩陣表示示意圖，本圖摘錄自 Ref.[1]

W 矩陣的維度是 $3MT \times MT$ ，將原本的靜態參數向量 c 轉換到參數向量 O ，參數維度也從 $MT \times 1$ 擴大到 $3MT \times 1$ 。

4.2.2 參數生成演算法

在合成階段，文字經由和訓練階段相同的文本分析器後得，到相對應的前後文相關語言參數序列 S_1, S_2, \dots, S_N ，經過 HMM 模型預測後得到模型序列 $\lambda_1, \lambda_2, \dots, \lambda_N$ ，以及模型參數向量 O_1, O_2, \dots, O_N ， O_n 是由平均值向量 μ_n (Mean vector) 及共變異數向量 U_n (Covariance matrix) 所組合而成， μ_n 和 U_n 的維度包含靜態參數、一階動態參數以及二階動態參數的維度，是訓練階段所抽取出的參數維度之三倍。在合成語音前，必須將 O_n 還原成對應的 C_n ，才能得到頻譜或音高軌跡，此轉換方式稱為參數生成演算法，如圖 4-5 表示。

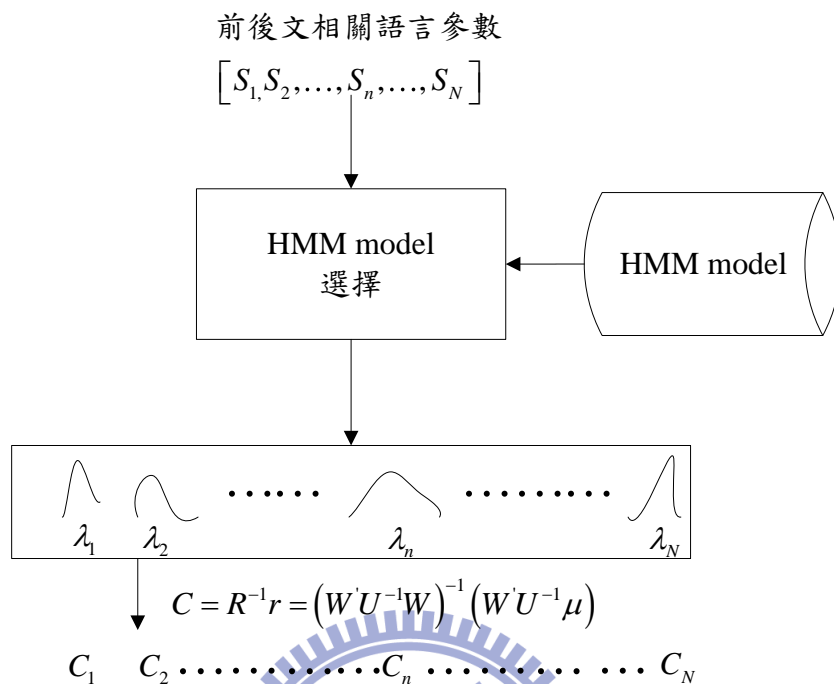


圖 4-4 參數生成示意圖

合成階段將合成單元參數串接成 O ， $O = \{O_1, O_2, \dots, O_N\}$ ，接著產生整個輸入字串的 HMM 模型 $\Lambda = \{\lambda_1, \lambda_2, \dots, \lambda_N\}$ ，模型的平均值向量為 $\Lambda_\mu [\mu_1, \mu_2, \dots, \mu_N]^T$ ，共變異數矩陣 $\Lambda_U = \text{diag}[U_1, U_2, \dots, U_N]$ ，我們需要找出使得已知模型 Λ 產生 O 中機率最大的 C ，也就是找到最佳的 C 使 $P(O|\Lambda)$ 有最大機率值。其中 $O=WC$ ，由 $P(O|\Lambda)$ 取對數值得到：

$$\log P(O|\Lambda) = -\frac{1}{2}(O - \Lambda_U)' \Lambda_U^{-1}(O - \Lambda_U) - K \tag{4-5}$$

其中 $K = \frac{1}{2} \log |\Lambda_U| - \frac{3}{2} MN \log 2\pi$ ，和 O 不相關。

將式子對 C 偏微分，在偏微分為零的地方為極大值：

$$\frac{\partial \log P(WC|\Lambda)}{\partial C} = 0 \quad (4-6)$$

可以得到：

$$W'U^{-1}WC = W'U^{-1}\mu \quad (4-7)$$

最後我們得到： $C = R^{-1}r$

其中 $R = W'U^{-1}W$ 而 $r = W'U^{-1}\mu$ 。

4.3 切割位置

進行語音合成的訓練之前，必須先將語音標示出各合成單元的邊界資訊，才能正確地抽出各合成單元的頻譜、音高及音長之特徵值。本實驗所採用的合成單元為聲母與韻母。我們由 HTK (Hidden Markov Toolkit) 訓練出聲母及韻母的 HMM 模型，對所有語料進行強制切割 (Forced Alignment)。參數設定為 38 維，其中包含 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。

由 HTK 得到初步切割位置後，我們使用本實驗室之高解析度自動語音切割系統，其架構是根據聲/韻母的層級來分類，並依照各個分類彼此之間可能的轉換狀態訂定目標函數，再使用類神經網路之多層感知器架構，以半監督式之模型訓練方法建立起端點偵測器模型，將 HMM 的自動切割結果以端點偵測器所產生的概似度經正規化後作為分數，進行維特比搜尋並限制搜尋範圍在初始切割位置前後 100 毫秒之內，最後得到本論文之聲/韻母層級自動語音切割位置。

4.4 文本相關資訊

在將語料及其對應文字送入訓練以及將文字輸入合成語音前，我們要將文字檔轉為合成單元及其相對應之文本相關資訊；文本相關資訊包含聲學模型之合成單元以及前後合成單元的特徵資訊，在 HTS 訓練階段時根據文本相關資訊和決策樹問題集進行比對分群，合成部分則由決策樹取得參數而合成語音。本論文所採用的合成單元為聲母/韻母/停頓(short pause, sp)資訊，表 4-1 為文本相關資訊和說明：


表 4-1 文本相關資訊格式

	聲母/韻母/停頓(initial/final/sp)
p1	前一個聲母/韻母/停頓
p2	當前聲母/韻母/停頓
p3	下一個聲母/韻母/停頓
	音節(Syllable)
t1	前一個音節聲調
t2	當前音節聲調
t3	下一個音節聲調
w1	當前音節在詞中的位置(從前面數來)
w2	當前音節在詞中的位置(從後面數來)
w6	音節位置屬於詞間或詞中
s1	當前音節在句中的位置(從前面數來)
s2	當前音節在句中的位置(從後面數來)
PM	音節後是否有標點符號 1：「，」 2：「。」 3：「、」 4：其他標點符號 5：無標點符號
	詞(Word)
w3	前一個詞由幾個字組成
w4	當前詞由幾個字組成

w5	下一個詞由幾個字組成
POS1	前一個詞的詞性
POS2	當前詞的詞性
POS3	下一個詞的詞性
	句(Sentence)
s3	前一個句子由幾個字組成
s4	當前句子由幾個字組成
s5	下一個句子由幾個字組成

4.5 狀態合併決策樹問題集

決策樹問題集紀錄資料分群規則，HTS 建立決策樹時會依據問題集進行分群。為了達到最佳合併分裂結果，我們考慮了合成單元相關、字相關、詞相關、句相關以及聲調相關五類問題。

- 
- (一) 合成單元相關：考慮當前合成單元及其前後各一個合成單元
 - 聲母發音方法：塞音、非塞音、摩擦音、邊音、清音、濁音...
 - 聲母發音部位：唇間、舌尖、牙喉...
 - 韻母發音類別：韻母是以 a、e、i、o、u、ii 為基礎
 - (二) 字相關：考慮當前字
 - 字在詞中的順序（正序、倒序）
 - 字在句中的順序（正序、倒序）
 - (三) 詞相關：考慮當前詞以及其後詞
 - 詞中字元個數
 - (四) 句相關：考慮當前句以及其前後句
 - 句中字元個數
 - (五) 聲調相關：考慮當前合成單元及其前後各一個聲調
 - 聲調 1、2、3、4、5、8

4.6 停頓預估

我們有了斷詞系統和 HTS 所訓練出的 HMM 模型後，在 TTS 系統實作時，仍然需要停頓資訊的預估，才能合成出更加自然流暢的語音。我們分析語料庫中音檔和其對應的文字檔發現在文字轉語音的過程中，可以由文字直接預估語音的停頓位置；例如在有標點符號的句子和句子之間，通常會有停頓的發生。有良好效能的停頓預估能讓合成語音更加流利，本研究使用條件隨機域模型訓練停頓預估模型，以下將介紹模型特徵樣版的選擇，以及系統效能之評估。

4.6.1 停頓預估模型

一般人在唸一段語音時幾乎不會在詞內做一個明顯的停頓，因此本研究在停頓預估模型的訓練單元是以詞為單位。而我們從語料庫中發現，詞間是否有停頓和詞性標籤有很大的關係，例如在關係連接詞或是代名詞的前後通常都會有韻律停頓，我們舉例如下：

例如：故所(Cbb)「停頓」對(P)伯公(Nb)個(DE)稱呼(Na)

一定(D)愛(D)用(P)佢(Nh)「停頓」本身(Nh)個(DE)語言(Na)來(D)表現(VC)

而詞長也是影響詞間是否有停頓很大的因素之一，通常長詞之後的停頓機率較短詞高；最靠近詞間的聲母、韻母類別也是我們在此所考慮的重點。我們利用條件隨機域訓練預估詞後是否有停頓的模型，本研究的特徵及特徵樣版如表 4-2 和表 4-3 所表示：

表 4-2 停頓預估模型之特徵

特徵	含意
PM_n	詞後是否有標點符號
WL_n	詞長
POS_n	詞性
Wi_n	詞從前面數來第一個字的聲母發音類別 1 : b、d、g 2 : p、t、k 3 : m、n、ng 4 : z 5 : c 6 : f、s、h
Wf_n	詞從前面數來最後一個字的韻母發音類別 1 : a、ai、au、am、an、ang、ap、at、ak 2 : e、eu、ieu、iau、im、em、en、ep、et 3 : i、ie、ia、io、iu、iui、ioi、iem、iam、in、ien、ion、iun、iang、iong、iung、ip、iep、iap、it、iet、iut、iak、iok、iuk、iing、iui 4 : o、oi、on、ong、ot、ok 5 : ue、ua、u、uai、ui、uen、uan、un、ung、uang、uet、uat、ut、uk、uak 6 : ii、iim、iin、iip、iit 7 : FNULL

表 4-3 停頓預估模型之特徵樣版

特徵樣板	含意
$PM_{n-1}, PM_n, PM_{n+1}, (PM_{n-1}PM_nPM_{n+1}) (PM_{n-2}PM_{n-1}PM_n)$	詞後是否有 PM 資訊
$WL_{n-2}, WL_{n-1}, WL_n, WL_{n+1}, WL_n, (WL_{n-2}WL_{n-1}WL_nWL_{n+1}WL_n)$	詞長資訊
$POS_{n-1}, POS_n, POS_{n+1}$	POS 資訊
Wi_{n+1}	下一個詞第一字聲母發音類別資訊
Wf_n	當前詞最後一字韻母發音類別資訊
$(Wi_{n+1}Wf_n), (POS_nWi_{n+1}Wf_n)$	組合資訊

第五章 實驗結果

本論文為客語文字轉語音系統之實作，TTS 系統最主要的兩大部分為斷詞單元以及基於 HMM 之語音合成器，本章將針對這兩部分做實驗設計以評估其效能。

5.1 斷詞系統之效能分析

我們在客語斷詞系統中，結合中文分詞與詞性標記模型、客語外掛詞典以及客語構詞規則。本節將介紹訓練中文斷詞模型之語料，並先對中文斷詞做效能分析；再結合客語外掛詞典和構詞規則後對客語語料做測試以評估其系統效能。

5.1.1 中文分詞和詞性標記系統之效能分析

本節介紹中文分詞及詞性標記系統之訓練語料、測試語料、系統所使用的詞典、訓練模型特徵數量，以及如何分析系統效能和實驗結果。

5.1.1.1 訓練、測試語料及中文詞典

本實驗所採用的中文斷詞訓練及測試語料庫為 SIGHAN 在中文分詞競賽中所提供標記正確分詞結果的語料。SIGHAN (Special Internet Group of the Association for Computational Linguistics)舉辦的會議是國際間關於中文分詞最有名的研討會，其中訓練語料共有 8350369 個單元，測試語料共有 196867 個單元。而由訓練語料以及特徵樣版所產生的總特徵數共為 3882102 個。

而詞性標記系統所使用的訓練及測試語料庫為中央研究院平衡語料庫 3.0 版，此

語料庫共有 9286 篇文章，總詞數為 5841942 個。我們從中取出 4426886 個詞做為訓練語料，剩下 484070 詞為測試語料。而詞性標籤共有 46 種，其 46 種詞性關係及詞性意義的對應關係在附錄三表示，本研究在詞性標記訓練部分共產生 5176104 個特徵。

我們在此系統中所採用的內部詞典是向中研院中文處理小組所發行的中央研究院平衡語料庫詞典，由中研院資訊所及語言所對語言進行分析所設計的。詞典內容包含詞以及詞性組合標籤，詞典的詞數統計結果如表 5-1：

表 5-1 中文內部詞典統計表

	詞數
一字詞	7153
二字詞	55662
三字詞	13696
四字詞	11237
五字詞	252
六字詞	114
七字詞	23
八字詞	1896
總計	88142

5.1.1.2 評估方法及實驗結果

一、斷詞效能評估

我們使用精確率(Precision)、召回率(Recall)及 F 分數(F-score)來作為評估中文斷詞實驗系統的效能，此三種方法定義如下式表示：

$$\text{精確率} = \frac{\text{斷詞器正確斷出的詞數}}{\text{斷詞器斷出的總詞數}} \quad (5-1)$$

$$\text{召回率} = \frac{\text{斷詞器正確斷出的詞數}}{\text{參考答案的總詞數}} \quad (5-2)$$

$$F\text{分數} = \frac{2 \times \text{精確率} \times \text{召回率}}{\text{精確率} + \text{召回率}} \quad (5-3)$$

我們在此舉例說明，句子「電影明星永遠是人們視覺的焦點。」，其實際的評估範例如下列表示：

<p>參考答案：</p> <p>電影明星 永遠 是 人們 視覺 的 焦點 。</p> <p>分詞系統輸出結果：</p> <p>電影 明星 永遠 是 人們 視覺 的 焦點 。</p>
<p>分詞系統正確斷出的詞數：6，分詞系統斷出的總詞數：8，參考答案的總詞數：7</p> <p>精確率：0.75，召回率：0.86，F分數：0.80</p>

圖 5-1 中列出了參考答案詞彙序列和分詞系統輸出結果，參考答案的總詞數為 7 個，斷詞系統斷出的詞數有 8 個，其中有 6 個詞和參考答案相同，我們由式子(6-1)到(6-3)來評估此範例之系統效能，得到精確率為 0.75，召回率為 0.86，F 分數為 0.80。

本論文的中文分詞系統之精確率、召回率以及 F 分數如表 5-2 所示，由表中可以看出我們在中文分詞系統對中文分詞的效能上已有很好的結果：

表 5-2 中文分詞實驗結果

	精確率	召回率	F 分數
外部測試語料	95.95%	96.79%	96.37
內部測試語料	98.94%	98.82%	98.88

二、詞性標記效能評估

而在詞性標記的實驗中，我們將已有標準分詞結果之中文詞彙序列輸入詞性標記系統，系統將每個詞彙標記出最佳對應詞性標籤輸出；我們使用正確率來作為評估詞性分析系統的效能，正確率之定義如下：

$$\text{正確率} = \frac{\text{系統標記詞性和參考答案相同的詞數}}{\text{總詞數}} \quad (5-4)$$

我們在詞性標記系統測試中，所得到的內部測試正確率 96.01%，而外部測試結果為 94.73%，由實驗數據顯示，我們在中文詞性標記系統已有不錯的效能。

得到了整體正確率數據後，我們對於個別的詞性標記分析其效能，本實驗採用精確率(Precision)、召回率(Recall)及 F 分數(F-score)來評估，其定義方式如下：

$$\text{精確率} = \frac{\text{系統正確標出的總詞性數(單一詞性)}}{\text{系統標出的總詞性數(單一詞性)}} \quad (5-5)$$

$$\text{召回率} = \frac{\text{系統正確斷出的總詞性數(單一詞性)}}{\text{參考答案的總詞性數(單一詞性)}} \quad (5-6)$$

$$F \text{ 分數} = \frac{2 \times \text{精確率} \times \text{召回率}}{\text{精確率} + \text{召回率}} \quad (5-7)$$

在本實驗中 46 個詞性的精確率、召回率、F 分數以及出現次數如表 5-3 所示，其中如「Dfb」(動詞後程度副詞)，由於資料量較少，因此容易被判斷為「Dfa」(動詞後程度副詞)；「Neqb」(後置數量定詞)的情形也跟「Dfb」相同，由於語料的不足容易誤判為「Neqa」(數量定詞)。而其餘詞性在詞性標記方面則都有很好的效能表現。

表 5-3 中文詞性標記實驗結果

詞性	精確率	召回率	F 分數	出現次數	詞性	精確率	召回率	F 分數	出現次數
A	87.89%	77.14%	82.16	2825	Neu	97.64%	99.05%	98.34	14168
Caa	88.84%	93.03%	90.88	6082	Nf	96.56%	97.91%	97.23	14820
Cab	97.53%	98.91%	98.21	930	Ng	88.54%	90.96%	89.73	7406
Cba	97.16%	98.84%	97.99	176	Nh	98.84%	99.55%	99.19	14851
Cbb	96.34%	96.55%	96.45	12000	P	92.67%	94.32%	93.49	21462
D	96.07%	96.44%	96.26	46023	SHI	100.00%	100.00%	100	7705
DE	99.23%	98.71%	98.97	30067	T	90.06%	94.10%	92.04	5210
Da	96.76%	98.33%	97.54	2190	VA	93.76%	90.27%	91.98	9761
Dfa	97.29%	98.01%	97.65	4870	VAC	88.11%	90.97%	89.52	286
Dfb	79.09%	69.60%	74.04	110	VB	89.20%	86.21%	87.68	1009
Di	96.63%	96.63%	96.63	4450	VC	91.80%	92.49%	92.14	32306
Dk	74.58%	90.15%	81.63	712	VCL	89.25%	80.17%	84.47	2745
FW	92.28%	95.42%	93.83	3096	VD	91.69%	86.78%	89.17	1769
I	92.55%	86.35%	89.35	376	VE	92.35%	94.61%	93.47	11686
Na	94.97%	94.99%	94.98	102185	VF	92.26%	93.68%	92.96	1640
Nb	88.68%	85.13%	86.87	8859	VG	90.65%	84.43%	87.43	4492
Nc	94.05%	95.53%	94.79	24554	VH	93.48%	92.49%	92.98	29420
Ncd	85.66%	83.55%	84.6	4478	VHC	93.39%	95.71%	94.54	2510
Nd	97.55%	94.25%	95.87	10126	VI	96.63%	95.13%	95.88	445
Nep	99.26%	97.69%	98.47	6733	VJ	94.10%	91.12%	92.58	7587
Neqa	97.05%	97.07%	97.06	5317	VK	93.67%	95.83%	94.74	5989
Neqb	86.54%	73.77%	79.65	52	VL	96.84%	95.65%	96.24	2655
Nes	96.38%	96.41%	96.4	3427	V_2	98.38%	99.51%	98.94	4510

5.1.2 客語分詞和詞性標記

本論文在上一節中評估了中文分詞與詞性標記系統的效能，我們接著加入客語外掛詞典及客語構詞規則，對客語文章進行分詞及詞性標記；本節將介紹客語測試語料以及設計實驗以評估客語斷詞系統之效能。

5.1.2.1 測試語料

本研究所採用的客語測試語料庫為龔萬灶老師所主筆的「阿啾箭个故鄉」一書，其中包含 42 篇文章，共有 63158 個單元(包含標點符號)；由余秀敏老師將此客語文語料翻譯成相對應之中文平行語料，先從中研院斷詞器得到此中文平行語料之分詞和詞性結果標記，再以人工校正的方法，將此分詞和詞性標籤資訊對應至客語語料，最後得到標記好詞性及分詞結果之客語文章，我們把這份語料當作本實驗的參考標準答案。而此平行語料中，國客語使用詞彙完全一樣的詞共占了總詞數的 48.57%。

在詞性標記測試中我們使用和分詞測試相同的語料，輸入為參考答案的分詞結果，由客語詞性標記系統標記出相對應之詞性，再與參考答案的詞性標籤相比較以分析其效能，此測試語料共有 36450 詞。



5.1.2.2 實驗方法

一、客語分詞實驗

為了測試結合中文條件隨機域、中文詞典以及客語詞典在客語分詞的表現，本研究設計了二個實驗進行測試，分別是(a)實驗 A：僅使用中文斷詞系統對客語文章進行斷詞。(b)實驗 B：加入客語詞典和構詞規則，並採用客語外部詞典優先於中文詞典

(a) 實驗 A：僅使用中文分詞系統進行客語文章分詞

在實驗 A 當中，僅用中文分詞系統為客語文章做分詞，分詞結果可以看出客語文章中
使用中文詞彙的比例，並做為實驗 B 的基本比較。

(b) 實驗 B：加入客語詞典和構詞規則，並採用客語外部詞典優先於中文詞典

實驗 B 在原本中文斷詞系統中加入客語外掛詞典，並在分詞階層中採用外掛詞典
優先的規則，得到分詞和詞性標記結果後後再加入客語構詞規則，將時間詞、重
疊詞等構出。

二、客語詞性標記實驗

我們將參考答案的正確分詞結果輸入至系統標記出每個詞的詞性，為了測試客語外部詞典對於處理客語詞性的效能，在此設計兩個實驗，(a)實驗 A：僅使用中文詞性標記系統標記客語詞性。(b)實驗 B：加入客語詞典至中文詞性標記系統標記客語文章詞性。

(a) 實驗 A：僅使用中文詞性標記系統標記客語詞性

在實驗 A 當中，我們只用中文詞性標記系統為客語文章標記詞性。

(b) 實驗 B：加入客語詞典和構詞規則，並採用客語外部詞典優先於中文詞典

實驗 B 在系統中加入客語外掛詞典，收錄客語獨有的詞彙，並加入發生客語借字情況之一字詞與其詞性組合標籤。

5.1.2.3 實驗結果分析

一、客語分詞實驗

實驗 A 及實驗 B 的結果列於表 5-4，我們發現實驗 A，只由中文斷詞系統對客語文章進行斷詞的 F 分數有 74.03，而語料庫中國客語用詞完全相同的比例卻只有 48.57%，這不僅顯示出在客語文章中使用國語詞有很高的比例，更可以發現客語文章的文法結構與中文相似。而加入了客語外掛詞典後，很明顯可以看到實驗 B 的 F 分數較實驗 A 提升了 8.84，可以看出在加入客語外掛詞典後對斷詞效能的提升。

表 5-4 客語分詞實驗結果

	精確率	召回率	F 分數
實驗 A	71.09%	77.22%	74.03
實驗 B	83.96%	81.80%	82.87

我們詳細比較實驗 A 和實驗 B 的斷詞輸出結果，發現到許多實驗 A 無法解決的搶詞情形在實驗 B 已獲得改善，以下舉出兩個例子說明：

例子 1：實驗 B 在外掛詞典中收錄了客語詞「朝晨」及「打早」

實驗 A 斷詞結果	客家人 在 年初 一朝 晨 打 早
實驗 B 斷詞結果	客家人 在 年初一 朝晨 打早
參考斷詞結果	客家人 在 年 初一 朝晨 打早
中文平行語料斷詞結果	客家人 在 大年 初一 清晨 一大早

我們列出了實驗 A、實驗 B、參考答案及中文平行語料之分詞結果，實驗 A 中由於未加入客語詞「朝晨」及「打早」，造成中文詞搶客語詞的情況，此兩詞在實驗 B 的分詞輸出則已可正確斷出。

例子 2：實驗 B 在外掛詞典中收錄了客語詞「祖公祖婆」及「共下」

實驗 A 斷詞結果	請 祖公 祖婆 共 下來 過年
實驗 B 斷詞結果	請 祖公祖婆 共下 來 過年
參考斷詞結果	請 祖公祖婆 共下 來 過年
中文平行語料斷詞結果	請 祖先 一起 來 過年

實驗 A 的客語字串「共下來」被斷成中文詞「共」和「下來」，再將「共下」加入外掛詞典後，可正確斷出「共下」一詞。

二、客語詞性標記實驗

實驗 A 及實驗 B 對已給定的客語文章分詞結果標記詞性，正確率如表 5-5 所表示；由中文詞性標記系統直接標記客語文章詞性的正確率為 59.41%，而加入外掛詞典後正確率上升了 17.73%，由於在詞性標記階段，條件隨機域會依據詞典內的詞性組合給予每段文句最佳的配對詞性標籤，而未加入客語詞典時，系統無法找到其對應之詞性標籤，因此在實驗 B 的正確率較實驗 A 上升了許多。

表 5-5 客語詞性標記實驗結果

	正確率
實驗 A	59.41%
實驗 B	77.14%

而每個詞性的精確率、召回率及 F 分數列於表 5-6，我們由表 5-6 發現詞性「SHI」

的 F 分數由 0 上升到 93.49，這是由於在客語文章中只有「係」的詞性為「SHI」，而在未加入客語詞典時，「係」的詞性會被標記為中文詞典中「係」的詞性「VG」，因此無法得到正確詞性標籤造成 F 分數為 0；我們在外掛詞典加入「係」與其對應詞性「SHI」後，則可正確標記出此單字詞之詞性；而在「V」（動詞）方面，例如「VA」（動作不及物動詞）、「VB」（動作類及物動詞）等的實驗分數，實驗 B 並沒有較實驗 A 有顯著的上升，此部分是由於我們收錄詞典上的詞性標記錯誤，以及參考答案標記詞性不準確所造成的。

表 5-6 客語詞性標記實驗結果

詞性	實驗 A				實驗 B			
	精確率	召回率	F 分數	系統標記出現次數	精確率	召回率	F 分數	系統標記出現次數
A	24.42%	77.78	37.17	172	31.25%	90.91%	46.51	32
Caa	5.41%	4.08%	4.65	70	41.18%	72.41%	52.5	202
Cab	21.82%	92.31%	35.29	207	6.12%	75.00%	11.32	196
Cba	0.00%	0.00%	0	0	100.00%	50.00%	66.67	5
Cbb	84.00%	23.25%	36.42	141	85.04%	77.70%	81.2	504
D	81.94%	57.54%	67.61	3552	85.14%	82.89%	84	4949
Da	77.78%	18.42%	29.79	51	82.35%	85.71%	84	203
DE	98.61%	14.00%	24.53	274	97.09%	90.99%	93.94	1769
Dfa	75.00%	36.96%	49.51	388	89.19%	76.04%	82.09	733
Dfb	100.00%	12.50%	22.22	2	50.00%	14.29%	22.22	9
Di	46.88%	28.57%	35.5	243	65.91%	26.13%	37.42	174
Dk	100.00%	4.35%	8.33	2	100.00%	12.50%	22.22	6
I	90.91%	100.00%	95.24	21	75.00%	75.00%	75	17
Na	70.43%	84.79%	76.95	9531	81.78%	92.63%	86.87	8839
Nb	13.23%	59.11%	21.62	1721	69.77%	68.97%	69.36	340
Nc	48.85%	63.52%	55.23	996	63.74%	79.45%	70.73	719
Ncd	58.24%	14.36%	23.04	173	66.67%	59.26%	62.75	571
Nd	68.67%	52.42%	59.45	570	66.88%	66.88%	66.88	634
Nep	73.48%	43.11%	54.34	251	99.02%	75.94%	85.96	405
Neqa	24.36%	41.30%	30.65	889	34.89%	78.86%	48.38	1098

詞性	實驗 A				實驗 B			
	精確率	召回率	F 分數	系統標記 出現次數	精確率	召回率	F 分數	系統標記 出現次數
Neqb	0.00%	0.00%	0	0	0.00%	0.00%	0	0
Nes	9.65%	57.89%	16.54	217	66.67%	80.00%	72.73	25
Neu	48.45%	83.19%	61.24	369	77.59%	84.91%	81.08	229
Nf	12.08%	49.59%	19.43	1883	71.03%	63.87%	67.26	423
Ng	58.95%	47.06%	52.34	181	34.69%	58.62%	43.59	387
Nh	53.16%	31.58%	39.62	300	65.47%	85.38%	74.11	883
P	76.97%	75.08%	76.01	1664	79.42%	83.25%	81.29	1631
SHI	0.00%	0.00%	0	0	87.77%	100.00%	93.49	549
T	70.15%	76.22%	73.06	382	73.08%	84.44%	78.35	411
V_2	99.38%	100.00%	99.69	616	100.00%	99.38%	99.69	628
VA	65.60%	54.67%	59.64	1425	74.84%	63.25%	68.56	1274
VAC	14.29%	28.57%	19.05	27	66.67%	50.00%	57.14	13
VB	60.61%	50.63%	55.17	125	55.10%	64.29%	59.34	194
VC	59.20%	77.11%	66.98	4398	73.65%	80.81%	77.07	3525
VCL	57.82%	47.75%	52.31	279	59.46%	58.67%	59.06	292
VD	47.46%	37.33%	41.79	112	58.82%	24.39%	34.48	67
VE	79.93%	59.58%	68.27	540	92.52%	59.39%	72.34	581
VF	75.00%	68.75%	71.74	84	64.29%	75.00%	69.23	56
VG	15.70%	46.72%	23.51	688	75.61%	50.82%	60.78	162
VH	63.51%	65.62%	64.55	2483	84.09%	72.75%	78.01	2361
VHC	66.67%	62.50%	64.52	112	80.77%	70.00%	75	103
VI	87.50%	53.85%	66.67	15	100.00%	33.33%	50	14
VJ	43.42%	67.69%	52.91	575	50.00%	75.96%	60.31	624
VK	80.00%	54.79%	65.04	190	94.12%	72.73%	82.05	269
VL	10.68%	44.12%	17.19	531	29.89%	74.29%	42.62	345

5.2 停頓預估模型之效能分析

我們在此節將介紹停頓預估模型的訓練及測試語料，以及設計實驗對停頓預估模型作效能評估。

5.2.1 訓練及測試語料

本論文中男性語料庫經由標記及修正切割位置後，得到各音節聲母、韻母及音節後停頓資訊，我們將切割位置中 25ms 以上的停頓保留，而 25ms 秒以下皆視為沒有停頓。其中訓練語料共有 36991 個詞，測試語料包含 4095 詞。

5.2.2 實驗設計

我們在此設計了兩個實驗，以評估由條件隨機域訓練停頓預估模型的效能，分別為(a)實驗 A：詞後有標點符號則預估為有停頓。(b)實驗 B：由條件隨機域訓練模型預估詞後是否有停頓。

(a) 實驗 A：詞後有標點符號則預估為有停頓

只將詞後有標點符號的詞間邊界視為有停頓，其於詞間皆視為沒有停頓。

(b) 實驗 B：由條件隨機域訓練模型預估出詞後是否有停頓

我們將語料標記出詞長、POS、詞後是否有標點符號、詞首聲母及詞尾韻母類別，

並由 CRF++訓練出條件隨機域模型，再對語料進行停頓預估。

5.2.3 實驗結果分析

我們使用正確率和以下四種機率值來作為評估詞性分析系統的效能，其定義如下：

$$\text{正確率} = \frac{\text{系統標記是否有停頓和參考答案相同的詞數}}{\text{總詞數}} \quad (5-8)$$

$$P(1|1) = \frac{\text{預估結果有停頓}}{\text{正確答案有停頓}} \quad (5-9)$$

$$P(1|0) = \frac{\text{預估結果有停頓}}{\text{正確答案沒有停頓}} \quad (5-10)$$

$$P(0|0) = \frac{\text{預估結果沒有停頓}}{\text{正確答案沒有停頓}} \quad (5-11)$$

$$P(0|1) = \frac{\text{預估結果有停頓}}{\text{正確答案有停頓}} \quad (5-12)$$

實驗 A、實驗 B 內部及外部測試語料的實驗正確率如表 5-7 表示，而其四種機率結果列在 5-8。實驗 A、B 的 $P(1|0)$ 明顯比 $P(0|1)$ 低，這樣的結果對於文字轉語音中的停頓預估單元來說是合理的，在沒有停頓的情況下預估為有停頓的錯誤會造成合成語音的不自然；而實驗 B 的 $P(1|1)$ 較實驗 A 明顯上升許多；而在整體正確率方面，實驗 B 的內部測試及外部測試均較實驗 A 的正確率高，顯示出加入條件隨機域模型的停頓預估比只在標點符號處預估為停頓的效能提高。

表 5-7 實驗正確率

	正確率
實驗 A	85.71%
實驗 B 內部測試語料	87.04%
實驗 B 外部測試語料	88.04%

表 5-8 實驗之四種機率值

	實驗 A	實驗 B 內部測試語料	實驗 B 外部測試語料
$P(0 0)$	98.51% $\left(\frac{22970\text{個}}{23317\text{個}}\right)$	94.83% $\left(\frac{22111\text{個}}{23317\text{個}}\right)$	94.14% $\left(\frac{2409\text{個}}{2559\text{個}}\right)$
$P(1 0)$	1.49% $\left(\frac{347\text{個}}{23317\text{個}}\right)$	5.17% $\left(\frac{1206\text{個}}{23317\text{個}}\right)$	5.86% $\left(\frac{150\text{個}}{2559\text{個}}\right)$
$P(1 1)$	61.47% $\left(\frac{7569\text{個}}{12314\text{個}}\right)$	75.18% $\left(\frac{9258\text{個}}{12314\text{個}}\right)$	74.03% $\left(\frac{1035\text{個}}{1398\text{個}}\right)$
$P(0 1)$	38.53% $\left(\frac{4745\text{個}}{12314\text{個}}\right)$	24.82% $\left(\frac{3056\text{個}}{12314\text{個}}\right)$	25.97% $\left(\frac{363\text{個}}{1398\text{個}}\right)$

5.3 客語語音合成實驗結果與分析

本節將介紹 HMM 語音合成之訓練、測試語料，以及對其合成方法做效能分析，包括客觀評估。

5.3.1 實驗與測試語料

本論文採用的男性語料庫語音檔有 458 個，共有 63158 個音節，平均發音速度為 0.24syllable/second，我們從中取出 407 個音檔做為訓練語料，其餘 51 個當作測試語料。女性語料庫中共有 292 個語音檔，總音節數為 42337 個，平均發音速度為 0.22syllable/second，其中 262 個音檔為訓練語料，測試語料共有 30 個音檔。

5.3.2 客觀評估

我們採用客觀評估對模型的效能進行評估，評估的語料以及標準如表 5-9 所示：

表 5-9 客觀評估

	語料/受測者	評估標準
客觀評估	內部測試語料	RMSE
	外部測試語料	

我們採用均方根誤差(Root Mean Square Error, RMSE)作為本實驗的評估標準，計算目標和其預測的音高軌跡、音節長度以及頻譜能量之間的誤差，例如目標之音高軌跡為 $S = (s_1, s_2, \dots, s_N)$ ，而預測音高軌跡是 $T = (t_1, t_2, \dots, t_N)$ ，則 RMSE 之計算方式如下式表示：

$$RMSE(S,T) = \sqrt{\frac{\sum_{i=1}^N (s_i - t_i)^2}{N}} \quad (5-13)$$

本實驗分別對男性及女性語料庫，計算在內部測試語料和外部測試語料的音高軌跡和音長之 RMSE 值，如表 5-8 及表 5-9 所示。我們由此表看出男性語料模型在預測音高軌跡有不錯的表現，比女性語料模型佳，這是由於女性在講話時，音高軌跡範圍較廣，而且本實驗的女性發音人聲音特性較沙啞，在原始語料音高參數求取上較不準確，因此在基頻的 RMSE 值較男性語料高。

表 5-10 音長之 RMSE 值

	內部測試語料	外部測試語料
女性語料庫	37.35(Hz)	45.64(Hz)
男性語料庫	16.08(Hz)	19.4(Hz)

表 5-11 基頻之 RMSE 值

	內部測試語料	外部測試語料
女性語料庫	37.31(ms)	33.13(ms)
男性語料庫	45.13(ms)	47.57(ms)

5.4 文字轉語音系統效能分析

在此節中，我們由整個 TTS 系統所合成出的語音設計實驗，並採用主觀評分以評估本論文之客語 TTS 效能。

5.4.1 實驗方法

我們在此設計了兩個實驗，以評估客語 TTS 系統效能，以及停頓預估的有無對 TTS 系統效能的影響。這兩個實驗分別為(a)實驗 A：客語 TTS 系統中不加入停頓預估

單元。(b)實驗 B：客語 TTS 系統中加入停頓預估單元。

(a) 實驗 A：客語 TTS 系統中不加入停頓預估單元

我們將文字輸入經由斷詞系統、文本分析器以及基於 HMM 之語音合成器合成語音輸出。

(b) 實驗 B：客語 TTS 系統中加入停頓預估單元

我們將文字輸入經由斷詞系統、文本分析器、停頓預估系統以及基於 HMM 之語音合成器合成語音輸出。

5.4.2 主觀評分

在主觀評估部分，使用平均主觀分數 (Mean Opinion Score, MOS) 當做評估標準；評估的語料以及標準如表 5-12 所示：

表 5-12 主觀評估

	語料	受測者	評估標準
主觀評估	男性語料模型	男性：5 位 女性：5 位	MOS
	實驗 A：10 句		
	實驗 B：10 句		
	女性語料模型		
	實驗 A：10 句		
	實驗 B：10 句		

在男、女性語料庫實驗 A 及實驗 B 合成各 10 句語音，此文句不為內部訓練語料，由 5 位男性和 5 位女性聽得懂客語的受測者進行評估實驗，實驗時四組語音以先後次序播放，並要求受測者根據自己的觀感給予合成語音在自然度上的分數，評分標準如表 5-13：

表 5-13MOS 評分標準

評等	分數	說明
優	5	合成語音非常自然
良	4	合成語音自然
可	3	合成語音表現普通
差	2	合成語音不自然
劣	1	合成語音非常不自然

我們在 MOS 所得到男性合成語音的分數如圖 5-1 所表示：

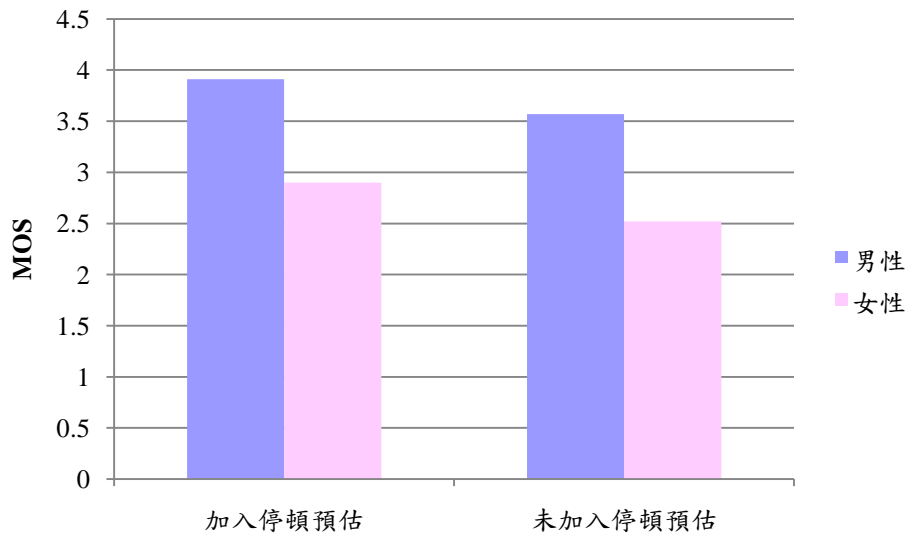


圖 5-1 實驗之 MOS 比較

我們在男性模型的 MOS 表現都較女性 MOS 表現良好，這與客觀評估女性模型在基頻的 RMSE 表現較差相對應，而加入停頓預估系統的實驗 B 分數不管是男性模型或女性模型都比實驗 A 分數高，顯示出我們在加入了停頓預估單元後，能合成出較自然的語音。

第六章 結論與未來展望

6.1 結論

本論文中，我們實現了單機版的四縣客語文字轉語音系統，其系統包含分詞與詞性標記單元、停頓預估單元以及基於HMM之語音合成器。在整個實作過程中，我們得到了以下結論：

1. 客語語料的取得是非常不容易的，因此我們在客語斷詞系統部分無法訓練客語專屬分詞及詞性標記條件隨機域模型，本論文在中文斷詞系統中加入客語外掛詞典以及客語規則，完成一套客語斷詞系統。
2. 本論文在TTS系統中加入了預估詞後是否有停頓的單元，停頓預估的加入能讓合成語音在韻律表現上更加自然。
3. 由基於HMM之語音合成器所合成的客語語音在客觀及主觀評估上都有不錯的表現，這對於客語TTS系統應用在本土語言學習中是一大幫助。

6.2 未來展望

針對本實驗使用的方法及技術，有以下幾點可以提供未來繼續深入探討並改進的方向：

1. 雖然目前在客語詞典的收錄已有一定的數量，但是跟中文詞典和文章數比起來還是非常少，我們希望能在客語詞典中收錄中文對照翻譯之資訊，進而從斷詞系統中訓練出國客語之雙語斷詞系統。
2. 國客語之雙語合成、或是輸入中文字而系統合成客語語音，這都是未來的趨勢以及我們所要努力的方向。

文獻回顧

- 【1】 K. Tokuda, T. Kobayashi and S. Imai, “Speech parameter generation from HMM using dynamic features,” *Proc. of ICASSP*, 1995.
- 【2】 龔萬灶，2004，“臺灣客家文學選集小說”。
- 【3】 Hai Zhao, Chang-Ning Huang, Mu Li, and Bao-Liang Lu. Effective tag set selection in Chinese word segmentation via conditional random field modeling. *PACLIC-20*, pp. 87–94, 2006.
- 【4】 T. Yoshimura, K. Tokuda, T. Masuko, T. Kobayashi and T. Kitamura, “Duration Modeling in HMMbased Speech Synthesis System,” *Proc. of ICSLP*, vol.2, pp.29–32, 1998.
- 【5】 K. Tokuda, T. Masuko, N. Miyazaki and T. Kobayashi, “Hidden Markov Models Based on MultiSpace Probability Distribution for Pitch Pattern Modeling,” *Proc. of ICASSP*, 1999.
- 【6】 Heiga Zen and Tomoki Toda, “An Overview of Nitech HMM-based Speech Synthesis System for Blizzard Challenge 2005,” *Proc. of InterSpeech 2005*, pp. 93-96
- 【7】 J. Lafferty, A. McCallum, and F. Pereira. “Conditional random fields: Probabilistic models for segmenting and labeling sequence data,” *Proc. of ICML*, pp.282-289, 2001.
- 【8】 Hai Zhao, Chang-Ning Huang, and Mu Li, “An Improved Chinese Word Segmentation System with Conditional Random Field,” *Proc. of SIGHAN-5*, pp.162-165, Sydney, Australia, July 22-23, 2006.

附錄一

四縣腔客語 671 個音節、17 個聲母、71 個韻母對照表

音碼	拼音	子音	母音	音碼	通用	子音	母音	音碼	通用	子音	母音
1	bi	b	i	30	bok	b	ok	59	piak	p	iak
2	be	b	e	31	buk	b	uk	60	pok	p	ok
3	ba	b	a	32	pi	p	i	61	piok	p	iok
4	bia	b	ia	33	pa	p	a	62	puk	p	uk
5	bo	b	o	34	pia	p	ia	63	mi	m	i
6	bu	b	u	35	po	p	o	64	me	m	e
7	biu	b	iu	36	pu	p	u	65	ma	m	a
8	bai	b	ai	37	piu	p	iu	66	mia	m	ia
9	boi	b	oi	38	pai	p	ai	67	mo	m	o
10	beu	b	eu	39	poi	p	oi	68	mu	m	u
11	bau	b	au	40	peu	p	eu	69	mai	m	ai
12	bin	b	in	41	pau	p	au	70	moi	m	oi
13	ben	b	en	42	pin	p	in	71	meu	m	eu
14	bien	b	ien	43	pen	p	en	72	mau	m	au
15	ban	b	an	44	pien	p	ien	73	miau	m	iau
16	bun	b	un	45	pan	p	an	74	min	m	in
17	bang	b	ang	46	pon	p	on	75	men	m	en
18	biang	b	iang	47	pun	p	un	76	mien	m	ien
19	bong	b	ong	48	pang	p	ang	77	man	m	an
20	biong	b	iong	49	piang	p	iang	78	mun	m	un
21	bung	b	ung	50	pong	p	ong	79	mang	m	ang
22	bit	b	it	51	piong	p	iong	80	miang	m	iang
23	bet	b	et	52	pung	p	ung	81	mong	m	ong
24	biet	b	iet	53	pit	p	it	82	miong	m	iong
25	bat	b	at	54	pet	p	et	83	mung	m	ung
26	bot	b	ot	55	piet	p	iet	84	mit	m	it
27	but	b	ut	56	pat	p	at	85	met	m	et
28	bak	b	ak	57	put	p	ut	86	miet	m	iet
29	biak	b	iak	58	pak	p	ak	87	mat	m	at

音碼	通用	子音	母音	音碼	通用	子音	母音	音碼	通用	子音	母音
88	mut	m	ut	117	ve	v	e	146	deu	d	eu
89	mak	m	ak	118	va	v	a	147	dau	d	au
90	mok	m	ok	119	vo	v	o	148	diau	d	iau
91	muk	m	uk	120	vu	v	u	149	dem	d	em
92	fi	f	i	121	vai	v	ai	150	dam	d	am
93	fe	f	e	122	voi	v	oi	151	diam	d	iam
94	fa	f	a	123	van	v	an	152	din	d	in
95	fo	f	o	124	von	v	on	153	den	d	en
96	fu	f	u	125	vun	v	un	154	dien	d	ien
97	fai	f	ai	126	vang	v	ang	155	dan	d	an
98	foi	f	oi	127	vong	v	ong	156	don	d	on
99	feu	f	eu	128	vung	v	ung	157	dun	d	un
100	fam	f	am	129	vet	v	et	158	dang	d	ang
101	fin	f	in	130	vat	v	at	159	dong	d	ong
102	fen	f	en	131	vut	v	ut	160	diong	d	iong
103	fan	f	an	132	vak	v	ak	161	dung	d	ung
104	fon	f	on	133	vok	v	ok	162	dep	d	ep
105	fun	f	un	134	vuk	v	uk	163	dap	d	ap
106	fang	f	ang	135	di	d	i	164	diap	d	iap
107	fong	f	ong	136	de	d	e	165	dit	d	it
108	fung	f	ung	137	da	d	a	166	det	d	et
109	fap	f	ap	138	dia	d	ia	167	diet	d	iet
110	fit	f	it	139	do	d	o	168	dat	d	at
111	fet	f	et	140	dio	d	io	169	dot	d	ot
112	fat	f	at	141	du	d	u	170	dut	d	ut
113	fok	f	ok	142	diu	d	iu	171	dak	d	ak
114	fut	f	ut	143	dai	d	ai	172	diak	d	iak
115	fuk	f	uk	144	doi	d	oi	173	dok	d	ok
116	vi	v	i	145	dui	d	ui	174	diok	d	iok

音碼	通用	子音	母音	音碼	通用	子音	母音	音碼	通用	子音	母音
175	duk	d	uk	204	tet	t	et	233	nap	n	ap
176	ti	t	i	205	tiet	t	iet	234	nit	n	it
177	te	t	e	206	tat	t	at	235	net	n	et
178	ta	t	a	207	tot	t	ot	236	nat	n	at
179	to	t	o	208	tut	t	ut	237	not	n	ot
180	tu	t	u	209	tak	t	ak	238	nak	n	ak
181	tiu	t	iu	210	tok	t	ok	239	nuk	n	uk
182	tai	t	ai	211	tuk	t	uk	240	li	l	i
183	toi	t	oi	212	ni	n	i	241	le	l	e
184	tui	t	ui	213	ne	n	e	242	la	l	a
185	teu	t	eu	214	na	n	a	243	lo	l	o
186	tiau	t	iau	215	no	n	o	244	lio	l	io
187	tiem	t	iem	216	nu	n	u	245	lu	l	u
188	tam	t	am	217	niu	n	iu	246	liu	l	iu
189	tiam	t	iam	218	nai	n	ai	247	lai	l	ai
190	tin	t	in	219	nui	n	ui	248	loi	l	oi
191	ten	t	en	220	neu	n	eu	249	lui	l	ui
192	tien	t	ien	221	nau	n	au	250	leu	l	eu
193	tan	t	an	222	niau	n	iau	251	lau	l	au
194	ton	t	on	223	nem	n	em	252	liau	l	iau
195	tun	t	un	224	nam	n	am	253	lim	l	im
196	tang	t	ang	225	nin	n	in	254	lem	l	em
197	tong	t	ong	226	nen	n	en	255	lam	l	am
198	tiong	t	iong	227	nan	n	an	256	liam	l	iam
199	tung	t	ung	228	non	n	on	257	lin	l	in
200	tep	t	ep	229	nun	n	un	258	lien	l	ien
201	tap	t	ap	230	nang	n	ang	259	lan	l	an
202	tiap	t	iap	231	nong	n	ong	260	lon	l	on
203	tit	t	it	232	nung	n	ung	261	lion	l	ion

音碼	通用	子音	母音	音碼	通用	子音	母音	音碼	通用	子音	母音
262	lun	l	un	291	zio	z	io	320	zep	z	ep
263	lang	l	ang	292	zu	z	u	321	zap	z	ap
264	liang	l	iang	293	ziu	z	iu	322	ziap	z	iap
265	long	l	ong	294	zai	z	ai	323	ziit	z	iit
266	liong	l	iong	295	zoi	z	oi	324	zit	z	it
267	lung	l	ung	296	zui	z	ui	325	zet	z	et
268	liung	l	iung	297	zeu	z	eu	326	ziet	z	iet
269	lip	l	ip	298	zau	z	au	327	zat	z	at
270	lep	l	ep	299	ziau	z	iau	328	zot	z	ot
271	lap	l	ap	300	ziim	z	iim	329	zut	z	ut
272	liap	l	iap	301	zim	z	im	330	zak	z	ak
273	lit	l	it	302	zem	z	em	331	ziak	z	iak
274	let	l	et	303	zam	z	am	332	zok	z	ok
275	liet	l	iet	304	ziam	z	iam	333	ziok	z	iok
276	lat	l	at	305	ziin	z	iin	334	zuk	z	uk
277	lot	l	ot	306	zin	z	in	335	ziuk	z	iuk
278	lut	l	ut	307	zen	z	en	336	cii	c	ii
279	lak	l	ak	308	zien	z	ien	337	ci	c	i
280	liak	l	iak	309	zan	z	an	338	ce	c	e
281	lok	l	ok	310	zon	z	on	339	ca	c	a
282	liok	l	iok	311	zun	z	un	340	cia	c	ia
283	luk	l	uk	312	zang	z	ang	341	co	c	o
284	liuk	l	iuk	313	ziang	z	iang	342	cio	c	io
285	zii	z	ii	314	zong	z	ong	343	cu	c	u
286	zi	z	i	315	ziong	z	iong	344	ciu	c	iu
287	ze	z	e	316	zung	z	ung	345	cai	c	ai
288	za	z	a	317	ziung	z	iung	346	coi	c	oi
289	zia	z	ia	318	ziip	z	iip	347	cui	c	ui
290	zo	z	o	319	zip	z	ip	348	ceu	c	eu

音碼	通用	子音	母音	音碼	通用	子音	母音	音碼	通用	子音	母音
349	cau	c	au	378	cut	c	ut	407	sen	s	en
350	ciau	c	iau	379	cak	c	ak	408	sien	s	ien
351	ciim	c	iim	380	ciak	c	iak	409	san	s	an
352	cim	c	im	381	cok	c	ok	410	son	s	on
353	cam	c	am	382	ciok	c	iok	411	sun	s	un
354	ciam	c	iam	383	cuk	c	uk	412	sang	s	ang
355	ciin	c	iin	384	ciuk	c	iuk	413	siang	s	iang
356	cin	c	in	385	sii	s	ii	414	song	s	ong
357	cen	c	en	386	si	s	i	415	siong	s	iong
358	cien	c	ien	387	se	s	e	416	sung	s	ung
359	can	c	an	388	sa	s	a	417	siung	s	iung
360	con	c	on	389	sia	s	ia	418	siip	s	iip
361	cion	c	iong	390	so	s	o	419	sip	s	ip
362	cun	c	un	391	sio	s	io	420	sep	s	ep
363	cang	c	ang	392	su	s	u	421	sap	s	ap
364	ciang	c	iang	393	siu	s	iu	422	siap	s	iap
365	cong	c	ong	394	sai	s	ai	423	siit	s	iit
366	ciong	c	iong	395	soi	s	oi	424	sit	s	it
367	cung	c	ung	396	sui	s	ui	425	set	s	et
368	ciung	c	iung	397	seu	s	eu	426	siet	s	iet
369	cip	c	ip	398	sau	s	au	427	sat	s	at
370	cap	c	ap	399	siau	s	iau	428	sot	s	ot
371	ciap	c	iap	400	siim	s	iim	429	sut	s	ut
372	ciit	c	iit	401	sim	s	im	430	sak	s	ak
373	cit	c	it	402	sem	s	em	431	siak	s	iak
374	cet	c	et	403	sam	s	am	432	sok	s	ok
375	ciet	c	iet	404	siam	s	iam	433	siok	s	iok
376	cat	c	at	405	siin	s	iin	434	suk	s	uk
377	cot	c	ot	406	sin	s	in	435	siuk	s	iuk

音碼	通用	子音	母音	音碼	通用	子音	母音	音碼	通用	子音	母音
436	gi	g	i	465	gang	g	ang	494	kia	k	ia
437	ge	g	e	466	giang	g	iang	495	kua	k	ua
438	gie	g	ie	467	gong	g	ong	496	ko	k	o
439	ga	g	a	468	giong	g	iong	497	kio	k	io
440	gia	g	ia	469	gung	g	ung	498	ku	k	u
441	gua	g	ua	470	giung	g	iung	499	kiu	k	iu
442	go	g	o	471	guang	g	uang	500	kai	k	ai
443	gio	g	io	472	gip	g	ip	501	kuai	k	uai
444	gu	g	u	473	giep	g	iep	502	koi	k	oi
445	giu	g	iu	474	gap	g	ap	503	kioi	k	ioi
446	guai	g	uai	475	giap	g	iap	504	kui	k	ui
447	goi	g	oi	476	git	g	it	505	keu	k	eu
448	gui	g	ui	477	giet	g	iet	506	kau	k	au
449	gau	g	au	478	guet	g	uet	507	kieu	k	ieu
450	gieu	g	ieu	479	gat	g	at	508	kiau	k	iau
451	giau	g	iau	480	guat	g	uat	509	kim	k	im
452	gim	g	im	481	got	g	ot	510	kam	k	am
453	giem	g	iem	482	gut	g	ut	511	kiam	k	iam
454	gam	g	am	483	gak	g	ak	512	kiem	k	iem
455	giam	g	iam	484	giak	g	iak	513	kin	k	in
456	gin	g	in	485	gok	g	ok	514	ken	k	en
457	gen	g	en	486	giok	g	iok	515	kien	k	ien
458	gien	g	ien	487	guk	g	uk	516	kan	k	an
459	guen	g	uen	488	giuk	g	iuk	517	kuan	k	uan
460	gan	g	an	489	guak	g	uak	518	kon	k	on
461	guan	g	uan	490	ki	k	i	519	kun	k	un
462	gon	g	on	491	kie	k	ie	520	kiun	k	iun
463	gun	g	un	492	kue	k	ue	521	kang	k	ang
464	giun	g	iun	493	ka	k	a	522	kiang	k	iang

音碼	通用	子音	母音	音碼	通用	子音	母音	音碼	通用	子音	母音
523	kong	k	ong	552	nguai	ng	uai	581	ngok	ng	ok
524	kiong	k	iong	553	ngoi	ng	oi	582	ngiok	ng	iok
525	kung	k	ung	554	ngui	ng	ui	583	ngiuk	ng	iuk
526	kiung	k	iung	555	ngau	ng	au	584	hi	h	i
527	kip	k	ip	556	ngieu	ng	ieu	585	he	h	e
528	kep	k	ep	557	ngiau	ng	iau	586	ha	h	a
529	kap	k	ap	558	ngim	ng	im	587	hia	h	ia
530	kiap	k	iap	559	ngiem	ng	iem	588	ho	h	o
531	kit	k	it	560	ngam	ng	am	589	hio	h	io
532	kiet	k	iet	561	ngiam	ng	iam	590	hiu	h	iu
533	kat	k	at	562	ngin	ng	in	591	hai	h	ai
534	kut	k	ut	563	ngien	ng	ien	592	hoi	h	oi
535	kiut	k	iut	564	ngan	ng	an	593	heu	h	eu
536	kak	k	ak	565	nguan	ng	uan	594	hau	h	au
537	kiak	k	iak	566	ngion	ng	ion	595	hieu	h	ieu
538	kok	k	ok	567	ngiun	ng	iun	596	hiau	h	iau
539	kiok	k	iok	568	ngang	ng	ang	597	him	h	im
540	kuk	k	uk	569	ngiang	ng	iang	598	hem	h	em
541	kiuk	k	iuk	570	ngong	ng	ong	599	ham	h	am
542	kuak	k	uak	571	ngiong	ng	iong	600	hiam	h	iam
543	ngi	ng	i	572	ngiung	ng	iung	601	hin	h	in
544	ngie	ng	ie	573	ngip	ng	ip	602	hen	h	en
545	nga	ng	a	574	ngap	ng	ap	603	hien	h	ien
546	ngia	ng	ia	575	ngiap	ng	iap	604	han	h	an
547	ngo	ng	o	576	ngit	ng	it	605	hon	h	on
548	ngio	ng	io	577	ngiet	ng	iet	606	hiun	h	iun
549	ngu	ng	u	578	ngat	ng	at	607	hang	h	ang
550	ngiu	ng	iu	579	ngut	ng	ut	608	hong	h	ong
551	ngai	ng	ai	580	ngiak	ng	iak	609	hiong	h	iong

音碼	通用	子音	母音	音碼	通用	子音	母音	音碼	通用	子音	母音
610	hiung	h	iung	639	am	INULL	am	668	ng	ng	FNULL
611	hiing	h	iing	640	iam	INULL	iam	669	iui	INULL	iui
612	hip	h	ip	641	in	INULL	in	670	ngung	ng	ung
613	hap	h	ap	642	en	INULL	en	671	nguk	ng	uk
614	hiap	h	iap	643	ien	INULL	ien				
615	hit	h	it	644	an	INULL	an				
616	het	h	et	645	on	INULL	on				
617	hiet	h	iet	646	iun	INULL	iun				
618	hat	h	at	647	ang	INULL	ang				
619	hot	h	ot	648	iang	INULL	iang				
620	hak	h	ak	649	ong	INULL	ong				
621	hok	h	ok	650	iong	INULL	iong				
622	hiuk	h	iuk	651	iung	INULL	iung				
623	i	INULL	i	652	ip	INULL	ip				
624	e	INULL	e	653	ep	INULL	ep				
625	ie	INULL	ie	654	ap	INULL	ap				
626	a	INULL	a	655	iap	INULL	iap				
627	ia	INULL	ia	656	it	INULL	it				
628	o	INULL	o	657	et	INULL	et				
629	io	INULL	io	658	iet	INULL	iet				
630	iu	INULL	iu	659	at	INULL	at				
631	ai	INULL	ai	660	ot	INULL	ot				
632	oi	INULL	oi	661	iut	INULL	iut				
633	eu	INULL	eu	662	ak	INULL	ak				
634	au	INULL	au	663	iak	INULL	iak				
635	ieu	INULL	ieu	664	ok	INULL	ok				
636	iau	INULL	iau	665	iok	INULL	iok				
637	im	INULL	im	666	iuk	INULL	iuk				
638	em	INULL	em	667	m	m	FNULL				

附錄二

新增至客語詞典之一字多義一字詞表

客語詞	詞性組合	客語詞	詞性組合
個	DE	肚	Ncd_Ng
愛	D_VC_VE	當	Dfa_P
係	SHI	儕	Na_Nf
等	Di_VC	同	Caa_P
該	Dk_Ncd_Nep_T	歸	Nb_Neqa
異	Dfa	畜	Na_VC
到	Caa_P_VCL	摻	Caa_P
正	Da_D	屋	Na_Nc_Nf
項	Ncd_Nf_Ng	適	P_VJ
食	Na_VC	歇	VA_VCL
好	D_VH	驚	Na_VK
逐	D_Nes	企	Nc_VA
諉	Di_T_VJ	識	D_VJ
盡	Dfa	喊	P_VA_VC_VE_VF_VG_VL
分	P_VD_VL	供	VC_VH
擺	Na_Nf_VC	奈	Ncd_Nep_VJ_T
著	P_VC	遽	D_VH
恁	D_Dfa_VD	忒	Dfa_Di_VCL_VH
兜	Dfb_Nf	仰	D_VC_VH
仔	DE_Na_Nf		

附錄三

詞類表

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

