

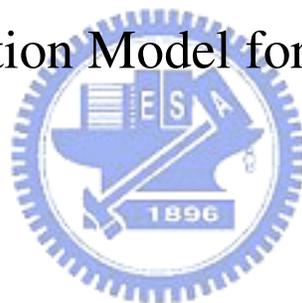
國立交通大學

電信工程學系碩士班

碩士論文

漢語語音音節長度模型之研究

A Syllable Duration Model for Mandarin Speech



研究生：簡家勇

指導教授：王逸如 博士

中華民國九十六年九月



漢語語音音節長度模型之研究
A Syllable Duration Model for Mandarin Speech

研究生：簡家勇

Student : Chia-Yung Chien

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

碩士論文



Submitted to Department of Communication Engineering

College of Electrical and Computer Engineering

National Chiao Tung University

in Partial Fulfillment of the Requirements

for the Degree of

Master of Science

in Communication Engineering

September 2007

Hsinchu, Taiwan, Republic of China

中華民國九十六年九月



漢語語音音節長度模型之研究

研究生：簡家勇

指導教授：王逸如 博士

國立交通大學電信工程學系

中文摘要

在本論文中我們對單一語者的閱讀語音說話模式建立了一個以音節為單位的音長模型，在模型中我們考慮了數個會影響到音長的因素，包含了每一個語者整體說話之平均音長特性的說話風格、低階音節層次的基本音節和聲調與代表著高階語言層次音長特性的韻律狀態。而對於模型的建立方面，我們以一個線性相加模型來納入這些影響因素，並且以循序最佳化程序的方式搭配著最大概似法則來訓練模型直到收斂。而後我們對模型所產生的參數做了分析，以觀察其是否符合我們所認知的語音特性，而對於韻律狀態方面我們除了分析其自身的特性，也搭配著在[11]中作者以韻律訊息中的音高訊息和停頓訊息與一些語言參數，對語料庫所標示的音節間之停頓類型的 Break Type 做了相關性分析，而 Break Type 也就是在[2]中作者所述的漢語語音階層韻律結構的標記。接著我們也將[11]中對語料庫的每一個音節所標示的音高韻律狀態和本論文的音長韻律狀態做了相關性分析，以觀察高階語言層次的音高和音長訊息之間的相關性。最後我們對去掉掉語者、基本音節和聲調之影響的音長(其亦代表著高階語言層次音長特性)，搭配著[2]中所述的漢語語音階層韻律結構建立一個高階語言層次階層音長模型。而對於模型的建立方面，我們同樣以一個線性相加模型來納入這些階層的影響因素，並且以循序最佳化程序的方式搭配著最小平方誤差法則來訓練模型直到收斂。而後我們對模型所產生的參數做了分析，並且呈現出各個階層的音長軌跡，以觀察其是否符合我們所認知的語音特性。



A Syllable Duration Model for Mandarin Speech

Student : Chia-Yung Chien

Advisor : Dr. Yih-Ru Wang

Department of Communication Engineering
National Chiao Tung University

Abstract

In this thesis, we construct a syllable duration model for single-speaker Mandarin read speech. In this model, several affecting factors which control the variation of syllable duration are considered. They include the global mean of each speaker, base-syllable, tone, and prosodic state which represents the influences from all high-level linguistic features. Those affecting factors are combined additively. The model is trained by a sequential optimization procedure with maximum likelihood criterion. After well training, we analyze the model parameters to see whether they conform to our a priori knowledge about Mandarin speech prosody. The relationship between the prosodic state and the inter-syllable break type, labeled by [11], are also explored. Besides, the collocation of the prosodic state and the pitch prosodic state derived by [11] are analyzed. Lastly, we construct a hierarchial duration model using the features combining the prosodic-state affecting factor and the residuals. The model is composed of three layers: PW, MIPPH, and MPPH. The affecting patterns of these three layers are length-dependent and assumed to be combined additively. The model is trained by a sequential optimization procedure with minimum squared error criterion. After well training, we analyze the resulting affecting patterns to see whether they conform to our a priori knowledge about Mandarin speech prosody. Experimental results confirmed that they conformed to those explored by Tseng [2].



誌謝

終於完成了這一本論文，回想過去這一路走來並不是挺順利的，不過終究是走過了這一段人生旅程。首先要感謝陳信宏老師和王逸如老師的指導，才能順利完成這一本論文；接著要感謝江振宇學長提供了我一個方向，讓我可以朝著這個方向進行研究，進而完成這本論文；再來要感謝實驗室的戰友們一同努力奮鬥，特別是和世帆與振豐彼此互相鼓舞、打拼；再者要感謝學弟們的陪伴，尤其是啟風，一起修課時麻煩了你不少，謝啦；最後要感謝我的家人和親人，支持我完成碩士學業，陪伴我度過種種難關。





目錄

中文摘要.....	I
英文摘要.....	III
誌謝.....	V
目錄.....	VII
表目錄.....	XIII
圖目錄.....	XVII
第一章 緒論.....	1
1.1 研究動機.....	1
1.2 研究方向.....	1
1.3 章節概要說明.....	1
第二章 漢語語音特性和漢語 TTS 系統概述.....	5
2.1 漢語語音特性.....	5
2.2 漢語 TTS 系統概述.....	7
2.2.1 文字分析器.....	8
2.2.2 韻律產生器.....	8
2.2.3 韻律模型.....	8
2.2.3.1 規則法.....	9
2.2.3.2 資料驅動法.....	9
2.2.3.2.1 RNN 韻律模型.....	9
2.2.4 語音合成器.....	11
2.2.5 411 基本音節語音波形樣本資料庫.....	11
第三章 音長模型的原理和數學說明.....	13
3.1 研究目標.....	13
3.2 文獻回顧.....	13
3.2.1 漢語語音階層韻律模型.....	13

3.2.2 潛藏韻律模擬法	13
3.3 語音資料庫的介紹	15
3.4 音長模型的原理說明	16
3.4.1 語者類型的音長特性	17
3.4.2 基本音節類型的音長特性	17
3.4.3 聲調類型的音長特性	18
3.4.4 高階語言層次參數的音長特性	18
3.5 音長模型的數學說明	18
3.5.1 目標函數的數學說明	20
3.5.2 語者類型影響因素的數學說明	21
3.5.3 聲調類型影響因素的數學說明	22
3.5.4 基本音節類型影響因素的數學說明	22
3.5.5 韻律狀態類型影響因素的數學說明	23
3.5.6 殘餘音長之變異數的數學說明	24
3.6 音長模型之建立的流程說明	25
3.6.1 模型參數的初始值設定	25
3.6.1.1 聲調類型影響因素的初始值設定	25
3.6.1.2 基本音節類型影響因素的初始值設定	25
3.6.1.3 韻律狀態類型的初始設定	25
3.6.1.3.1 韻律狀態類型影響因素的初始值設定	26
3.6.1.3.2 每一個音節所屬的韻律狀態類型的初始標示	26
3.6.1.4 殘餘音長之變異數的初始值設定	26
3.6.1.5 目標函數之概似值的初始值計算	26
3.6.2 模型的訓練方式	26
3.6.2.1 聲調類型影響因素的更新數學式	27
3.6.2.2 基本音節類型影響因素的更新數學式	27

3.6.2.3 韻律狀態類型的更新	27
3.6.2.3.1 更新每一個音節所屬的韻律狀態類型標示	27
3.6.2.3.2 韻律狀態類型影響因素的更新數學式	28
3.6.2.4 殘餘音長之變異數的更新數學式	28
3.6.2.5 目標函數之概似值的更新數學式	28
3.6.2.6 模型訓練的收斂條件	28
3.6.2.7 模型訓練的流程	28
第四章 音長模型的實驗結果分析	31
4.1 影響因素	31
4.1.1 語者類型影響因素	31
4.1.2 聲調類型影響因素	31
4.1.2.1 實驗結果分析	31
4.1.3 基本音節類型影響因素	32
4.1.3.1 實驗結果分析	32
4.1.3.1.1 和實際音長的比較	32
4.1.3.1.2 決策樹分析	32
4.1.3.1.2.1 問題集	32
4.1.3.1.2.2 分裂準則	35
4.1.3.1.2.3 樹的長成	35
4.1.3.1.2.4 基本音節類型影響因素的決策樹分析	36
4.1.4 韻律狀態類型影響因素	38
4.1.4.1 實驗結果分析	39
4.2 殘餘音長之變異數	39
4.2.1 初始值	39
4.2.2 收斂值	39
4.2.3 實驗結果分析	39

4.3 目標函數之概似值.....	40
4.3.1 初始值.....	40
4.3.2 收斂值.....	40
4.3.3 更新曲線.....	40
4.3.4 實驗結果分析.....	40
4.4 音長分佈和音長軌跡.....	41
4.4.1 音長分佈.....	41
4.4.1.1 實驗結果分析.....	43
4.4.2 音長軌跡.....	43
4.4.2.1 實驗結果分析.....	47
第五章 音節之韻律狀態類型的分析.....	49
5.1 單連文(Unigram)機率.....	49
5.1.1 實驗結果分析.....	50
5.2 雙連文(Bigram)機率.....	50
5.2.1 實驗結果分析.....	52
5.3 和 Break Type 的相關性觀察.....	53
5.3.1 Break Type 的介紹.....	53
5.3.1.1 漢語語音階層韻律結構.....	53
5.3.1.2 Break Type 的標記.....	54
5.3.1.3 Break Type 於本論文的應用.....	56
5.3.2 音長軌跡和 Break Type 的相關性觀察.....	57
5.3.2.1 實驗結果分析.....	58
5.3.3 和前 Break Type 的雙連文機率.....	58
5.3.3.1 實驗結果分析.....	61
5.3.4 和後 Break Type 的雙連文機率.....	61
5.3.4.1 實驗結果分析.....	63

5.3.5 和前後 Break Type 的結合個數.....	63
5.3.5.1 實驗結果分析.....	65
5.4 音長和音高韻律狀態類型的相關性觀察.....	65
5.4.1 實驗結果分析.....	67
第六章 高階音長模型的原理和數學說明.....	69
6.1 高階語言層次韻律訊息.....	69
6.1.1 音高訊息.....	69
6.1.2 音量訊息.....	69
6.1.3 音長訊息.....	70
6.1.4 停頓訊息.....	70
6.2 高階音長模型的原理說明.....	70
6.2.1 漢語語法結構.....	71
6.2.2 漢語韻律結構.....	71
6.2.3 高階語言層次階層音長軌跡現象.....	71
6.2.4 高階語言層次階層音長軌跡示意圖.....	76
6.3 高階音長模型的數學說明.....	78
6.3.1 目標函數的數學說明.....	79
6.3.2 PW 類型影響因素的數學說明.....	79
6.3.3 MIPPH 類型影響因素的數學說明.....	80
6.3.4 MPPH 類型影響因素的數學說明.....	81
6.4 高階音長模型之建立的流程說明.....	82
6.4.1 模型參數的初始值設定.....	82
6.4.1.1 PW 類型影響因素的初始值設定.....	82
6.4.1.2 MIPPH 類型影響因素的初始值設定.....	82
6.4.1.3 MPPH 類型影響因素的初始值設定.....	83
6.4.1.4 目標函數的初始值計算.....	83

6.4.2 模型的訓練方式	83
6.4.2.1 PW 類型影響因素的更新數學式.....	83
6.4.2.2 MIPPH 類型影響因素的更新數學式.....	84
6.4.2.3 MPPH 類型影響因素的更新數學式	84
6.4.2.4 目標函數的更新數學式	84
6.4.2.5 模型訓練的收斂條件	84
6.4.2.6 模型訓練的流程.....	85
第七章 高階音長模型的實驗結果分析.....	87
7.1 影響因素	87
7.1.1 PW 類型影響因素	87
7.1.1.1 實驗結果分析.....	89
7.1.2 MIPPH 類型影響因素	89
7.1.2.1 實驗結果分析.....	93
7.1.3 MPPH 類型影響因素	93
7.1.3.1 實驗結果分析.....	101
7.2 目標函數	101
7.2.1 初始值.....	101
7.2.2 收斂值.....	102
7.2.3 更新曲線.....	102
7.2.4 實驗結果分析	102
7.3 音長分佈和音長軌跡.....	102
7.3.1 音長分佈	103
7.3.1.1 實驗結果分析.....	104
7.3.2 音長軌跡.....	104
7.3.2.1 實驗結果分析.....	106
第八章 結論與未來展望.....	107



8.1 結論	107
8.2 未來展望	107
參考文獻	109
附錄一	111
附錄二	117
附錄三	119
附錄四	121
附錄五	127
附錄六	133
附錄七	139
附錄八	141





表目錄

表二.1：漢語聲母之發音方法及發音部位對應表。	6
表二.2：漢語韻母之發音方法及發音部位對應表。	6
表三.1：語料庫中每一種聲調類型的音節個數。	15
表三.2：語料庫中每一種聲調類型之音節音長的平均值。	15
表三.3：語料庫中一字詞到十二字詞的個數。	16
表四.1： γ_i 的值。	31
表四.2： γ_{ps} 的值。	38
表五.1： $C(ps)$ 。	49
表五.2： $P(ps)$ 。	50
表五.3： $C(ps_{n-1}, ps_n)$ 。	51
表五.4： $P(ps_n ps_{n-1})$ ，其單位為 10^{-3} 。	52
表五.5：Break Type 和 b 值的對應表。	56
表五.6：改變順序後的 Break Type 和 b 值對應表。	56
表五.7： $C(b_{n-1})$ 。	59
表五.8： $C(b_{n-1}, ps_n)$ 。	59
表五.9： $P(ps_n b_{n-1})$ ，其單位為 10^{-4} 。	60
表五.10： $C(b_n)$ 。	61
表五.11： $C(ps_n, b_n)$ 。	62
表五.12： $P(ps_n b_n)$ ，其單位為 10^{-4} 。	62
表五.13： $C(b_{n-1}, ps_n, b_n)$ 。	64
表五.14： $C(pps)$ 。	66
表五.15： $C(pps, dps)$ 。	66



圖目錄

圖二.1：漢語四個聲調(輕聲除外)的音高軌跡和時間的關係圖。	6
圖二.2：漢語 TTS 系統的基本架構圖。	7
圖二.3：RNN 韻律模型的基本架構圖。	10
圖三.1：語料庫中每一種聲調類型之音節音長的平均值。	16
圖三.2：音長模型的訓練流程方塊圖。	30
圖四.1： γ_t 的值。	31
圖四.2：節點分裂的示意圖。	36
圖四.3： γ_{bs} 的決策樹分析。	37
圖四.4： γ_{ps} 的值。	39
圖四.5：目標函數 Q 之概似值的更新曲線圖。	40
圖四.6： SD 的分佈圖。	41
圖四.7： $SD1$ 的分佈圖。	42
圖四.8： $SD2$ 的分佈圖。	42
圖四.9： $SD3$ 的分佈圖。	42
圖四.10： $SD4$ 的分佈圖。	43
圖四.11：語料庫中一段句子的 SD 、 $SD1$ 、 $SD2$ 、 $SD3$ 和 $SD4$ 軌跡圖。	44
圖四.12：相同一段句子的 SD 軌跡和 LWB 與 PM 之間的關係圖。	44
圖四.13：另一段句子的 SD 軌跡和 LWB 與 PM 之間的關係圖。	45
圖四.14：第三段句子的 SD 軌跡和 LWB 與 PM 之間的關係圖。	45
圖四.15：圖四.12 的 $SD3$ 軌跡，並且標記上每一個音節所屬的韻律狀態類型。	46
圖四.16：圖四.13 的 $SD3$ 軌跡，並且標記上每一個音節所屬的韻律狀態類型。	47
圖四.17：圖四.14 的 $SD3$ 軌跡，並且標記上每一個音節所屬的韻律狀態類型。	47
圖五.1： $C(ps)$ 。	50
圖五.2： $P(ps)$ 。	50
圖五.3： $C(ps_{n-1}, ps_n)$ 。	51

圖五.4： $P(ps_n ps_{n-1})$ 。	52
圖五.5： 漢語語音階層韻律結構。	53
圖五.6： 音長和四個影響因素與 Break Type 的數學符號關係圖。	56
圖五.7： 將圖四.15 標記上 Break Type。	57
圖五.8： 將圖四.16 標記上 Break Type。	57
圖五.9： 將圖四.17 標記上 Break Type。	58
圖五.10： $C(b_{n-1})$ 。	59
圖五.11： $C(b_{n-1}, ps_n)$ 。	60
圖五.12： $P(ps_n b_{n-1})$ 。	60
圖五.13： $C(b_n)$ 。	61
圖五.14： $C(ps_n, b_n)$ 。	62
圖五.15： $P(ps_n b_n)$ 。	63
圖五.16： $C(b_{n-1}, ps_n, b_n)$ 。	65
圖五.17： $C(pps)$ 。	66
圖五.18： $C(pps, dps)$ 。	67
圖六.1： 音長的漢語語音階層韻律結構。	71
圖六.2： 將圖五.7 去除掉 B_0 和 B_1 ，並且將 B_2-1 和 B_2-2 結合為 B_2 。	72
圖六.3： 將圖五.8 去除掉 B_0 和 B_1 ，並且將 B_2-1 和 B_2-2 結合為 B_2 。	72
圖六.4： 將圖五.9 去除掉 B_0 和 B_1 ，並且將 B_2-1 和 B_2-2 結合為 B_2 。	73
圖六.5： 將圖六.2 去除掉 B_2 。	74
圖六.6： 將圖六.3 去除掉 B_2 。	74
圖六.7： 將圖六.4 去除掉 B_2 。	75
圖六.8： 將圖六.5 去除掉 B_3 。	75
圖六.9： 將圖六.6 去除掉 B_3 。	76
圖六.10： 將圖六.7 去除掉 B_3 。	76
圖六.11： 漢語語音階層韻律單位之音長軌跡 Pattern 示意圖。(a).原始音長軌跡，	

(b).MPPH 之音長軌跡 Pattern , (c). MIPPH 之音長軌跡 Pattern , (d). PW 之音長軌跡 Pattern 。	77
圖六.12：一句 13 個字的 MPPH 與 MIPPH 和 PW 之間的關係圖。	78
圖六.13：高階音長模型的訓練流程方塊圖。	86
圖七.1： γ_{PW} 的值(一字~四字)。	87
圖七.2： γ_{PW} 的值(五字~八字)。	88
圖七.3： γ_{PW} 的值(九字~十二字)。	88
圖七.4： γ_{PW} 的值(十三字~十五字)。	89
圖七.5： γ_{MIPPH} 的值(一字~四字)。	90
圖七.6： γ_{MIPPH} 的值(五字~八字)。	90
圖七.7： γ_{MIPPH} 的值(九字~十二字)。	91
圖七.8： γ_{MIPPH} 的值(十三字~十六字)。	91
圖七.9： γ_{MIPPH} 的值(十七字~二十字)。	92
圖七.10： γ_{MIPPH} 的值(二十一字~二十四字)。	92
圖七.11： γ_{MIPPH} 的值(二十五字、二十八字)。	93
圖七.12： γ_{MPPH} 的值(一字~四字)。	94
圖七.13： γ_{MPPH} 的值(五字~八字)。	94
圖七.14： γ_{MPPH} 的值(九字~十二字)。	95
圖七.15： γ_{MPPH} 的值(十三字~十六字)。	95
圖七.16： γ_{MPPH} 的值(十七字~二十字)。	96
圖七.17： γ_{MPPH} 的值(二十一字~二十四字)。	96
圖七.18： γ_{MPPH} 的值(二十五字~二十八字)。	97
圖七.19： γ_{MPPH} 的值(二十九字~三十二字)。	97
圖七.20： γ_{MPPH} 的值(三十三字~三十六字)。	98
圖七.21： γ_{MPPH} 的值(三十七字~四十字)。	98
圖七.22： γ_{MPPH} 的值(四十一字~四十四字)。	99

圖七.23： γ_{MPPH} 的值(四十五字~四十八字)。	99
圖七.24： γ_{MPPH} 的值(四十九字~五十二字)。	100
圖七.25： γ_{MPPH} 的值(五十四字、五十九字、六十一字、六十二字)。	100
圖七.26： γ_{MPPH} 的值(六十三字)。	101
圖七.27： 目標函數 Q'' 的更新曲線圖。	102
圖七.28： $SD3$ 的分佈圖。	103
圖七.29： $SD3_1$ 的分佈圖。	103
圖七.30： $SD3_2$ 的分佈圖。	104
圖七.31： $SD3_3$ 的分佈圖。	104
圖七.32： 圖四.15 的 $SD3$ 、 $SD3_1$ 、 $SD3_2$ 和 $SD3_3$ 軌跡圖。	105
圖七.33： 圖四.16 的 $SD3$ 、 $SD3_1$ 、 $SD3_2$ 和 $SD3_3$ 軌跡圖。	105
圖七.34： 圖四.17 的 $SD3$ 、 $SD3_1$ 、 $SD3_2$ 和 $SD3_3$ 軌跡圖。	106



第一章 緒論

本章首先對於論文的研究動機做呈述，而後對於研究方向做闡述，最後再對於論文的各個章節做概要說明。

1.1 研究動機

語言是人類最自然的訊息傳遞模式，人和人之間是如此，但是人和機器之間長久以來卻是透過其它的媒介來做訊息傳遞，例如人類使用滑鼠和鍵盤來控制電腦。因此長久以來科學家便不斷地投身研究希望能夠創造出以語音為溝通媒介的語音人機介面，而要使機器能夠如人類般做訊息的溝通首要的當然是給機器一對耳朵和一張嘴巴。一對耳朵要使得機器能夠聽懂人類說的話，此即為語音辨識技術；一張嘴巴要使得機器能夠說出它想說的話，此即為語音合成技術，亦或稱為文字轉語音合成(Text-to-Speech Synthesis, TTS)技術。而本論文的研究動機便是在漢語(Mandarin)的 TTS 技術上，希望能夠對於交通大學電信工程學系語音處理實驗室所發展的漢語 TTS 系統的合成語音品質能夠有所提升。

1.2 研究方向

一套 TTS 系統要想合成出流利的語句，韻律產生器便是關鍵的所在，而韻律產生器的優劣又是取決於所建立的韻律模型，因此本論文的研究方向便是在於韻律模型中的音長訊息。我們所針對的音長單位為音節，所針對的說話模式為閱讀語音模式。而對於音節音長的影響方面基本音節和聲調是早已眾所皆知，再者每一個人說話也有其之說話速度快慢的風格，其也就影響到了每一個音節的音長。除了這些低階音節層次參數的影響之外，我們的研究方向也著重於高階語言層次參數對於音長的影響。每一個音節的音長在不同詞中的位置會不同，在不同句中的位置亦會不同，也會隨著週遭是否有標點符號而不同，此即高階語言層次參數對於音長的影響。而高階語言層次的音高和音長訊息是否有相關性，也會在本論文中對其做分析。最後本論文會針對代表著高階語言層次的音長建立一個階層式的高階音長模型，以呈現出在各個階層中的高階音長軌跡特性。

1.3 章節概要說明

本論文的内容共分为七章，我们将各章的内容以条列式的方式做概略性地呈述。

第一章 緒論：

本章首先對於論文的研究動機做呈述，而後對於研究方向做闡述，最後再對於論文的各個章節做概要說明。

第二章 漢語語音特性和漢語 TTS 系統概述：

本章對於漢語語音特性和本實驗室過去所發展的漢語 TTS 系統做概略性地介紹，而對於漢語 TTS 系統尤其針對韻律產生器的部份將做較詳細的說明，因為韻律方面是本論文的研究重點。

第三章 音長模型的原理和數學說明：

在這章中首先對於本論文的研究目標做闡述，而後再回顧一些過去在此研究方向的相關文獻內容，最後針對所要建立的音長模型做原理和數學的說明。

第四章 音長模型的實驗結果分析：

本章對於所建立的音長模型在符合收斂條件狀態下分析其參數，以觀察是否符合我們所知道的各種語音特性，藉此來判斷所建立的模型是否正確。

第五章 音節之韻律狀態類型的分析：

本章對於音長模型在收斂狀態下對語料庫中每一個音節所標示的韻律狀態類型做分析，我們希望能從中觀察出其代表著高階語言層次之音長特性的現象。我們也使用參考文獻[11]裡作者以韻律訊息中的音高訊息和停頓訊息以及一些語言參數對本語料庫所標示的音節間之停頓類型的 Break Type，來觀察其和音節之韻律狀態類型的相關性。最後我們也使用文獻裡作者對語料庫中每一個音節所標示的音高韻律狀態類型，來觀察其和本論文的音長韻律狀態類型之相關性。

第六章 高階音長模型的原理和數學說明：

本章依據第五章所述的漢語語音階層韻律結構搭配著所觀察到的音長軌跡現象建立一個高階音長模型，以數學的方式呈現出每一個階層之音長軌跡的

Pattern。本章內容首先對於高階語言層次的韻律訊息做介紹，而後對於模型的原理和數學做說明。

第七章 高階音長模型的實驗結果分析：

本章對於所建立的高階音長模型在符合收斂條件狀態下分析其參數，以觀察是否符合我們所認知的語音特性，藉此來判斷所建立的模型是否正確，並且呈現出各個漢語語音階層韻律結構的階層之高階音長軌跡 Pattern。

第八章 結論與未來展望：

本章首先對於本論文所完成的工作做整體性的回顧和結論，而後提出幾個方向以在未來能夠做更進一步的研究。





第二章 漢語語音特性和漢語 TTS 系統 概述

本章對於漢語語音特性和本實驗室過去所發展的漢語 TTS 系統做概略性地介紹，而對於漢語 TTS 系統尤其針對韻律產生器的部份將做較詳細的說明，因為韻律方面是本論文的研究重點。

2.1 漢語語音特性

漢語的音系特點不同於英語為音節(Syllable)界線分明和音節帶有聲調，其發音單位(Pronunciation Unit)是帶有聲調(Tone)的基本音節(Base-Syllable)。漢語在語言上的單位為字(Character)而在語音上的單位是音節，也就是字的發音一個字對應一個音節。漢語文字常見的有 12000 多個，但在發音上同音字很多，總共的音節約略是 1300 個。漢語的音節包含基本音節和聲調兩個部份，其分別有 411 個基本音節類型和五個聲調類型，我們將漢語 411 個基本音節碼與對應的注音符號、22 個聲母碼、40 個韻母碼列於附錄一。

音節的基本音節包含聲母(亦可稱為子音或輔音)和韻母(亦可稱為母音或元音)兩個部份，聲母是位於音節開頭的位置因此其英文為 Initial 也就是首音的意思，如果一個音節是屬於沒有聲母者則稱其為空聲母的音節。連同空聲母漢語共有 22 個聲母，我們將漢語 22 個聲母碼與對應的注音符號列於附錄二，各聲母之發音方法、發音部位、清音或是濁音和送氣與否的對應表如表二.1；韻母是位於音節尾端的位置因此英文稱為 Final 也就是尾音的意思，音節中聲母之後的部份即為韻母，如果一個音節是屬於沒有韻母者則稱其為空韻母的音節。連同空韻母漢語共有 40 個韻母其中兩個為不同類型的空韻母，我們將漢語 40 個韻母碼與對應的注音符號列於附錄三，韻母之發音方法、發音部位和組成結構的對應表如表二.2 所示。

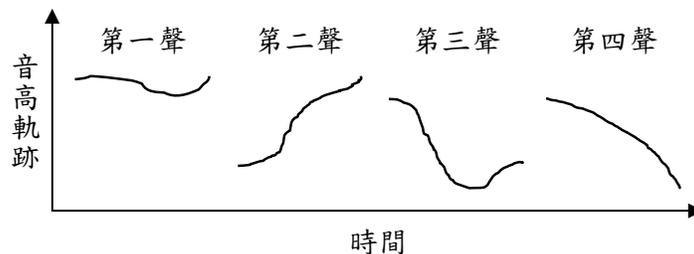
表二.1：漢語聲母之發音方法及發音部位對應表。

發音方法	清音/濁音	送氣/不送氣	發音部位					
			唇	齒	齒齦	捲舌	齶	軟齶
爆破音	清音	不送氣	ㄅ		ㄆ			ㄏ
		送氣	ㄆ		ㄆ			ㄏ
鼻音	濁音		ㄇ		ㄋ			
摩擦音	清音		ㄈ	ㄇ		ㄨ	ㄒ	
	濁音					ㄣ		
邊音	濁音				ㄌ			
塞擦音	清音	不送氣		ㄗ		ㄘ	ㄙ	
		送氣		ㄗ		ㄘ	ㄙ	

表二.2：漢語韻母之發音方法及發音部位對應表。

發音方法	單韻母	複韻母	聲隨韻母
開口呼	ㄚ ㄛ ㄜ ㄝ 兒	ㄚ ㄛ ㄜ ㄝ	ㄛ ㄜ ㄝ ㄚ
齊齒呼	一	一ㄚ 一ㄛ 一ㄜ 一ㄝ 一ㄞ 一ㄟ	一ㄛ 一ㄜ 一ㄝ 一ㄚ
合口呼	ㄨ	ㄨㄚ ㄨㄛ ㄨㄜ ㄨㄝ ㄨㄞ ㄨㄟ	ㄨㄛ ㄨㄜ ㄨㄝ ㄨㄚ
撮口呼	ㄨ	ㄨㄛ ㄨㄜ	ㄨㄛ ㄨㄜ ㄨㄝ

音節具有聲調是漢語的一項重要特徵，漢語有五種聲調，分別是第一聲、第二聲、第三聲、第四聲和第五聲(輕聲)，聲調之主要特徵是由音高軌跡(Pitch Contour)形狀、聲音響度(Loudness)和音長(Duration)來呈現，我們將漢語四個聲調(輕聲除外)的音高軌跡和時間的關係繪出如下的圖二.1所示。



圖二.1：漢語四個聲調(輕聲除外)的音高軌跡和時間的關係圖。

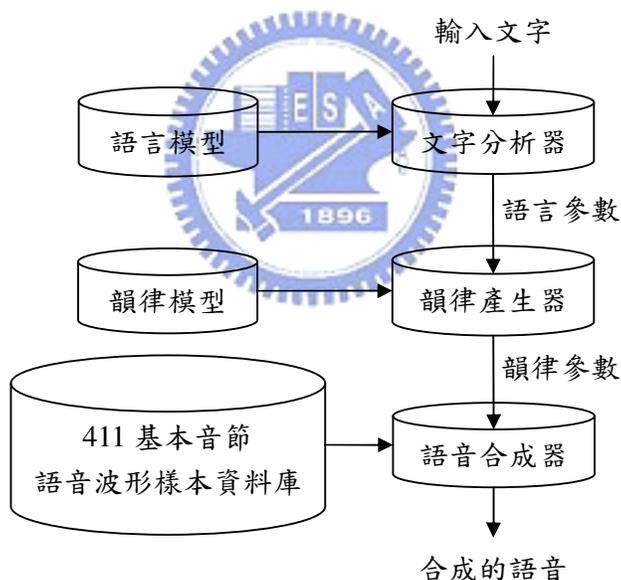
至於第五聲，其音高軌跡非常容易受到前後音節聲調的影響，因此它沒有固定的音高軌

跡形狀，而且它的音長明顯比其他四個聲調短很多，音量也小很多，因此我們可以很容易地以音長和音量來判斷一個音節是否為第五聲。

依照漢語的聲母、韻母和聲調之音系結構，可能的音節組合形式共有 $22 \times 40 \times 5 = 4400$ 種，但是實際上漢語有嚴格的聲母和韻母的組合規則並不是每一種聲母和韻母的組合皆為合法的，而且也不是每一種基本音節組合都有五種聲調，因此漢語僅有 1300 種左右的聲母、韻母和聲調之組合音節。

2.2 漢語 TTS 系統概述

本實驗室長期以來致力於語音辨識技術和語音合成技術的研究，而對於語音合成技術方面本實驗室自行發展了一套漢語 TTS 系統，它是運作於電腦上的軟體程式，其基本架構如圖二.2 所示。



圖二.2：漢語 TTS 系統的基本架構圖。

這套系統是由四個主要的部份所組成，分別是文字分析器、韻律產生器、語音合成器和 411 基本音節語音波形樣本資料庫。文字分析器的功能是將輸入文字轉換成為語言參數和合成音節的 411 基本音節碼，而後將語言參數送入韻律產生器以產生每個音節的韻律訊息，合成音節的 411 基本音節碼則送入 411 基本音節語音波形樣本資料庫以取出對應的 411 基本音節類型之樣本波形。最後將每一個音節的韻律訊息和對應的 411 基本音節類型之樣本波形送

入語音合成器，如此便可以產生每一個音節的合成語音。我們將在以下四個小節中分別對漢語 TTS 系統的四個主要部份做進一步的概述，其中將特別著墨於韻律訊息產生器，因其為本論文的研究重點。

2.2.1 文字分析器

文字分析器的功能是要將一段文字做剖析(Parsing)，以將所含的詞彙及構詞資訊抽取出來送入韻律產生器。不同於英語般漢語的文句在詞(一般所指的詞是指 Lexical Word，有別於在本論文的後段部分所提及的 Prosodic Word 我們將其翻譯為韻律詞。)和詞之間不會以空白表示，因此文字分析器需要對漢語文句進行斷詞的動作，而斷詞如果想要更準確語言模型便是一項不可少的要素。在完成了文字分析器的工作之後便可以得到一些語言參數，這些語言參數由低層次到高層次我們將其條列呈述如下：

1. 音節層次：基本音節類型、聲調類型和音節於詞中的位置等；
2. 詞層次：詞長和詞類(Part of Speech, POS)等；
3. 語法片語(Syntactic Phrase)層次：如動詞片語和名詞片語等；
4. 語法樹(Syntactic Tree)層次；
5. 語意(Semantic)層次。

2.2.2 韻律產生器

TTS 系統想要能夠合成出自然流暢的語音其關鍵便是在於韻律產生器所產生的韻律訊息，一般而言韻律訊息包含了音高的高低、音量(Energy)的大小、音長的長短和停頓(Pause)的長短。而韻律訊息的產生是要依據事先所建立好的韻律模型，我們將在下一個小節中對於韻律模型做概略性地介紹。

2.2.3 韻律模型

在目前的語音合成技術領域裡，對於韻律模型的建立方式可以分為兩種：規則法(Rule-Based Approach)和資料驅動法(Data-Driven Approach)，我們將在以下兩個小節中分別對其做概略性的說明。

2.2.3.1 規則法

規則法是利用一些已經觀察到的韻律現象再輔以一些統計結果而設計出的一套韻律規則，這些規則通常和語言層面的訊息有很大的關聯性，必須先對所要發音的文句做詳盡的分析而後才套用適當的規則。一般常見的規則我們可以舉出數個例子，首先是在各個音節的音高軌跡決定上使用變調規則(Tone-Sandhi Rule)；再來是在考慮不同的音節、聲母、韻母以及聲調的情形之下音節音長的分佈狀況；再者是觀察文句中的標點符號類型以決定整體句子的抑揚頓挫(其也就是語調)；最後是分析各個詞的詞類、語法結構和語意等。在特定文句的語法結構下套用特定的韻律規則，這種韻律模型的建構動輒需要數十個甚至上百個韻律規則，並且需要大量的語言學知識和龐大的人工處理才能得到，因此這種方法並不適用於龐大的語料庫，且其所得到的韻律模型在調適性上有很大的限制，亦即無法依據不同的語料特性做適度的調整。

2.2.3.2 資料驅動法

資料驅動法是利用語料本身的韻律參數以一些數學模型來自動建立韻律模型，且由語料本身的韻律特性自動地學習出韻律規則，以建立語音和文字之間韻律參數的關聯性。此種方法最大的好處是我們不需要定義太多的語言行為，也不必瞭解太深入的語言學知識，只需要藉由適當的參數選取以及套用適合的數學模型，就可以由大量的語料自動地產生出效能不錯的韻律模型。

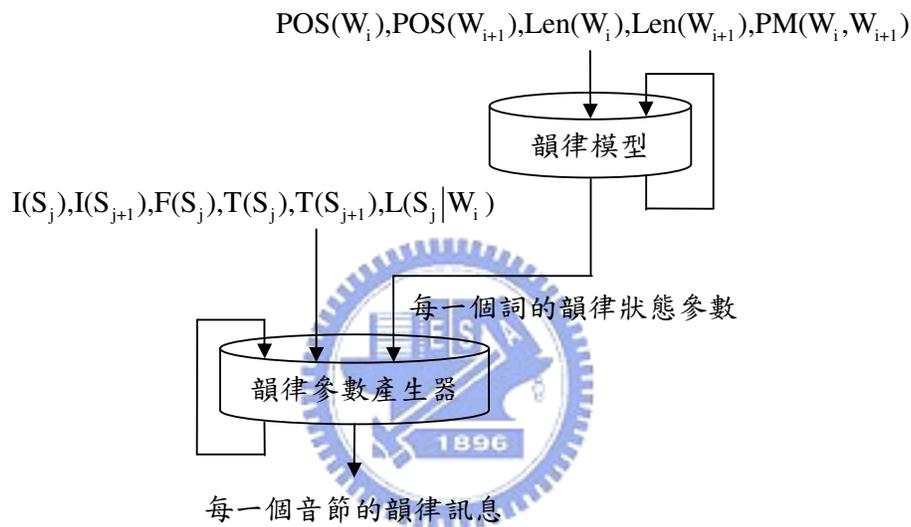
資料驅動法所使用到的數學模型最主要是一些圖像識別(Pattern Recognition)用途的數學工具，如決策樹(Decision Tree)和類神經網路(Artificial Neural Network, ANN)，另外還有使用於統計學上的線性迴歸法(Linear Regression)。我們將在下一個小節中對於以類神經網路為數學模型的方式做進一步的介紹，此為本實驗室所發展出來的韻律模型“遞迴式類神經網路(Recurrent Neural Network, RNN)韻律模型”。

2.2.3.2.1 RNN 韻律模型

RNN 模型的概念是以一組複雜的網路模型來模擬人腦的學習與記憶功能，其學習是採

用漸進式的修正錯誤與更新記憶方式，模型經過長時間且大量的資料學習訓練可以獲得不錯的效能。這個方法可以避免規則法所需的人工分析歸納龐大的韻律規則，且亦能夠兼顧到文句整體的韻律訊息。

圖二.3 為本實驗室所開發出來的 RNN 韻律模型的基本架構圖。RNN 為 ANN 架構型態的一種用以模擬人類韻律訊息產生的方式，其輸入的參數分為兩個部份，一為音節層次語言參數(Syllable-Level Linguistic Feature)，另一為詞層次語言參數(Word-Level Linguistic Feature)。



圖二.3：RNN 韻律模型的基本架構圖。

我們將圖中的音節層次語言參數符號條列說明如下：

- $I(S_j)$ 和 $I(S_{j+1})$ 為現在之音節 S_j 和下一個音節 S_{j+1} 的聲母類型；
- $F(S_j)$ 為現在之音節 S_j 的韻母類型；
- $T(S_j)$ 和 $T(S_{j+1})$ 為現在之音節 S_j 和下一個音節 S_{j+1} 的聲調類型；
- $L(S_j | W_i)$ 為現在之音節 S_j 在現在的詞 W_i 裡面的位置。

我們亦將圖中的詞層次語言參數符號條列說明如下：

- $POS(W_i)$ 和 $POS(W_{i+1})$ 為現在的詞 W_i 和下一個詞 W_{i+1} 的詞類；
- $Len(W_i)$ 和 $Len(W_{i+1})$ 為現在的詞 W_i 和下一個詞 W_{i+1} 的詞長；
- $PM(W_i, W_{i+1})$ 和為現在的詞 W_i 和下一個詞 W_{i+1} 中間的標點符號。

輸出的參數包含了音節的四個音高軌跡參數、一個聲母音長參數、一個韻母音長參數、一個音節間停頓長度參數和一個能量位準(Energy Level)參數，其中四個音高軌跡參數為以四個正交函數對每一個音節的音高軌跡做正交化展開的係數，詳細敘述請參閱[參考文獻\[1\]](#)。

此 RNN 韻律模型可分為兩個部份，一為韻律模型(Prosodic Model)，另一為韻律參數產生器(Prosodic Parameter Generator)。韻律模型藉由輸入詞層次語言參數以分析與探索隱藏於文句中的韻律片語(Prosodic Phrase)結構，並且以詞為單位輸出每個詞的韻律狀態(Prosodic State)參數；韻律參數產生器藉由輸入音節層次語言參數以及韻律模型所產生的韻律狀態參數，來產生以音節為單位的韻律訊息。

2.2.4 語音合成器

本實驗室所發展的漢語 TTS 系統對於語音的合成方法是採用現今世界上被廣為採用的音高同步疊加(Pitch Synchronous Overlap and Add, PSOLA)法，因為其普遍性和並非我們在本論文的研究方向，我們不對其作贅述。

2.2.5 411 基本音節語音波形樣本資料庫

在 411 基本音節語音波形樣本資料庫裡，我們預先錄製了 411 個基本音節的語音波形，以做為語音合成器對於每一個欲合成之音節的樣本波形。而語音合成器也根據由韻律產生器所產生的韻律訊息來調整每一個音節的韻律，如此即完成了每一個音節的語音合成。



第三章 音長模型的原理和數學說明

在這章中首先對於本論文的研究目標做闡述，而後再回顧一些過去在此研究方向的相關文獻內容，最後針對所要建立的音長模型做原理和數學的說明。

3.1 研究目標

在第二章中我們已經對於本實驗室過去所發展出來的 RNN 韻律產生器做了概略性的介紹，這類以資料驅動的方式所建立的韻律模型雖然能夠得到不錯的效能，但是其所產生的韻律參數和語言參數的相關性卻是非常的模糊，我們無法明確地觀察出哪些語言參數對於哪些韻律參數所造成的何種影響。簡單地說，它就如同黑盒子般我們所能知道的只有輸入參數和輸出參數，而對於黑盒子內部做了何種的動作卻無法得知。這樣的模型也許比較適用於簡單的 TTS 系統，但是若要更進一步地考慮更高層次的語言參數如語法和語意等與韻律訊息之間的關係時，就變得非常難以分析。

為了解決上述的問題，有兩種方法被提出，分別為「漢語語音階層韻律模型」和「潛藏韻律模擬法(Latent Prosody Modeling Method)」，我們將在以下的小節中對於這兩種方法做一些相關文獻上的回顧。

3.2 文獻回顧

在這一個節中，我們對於「漢語語音階層韻律模型」和「潛藏韻律模擬法」這兩種方法做一些相關文獻上的回顧。

3.2.1 漢語語音階層韻律模型

此方法是以人工聽覺做為判斷準則的方式，將聽覺上感覺有各種停頓類型的位置分別予以各種 Break Type 標記(Break Type Labeling)，以此來建構漢語語音階層韻律模型，而後利用 Linear Model 來找到這些韻律單位的 Pattern，其包含了 Tone Pattern 和 Duration Pattern 等，詳細敘述請參閱[參考文獻\[2\]](#)。

3.2.2 潛藏韻律模擬法

漢語語音階層韻律模型的建立是以聽者的聽覺為主觀的判斷準則，因此其需要大量的人工來對語料庫做標記，並且不同的聽者有不同的聽覺感受進而造成標記上的一致性。為了能夠節省大量的人力並且希望達到標記上的一致性，由本實驗室所發展出來的一套以電腦對語料庫做自動 Break Type 標記的「潛藏韻律模擬法」能夠達成這個目的。

「潛藏韻律模擬法」是假設韻律的形成是受到各種固有的和潛藏的(Latent)影響因素的影響，而這些影響因素可以藉由統計方法(Statistical Method)將其分離出來，並且可將各個影響因素和語言參數做相關性的分析。而這些影響因素可分為低層次語言參數以及韻律狀態兩類，其中低層次語言參數包含了基本音節類型和聲調類型，而韻律狀態這項影響因素我們將其視為是一個潛藏的變數(Latent Variable)，其代表著所有高層次語言參數對韻律的影響，例如詞、語法片語和語意等。最後我們將所觀察到的韻律參數以這些影響因素的相加或相乘來表示，我們稱其為潛藏韻律模型(Latent Prosody Model)，如下的數學式所述：

$$Z_n = X_n + \gamma_{bs_n} + \gamma_{t_n} + \gamma_{ps_n} \quad (3.1)$$

$$Z_n = X_n \gamma_{bs_n} \gamma_{t_n} \gamma_{ps_n} \quad (3.2)$$

，其中 Z_n 代表語料庫中第 n 個音節的韻律參數； γ_{bs_n} 代表基本音節類型影響因素； γ_{t_n} 代表聲調類型影響因素； γ_{ps_n} 代表韻律狀態類型影響因素； X_n 代表 Z_n 對上述所有影響因素做了正規化之後的殘餘值(Residual)。接著我們使用了一些演算法如 EM Algorithm 等來推導出各個影響因素的數學式，以得到各個影響因素對韻律的影響程度，詳細敘述請參閱[參考文獻\[3 ~ 11\]](#)。

在[參考文獻\[11\]](#)中，作者利用韻律訊息中的音節音高、停頓長度和 Pause Energy Dip 與各種層次的語言參數，以潛藏韻律模擬法對於本論文所使用的語料庫進行韻律標記(Prosody Labeling)的動作，亦即建立一套以模型為基礎的(Model-Based)自動韻律標記方法對語料庫自動地標記上韻律線索(Prosody Cues)，其中韻律線索為各音節間之停頓類型的 Break Type 和各音節的音高韻律狀態(Pitch Prosodic State)，而音節之韻律狀態的定義為音節在一個韻律片語中的狀態標籤(Tag)。在文獻中作者導入四個數學模型來描述韻律標記和語音的韻律參數與文字的語言參數之間的關係，而後再以循序最佳化程序(Sequential Optimization Procedure)來估計出這四個模型的參數並且完成韻律標記。經過在同於本論文所使用的語料庫上的實驗證

實絕大部分的韻律標記是有意義的，且其所估計的四個模型之參數是符合我們所認知的各種語音特性。在本論文中我們將會以此篇文獻的 Break Type 和音高韻律狀態來對於所建立的音長模型進行各種分析，因此我們對其多加著墨。

3.3 語音資料庫的介紹

在本論中我們使用一套漢語語音資料庫來建立我們所提出的音長模型，這一套語料庫的文字內容是取自於中央研究院的 Treebank 語料庫[12]，全部都是由包含數個 Sentences 的 Short Paragraph 所組成，整套語料庫共有 379 段 Short Paragraphs，共含有 52192 個音節。

對於這一套語料庫的語音內容方面，我們是請一位女性的專業廣播員以漢語來朗讀所有的 Short Paragraphs，以取樣頻率 16 kHz 及 16 位元資料格式錄音，而後將錄製好的語料先以 HTK 對其做自動音節切割，再輔以人工的方式對其做音節切割位置修正，以將每一個音節切取出來而得到每一個音節的音長資訊。另外，我們對文字部份以人工的方式做破音字修正。

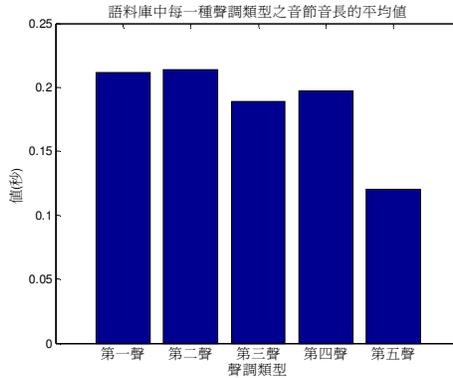
而對於這一套語料庫的統計資訊方面，我們首先將所有音節音長的平均值做了統計其為 0.19836(秒)；再來將每一種基本音節類型之個數做了統計如附錄四所示；而後將每一種基本音節類型之音節音長的平均值做了統計如附錄五所示；接著將每一種聲調類型之個數做了統計如表三.1 所示；最後將每一種聲調類型之音節音長的平均值做了統計如表三.2 和圖三.1 所示。

表三.1：語料庫中每一種聲調類型的音節個數。

	聲調一	聲調二	聲調三	聲調四	聲調五
個數	9744	12529	9508	17692	2719

表三.2：語料庫中每一種聲調類型之音節音長的平均值。

	聲調一	聲調二	聲調三	聲調四	聲調五
值(秒)	0.2115	0.2136	0.1893	0.1973	0.1201



圖三.1：語料庫中每一種聲調類型之音節音長的平均值。

我們依據語料庫中由人工所標記的詞邊界撰寫了程式來找出語料庫中最多字數的詞，其為十二字詞。也撰寫了程式來統計出語料庫中一字詞到十二字詞的個數以及其所在的位置，一字詞到十二字詞的個數列出如表三.3 所示。

表三.3：語料庫中一字詞到十二字詞的個數。

	一字詞	二字詞	三字詞	四字詞	五字詞	六字詞	七字詞	八字詞	九字詞
個數	11724	15531	2113	635	63	12	10	1	2

十字詞	十一字詞	十二字詞
1	2	1

3.4 音長模型的原理說明

在本論文中我們對於漢語語音建立一個以音節為單位的音長模型，一般對於音長模型的建立其單位必須是語音片段(Speech Segment)，例如 HMM 的狀態(State)、音素(Phone)、聲母、韻母、音節甚至是詞，而因為漢語是一個以基本音節和聲調為發音基礎的語言，因此我們選擇了音節當作音長模型的單位。

我們所提出的音長模型之想法是希望能夠將音長以數個相加的影響因素來控制其壓縮(Compression)和延伸(Expansion)，亦即我們可以將影響音長的數個主要原因分離並且抽取出來。而我們所假設的各個影響因素包含了語者類型(Speaker Type)、基本音節類型(Base-Syllable Type)、聲調類型(Tone Type)和韻律狀態類型(Prosody State Type)，其中我們將韻律狀態類型的數目定為 16 個，而之所以將其定為 16 個是因為在參考文獻[8]中作者已經對

於殘餘音長 SD' (其定義我們將在之後的 3.5 節中做說明)的變異數 R 和韻律狀態類型數目的關係做了說明。在文獻中提到殘餘音長之變異數將隨著韻律狀態類型數目的增加而下降，這也代表了我們的音長模型將隨著韻律狀態類型數目的增加而更趨準確。但是在文獻中也有提到殘餘音長之變異數並不是一味地隨著韻律狀態類型數目的增加而下降，其將會到達一個飽和值而這個值就是 16，因此我們選擇了這個能夠讓我們的模型達到不錯的效能，卻又不會使得模型過於複雜的韻律狀態類型數目。

我們將在以下的四個小節中分別就語者類型、基本音節類型、聲調類型和高階語言層次參數對於音長的影響做進一步的說明。

3.4.1 語者類型的音長特性

每一個語者有其之獨特的說話風格，而對於音長的差異方面即是每一個人有其各自的說話速度(Speaking Rate)，因此我們在本論文中將每一個語者的所有音節音長做平均，以視其為每一個語者的語者類型影響因素。而在本論文所使用的語料庫由於只有一個語者，因此語者類型影響因素為一個定值。

3.4.2 基本音節類型的音長特性

音長的差異和基本音節的音素結構(Phonetic Structure)是有著極大的關聯性，漢語的基本音節是有著非常規則的音素結構，其可以少到只由單一韻母所組成例如“衣”，也可以多到由聲母、介音、韻母和韻尾四個部份都備齊來組成例如“香”，因此 411 個基本音節的音長不盡相同。再更詳細地說明每一種基本音節類型是由聲母和韻母或者僅其中之一所組成，而對於韻母方面韻母核心(Nucleus)是必須要有的，它又可以更進一步的區分為含有介音(Medial)與否和含有鼻音結尾(Nasal Ending)與否，因此基本音節將可能是由一到四個音素所組成，一般來說音長會隨著組成音素的增加而增加，因此聲母、介音、韻母和韻尾四個部份都備齊的音節其音長比較長。至於聲母和韻母類型對於音長的影響方面，一般而言僅由單韻母(Single Vowel)所組成的音節其音長一般而言是最短的，再者沒有聲母(亦即為空聲母)或聲母為塞音(Stop)或者沒有鼻音結尾的音節，其音長一般而言是較短的；而聲母為摩擦音(Fricative)或塞擦音(Affricate)或者韻母含有鼻音結尾的音節，其音長一般而言是較長的。

以上所述有關於基本音節類型對於音長的影響程度，我們可以由附錄五中的數據觀察到確實是符合上述的語音特性。而由於基本音節對於音長的影響甚鉅，因此我們將基本音節類型視為一個影響音長的影響因素。

3.4.3 聲調類型的音長特性

音長的差異和聲調類型是有著相當程度的關聯性，最明顯的莫過於聲調五的音長會比其它聲調短許多。而我們再由圖三.1 可以觀察到聲調一略短於聲調二、聲調三略短於聲調四，而聲調一和聲調二皆長於聲調三和聲調四。就經驗上而言除了聲調五以外，其它四個聲調的長短和語者的說話風格(Speaking Style)是有著極大的關聯性。而由於聲調對於音長的影響亦不小，因此我們將聲調類型視為一個影響音長的影響因素。

3.4.4 高階語言層次參數的音長特性

除了上述的低階音節層次參數之外，高階語言層次參數，例如詞層次(Word-Level)和語法層次(Syntactic-Level)也會嚴重地影響到音長。在本論文中我們以韻律狀態來描述高階語言層次參數對音長的影響，也就是音節在一句韻律片語中的高階語言層次之音長狀態。音長將隨著音節在韻律片語位置的不同而有所差異，其中最明顯的現象便是一句韻律片語的最後一個音節的音長會被拉長，這也就是所謂的 Lengthening Effect。由於音長訊息是一種韻律上的訊息，因此音長訊息應該要符合韻律結構(Prosodic Structure)而非語法結構(Syntactic Structure)。我們把韻律狀態視為是一種潛藏的韻律訊息(Latent Prosodic Information)，在過去對於使用韻律狀態的難處為不易對語料庫以人工的方式將所有音節標記韻律狀態類型，因此在本論文中我們以電腦輔助的方式將潛藏的韻律狀態標記出來，如此我們便能夠把文句的高階語言層次韻律訊息擷取出來以作為 TTS 系統的韻律訊息產生之用。

3.5 音長模型的數學說明

我們將所提出的數學模型表示成一個參數最佳化的問題如下面的數學式所述，首先注意在以下的數學式中並未包含語者類型影響因素 γ_s ，這是因為在我們所使用的語料庫中僅有一位語者，因此 γ_s 為一常數，詳細敘述將在之後的 3.5.1 節中做說明。再者注意我們在建立音

長模型的過程中，有別於一般對音節的敘述是先基本音節而後才聲調的順序，我們是先處理聲調而後才處理基本音節，這是因為聲調相較於基本音節是更確定的因子，亦即語料庫中每一種聲調類型的個數遠大於每一種基本音節類型的個數。

$$\begin{aligned}
 PS^* &= \arg \max_{PS} P(PS | SD, T, BS) \\
 &= \arg \max_{PS} \frac{P(PS, SD, T, BS)}{P(SD, T, BS)} \\
 &= \arg \max_{PS} \frac{P(SD | T, BS, PS) P(T, BS, PS)}{P(SD, T, BS)} \\
 &\approx \arg \max_{PS} P(SD | T, BS, PS)
 \end{aligned} \tag{3.3}$$

，上式所代表的即為一個最大事後(Maximum A Posteriori, MAP)機率的問題，亦即在我們已經給定了所要得到的結果和部份條件後，回過頭去以將概似機率最大化的方式來求得未知的參數。在這裡所給定的得到結果為音長，部份條件為聲調類型和基本音節類型，而所要求的未知參數為韻律狀態類型。

我們將上式中各個數學符號的定義條列說明如下：

- $SD = \{sd_{k,n} | n=1, 2, \dots, N_k; k=1, 2, \dots, K\}$ ， $sd_{k,n}$ 代表語料庫中第 k 個 Utterance 裡第 n 個音節的音長；
- $T = \{t_{k,n} | n=1, 2, \dots, N_k; k=1, 2, \dots, K\}$ ， $t \in \{1, 2, \dots, 5\}$ ， $t_{k,n}$ 代表所屬的聲調類型；
- $BS = \{bs_{k,n} | n=1, 2, \dots, N_k; k=1, 2, \dots, K\}$ ， $bs \in \{1, 2, \dots, 411\}$ ， $bs_{k,n}$ 代表所屬的基本音節類型；
- $PS = \{ps_{k,n} | n=1, 2, \dots, N_k; k=1, 2, \dots, K\}$ ， $ps \in \{1, 2, \dots, 16\}$ ， $ps_{k,n}$ 代表所屬的韻律狀態類型，

其中 N_k 為第 k 個 Utterance 的音節個數、 K 為 Utterance 的個數，在本論文中 $K=379$ ， N 為語料庫中所有的音節個數

$$\begin{aligned}
 N &= \sum_{k=1}^K \sum_{n=1}^{N_k} 1 \\
 &= 52192
 \end{aligned} \tag{3.4}$$

(3.3)式可再進一步表示如下

$$PS^* = \arg \max_{PS} P(sd_{1,1}, sd_{1,2}, \dots, sd_{K,N_K} | t_{1,1}, t_{1,2}, \dots, t_{K,N_K}; bs_{1,1}, bs_{1,2}, \dots, bs_{K,N_K};$$

$$ps_{1,1}, ps_{1,2}, \dots, ps_{K,N_K}) \quad (3.5)$$

在上式中我們合理的假設每一個音節的音長和聲調類型、基本音節類型及韻律狀態類型是互為不相關的，因此可表示為互相獨立，因此上式可再進一步表示如下：

$$\begin{aligned} PS^* &= \arg \max_{PS} P(sd_{1,1} | t_{1,1}, bs_{1,1}, ps_{1,1}) P(sd_{1,2} | t_{1,2}, bs_{1,2}, ps_{1,2}), \dots, \\ &P(sd_{K,N_K} | t_{K,N_K}, bs_{K,N_K}, ps_{K,N_K}) \\ &= \arg \max_{PS} \prod_{k=1}^K \prod_{n=1}^{N_k} P(sd_{k,n} | t_{k,n}, bs_{k,n}, ps_{k,n}) \end{aligned} \quad (3.6)$$

上式中 $P(sd_{k,n} | t_{k,n}, bs_{k,n}, ps_{k,n})$ 即為我們所要建立的音長模型，它是描述音長 $sd_{k,n}$ 和聲調類型 $t_{k,n}$ 、基本音節類型 $bs_{k,n}$ 與韻律狀態類型 $ps_{k,n}$ 的關係。我們以語者類型影響因素 $\gamma_{s_{k,n}}$ 、聲調類型影響因素 $\gamma_{t_{k,n}}$ 、基本音節類型影響因素 $\gamma_{bs_{k,n}}$ 和韻律狀態類型影響因素 $\gamma_{ps_{k,n}}$ 與殘餘音長 $sd_{k,n}^r$ 的線性相加方式來代表音長 $sd_{k,n}$ ，如下式所示：

$$sd_{k,n} = sd_{k,n}^r + \gamma_{s_{k,n}} + \gamma_{t_{k,n}} + \gamma_{bs_{k,n}} + \gamma_{ps_{k,n}} \quad (3.7)$$

其中 $sd_{k,n}^r$ 代表音長扣除了上述四個影響因素後的殘餘值(Residual)，亦即為原始音長對四個影響因素做正規化後的值，我們稱其為殘餘音長，我們合理的令其機率分佈為平均值0、變異數 R 的高斯分佈，亦即 $SD^r \sim N(sd_{k,n}^r; 0, R)$ 。

由(3.7)式和 $SD^r \sim N(sd_{k,n}^r; 0, R)$ ，我們可以得到 SD 的機率分佈如下式所述

$$\begin{aligned} P(sd_{k,n} | t_{k,n}, bs_{k,n}, ps_{k,n}) &= N(sd_{k,n}; \gamma_{s_{k,n}} + \gamma_{t_{k,n}} + \gamma_{bs_{k,n}} + \gamma_{ps_{k,n}}, R) \\ &= \frac{1}{\sqrt{2\pi R}} \exp \left[-\frac{(sd_{k,n} - m)^2}{2R} \right] \end{aligned} \quad (3.8)$$

，其中 $m = \gamma_{s_{k,n}} + \gamma_{t_{k,n}} + \gamma_{bs_{k,n}} + \gamma_{ps_{k,n}}$ 。

3.5.1 目標函數的數學說明

於(3.6)式中

$$\prod_{k=1}^K \prod_{n=1}^{N_k} P(sd_{k,n} | t_{k,n}, bs_{k,n}, ps_{k,n})$$

即為所謂的概似函數(Likelihood Function)，而我們為了運算上的方便，對其做自然對數(Natural Logarithm, \ln)運算，並令其為目標函數 Q 得到

$$\begin{aligned}
Q &= \ln \left[\prod_{k=1}^K \prod_{n=1}^{N_k} P(sd_{k,n} | t_{k,n}, bs_{k,n}, ps_{k,n}) \right] \\
&= \sum_{k=1}^K \sum_{n=1}^{N_k} \ln \left[P(sd_{k,n} | t_{k,n}, bs_{k,n}, ps_{k,n}) \right] \\
&= \sum_{k=1}^K \sum_{n=1}^{N_k} \ln \left[N(sd_{k,n}; \gamma_{s_{k,n}} + \gamma_{t_{k,n}} + \gamma_{bs_{k,n}} + \gamma_{ps_{k,n}}, R) \right] \\
&= \sum_{k=1}^K \sum_{n=1}^{N_k} \ln \left\{ \frac{1}{\sqrt{2\pi R}} \exp \left[-\frac{(sd_{k,n} - m)^2}{2R} \right] \right\} \\
&= \sum_{k=1}^K \sum_{n=1}^{N_k} \left[-\frac{1}{2} \ln(2\pi) - \frac{1}{2} \ln(R) - \frac{(sd_{k,n} - m)^2}{2R} \right] \\
&= -\frac{N}{2} \ln(2\pi) - \frac{N}{2} \ln(R) - \frac{1}{2R} \sum_{k=1}^K \sum_{n=1}^{N_k} (sd_{k,n} - m)^2
\end{aligned} \tag{3.9}$$

將上式帶入 $m = \gamma_{s_{k,n}} + \gamma_{t_{k,n}} + \gamma_{bs_{k,n}} + \gamma_{ps_{k,n}}$ 以利往後的數學式推導，得到

$$Q = -\frac{N}{2} \ln(2\pi) - \frac{N}{2} \ln(R) - \frac{1}{2R} \sum_{k=1}^K \sum_{n=1}^{N_k} (sd_{k,n} - \gamma_{s_{k,n}} - \gamma_{t_{k,n}} - \gamma_{bs_{k,n}} - \gamma_{ps_{k,n}})^2 \tag{3.10}$$

由以上的敘述和結合(3.3)式，我們可以將欲求解的問題變成求解最大概似(Maximum Likelihood, ML)函數的問題，如下式所述：

$$PS^* = \arg \max_{PS} Q \tag{3.11}$$

因此我們使用了最大概似法則(ML Criterion, MLC)，來推導出往後的數學式和建立模型。

3.5.2 語者類型影響因素的數學說明

由於在我們所使用的語料庫中只有一位語者，因此 γ_s 對於整個語料庫中的所有音節並無不同，而我們以同一個語者所發出的所有音節音長之平均值亦即全域平均值(Global Mean)來代表 γ_s ，如下的數學式所述：

$$\gamma_s = \frac{1}{N} \sum_{k=1}^K \sum_{n=1}^{N_k} sd_{k,n} \tag{3.12}$$

在上式中我們以同一個語者所發出的所有音節音長之平均值來表示每一個音節的語者類型影響因素是符合語音特性上的意義的，在經過大量的樣本資料統計平均下，每一個語者

其固有的平均音節音長之發音長度特性可以明確地被估計出來，而其也代表了這個語者整體說話之音長特性的說話風格。

3.5.3 聲調類型影響因素的數學說明

我們將 Q 對 γ_t 做偏微分並令其等於零以求得 Q 對 γ_t 的極大值之所在的 γ_t 值，數學式的推導如下所述：

$$\begin{aligned}
\frac{\partial Q}{\partial \gamma_t} &= 0 \\
\Rightarrow -\frac{1}{2R} \sum_{k=1}^K \sum_{n=1}^{N_k} (2)(sd_{k,n} - \gamma_{s_{k,n}} - \gamma_{t_{k,n}} - \gamma_{bs_{k,n}} - \gamma_{ps_{k,n}})(-1)\delta(t_{k,n} = t) &= 0 \\
\Rightarrow \sum_{k=1}^K \sum_{n=1}^{N_k} (sd_{k,n} - \gamma_{s_{k,n}} - \gamma_{t_{k,n}} - \gamma_{bs_{k,n}} - \gamma_{ps_{k,n}})\delta(t_{k,n} = t) &= 0 \\
\Rightarrow \sum_{k=1}^K \sum_{n=1}^{N_k} (sd_{k,n} - \gamma_{s_{k,n}} - \gamma_{bs_{k,n}} - \gamma_{ps_{k,n}})\delta(t_{k,n} = t) - \sum_{k=1}^K \sum_{n=1}^{N_k} \gamma_t \delta(t_{k,n} = t) &= 0 \\
\Rightarrow \sum_{k=1}^K \sum_{n=1}^{N_k} (sd_{k,n} - \gamma_{s_{k,n}} - \gamma_{bs_{k,n}} - \gamma_{ps_{k,n}})\delta(t_{k,n} = t) - \gamma_t \sum_{k=1}^K \sum_{n=1}^{N_k} \delta(t_{k,n} = t) &= 0 \\
\Rightarrow \sum_{k=1}^K \sum_{n=1}^{N_k} (sd_{k,n} - \gamma_{s_{k,n}} - \gamma_{bs_{k,n}} - \gamma_{ps_{k,n}})\delta(t_{k,n} = t) = \gamma_t \sum_{k=1}^K \sum_{n=1}^{N_k} \delta(t_{k,n} = t) \\
\Rightarrow \gamma_t = \frac{\sum_{k=1}^K \sum_{n=1}^{N_k} (sd_{k,n} - \gamma_{s_{k,n}} - \gamma_{bs_{k,n}} - \gamma_{ps_{k,n}})\delta(t_{k,n} = t)}{\sum_{k=1}^K \sum_{n=1}^{N_k} \delta(t_{k,n} = t)}, t \in \{1, 2, \dots, 5\} & \quad (3.13)
\end{aligned}$$

其中

$$\delta(t_{k,n} = t) = \begin{cases} 1, & \text{if } t_{k,n} = t \\ 0, & \text{if } t_{k,n} \neq t \end{cases} \quad (3.14)$$

上式所代表的意義為將語料庫中所有的 sd 減去其所屬的 γ_s 、 γ_{bs} 和 γ_{ps} 後也就是 $sd_{k,n} - \gamma_{s_{k,n}} - \gamma_{bs_{k,n}} - \gamma_{ps_{k,n}}$ ，再對每一種聲調類型分別相加起來後除以其之個數。

3.5.4 基本音節類型影響因素的數學說明

我們將 Q 對 γ_{bs} 做偏微分並令其等於零以求得 Q 對 γ_{bs} 的極大值之所在的 γ_{bs} 值，數學式的推導如下所述：

$$\frac{\partial Q}{\partial \gamma_{bs}} = 0$$

$$\begin{aligned}
&\Rightarrow -\frac{1}{2R} \sum_{k=1}^K \sum_{n=1}^{N_k} (2)(sd_{k,n} - \gamma_{s_{k,n}} - \gamma_{t_{k,n}} - \gamma_{bs_{k,n}} - \gamma_{ps_{k,n}})(-1)\delta(bs_{k,n} = bs) = 0 \\
&\Rightarrow \sum_{k=1}^K \sum_{n=1}^{N_k} (sd_{k,n} - \gamma_{s_{k,n}} - \gamma_{t_{k,n}} - \gamma_{bs_{k,n}} - \gamma_{ps_{k,n}})\delta(bs_{k,n} = bs) = 0 \\
&\Rightarrow \sum_{k=1}^K \sum_{n=1}^{N_k} (sd_{k,n} - \gamma_{s_{k,n}} - \gamma_{t_{k,n}} - \gamma_{ps_{k,n}})\delta(bs_{k,n} = bs) - \sum_{k=1}^K \sum_{n=1}^{N_k} \gamma_{bs_{k,n}} \delta(bs_{k,n} = bs) = 0 \\
&\Rightarrow \sum_{k=1}^K \sum_{n=1}^{N_k} (sd_{k,n} - \gamma_{s_{k,n}} - \gamma_{t_{k,n}} - \gamma_{ps_{k,n}})\delta(bs_{k,n} = bs) - \gamma_{bs} \sum_{k=1}^K \sum_{n=1}^{N_k} \delta(bs_{k,n} = bs) = 0 \quad (3.15) \\
&\Rightarrow \sum_{k=1}^K \sum_{n=1}^{N_k} (sd_{k,n} - \gamma_{s_{k,n}} - \gamma_{t_{k,n}} - \gamma_{ps_{k,n}})\delta(bs_{k,n} = bs) = \gamma_{bs} \sum_{k=1}^K \sum_{n=1}^{N_k} \delta(bs_{k,n} = bs) \\
&\Rightarrow \gamma_{bs} = \frac{\sum_{k=1}^K \sum_{n=1}^{N_k} (sd_{k,n} - \gamma_{s_{k,n}} - \gamma_{t_{k,n}} - \gamma_{ps_{k,n}})\delta(bs_{k,n} = bs)}{\sum_{k=1}^K \sum_{n=1}^{N_k} \delta(bs_{k,n} = bs)}, bs \in \{1, 2, \dots, 411\}
\end{aligned}$$

其中

$$\delta(bs_{k,n} = bs) = \begin{cases} 1, & \text{if } bs_{k,n} = bs \\ 0, & \text{if } bs_{k,n} \neq bs \end{cases} \quad (3.16)$$

上式所代表的意義為將語料庫中所有的 sd 減去其所屬的 γ_s 、 γ_t 和 γ_{ps} 後也就是 $sd_{k,n} - \gamma_{s_{k,n}} - \gamma_{t_{k,n}} - \gamma_{ps_{k,n}}$ ，再對每一種基本音節類型分別相加起來後除以它的個數。

3.5.5 韻律狀態類型影響因素的數學說明

我們將 Q 對 γ_{ps} 做偏微分並令其等於零，以求得 Q 對 γ_{ps} 的極大值之所在的 γ_{ps} 值，數學式的推導如下所述：

$$\begin{aligned}
&\frac{\partial Q}{\partial \gamma_{ps}} = 0 \\
&\Rightarrow -\frac{1}{2R} \sum_{k=1}^K \sum_{n=1}^{N_k} (2)(sd_{k,n} - \gamma_{s_{k,n}} - \gamma_{t_{k,n}} - \gamma_{bs_{k,n}} - \gamma_{ps_{k,n}})(-1)\delta(ps_{k,n} = ps) = 0 \\
&\Rightarrow \sum_{k=1}^K \sum_{n=1}^{N_k} (sd_{k,n} - \gamma_{s_{k,n}} - \gamma_{t_{k,n}} - \gamma_{bs_{k,n}} - \gamma_{ps_{k,n}})\delta(ps_{k,n} = ps) = 0 \\
&\Rightarrow \sum_{k=1}^K \sum_{n=1}^{N_k} (sd_{k,n} - \gamma_{s_{k,n}} - \gamma_{t_{k,n}} - \gamma_{bs_{k,n}})\delta(ps_{k,n} = ps) - \sum_{k=1}^K \sum_{n=1}^{N_k} \gamma_{ps_{k,n}} \delta(ps_{k,n} = ps) = 0 \\
&\Rightarrow \sum_{k=1}^K \sum_{n=1}^{N_k} (sd_{k,n} - \gamma_{s_{k,n}} - \gamma_{t_{k,n}} - \gamma_{bs_{k,n}})\delta(ps_{k,n} = ps) - \gamma_{ps} \sum_{k=1}^K \sum_{n=1}^{N_k} \delta(ps_{k,n} = ps) = 0 \\
&\Rightarrow \sum_{k=1}^K \sum_{n=1}^{N_k} (sd_{k,n} - \gamma_{s_{k,n}} - \gamma_{t_{k,n}} - \gamma_{bs_{k,n}})\delta(ps_{k,n} = ps) = \gamma_{ps} \sum_{k=1}^K \sum_{n=1}^{N_k} \delta(ps_{k,n} = ps)
\end{aligned}$$

$$\Rightarrow \gamma_{ps} = \frac{\sum_{k=1}^K \sum_{n=1}^{N_k} (sd_{k,n} - \gamma_{s_{k,n}} - \gamma_{t_{k,n}} - \gamma_{bs_{k,n}}) \delta(ps_{k,n} = ps)}{\sum_{k=1}^K \sum_{n=1}^{N_k} \delta(ps_{k,n} = ps)}, ps \in \{1, 2, \dots, 16\} \quad (3.17)$$

其中

$$\delta(ps_{k,n} = ps) = \begin{cases} 1, & \text{if } ps_{k,n} = ps \\ 0, & \text{if } ps_{k,n} \neq ps \end{cases} \quad (3.18)$$

上式所代表的意義為將語料庫中所有的 sd 減去其所屬的 γ_s 、 γ_t 和 γ_{bs} 後也就是 $sd_{k,n} - \gamma_{s_{k,n}} - \gamma_{t_{k,n}} - \gamma_{bs_{k,n}}$ ，再對每一種韻律狀態類型分別相加起來後除以其之個數。

3.5.6 殘餘音長之變異數的數學說明

我們將 Q 對 R 做偏微分並令其等於零以求得 Q 對 R 的極大值之所在的 R 值，數學式的推導如下所述：

$$\begin{aligned} \frac{\partial Q}{\partial R} &= 0 \\ \Rightarrow -\frac{N}{2R} + \frac{1}{2R^2} \sum_{k=1}^K \sum_{n=1}^{N_k} (sd_{k,n} - \gamma_{s_{k,n}} - \gamma_{t_{k,n}} - \gamma_{bs_{k,n}} - \gamma_{ps_{k,n}})^2 &= 0 \\ \Rightarrow \frac{N}{2R} &= \frac{1}{2R^2} \sum_{k=1}^K \sum_{n=1}^{N_k} (sd_{k,n} - \gamma_{s_{k,n}} - \gamma_{t_{k,n}} - \gamma_{bs_{k,n}} - \gamma_{ps_{k,n}})^2 \\ \Rightarrow R &= \frac{\sum_{k=1}^K \sum_{n=1}^{N_k} (sd_{k,n} - \gamma_{s_{k,n}} - \gamma_{t_{k,n}} - \gamma_{bs_{k,n}} - \gamma_{ps_{k,n}})^2}{N} \end{aligned} \quad (3.19)$$

上式所代表的意義為將語料庫中所有的 sd 減去其所屬的 γ_s 、 γ_t 、 γ_{bs} 和 γ_{ps} 後也就是 $sd_{k,n} - \gamma_{s_{k,n}} - \gamma_{t_{k,n}} - \gamma_{bs_{k,n}} - \gamma_{ps_{k,n}}$ ，再將所有的值全部加起來後除以語料庫的音節個數。

上述對於 R 的推導結果和直接對(3.7)式所得到的 $sd_{k,n}^r$

$$sd_{k,n}^r = sd_{k,n} - \gamma_{s_{k,n}} - \gamma_{t_{k,n}} - \gamma_{bs_{k,n}} - \gamma_{ps_{k,n}} \quad (3.20)$$

計算其變異數所得到的結果完全相同。

$$\begin{aligned} R &= E \left[\left(sd_{k,n}^r \right)^2 \right] - \left(E \left[sd_{k,n}^r \right] \right)^2 \\ &= E \left[\left(sd_{k,n}^r \right)^2 \right], \text{ 因為 } E \left[sd_{k,n}^r \right] = 0 \\ &= E \left[\left(sd_{k,n} - \gamma_{s_{k,n}} - \gamma_{t_{k,n}} - \gamma_{bs_{k,n}} - \gamma_{ps_{k,n}} \right)^2 \right] \end{aligned}$$

$$= \frac{\sum_{k=1}^K \sum_{n=1}^{N_k} (sd_{k,n} - \gamma_{s_{k,n}} - \gamma_{t_{k,n}} - \gamma_{bs_{k,n}} - \gamma_{ps_{k,n}})^2}{N} \quad (3.21)$$

3.6 音長模型之建立的流程說明

在完成了上述的數學說明之後接著要建立模型，我們將在以下的兩個小節分別對於模型參數的初始值設定和模型的訓練方式做說明。

3.6.1 模型參數的初始值設定

在訓練模型之前必須要先設定好各個模型參數的初始值，我們將在以下的數個小節中對於各個影響因素的初始值設定做說明。

3.6.1.1 聲調類型影響因素的初始值設定

γ_t 的初始值設定如下的數學式所述：

$$\gamma_t = \frac{\sum_{k=1}^K \sum_{n=1}^{N_k} (sd_{k,n} - \gamma_{s_{k,n}}) \delta(t_{k,n} = t)}{\sum_{k=1}^K \sum_{n=1}^{N_k} \delta(t_{k,n} = t)}, t \in \{1, 2, \dots, 5\} \quad (3.22)$$

上式所代表的意義為將語料庫中所有的 sd 減去其所屬的 γ_s 後也就是 $sd_{k,n} - \gamma_{s_{k,n}}$ ，再對每一種聲調類型分別相加起來後除以其之個數。

3.6.1.2 基本音節類型影響因素的初始值設定

γ_{bs} 的初始值設定如下的數學式所述：

$$\gamma_{bs} = \frac{\sum_{k=1}^K \sum_{n=1}^{N_k} (sd_{k,n} - \gamma_{s_{k,n}} - \gamma_{t_{k,n}}) \delta(bs_{k,n} = bs)}{\sum_{k=1}^K \sum_{n=1}^{N_k} \delta(bs_{k,n} = bs)}, bs \in \{1, 2, \dots, 411\} \quad (3.23)$$

上式所代表的意義為將語料庫中所有的 sd 減去其所屬的 γ_s 和 γ_t 後也就是 $sd_{k,n} - \gamma_{s_{k,n}} - \gamma_{t_{k,n}}$ ，再對每一種基本音節類型分別相加起來後除以其之個數。

3.6.1.3 韻律狀態類型的初始設定

這個部份將分為兩個階段來進行，首先是 γ_{ps} 的初始值設定，再來是每一個音節所屬的韻律狀態類型之初始標示。

3.6.1.3.1 韻律狀態類型影響因素的初始值設定

我們已經將韻律狀態類型的數目定為 16 個，然後再將語料庫中所有的 sd 減去其所屬的 γ_s 、 γ_t 和 γ_{bs} 後也就是 $sd_{k,n} - \gamma_{s_{k,n}} - \gamma_{t_{k,n}} - \gamma_{bs_{k,n}}$ ，以 K-Means 演算法對其進行 16 個碼字 (Codeword) 的向量量化 (Vector Quantization, VQ)，因此我們可以得到一組 16 個碼字的碼本 (Codebook)。我們將這 16 個碼字的數值由小到大排序，這些排序後的碼字即為 16 個 γ_{ps} 的初始值。

3.6.1.3.2 每一個音節所屬的韻律狀態類型的初始標示

我們將語料庫中所有的 $sd_{k,n} - \gamma_{s_{k,n}} - \gamma_{t_{k,n}} - \gamma_{bs_{k,n}}$ 對於 16 個 γ_{ps} 的初始值找出距離最短的其中一個，此即為其所屬的韻律狀態類型。

3.6.1.4 殘餘音長之變異數的初始值設定

R 的初始值設定如下的數學式所述：

$$R = \frac{\sum_{k=1}^K \sum_{n=1}^{N_k} (sd_{k,n} - \gamma_{s_{k,n}} - \gamma_{t_{k,n}} - \gamma_{bs_{k,n}} - \gamma_{ps_{k,n}})^2}{N} \quad (3.24)$$

上式和(3.19)式 R 的更新數學式完全相同。

3.6.1.5 目標函數之概似值的初始值計算

我們在設定好各個模型參數的初始值之後便將其代入(3.10)式，以計算 Q 之概似值的初始值，我們將 Q 之概似值的計算數學式重述如下：

$$Q = -\frac{N}{2} \ln(2\pi) - \frac{N}{2} \ln(R) - \frac{1}{2R} \sum_{k=1}^K \sum_{n=1}^{N_k} (sd_{k,n} - \gamma_{s_{k,n}} - \gamma_{t_{k,n}} - \gamma_{bs_{k,n}} - \gamma_{ps_{k,n}})^2 \quad (3.25)$$

3.6.2 模型的訓練方式

在完成了上述模型參數的初始值設定之後接著便是要訓練模型，我們所採用的方式是以

反覆循序最佳化(Iteratively Sequential Optimization)的方式來不斷地訓練模型和更新模型參數，直到 Q 的概似值符合我們所訂下的收斂條件。我們將在以下的數個小節中呈現出 γ_t 、 γ_{bs} 、 γ_{ps} 、 R 和 Q 之概似值的更新數學式，以及模型訓練的收斂條件。

3.6.2.1 聲調類型影響因素的更新數學式

γ_t 的更新數學式為在之前所推導出來的(3.13)式，我們將其重述如下：

$$\gamma_t = \frac{\sum_{k=1}^K \sum_{n=1}^{N_k} (sd_{k,n} - \gamma_{s_{k,n}} - \gamma_{bs_{k,n}} - \gamma_{ps_{k,n}}) \delta(t_{k,n} = t)}{\sum_{k=1}^K \sum_{n=1}^{N_k} \delta(t_{k,n} = t)}, t \in \{1, 2, \dots, 5\} \quad (3.26)$$

3.6.2.2 基本音節類型影響因素的更新數學式

γ_{bs} 的更新數學式為在之前所推導出來的(3.15)式，我們將其重述如下所述：

$$\gamma_{bs} = \frac{\sum_{k=1}^K \sum_{n=1}^{N_k} (sd_{k,n} - \gamma_{s_{k,n}} - \gamma_{t_{k,n}} - \gamma_{ps_{k,n}}) \delta(bs_{k,n} = bs)}{\sum_{k=1}^K \sum_{n=1}^{N_k} \delta(bs_{k,n} = bs)}, bs \in \{1, 2, \dots, 411\} \quad (3.27)$$

3.6.2.3 韻律狀態類型的更新

這個部份和初始設定一樣需分成兩個階段來進行，首先是更新每一個音節所屬的韻律狀態類型標示，再來是更新 γ_{ps} 。

3.6.2.3.1 更新每一個音節所屬的韻律狀態類型標示

我們以更新後的 γ_t 和 γ_{bs} 與更新前的 γ_{ps} 來對每一個音節重新找到 16 個韻律狀態類型的其中之一使得代入單一音節的 Q 之概似值能夠最大，其中單一音節的 Q 之概似值為將上述的(3.10)式中的全部音節加總去掉，我們令其為 Q' 並對其帶入各個影響因素的新舊資訊，如下的數學式所述：

$$Q' = -\frac{1}{2} \ln(2\pi) - \frac{1}{2} \ln(R) - \frac{(sd_{k,n} - \gamma_{s_{k,n}} - \gamma_{t_{k,n}}^{New} - \gamma_{bs_{k,n}}^{New} - \gamma_{ps_{k,n}}^{Old})^2}{2R} \quad (3.28)$$

其中 $\gamma_{t_{k,n}}^{New}$ 為更新後的 $\gamma_{t_{k,n}}$ 、 $\gamma_{bs_{k,n}}^{New}$ 為更新後的 $\gamma_{bs_{k,n}}$ 和 $\gamma_{ps_{k,n}}^{Old}$ 為更新前的 $\gamma_{ps_{k,n}}$ ，而更新每一個音節所屬的韻律狀態類型標示如下的數學式所述：

$$ps^{Update} = \arg \max_{ps} Q', ps \in \{1, 2, \dots, 16\} \quad (3.29)$$

3.6.2.3.2 韻律狀態類型影響因素的更新數學式

γ_{ps} 的更新數學式為在之前所推導出來的(3.17)式，我們將其重述如下：

$$\gamma_{ps} = \frac{\sum_{k=1}^K \sum_{n=1}^{N_k} (sd_{k,n} - \gamma_{s_{k,n}} - \gamma_{t_{k,n}} - \gamma_{bs_{k,n}}) \delta(ps_{k,n} = ps)}{\sum_{k=1}^K \sum_{n=1}^{N_k} \delta(ps_{k,n} = ps)}, ps \in \{1, 2, \dots, 16\} \quad (3.30)$$

3.6.2.4 殘餘音長之變異數的更新數學式

R 的更新數學式為在之前所推導出來的(3.19)式，我們將其重述如下：

$$R = \frac{\sum_{k=1}^K \sum_{n=1}^{N_k} (sd_{k,n} - \gamma_{s_{k,n}} - \gamma_{t_{k,n}} - \gamma_{bs_{k,n}} - \gamma_{ps_{k,n}})^2}{N} \quad (3.31)$$

3.6.2.5 目標函數之概似值的更新數學式

我們在更新好各個模型參數之後便將其代入上述的(3.10)式以更新 Q 之概似值，我們將其之計算數學式重述如下：

$$Q = -\frac{N}{2} \ln(2\pi) - \frac{N}{2} \ln(R) - \frac{1}{2R} \sum_{k=1}^K \sum_{n=1}^{N_k} (sd_{k,n} - \gamma_{s_{k,n}} - \gamma_{t_{k,n}} - \gamma_{bs_{k,n}} - \gamma_{ps_{k,n}})^2 \quad (3.32)$$

3.6.2.6 模型訓練的收斂條件

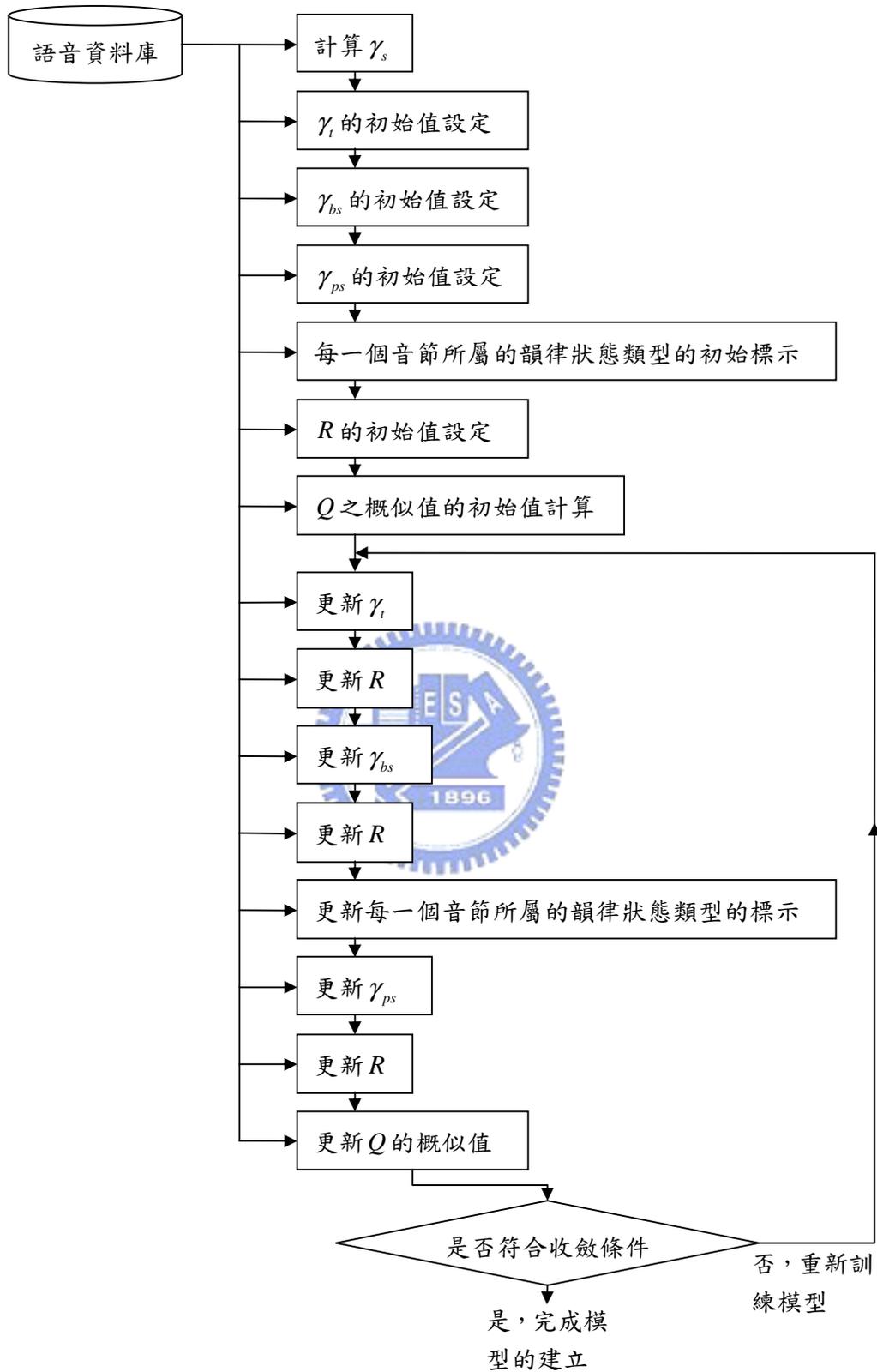
我們所訂下的模型訓練之收斂條件為“只要第 i 次更新之 Q 值和第 $i-1$ 次更新之 Q 值的差小於第 i 次更新之 Q 值乘上一個極小的數，我們令這個極小的數為 10^{-7} ，如此便符合收斂條件。”，其數學表示如下式所示：

$$Q(i) - Q(i-1) < Q(i) \times 10^{-7} \quad (3.33)$$

3.6.2.7 模型訓練的流程

我們所採用的模型訓練流程為每更新一個影響因素一次之後便更新 R 一次，而影響因素的更新順序為先更新 γ_i 再更新 γ_{bs} 最後更新 γ_{ps} ，因此一次完整的模型訓練包含了將三個影響因素各更新一次和 R 更新三次。而後再以上述的(3.32)式更新 Q 的概似值並且測試其是否符合上述的(3.33)式之收斂條件，如果符合便停止更新模型參數，否則繼續更新模型參數直到符合收斂條件為止，我們將模型訓練的流程方塊圖繪出如圖三.2 所示。





圖三.2：音長模型的訓練流程方塊圖。

第四章 音長模型的實驗結果分析

本章對於所建立的音長模型在符合收斂條件狀態下分析其參數，以觀察是否符合我們所認知的語音特性，藉此來判斷所建立的模型是否正確。

4.1 影響因素

我們將在這一個小節中呈現出各個影響因素的值。

4.1.1 語者類型影響因素

我們將 γ_s 的值呈現如下：

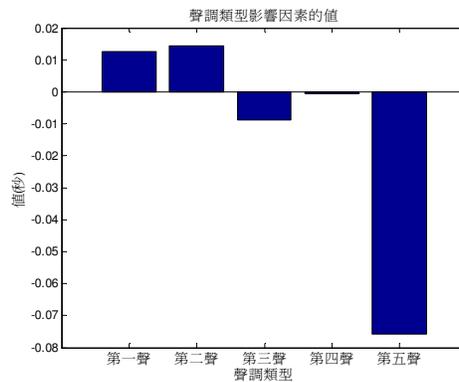
$$\gamma_s = 0.19836(\text{秒}) \quad (4.1)$$

4.1.2 聲調類型影響因素

我們將 γ_t 的值列出如表四.1 和繪出如圖四.1 所示。

表四.1： γ_t 的值。

	聲調一	聲調二	聲調三	聲調四	聲調五
值(秒)	0.0125	0.0142	-0.0087	-0.0006	-0.0757



圖四.1： γ_t 的值。

4.1.2.1 實驗結果分析

我們將表四.1 和圖四.1 與表三.2 和圖三.1 做比較，可以觀察到它們相互的關係是一致

的，因此模型所產生的 γ_i 是符合我們認知的語音特性。

4.1.3 基本音節類型影響因素

我們將 γ_{bs} 的值列於附錄六。

4.1.3.1 實驗結果分析

這個部份將分為兩個階段來進行，首先是和實際音長的比較，亦即和語料庫中每一種基本音節類型之音節音長的平均值做比較，如此的方法是再直觀不過，但卻因為基本音節類型的數量太多，因此無法明確地觀察出它們之間的相互關係。於是我們採用了第二個階段的決策樹(Decision Tree)分析，它是根據漢語中各個聲母和韻母的發音音長特性所設計的一系列問題，以將 γ_{bs} 做重點式的分類，來觀察分類後的平均值是否符合我們所知道的語音特性。

4.1.3.1.1 和實際音長的比較

我們比較附錄六和附錄五的值可以發現到其確實是有一致性的相互關係，最明顯的莫過於聲母為 / ㄅ(b) / 、 / ㄉ(d) / 和 / ㄍ(g) / 的音節其音長會明顯短許多。



4.1.3.1.2 決策樹分析

這個部份將分為三個階段來進行，第一個階段為設計一系列能夠將漢語中各個聲母和韻母的發音音長特性做明顯區別的問題集(Question Set)；第二個階段為訂定當節點的資料符合何種條件時才做節點分裂的的分裂準則(Splitting Criteria)；第三個階段為樹的長成(Growing Tree)；最後便是將決策樹分析應用在 γ_{bs} 上。

4.1.3.1.2.1 問題集

我們一共設計了 15 個問題來對 γ_{bs} 做決策樹分析，而這 15 個問題的設計是依據以下所述的三點考量，我們將其條列呈述如下：

1. 發音方式：例如是否為送氣音(Aspiration)；是否為帶聲的音(Voiced)；是否為塞音(Stop)；是否為擦音(Fricative)；

2. 基本音節的 **Phonetic Structure**：例如是否為單韻母(Single Vowel)；是否為複韻母(Compound Vowel)；是否為鼻音結尾韻母(Nasal Ending Vowel)；是否為含有介音(Medial)的韻母；

3. **韻母核心(Nucleus)的類型**：例如是否為開口(Open)韻母。

我們亦將這 15 個問題條列呈述如下，在注音符號右側的括號內容為其之漢語拼音，而附錄七為漢語 22 個聲母的注音符號與漢語拼音對照表，附錄八為漢語 40 個韻母的注音符號與漢語拼音對照表。最後要注意我們必須先將附錄六中由 MATLAB 所產生的 NaN 符號將其修改為 0，如此我們所撰寫的程式才能辨識出其所代表的意義，而對於在語料庫中未出現的基本音節類型我們便不對其做決策樹分析。

問題一：聲母是否為

/ 空(Null) / ?

問題二：聲母是否為

/ ㄅ(b) / 、 / ㄉ(d) / 、 / ㄍ(g) / ?

問題三：聲母是否為

/ ㄈ(f) / 、 / ㄏ(h) / 、 / ㄒ(x) / 、 / ㄕ(gh) / 、 / ㄙ(s) / ?

問題四：聲母是否為

/ ㄇ(m) / 、 / ㄋ(n) / 、 / ㄌ(l) / 、 / ㄋ(r) / ?

問題五：聲母是否為

/ ㄑ(q) / 、 / ㄔ(ch) / 、 / ㄔ(c) / ?

問題六：聲母是否為

/ ㄆ(p) / 、 / ㄊ(t) / 、 / ㄎ(k) / ?

問題七：聲母是否為

/ ㄐ(j) / 、 / ㄓ(zh) / 、 / ㄓ(z) / ?

問題八：韻母是否為單韻母(Single Vowel)的

/ ㄩ(yi) / 、 / ㄨ(wu) / 、 / ㄩ(yu) / 、 / ㄩ(a) / 、 / ㄩ(o) / 、

/ ㄛ(e) / 、 / ㄝ(eh) / 、 / ㄦ(er) / ？

問題九：韻母是否為複韻母(Compound Vowel)的

/ ㄞ(ai) / 、 / ㄟ(ei) / 、 / ㄠ(ao) / 、 / ㄡ(ou) / 、 / ㄚ(ya) / 、
/ ㄞ(ye) / 、 / ㄞ(yai) / 、 / ㄠ(yao) / 、 / ㄡ(you) / 、
/ ㄛ(yo) / 、 / ㄨㄚ(wa) / 、 / ㄨㄛ(wo) / 、 / ㄨㄞ(wai) / 、
/ ㄨㄟ(wei) / 、 / ㄩㄝ(yue) / ？

問題十：韻母是否為鼻音結尾韻母(Nasal Ending Vowel)的

/ ㄢ(an) / 、 / ㄣ(en) / 、 / ㄤ(ang) / 、 / ㄥ(eng) / 、
/ ㄢ(yan) / 、 / ㄣ(yin) / 、 / ㄤ(yang) / 、 / ㄥ(ying) / 、
/ ㄨㄢ(wan) / 、 / ㄨㄣ(wen) / 、 / ㄨㄤ(wang) / 、
/ ㄨㄥ(weng) / 、 / ㄩㄢ(yuan) / 、 / ㄩㄣ(yun) / 、
/ ㄩㄥ(yung) / ？

問題十一：韻母是否為含有介音(Medial)的

/ ㄚ(ya) / 、 / ㄞ(ye) / 、 / ㄞ(yai) / 、 / ㄠ(yao) / 、
/ ㄡ(you) / 、 / ㄢ(yan) / 、 / ㄣ(yin) / 、 / ㄤ(yang) / 、
/ ㄥ(ying) / 、 / ㄛ(yo) / 、 / ㄨㄚ(wa) / 、 / ㄨㄛ(wo) / 、
/ ㄨㄞ(wai) / 、 / ㄨㄟ(wei) / 、 / ㄨㄢ(wan) / 、 / ㄨㄣ(wen) / 、
/ ㄨㄤ(wang) / 、 / ㄨㄥ(weng) / 、 / ㄩㄝ(yue) / 、
/ ㄩㄢ(yuan) / 、 / ㄩㄣ(yun) / 、 / ㄩㄥ(yung) / ？

問題十二：韻母是否為開口(Open)韻母的

/ ㄚ(a) / 、 / ㄛ(o) / 、 / ㄛ(e) / 、 / ㄝ(eh) / 、 / ㄞ(ai) / 、
/ ㄟ(ei) / 、 / ㄠ(ao) / 、 / ㄡ(ou) / 、 / ㄢ(an) / 、 / ㄣ(en) / 、
/ ㄤ(ang) / 、 / ㄥ(eng) / 、 / ㄦ(er) / ？

問題十三：韻母是否為以 / ㄚ(yi) / 起始的

/ ㄚ(ya) / 、 / ㄞ(ye) / 、 / ㄞ(yai) / 、 / ㄠ(yao) / 、
/ ㄡ(you) / 、 / ㄢ(yan) / 、 / ㄣ(yin) / 、

/ 一尤(yang) / 、 / 一ㄥ(ying) / 、 / 一ㄛ(yo) / ?

問題十四：韻母是否為以 / ㄨ(wu) / 起始的

/ ㄨㄚ(wa) / 、 / ㄨㄛ(wo) / 、 / ㄨㄞ(wai) / 、 / ㄨㄟ(wei) / 、

/ ㄨㄢ(wan) / 、 / ㄨㄣ(wen) / 、 / ㄨㄤ(wang) / 、

/ ㄨㄥ(weng) / ?

問題十五：韻母是否為以 / ㄩ(yu) / 起始的

/ ㄩㄝ(yue) / 、 / ㄩㄢ(yuan) / 、 / ㄩㄣ(yun) / 、 / ㄩㄥ(yung) / ?

4.1.3.1.2.2 分裂準則

我們所訂定的節點分裂條件之判斷準則有兩點，首先是以節點資料的自然對數概似值(Log-Likelihood, LL)來做依據，我們可以將每個節點資料的機率分佈以一個高斯分佈(我們令其符號為 $X \sim N(x; m, v)$ ，其中 X 代表節點資料的隨機變數， m 代表 X 的平均值， v 代表 X 的變異數。)來表示，因此我們可以得到自然對數概似值的計算如下：


$$\begin{aligned} LL(X) &= \ln \left[\prod_x N(x; m, v) \right] \\ &= \sum_x \ln [N(x; m, v)] \\ &= \sum_x \ln \left\{ \frac{1}{\sqrt{2\pi v}} \exp \left[-\frac{(x-m)^2}{2v} \right] \right\} \\ &= \sum_x \left[-\frac{1}{2} \ln(2\pi) - \frac{1}{2} \ln(v) - \frac{(x-m)^2}{2v} \right] \\ &= -\frac{N}{2} \ln(2\pi) - \frac{N}{2} \ln(v) - \frac{1}{2v} \sum_x (x-m)^2 \end{aligned} \tag{4.2}$$

，其中 N 代表節點資料的個數；再來是以節點資料的個數來做依據。

4.1.3.1.2.3 樹的長成

我們將節點分裂的示意圖繪出如圖四.2 所示：

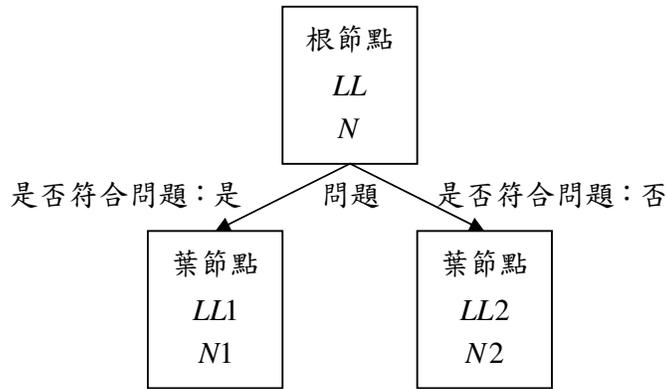


圖 4.2：節點分裂的示意圖。

我們在每一個節點詢問每一個問題後，計算根節點(Root Node)和兩個葉節點(Leaf Node)之概似值的差異，也就是 $LL1+LL2-LL$ ，繼而從中找出概似值之差異最大的問題，將其視為此節點所詢問的問題，其中要注意對每一筆資料所詢問過的問題便不能再詢問。重複上述的模式樹一直長成下去直到所有的節點分裂符合我們所訂定的停止分裂條件，而我們所訂定的停止分裂條件有以下兩點，首先是根節點的概似值減去分裂後的兩個葉節點之概似值的和需小於一個預設的數值；再來是分裂後的兩個葉節點之資料個數皆須小於一個預設的數量。我們將上述兩點以下述的數學式表示之

$$LL1 + LL2 - LL < LL_{Threshold} \quad (4.3)$$

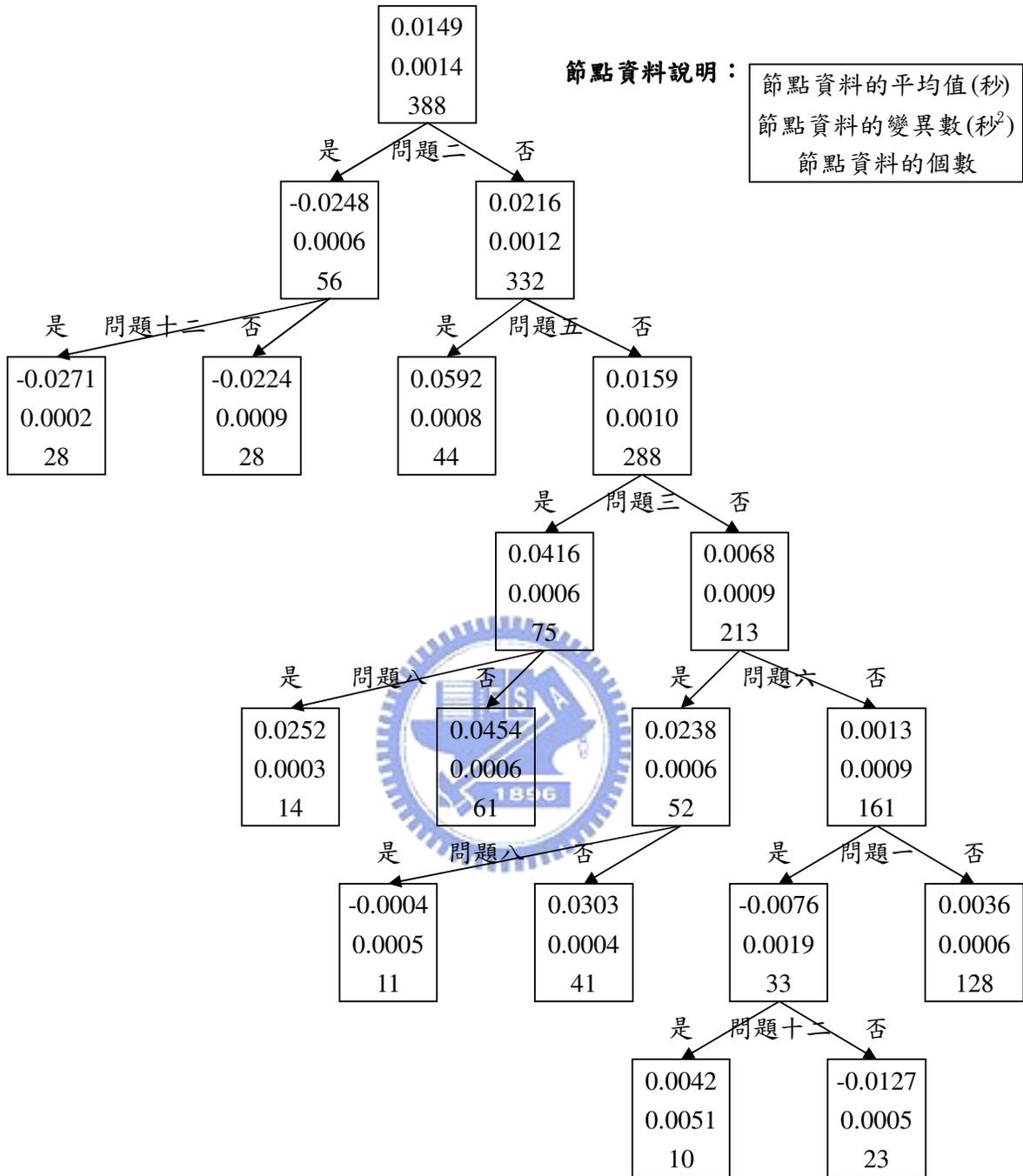
$$N1 < N_{Threshold} \quad (4.4)$$

$$N2 < N_{Threshold} \quad (4.5)$$

其中 $LL_{Threshold}$ 代表節點分裂的概似值之臨界值(Threshold)， $N_{Threshold}$ 代表節點分裂的資料個數之臨界值，當所有節點皆滿足停止分裂條件的任何一點時即完成了樹的長成。

4.1.3.1.2.4 基本音節類型影響因素的決策樹分析

我們將決策樹分析應用在 γ_{bs} 上，如圖 4.3 所示：



圖四.3： γ_{bs} 的決策樹分析。

其中 $LL_{Threshold}$ 和 $N_{Threshold}$ 的值我們設為 5，我們將由上圖中所觀察到的實驗結果條列示地分析如下：

I. 起始節點的資料個數為 388 而不是 411，這是因為在我們所使用的語料庫中有 23

個基本音節類型未出現，而未出現的類型我們便不對其做決策樹分析；

- II. 當基本音節類型符合問題二也就是聲母為 / ㄅ /、/ ㄆ /、/ ㄇ / 時，其音長會明顯較短；
- III. 當基本音節類型符合問題五也就是聲母為 / ㄎ /、/ ㄏ /、/ ㄎ / 時，其音長會明顯較長；
- IX. 當基本音節類型符合問題三也就是聲母為 / ㄐ /、/ ㄑ /、/ ㄒ /、/ ㄒ /、/ ㄌ / 時，其音長會明顯較長；
- V. 當基本音節類型符合問題六也就是聲母為 / ㄨ /、/ ㄨ /、/ ㄛ / 時，其音長會明顯較長；
- VI. 當基本音節類型符合問題八也就是韻母為單韻母時，其音長會明顯較短；
- VII. 當基本音節類型符合問題一也就是聲母為 / 空 / 時，其音長會明顯較短；
- VIII. 當基本音節類型符合問題十二也就是韻母為開口韻母時，其音長會明顯較長。

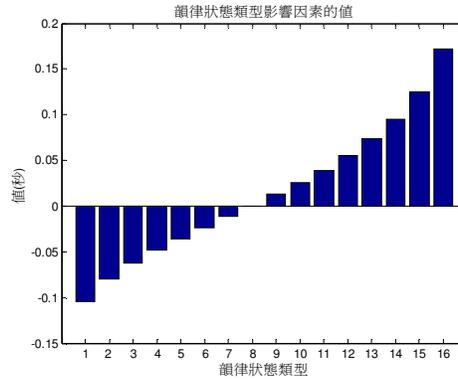
我們由以上的分析可以得知模型所產生的 γ_{bs} 是符合我們所認知的語音特性。

4.1.4 韻律狀態類型影響因素

我們將 γ_{ps} 的值列出如表四.2 和繪出如圖四.4 所示：

表四.2： γ_{ps} 的值。

	類型 1	類型 2	類型 3	類型 4	類型 5	類型 6	類型 7	類型 8
值(秒)	-0.1048	-0.0779	-0.0627	-0.0483	-0.0354	-0.0234	-0.0116	0.0004
類型 9	0.0128	0.0255	0.0393	0.055	0.0732	0.0951	0.1249	0.1723



圖四.4： γ_{ps} 的值。

4.1.4.1 實驗結果分析

我們以 16 個由小到大排序的韻律狀態類型影響因素來代表高階語言層次對於音長的影響，由以上的表和圖我們可以觀察到其數值和分佈是合理的。

4.2 殘餘音長之變異數

對於 R 我們除了呈現出其之收斂值外也必須要呈現出其初始值，以利於觀察模型的收斂性。

4.2.1 初始值

我們將 R 的初始值呈現出如下：

$$R = 2.2699 \times 10^{-5} (\text{秒}^2) \quad (4.6)$$

4.2.2 收斂值

我們將 R 的收斂值呈現出如下：

$$R = 2.1094 \times 10^{-5} (\text{秒}^2) \quad (4.7)$$

4.2.3 實驗結果分析

由以上所述我們可以觀察到收斂狀態的 R 值小於初始狀態的 R 值，此即代表了模型是朝著合理的趨勢在收斂。

4.3 目標函數之概似值

對於 Q 的概似值我們將呈現出初始值、收斂值和更新曲線，以利於觀察對於模型之評估的分數是否隨著模型的收斂而增加。

4.3.1 初始值

我們將 Q 之概似值的初始值呈現出如下：

$$Q = 2.0499 \times 10^5 \quad (4.8)$$

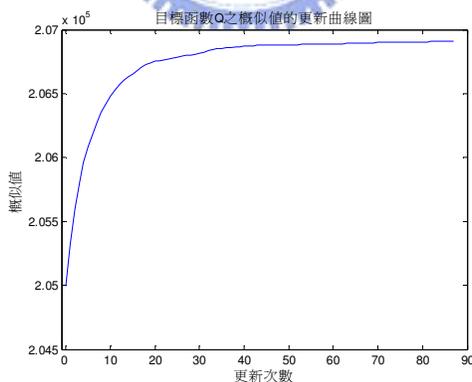
4.3.2 收斂值

我們將 Q 之概似值的收斂值呈現出如下：

$$Q = 2.0691 \times 10^5 \quad (4.9)$$

4.3.3 更新曲線

我們將 Q 之概似值的更新曲線繪出如圖四.5所示：



圖四.5：目標函數 Q 之概似值的更新曲線圖。

4.3.4 實驗結果分析

由以上所述，我們可以觀察到 Q 的概似值是隨著模型的收斂而增加，此即代表了模型是朝著合理的趨勢在收斂。

4.4 音長分佈和音長軌跡

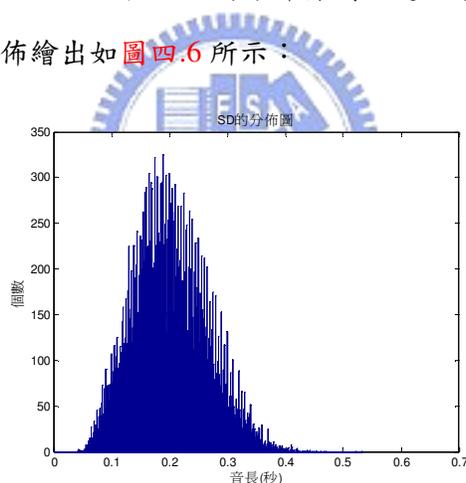
我們將在這一個小節中呈現出語料庫中所有音節的音長分佈以及從語料庫中擷取一段句子的音長軌跡，而為了方便對所有圖形的顯示，我們定義了以下的數學符號：

- SD : sd ;
- $SD1$: $sd - \gamma_s$;
- $SD2$: $sd - \gamma_s - \gamma_t$;
- $SD3$: $sd - \gamma_s - \gamma_t - \gamma_{bs}$;
- $SD4$: $sd - \gamma_s - \gamma_t - \gamma_{bs} - \gamma_{ps}$, 其也就是 sd^r 。

4.4.1 音長分佈

我們將 SD 、 $SD1$ 、 $SD2$ 、 $SD3$ 和 $SD4$ 的分佈條列呈述如下：

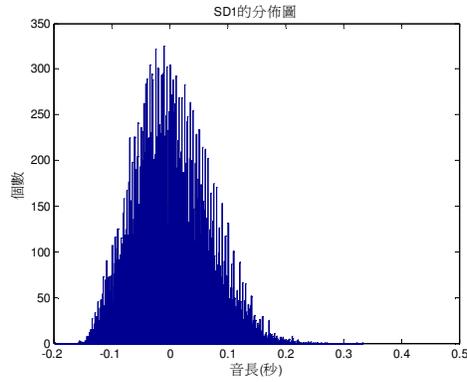
- 我們將 SD 的分佈繪出如圖四.6 所示：



圖四.6： SD 的分佈圖。

，其之變異數為 $383.96 \times 10^{-5} (\text{秒}^2)$ 。

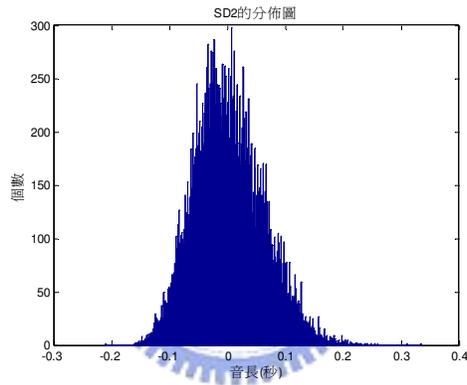
- 我們將 $SD1$ 的分佈繪出如圖四.7 所示：



圖四.7：SD1的分佈圖。

，其之變異數為 383.96×10^{-5} (秒²)。

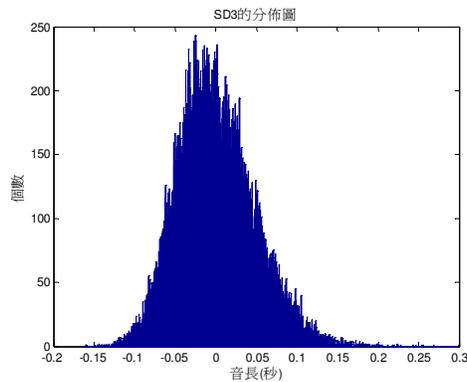
- 我們將SD2的分佈繪出如圖四.8所示：



圖四.8：SD2的分佈圖。

，其之變異數為 341.81×10^{-5} (秒²)。

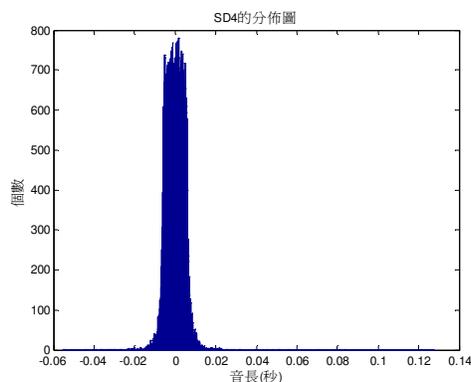
- 我們將SD3的分佈繪出如圖四.9所示：



圖四.9：SD3的分佈圖。

，其之變異數為 224.98×10^{-5} (秒²)。

- 我們將 $SD4$ 的分佈繪出如圖四.10 所示：



圖四.10： $SD4$ 的分佈圖。

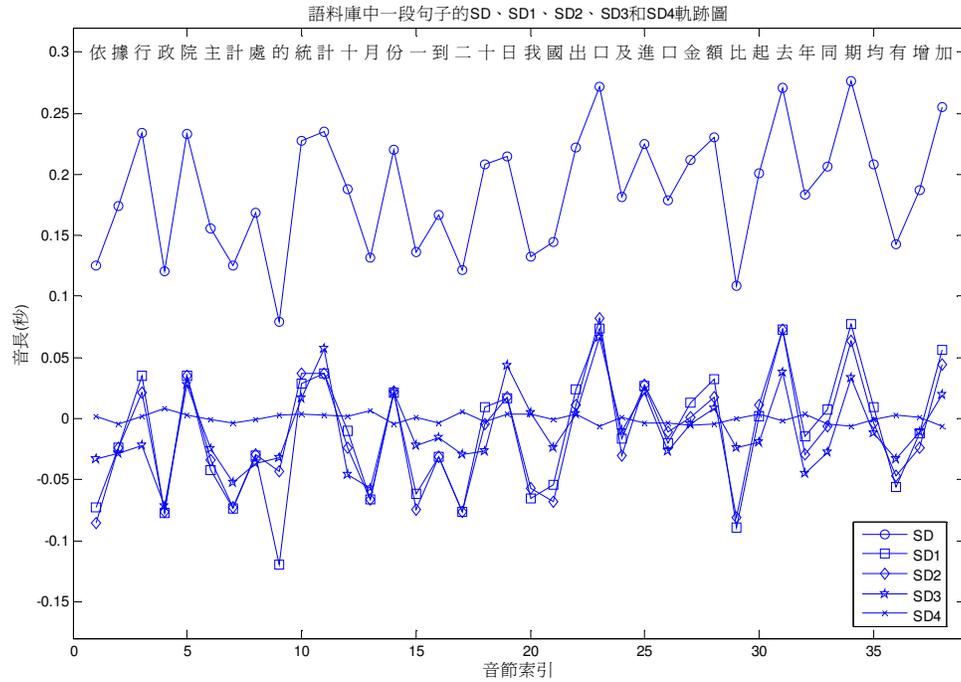
，其之變異數為 2.11×10^{-5} (秒²)。

4.4.1.1 實驗結果分析

我們由上述的圖可以觀察到首先音長分佈呈現出高斯(Gaussian)分佈的形狀；再來除了圖四.7 僅為平移外，圖四.8 到圖四.10 的變異數逐漸變小，亦即音長分佈漸趨集中，其所代表的意義為模型能夠將造成每一個音節之音長差異的因素分離並且抽取出來，使得音長分佈之變異數大幅地降低，由以上分析可以驗證我們的模型所造成的音長分佈變化是合理的。

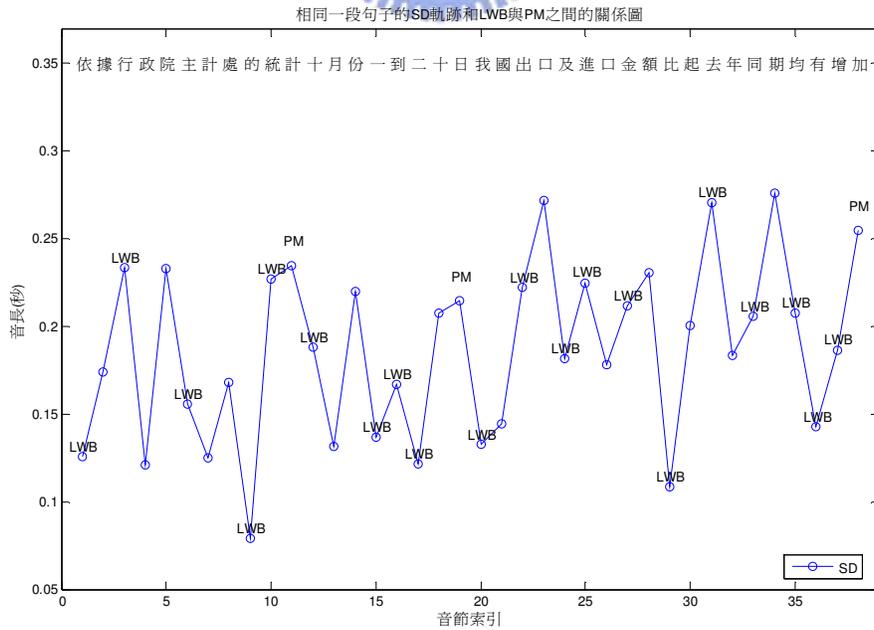
4.4.2 音長軌跡

首先我們呈現出語料庫中一段句子的 SD 、 $SD1$ 、 $SD2$ 、 $SD3$ 和 $SD4$ 軌跡，如圖四.11 所示：



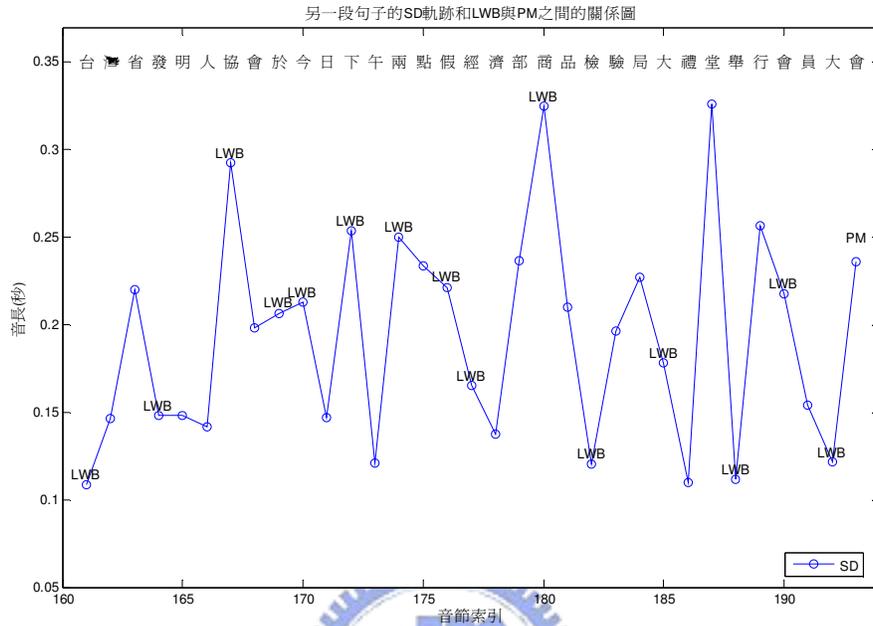
圖四.11：語料庫中一段句子的SD、SD1、SD2、SD3和SD4軌跡圖。

再來我們對相同一段句子呈現出其之SD軌跡和詞邊界(LWB)與標點符號(PM)之間的關係，如圖四.12所示：

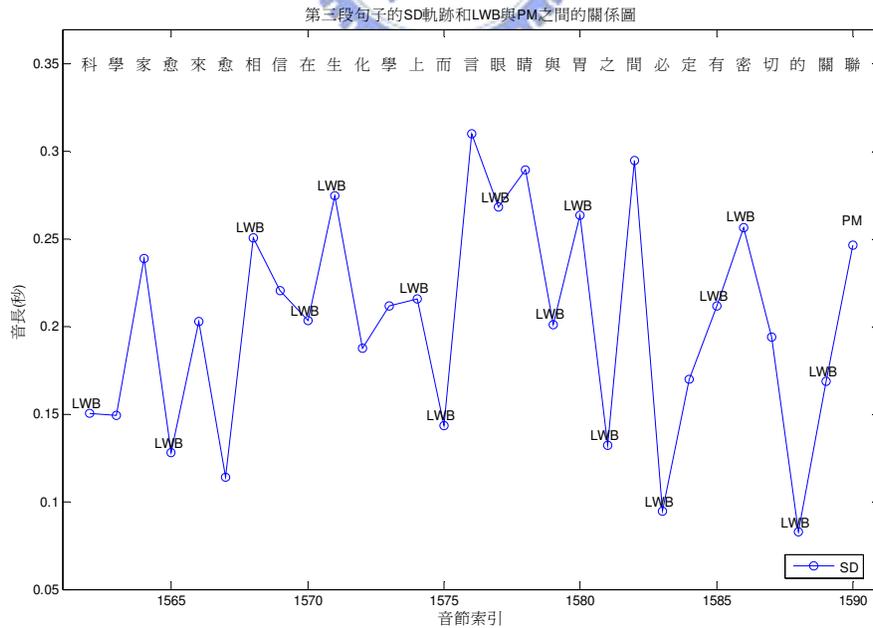


圖四.12：相同一段句子的SD軌跡和LWB與PM之間的關係圖。

而為了要對高階語言層次音長軌跡之特殊的 Pattern 多做觀察，因此我們再從語料庫中呈現出兩段句子的SD軌跡，如圖四.13和圖四.14所示：

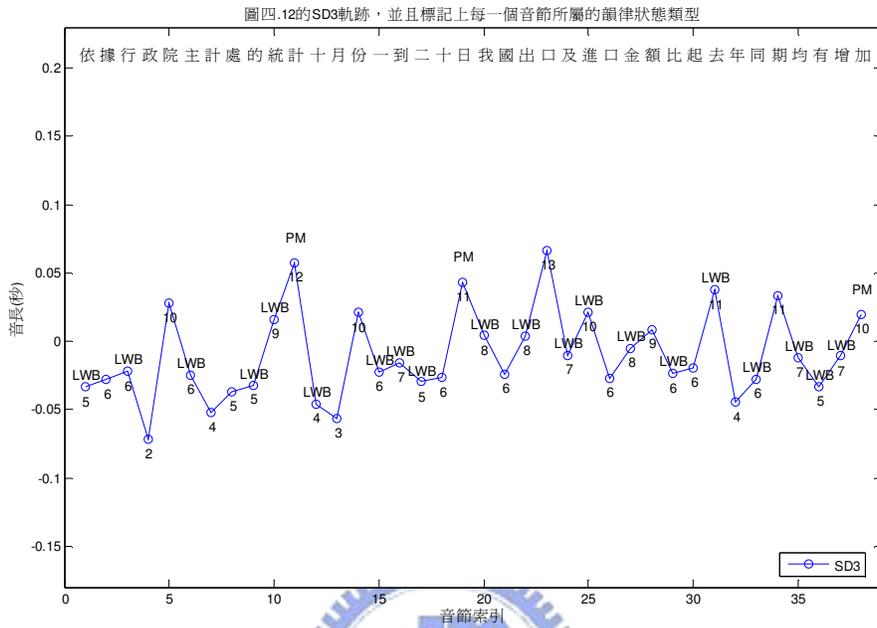


圖四.13：另一段句子的SD軌跡和LWB與PM之間的關係圖。

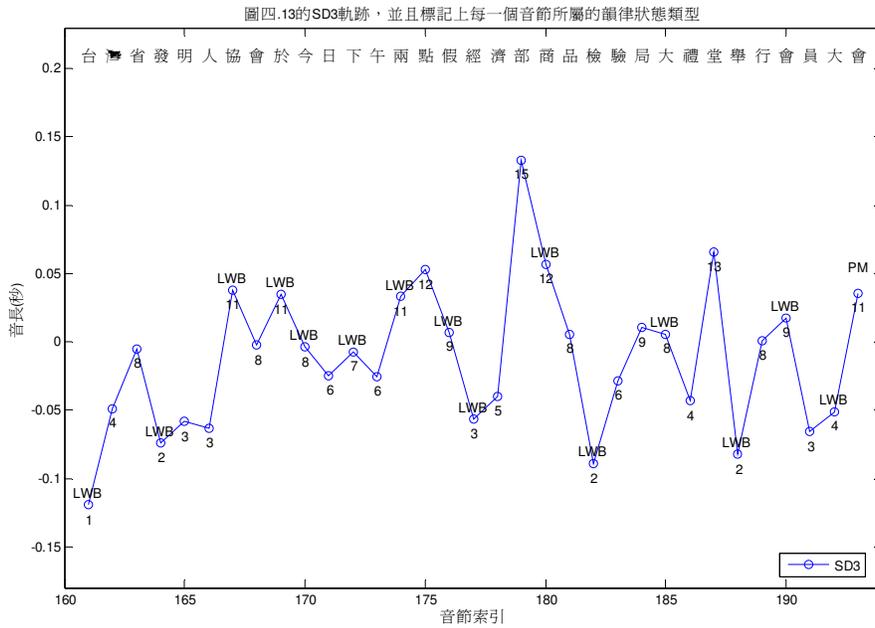


圖四.14：第三段句子的SD軌跡和LWB與PM之間的關係圖。

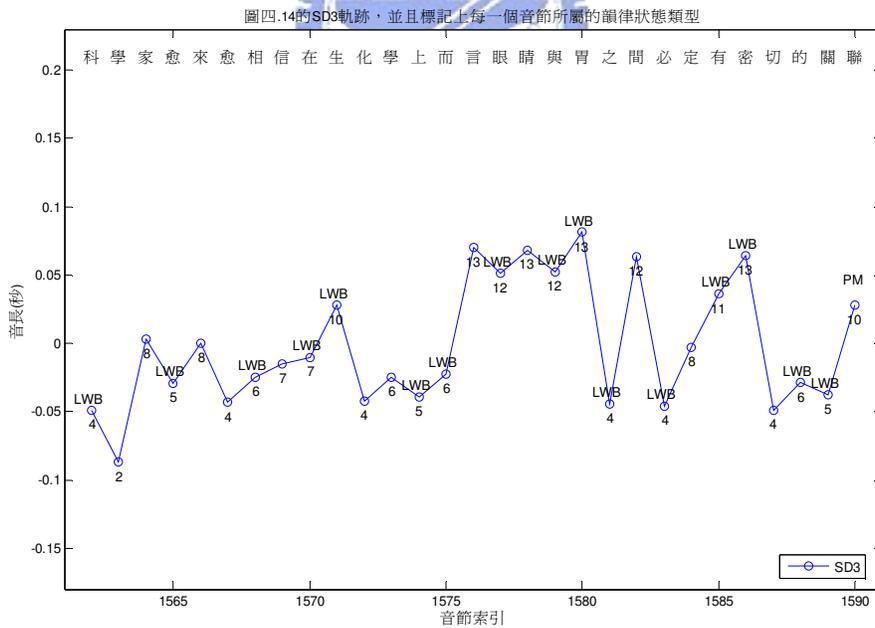
最後我們呈現出圖四.12、圖四.13 和圖四.14 的SD3軌跡，並且標記上每一個音節所屬的韻律狀態類型，如圖四.15、圖四.16 和圖四.17 所示：



圖四.15：圖四.12 的SD3軌跡，並且標記上每一個音節所屬的韻律狀態類型。



圖四.16：圖四.13 的 SD3 軌跡，並且標記上每一個音節所屬的韻律狀態類型。



圖四.17：圖四.14 的 SD3 軌跡，並且標記上每一個音節所屬的韻律狀態類型。

4.4.2.1 實驗結果分析

首先我們從圖四.11 可以觀察到音長軌跡除了減去 γ_s 只是平移外，減去其它的影響因素皆會使得音長軌跡逐漸平滑化，其之意義即為我們的模型能夠將造成每一個音節之音長差異的因素分離並且抽取出來，亦即驗證了我們的模型所造成的音長軌跡變化是合理的。

再來我們從圖四.12、圖四.13 和圖四.14 可以觀察到當一個音節其後跟隨著標點符號時，這個音節在其所屬的語法片語中有較長的音長，再來我們可以觀察到二字詞中第二個字的音長通常會比第一個字的音長還長，亦即音長軌跡呈現“∨”的 Pattern；而三字詞中第一個字和第三個字的音長通常會比第二個字的音長還長，亦即音長軌跡呈現“∨”的 Pattern。

最後於圖四.15、圖四.16 和圖四.17 中的 *SD3* 軌跡代表了高階語言層次音長軌跡，我們可以觀察到詞尾或句尾的音節通常含有較高的韻律狀態類型而詞中的音節通常含有較低的韻律狀態類型，這個現象告訴我們人類說話的韻律模式是以韻律片語為單位一個韻律片語可能是一個詞或包含多個詞，而由圖中我們亦可以觀察到所熟知的 Lengthening Effect。



第五章 音節之韻律狀態類型的分析

本章對於音長模型在收斂狀態下對語料庫中每一個音節所標示的韻律狀態類型做分析，我們希望能從中觀察出其代表著高階語言層次之音長特性的現象。我們也使用參考文獻 [11] 裡作者以韻律訊息中的音高(Pitch)訊息和停頓(Pause)訊息以及一些語言參數對本語料庫所標示的音節間之停頓類型的 Break Type，來觀察其和音節之韻律狀態類型的相關性。最後我們也使用文獻裡作者對語料庫中每一個音節所標示的音高韻律狀態類型，來觀察其和本論文的音長韻律狀態類型之相關性。

5.1 單連文(Unigram)機率

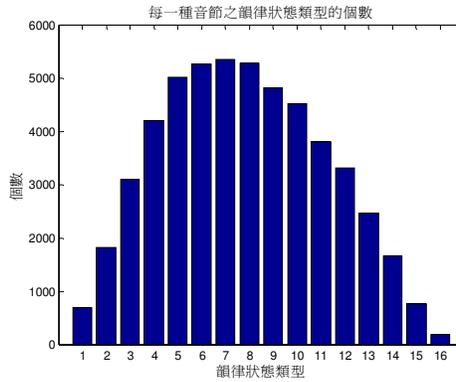
音節之韻律狀態類型的單連文機率 $P(ps)$, $ps \in \{1, 2, \dots, 16\}$ 計算方式如下的數學式所述

$$P(ps) = \frac{C(ps)}{\sum_{i=1}^{16} C(i)}, ps \in \{1, 2, \dots, 16\} \quad (5.1)$$

，其中 $C(ps)$, $ps \in \{1, 2, \dots, 16\}$ 為每一種音節之韻律狀態類型的個數(Count)，我們將其列出如表五.1 和繪出如圖五.1 所示：

表五.1： $C(ps)$ 。

	類型 1	類型 2	類型 3	類型 4	類型 5	類型 6	類型 7	類型 8
個數	691	1814	3089	4199	5013	5262	5340	5272
類型 9	類型 10	類型 11	類型 12	類型 13	類型 14	類型 15	類型 16	
4811	4515	3801	3307	2459	1665	767	187	

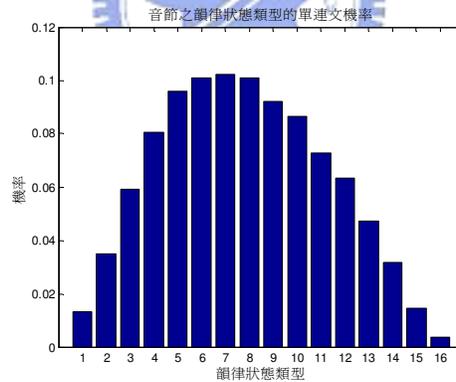


圖五.1： $C(ps)$ 。

我們將 $P(ps), ps \in \{1, 2, \dots, 16\}$ 列出如表五.2 和繪出如圖五.2 所示：

表五.2： $P(ps)$ 。

	類型 1	類型 2	類型 3	類型 4	類型 5	類型 6	類型 7	類型 8
機率	0.0132	0.0348	0.0592	0.0805	0.096	0.1008	0.1023	0.101
類型 9	類型 10	類型 11	類型 12	類型 13	類型 14	類型 15	類型 16	
0.0922	0.0865	0.0728	0.0634	0.0471	0.0319	0.0147	0.0036	



圖五.2： $P(ps)$ 。

5.1.1 實驗結果分析

由表五.2 和圖五.2 我們可以觀察到韻律狀態類型的單連文機率分佈為高斯分佈，其是符合我們所認知的語音特性。

5.2 雙連文(Bigram)機率

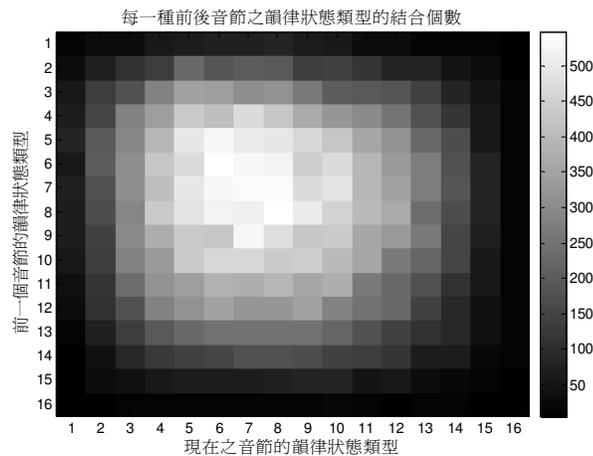
音節之韻律狀態類型的雙連文機率 $P(ps_n | ps_{n-1}), ps \in \{1, 2, \dots, 16\}$ 計算方式如下式所述：

$$P(ps_n | ps_{n-1}) = \frac{C(ps_{n-1}, ps_n)}{C(ps_{n-1})}, ps \in \{1, 2, \dots, 16\} \quad (5.2)$$

其中 $C(ps_{n-1}, ps_n), ps \in \{1, 2, \dots, 16\}$ 為每一種前後音節之韻律狀態類型的結合個數，我們將其列出如表五.3 和繪出如圖五.3 所示：

表五.3： $C(ps_{n-1}, ps_n)$ 。

$ps_{n-1} \backslash ps_n$	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
1	16	30	36	60	63	79	79	81	69	55	35	35	19	15	14	5
2	35	73	112	136	233	186	202	197	132	144	116	86	82	45	30	5
3	57	133	181	280	349	339	309	314	266	200	193	184	141	84	48	11
4	67	150	277	345	434	413	469	425	367	327	297	258	174	122	56	18
5	86	195	288	390	494	533	502	494	462	430	359	312	226	168	59	15
6	61	208	311	427	470	552	528	524	443	469	393	329	269	177	85	16
7	79	182	309	410	500	529	537	541	467	484	392	351	272	188	84	15
8	68	172	293	439	489	523	513	544	503	451	405	367	240	168	82	15
9	66	141	299	380	435	428	528	480	425	432	362	330	267	150	72	16
10	56	134	285	335	434	458	459	438	446	403	358	265	222	149	58	15
11	38	122	240	318	344	387	377	391	360	372	266	232	175	111	55	13
12	30	113	181	279	312	351	322	324	353	279	246	230	143	94	39	11
13	16	78	134	198	233	247	251	250	251	223	178	145	111	94	39	11
14	10	44	92	130	143	156	177	182	170	144	136	118	65	63	23	11
15	4	30	40	59	67	70	70	75	83	81	50	57	34	22	17	8
16	2	9	11	13	12	11	17	12	14	21	15	8	19	15	6	2

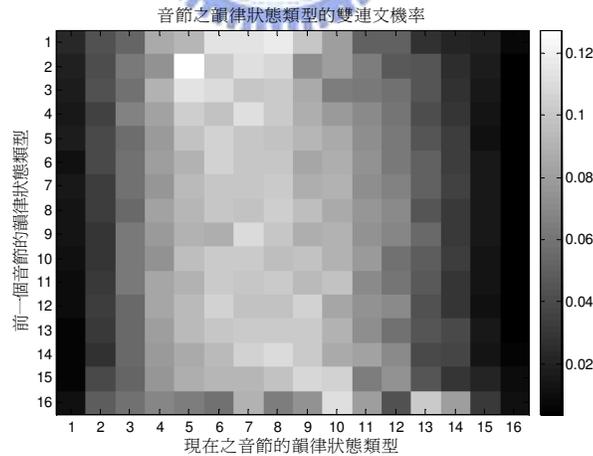


圖五.3： $C(ps_{n-1}, ps_n)$ 。

我們將 $P(ps_n | ps_{n-1})$, $ps \in \{1, 2, \dots, 16\}$ 列出如表五.4 和繪出如圖五.4 所示：

表五.4： $P(ps_n | ps_{n-1})$ ，其單位為 10^{-3} 。

$ps_{n-1} \backslash ps_n$	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
1	23	43	52	87	91	114	114	117	100	80	51	51	27	22	20	7
2	19	40	62	75	128	103	111	109	73	79	64	47	45	25	17	3
3	18	43	59	91	113	110	100	102	86	65	62	60	46	27	16	4
4	16	36	66	82	103	98	112	101	87	78	71	61	41	29	13	4
5	17	39	57	78	99	106	100	99	92	86	72	62	45	34	12	3
6	12	40	59	81	89	105	100	100	84	89	75	63	51	34	16	3
7	15	34	58	77	94	99	101	101	87	91	73	66	51	35	16	3
8	13	33	56	83	93	99	97	103	95	86	77	70	46	32	16	3
9	14	29	62	79	90	89	110	100	88	90	75	69	55	31	15	3
10	12	30	63	74	96	101	102	97	99	89	79	59	49	33	13	3
11	10	32	63	84	91	102	99	103	95	98	70	61	46	29	14	3
12	9	34	55	84	94	106	97	98	107	84	74	70	43	28	12	3
13	7	32	54	81	95	100	102	102	102	91	72	59	45	38	16	4
14	6	26	55	78	86	94	106	109	102	86	82	71	39	38	14	7
15	5	39	52	77	87	91	91	98	108	106	65	74	44	29	22	10
16	11	48	59	70	64	59	91	64	75	112	80	43	102	80	32	11



圖五.4： $P(ps_n | ps_{n-1})$ 。

5.2.1 實驗結果分析

由表五.4 和圖五.4 我們可以觀察到首先由前到後之韻律狀態類型的變化，大部份皆是

“低到中”、“中到中”和“高到中”，鮮少有“低到低”、“低到高”、“中到低”和“中到高”的情形出現，但是要注意“高到低”的機率不小其即是我們所熟知的詞尾和句尾之 Lengthening Effect，再者“高到高”尤其是 16 到 13 和 14 的機率亦不小其也是符合我們所認知的說話音長特性。

5.3 和 Break Type 的相關性觀察

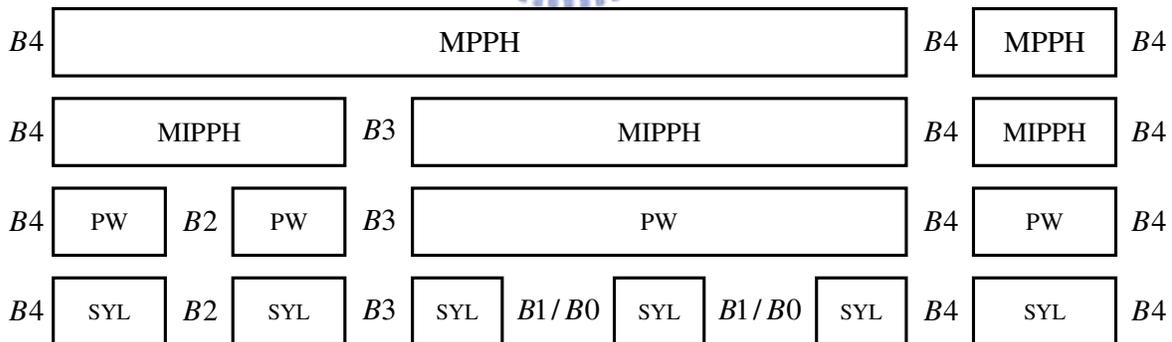
這一節中首先對於本論文所使用的 Break Type 做介紹，再來要觀察音節之韻律狀態類型和 Break Type 的相關性。

5.3.1 Break Type 的介紹

在 3.2.2 節中我們已經對於參考文獻[11]裡作者以「潛藏韻律模擬法」所建立的一套以模型為基礎之自動韻律標記方法，來對語料庫自動地標記上韻律線索做了初步的說明，而在本小節中我們將對於本論文所對其引用的部份做更進一步的說明。

5.3.1.1 漢語語音階層韻律結構

我們將參考文獻[11]中作者所述的漢語語音階層韻律結構繪出如圖五.5 所示：



圖五.5：漢語語音階層韻律結構。

圖中位於最底層的音節(Syllable, SYL)為最基本的韻律單位(Prosody Unit)；一個或多個 SYL 則構成位於第二層的韻律詞(Prosody Word, PW)；一個或多個 PW 則構成位於第三層的次要韻律片語(Minor Prosody Phrase, MIPPH)；一個或多個 MIPPH 則構成位於最上層的主要韻律片語(Major Prosody Phrase, MPPH)；圖中所示的 B0、B1、B2、B3 和 B4 皆為音節間之停頓

類型的 Break Type，我們將其條列說明如下：

B_0 ：Intra-PW 的音節間之停頓，音節間為 Tightly Couple；

B_1 ：Intra-PW 的音節間之停頓，音節間為 Normally Couple；

B_2 ：Inter-PW 的音節間之停頓，我們依停頓特性將其分為以下兩類，

B_2-1 ：音節間為 Minor Pitch Reset，

B_2-2 ：音節間為 Short Pause；

B_3 ：Inter-MIPPH 的音節間之停頓，音節間為 Medium Pause，其通常亦伴隨著 Medium Pitch Reset；

B_4 ：Inter-MPPH 的音節間之停頓，音節間為 Long Pause，其通常亦伴隨著 Large Pitch Reset。

5.3.1.2 Break Type 的標記

參考文獻[11]中作者將所提出的數學模型表示成一個參數最佳化的問題，如以下的數學式推導所述

$$\begin{aligned}
 B^*, PS^* &= \arg \max_{B, PS} P(B, PS | SP, PD, PED, L, T) \\
 &= \arg \max_{B, PS} P(B, PS, SP, PD, PED | L, T) \\
 &= \arg \max_{B, PS} P(SP, PD, PED | B, PS, L, T) P(B, PS | L, T) \\
 &\approx \arg \max_{B, PS} P(SP | B, PS, L, T) P(PD, PED | B, PS, L, T) P(B, PS | L) \\
 &\approx \arg \max_{B, PS} P(SP | B, PS, T) P(PD, PED | B, L) P(PS | B, L) P(B | L) \\
 &\approx \arg \max_{B, PS} P(SP | B, PS, T) P(PD, PED | B, L) P(PS | B) P(B | L)
 \end{aligned} \tag{5.3}$$

，我們將其中之各個參數條列說明：

- $B = \{b_{k,n} | n = 1, 2, \dots, N_k; k = 1, 2, \dots, K\}$, $b_{k,n} \in \{B_0, B_1, B_2-1, B_2-2, B_3, B_4\}$ ， $b_{k,n}$ 代表語料庫中第 k 個 Utterance 裡第 n 個音節後的 Break Type；
- $PS = \{ps_{k,n} | n = 1, 2, \dots, N_k; k = 1, 2, \dots, K\}$, $ps_{k,n} \in \{1, 2, \dots, 16\}$ ， $ps_{k,n}$ 代表語料庫中第 k 個 Utterance 裡第 n 個音節所屬的韻律狀態類型；
- $SP = \{sp_{k,n} | n = 1, 2, \dots, N_k; k = 1, 2, \dots, K\}$ ， $sp_{k,n}$ 代表語料庫中第 k 個 Utterance 裡第 n

個音節的四個音高軌跡正交展開係數向量；

- $PD = \{pd_{k,n} \mid n=1,2,\dots,N_k; k=1,2,\dots,K\}$ ， $pd_{k,n}$ 代表語料庫中第 k 個 Utterance 裡第 n 個音節後的停頓長度(Pause Duration)；
- $PED = \{ped_{k,n} \mid n=1,2,\dots,N_k; k=1,2,\dots,K\}$ ， $ped_{k,n}$ 代表語料庫中第 k 個 Utterance 裡第 n 個音節後的 Pause Energy Dip；
- $L = \{l_{k,n} \mid n=1,2,\dots,N_k; k=1,2,\dots,K\}$ ， $l_{k,n}$ 代表語料庫中第 k 個 Utterance 裡第 n 個音節後之 Pause 的前後語言參數(Contextual Linguistic Feature)；
- $T = \{t_{k,n} \mid n=1,2,\dots,N_k; k=1,2,\dots,K\}$ ， $t_{k,n}$ 代表語料庫中第 k 個 Utterance 裡第 n 個音節的聲調類型。

由(5.3)式的數學式推導可以得到四個數學模型，我們將其條列說明如下：

- $P(SP \mid B, PS, T) = \prod_{k=1}^K \prod_{n=1}^{N_k} P(sp_{k,n} \mid b_{k,n-1}, b_{k,n}, ps_{k,n}, t_{k,n-1}, t_{k,n}, t_{k,n+1})$ 為音高軌跡模型，代表著音高軌跡與其鄰近的 Prosodic Cues 和音節之聲調類型之間的關係；
- $P(PD, PED \mid B, L) = \prod_{k=1}^K \prod_{n=1}^{N_k} P(pd_{k,n}, ped_{k,n} \mid b_{k,n}, l_{k,n})$ 為停頓聲學(Pause Acoustic)模型，代表著停頓長度和 Pause Energy Dip 與 Break Type 和鄰近的語言參數之間的關係；
- $P(PS \mid B) = \prod_{k=1}^K P(ps_{k,1}) \prod_{n=2}^{N_k} P(ps_{k,n} \mid ps_{k,n-1}, b_{k,n-1})$ ，其中 $P(ps_{k,1})$ 為韻律狀態類型初始機率， $P(ps_{k,n} \mid ps_{k,n-1}, b_{k,n-1})$ 為韻律狀態類型轉移機率；
- $P(B \mid L) = \prod_{k=1}^K \prod_{n=1}^{N_k} P(b_{k,n} \mid l_{k,n})$ 為 Break-Syntax Model，代表著 Break Type 和鄰近的語言參數之間的關係。

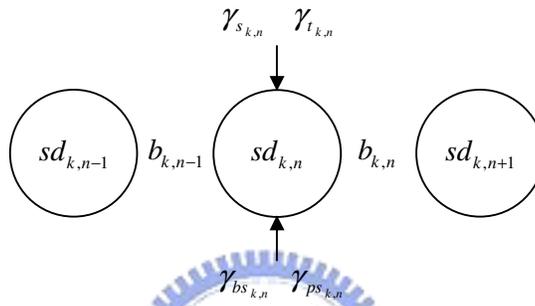
另一方面，作者使用了數種層次的語言參數來建立模型，我們將其條列說明如下：

- **音節層次**：聲母類型、詞間、詞內、...；
- **詞層次**：標點符號、詞長、Specific Word、POS、...；
- **Syntactic Tree Level**：Syntactic Phrase Length、Syntactic Phrase Boundary、...；
- **Sentence Level**：Sentence Length、...。

以上我們將參考文獻[11]中作者對於自動韻律標記方法所提出的數學模型與其所使用到的語言參數做了初步的說明，如此我們便可以知道在本論文之後的部份所使用到的語料庫之 Break Type 標記的由來，至於對其之更完整的敘述請參閱參考文獻[11]。

5.3.1.3 Break Type 於本論文的應用

我們將上述參考文獻[11]中的 Break Type 應用在本論文裡，圖五.6 為本論文中音長和四個影響因素與 Break Type 的數學符號關係圖。



圖五.6：音長和四個影響因素與 Break Type 的數學符號關係圖。

在參考文獻[11]中作者於程式裡將每一種 Break Type 以一個 b 值來代表，如表五.5 所示：

表五.5：Break Type 和 b 值的對應表。

Break Type	$B0$	$B1$	$B2-1$	$B3$	$B4$	$B2-2$
b	1	2	3	4	5	6

而為了在往後所製作的圖形上之視覺的直觀考量，我們更改了 $B3$ 、 $B4$ 和 $B2-2$ 的順序，如表五.6 所示：

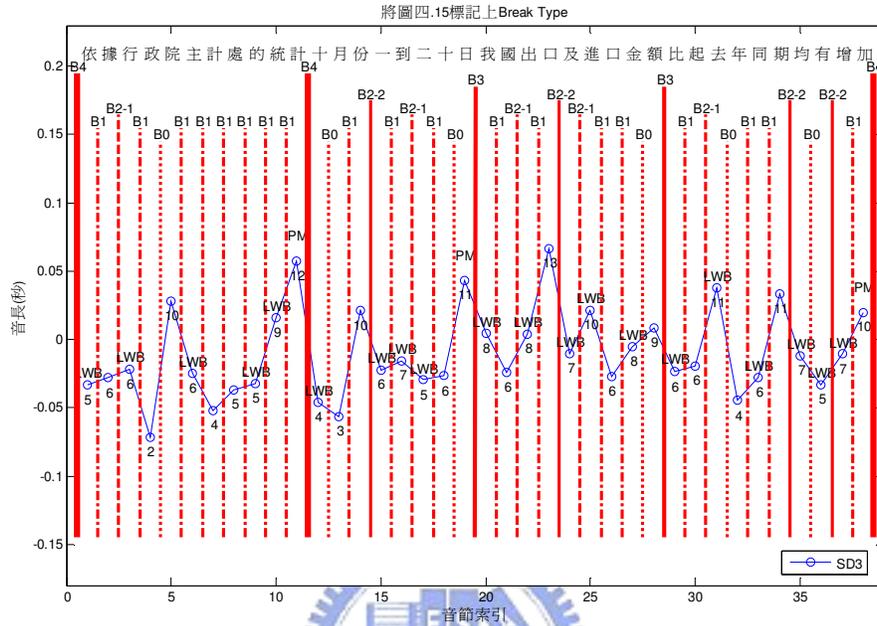
表五.6：改變順序後的 Break Type 和 b 值對應表。

Break Type	$B0$	$B1$	$B2-1$	$B2-2$	$B3$	$B4$
b	1	2	3	4	5	6

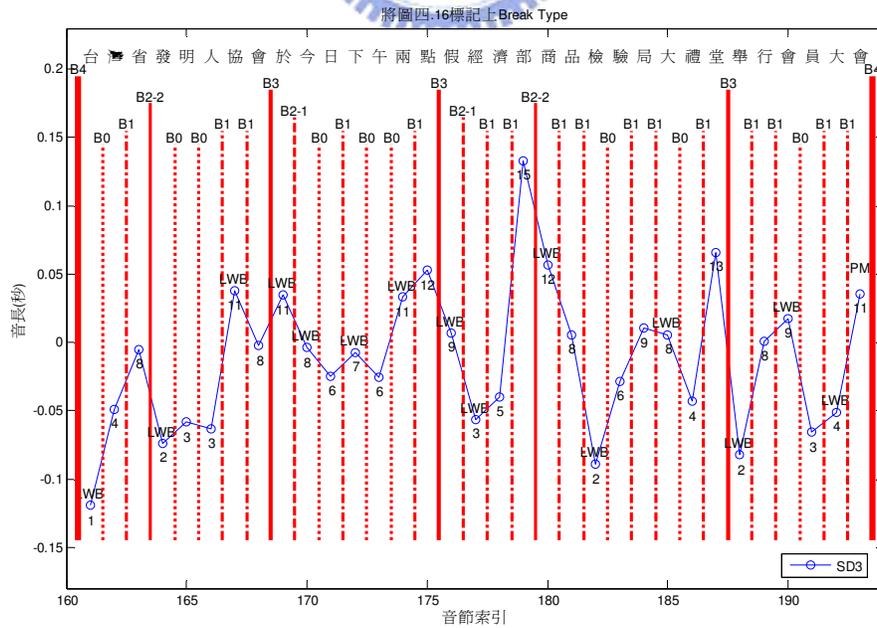
注意在語料庫中 379 個音檔的前後之 Break Type 我們將其設定為 $B4$ ，尤其留意第一個音節的前 Break Type 為 $B4$ 。

5.3.2 音長軌跡和 Break Type 的相關性觀察

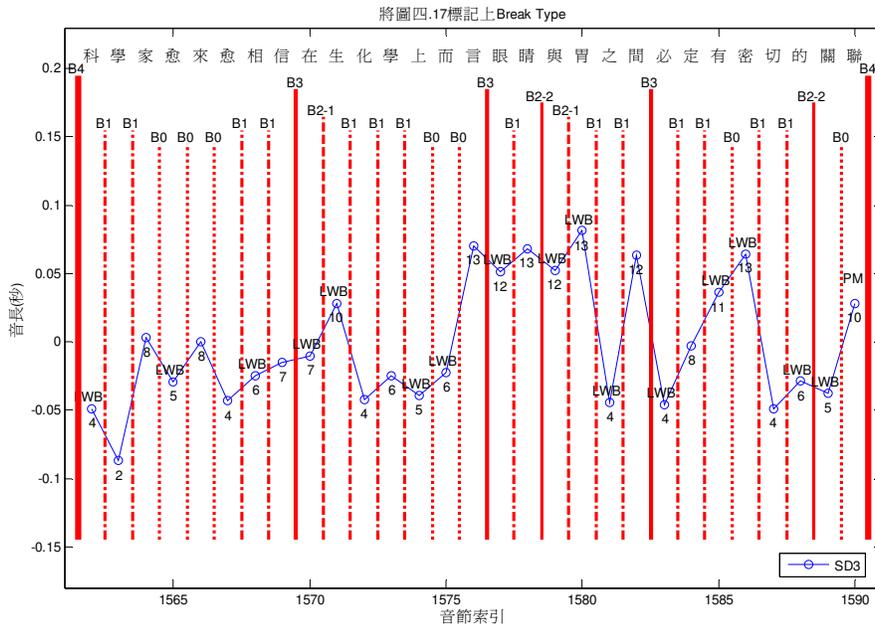
我們將圖四.13、圖四.14 和圖四.15 標記上 Break Type 如圖五.7、圖五.8 和圖五.9 所示：



圖五.7：將圖四.15 標記上 Break Type。



圖五.8：將圖四.16 標記上 Break Type。



圖五.9：將圖四.17標記上 Break Type。

5.3.2.1 實驗結果分析

由以上三圖我們可以觀察到在 4.4.2.1 節中所提及的音長軌跡之 Pattern 是以 $B2$ 、 $B3$ 和 $B4$ 為邊界來代表其現象，而 $B0$ 和 $B1$ 則無，換句話說我們在說話時的音長特性是以 PW 、 $MIPPH$ 和 $MPPH$ 為韻律單位，其分別有各自的 Pattern。對於這三種 Pattern 我們將在第六章中對其做詳細的說明，並且以數學方法再建立一個音長模型來將其顯現出來。

5.3.3 和前 Break Type 的雙連文機率

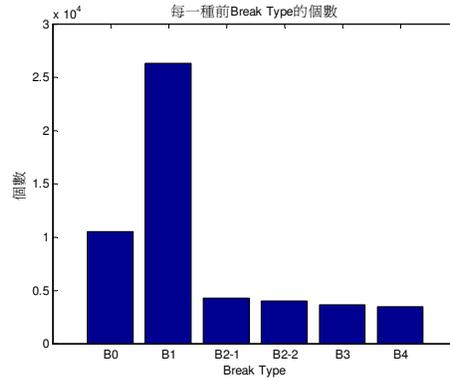
音節之韻律狀態類型和前 Break Type 的雙連文機率 $P(ps_n | b_{n-1})$, $ps \in \{1, 2, \dots, 16\}$, $b \in \{B0, B1, B2-1, B2-2, B3, B4\}$ 計算方式如下式所述：

$$P(ps_n | b_{n-1}) = \frac{C(b_{n-1}, ps_n)}{C(b_{n-1})}, ps \in \{1, 2, \dots, 16\}, b \in \{B0, B1, B2-1, B2-2, B3, B4\} \quad (5.3)$$

其中 $C(b_{n-1})$, $b \in \{B0, B1, B2-1, B2-2, B3, B4\}$ 為每一種前 Break Type 的個數，我們將其列出如表五.7 和繪出如圖五.10 所示：

表五.7 : $C(b_{n-1})$ 。

	$B0$	$B1$	$B2-1$	$B2-2$	$B3$	$B4$
個數	10499	26278	4213	4032	3682	3488

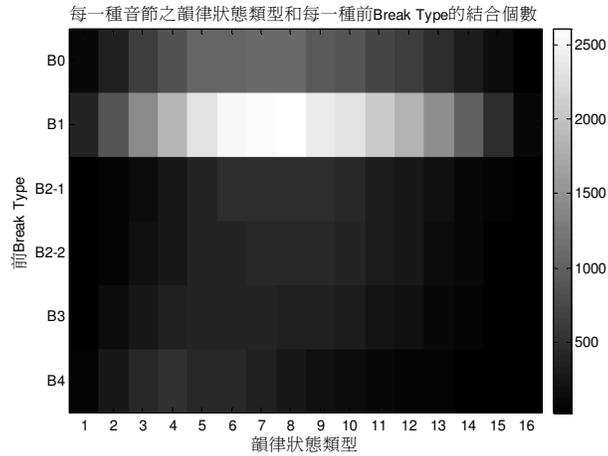


圖五.10 : $C(b_{n-1})$ 。

$C(b_{n-1}, ps_n)$, $ps \in \{1, 2, \dots, 16\}$, $b \in \{B0, B1, B2-1, B2-2, B3, B4\}$ 為每一種音節之韻律狀態類型和每一種前 Break Type 的結合個數，我們將其列出如表五.8 和繪出如圖五.11 所示：

表五.8 : $C(b_{n-1}, ps_n)$ 。

$b_{n-1} \setminus ps_n$	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
$B0$	123	337	648	855	1054	1054	1101	1075	935	894	743	649	493	331	164	43
$B1$	415	905	1405	1886	2328	2525	2550	2628	2408	2320	2085	1825	1441	990	466	101
$B2-1$	12	79	151	273	405	459	486	472	476	419	317	273	204	125	56	6
$B2-2$	28	85	201	274	391	397	453	441	428	410	307	288	163	109	39	18
$B3$	39	141	256	374	388	392	401	371	358	316	240	198	108	70	24	6
$B4$	74	267	428	537	447	435	349	285	206	156	109	74	50	40	18	13

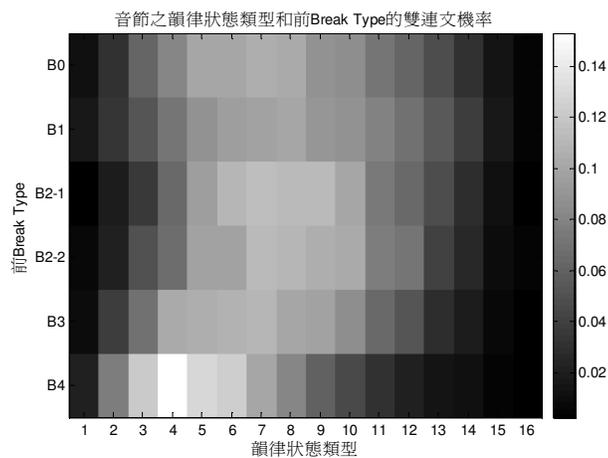


圖五.11： $C(b_{n-1}, ps_n)$ 。

我們將 $P(ps_n | b_{n-1})$, $ps \in \{1, 2, \dots, 16\}$, $b \in \{B0, B1, B2-1, B2-2, B3, B4\}$ 列出如表五.9 和繪出如圖五.12 所示：

表五.9： $P(ps_n | b_{n-1})$ ，其單位為 10^{-4} 。

$b_{n-1} \backslash ps_n$	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
B0	117	321	617	814	1004	1004	1049	1024	891	852	708	618	470	315	156	41
B1	158	344	535	718	886	961	970	1000	916	883	793	695	548	377	177	38
B2-1	28	188	358	648	961	1090	1154	1120	1130	995	752	648	484	297	133	14
B2-2	69	211	499	680	970	985	1124	1094	1062	1017	761	714	404	270	97	45
B3	106	383	695	1016	1054	1065	1089	1008	972	858	652	538	293	190	65	16
B4	212	765	1227	1540	1282	1247	1001	817	591	447	313	212	143	115	52	37



圖五.12： $P(ps_n | b_{n-1})$ 。

5.3.3.1 實驗結果分析

由表五.9 和圖五.12 我們可以觀察到首先當前 Break Type 為 $B4$ 時，通常是結束一段句子而重新另一段句子，此時一般伴隨著重新吸氣因此說話速度會較快；再來當前 Break Type 為 $B3$ 時，一個不吸氣但明顯的停頓產生，此時說話速度也會傾向加快；再者當前 Break Type 為 $B2-1$ 或 $B2-2$ 時，我們將其留在下一個小節“和後 Break Type 的雙連文機率”的實驗結果分析中，與“當後 Break Type 為 $B2-1$ 或 $B2-2$ 時”一併做分析；最後對於當前 Break Type 為 $B0$ 或 $B1$ 時其並沒有特別的訊息出現。

5.3.4 和後 Break Type 的雙連文機率

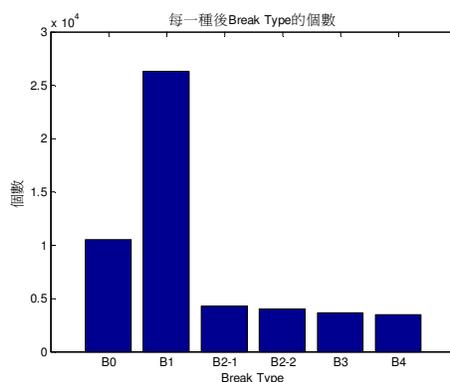
音節之韻律狀態類型和後 Break Type 的雙連文機率 $P(ps_n | b_n)$, $ps \in \{1, 2, \dots, 16\}$, $b \in \{B0, B1, B2-1, B2-2, B3, B4\}$ 計算方式如下式所述：

$$P(ps_n | b_n) = \frac{C(ps_n, b_n)}{C(b_n)}, ps \in \{1, 2, \dots, 16\}, b \in \{B0, B1, B2-1, B2-2, B3, B4\} \quad (5.4)$$

其中 $C(b_n)$, $b \in \{B0, B1, B2-1, B2-2, B3, B4\}$ 為每一種後 Break Type 的個數，我們將其列出如表五.10 和繪出如圖五.13 所示：

表五.10： $C(b_n)$ 。

	$B0$	$B1$	$B2-1$	$B2-2$	$B3$	$B4$
個數	10499	26278	4213	4032	3682	3488



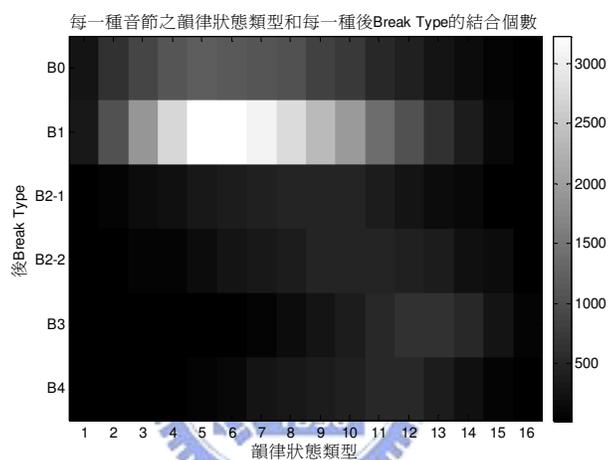
圖五.13： $C(b_n)$ 。

$C(ps_n, b_n)$, $ps \in \{1, 2, \dots, 16\}$, $b \in \{B0, B1, B2-1, B2-2, B3, B4\}$ 為每一種音節之韻律狀態類型和每

一種後 Break Type 的結合個數，我們將其列出如表五.11 和繪出如圖五.14 所示：

表五.11： $C(ps_n, b_n)$ 。

$b_n \backslash ps_n$	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
B0	293	653	911	1081	1181	1168	1113	1068	853	733	535	416	264	157	58	15
B1	349	1042	1926	2736	3219	3256	3095	2793	2353	1959	1384	1019	634	361	120	32
B2-1	40	100	184	249	322	388	446	508	468	493	364	280	191	110	50	20
B2-2	9	13	51	95	174	261	323	406	463	488	483	449	373	249	160	35
B3	0	1	4	8	24	37	97	158	278	406	518	630	625	538	295	63
B4	0	5	13	30	93	152	266	339	396	436	517	513	372	250	84	22

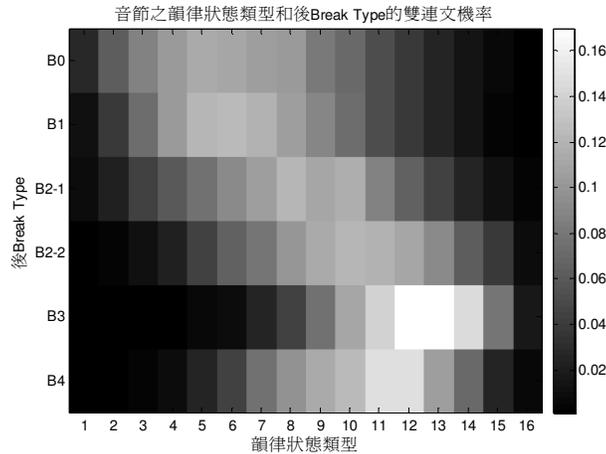


圖五.14： $C(ps_n, b_n)$ 。

我們將 $P(ps_n | b_n)$, $ps_n \in \{1, 2, \dots, 16\}$, $b_n \in \{B0, B1, B2-1, B2-2, B3, B4\}$ 列出如表五.12 和繪出如圖五.15 所示：

表五.12： $P(ps_n | b_n)$ ，其單位為 10^{-4} 。

$b_n \backslash ps_n$	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
B0	279	622	868	1030	1125	1113	1060	1017	812	698	510	396	251	150	55	14
B1	133	397	733	1041	1225	1239	1178	1063	895	745	527	388	241	137	46	12
B2-1	95	237	437	591	764	921	1059	1206	1111	1170	864	665	453	261	119	47
B2-2	22	32	126	236	432	647	801	1007	1148	1210	1198	1114	925	618	397	87
B3	0	3	11	22	65	100	263	429	755	1103	1407	1711	1697	1461	801	171
B4	0	14	37	86	267	436	763	972	1135	1250	1482	1471	1067	717	241	63



圖五.15： $P(p_{s_n} | b_n)$ 。

5.3.4.1 實驗結果分析

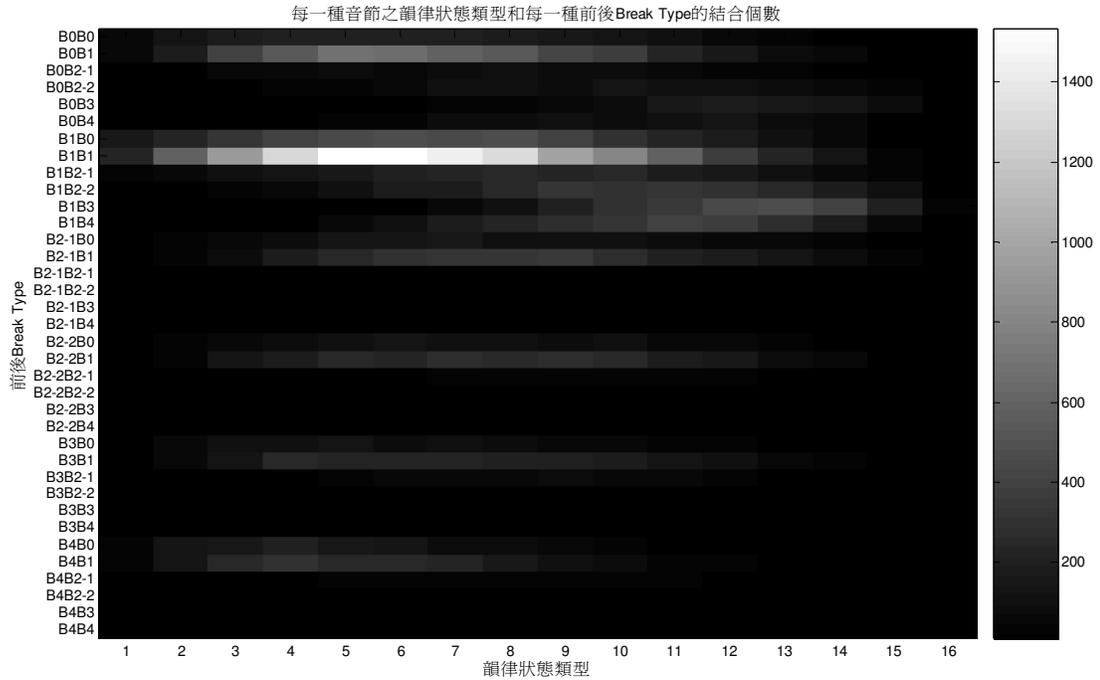
由表五.12和圖五.15我們可以觀察到首先當後 Break Type 為 B_3 或 B_4 時通常是一個語法片語或一段句子的結束，此時 Lengthening Effect 即顯現出來；再來當後 Break Type 為 B_2-1 或 B_2-2 時我們將其與表五.9和圖五.12做比較可以發現，Break Type 之前的韻律狀態類型會高於之後，此即顯現出了韻律詞的音長軌跡之 Pattern；最後對於當後 Break Type 為 B_0 或 B_1 時其並沒有特別的訊息出現。

5.3.5 和前後 Break Type 的結合個數

我們將每一種音節之韻律狀態類型和每一種前後 Break Type 的結合個數 $C(b_{n-1}, p_{s_n}, b_n)$, $p_{s_n} \in \{1, 2, \dots, 16\}$, $b_n \in \{B_0, B_1, B_2-1, B_2-2, B_3, B_4\}$ 列出如表五.13和繪出如圖五.16所示：

表五.13 : $C(b_{n-1}, ps_n, b_n)$ 。

$(b_{n-1}, b_n) \setminus ps_n$	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
B0 B0	53	126	180	203	212	205	208	178	156	124	98	54	44	15	6	1
B0 B1	62	186	401	546	689	664	603	550	426	364	231	158	93	61	17	7
B0 B2-1	6	18	50	57	75	64	86	102	88	89	57	34	26	14	6	4
B0 B2-2	2	6	12	35	41	59	97	108	95	127	99	109	88	49	36	9
B0 B3	0	0	1	0	8	15	26	43	68	96	149	173	155	126	80	15
B0 B4	0	1	4	14	29	47	81	94	102	94	109	121	87	66	19	7
B1 B0	156	241	333	390	450	475	447	476	388	313	218	181	101	59	23	6
B1 B1	224	580	923	1284	1538	1546	1435	1306	975	808	602	373	236	123	43	12
B1 B2-1	28	72	102	136	156	211	232	244	236	256	179	159	113	60	27	11
B1 B2-2	7	7	35	52	106	170	189	253	317	307	325	296	257	183	107	21
B1 B3	0	1	3	8	15	21	69	110	207	307	362	445	461	388	206	42
B1 B4	0	4	9	16	63	102	178	239	285	329	399	371	273	177	60	9
B2-1 B0	8	46	64	87	127	132	145	120	107	118	90	71	56	36	16	3
B2-1 B1	4	31	83	177	265	310	320	333	354	287	208	190	136	79	34	1
B2-1 B2-1	0	2	2	6	3	5	9	7	8	2	2	2	3	2	1	0
B2-1 B2-2	0	0	2	3	8	9	6	7	2	5	7	1	3	0	1	0
B2-1 B3	0	0	0	0	1	1	2	2	0	1	4	2	1	5	3	2
B2-1 B4	0	0	0	0	1	2	4	3	5	6	6	7	5	3	1	0
B2-2 B0	13	38	63	78	110	128	118	118	83	98	65	60	34	24	10	4
B2-2 B1	15	45	133	184	256	238	286	261	286	253	176	164	94	58	19	8
B2-2 B2-1	0	2	5	10	18	17	33	37	31	31	40	36	18	11	4	2
B2-2 B2-2	0	0	0	2	7	13	13	20	23	21	23	19	11	8	3	2
B2-2 B3	0	0	0	0	0	0	0	2	1	1	1	2	5	7	2	1
B2-2 B4	0	0	0	0	0	1	3	3	4	6	2	7	1	1	1	1
B3 B0	22	69	102	107	121	93	101	91	63	49	40	41	23	14	2	1
B3 B1	14	68	142	245	218	233	226	196	195	172	124	101	57	31	4	3
B3 B2-1	3	4	12	19	39	57	56	71	76	71	54	31	19	15	10	2
B3 B2-2	0	0	0	3	10	9	18	13	23	24	22	21	8	7	8	0
B3 B3	0	0	0	0	0	0	0	0	1	0	0	4	0	3	0	0
B3 B4	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0
B4 B0	41	133	169	216	161	135	94	85	56	31	24	9	6	9	1	0
B4 B1	30	132	244	300	253	265	225	147	117	75	43	33	18	9	3	1
B4 B2-1	3	2	13	21	31	34	30	47	29	44	32	18	12	8	2	1
B4 B2-2	0	0	2	0	2	1	0	5	3	4	7	3	6	2	5	3
B4 B3	0	0	0	0	0	0	0	1	1	1	2	4	3	9	4	3
B4 B4	0	0	0	0	0	0	0	0	0	1	1	7	5	3	3	5



圖五.16： $C(b_{n-1}, ps_n, b_n)$ 。

5.3.5.1 實驗結果分析

我們由表五.13和圖五.16可以觀察到鮮少有相鄰的 Break Type 為 $B2-1$ 、 $B2-2$ 、 $B3$ 和 $B4$ ，絕大部份其之前或之後皆會伴隨著 $B0$ 或 $B1$ 的出現，此現象即告訴了我們人類說話的音長韻律單位絕大部份是以 $B2-1$ 、 $B2-2$ 、 $B3$ 和 $B4$ 為邊界並且含有一個以上的音節。

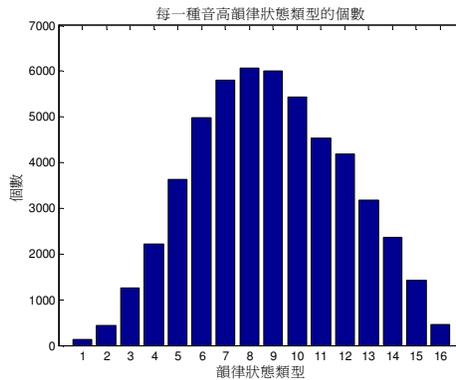
5.4 音長和音高韻律狀態類型的相關性觀察

我們將參考文獻[11]中作者以上述(5.3)式的數學模型對語料庫所標記的音高韻律狀態類型(Pitch ps , pps)，來對於本論文中的音長韻律狀態類型(Duration ps , dps)做相關性的比較，我們要統計出每一個音節其之每一種 pps 和 dps 的結合個數 $C(pps, dps)$, $pps \in \{1, 2, \dots, 16\}$, $dps \in \{1, 2, \dots, 16\}$ 。

首先我們將每一種音高韻律狀態類型的個數 $C(pps)$, $pps \in \{1, 2, \dots, 16\}$ 列出如表五.14和繪出如圖五.17所示：

表五.14： $C(pps)$ 。

	類型 1	類型 2	類型 3	類型 4	類型 5	類型 6	類型 7	類型 8
個數	127	441	1266	2218	3624	4993	5806	6067
類型 9	類型 10	類型 11	類型 12	類型 13	類型 14	類型 15	類型 16	
6005	5440	4525	4185	3191	2366	1426	452	

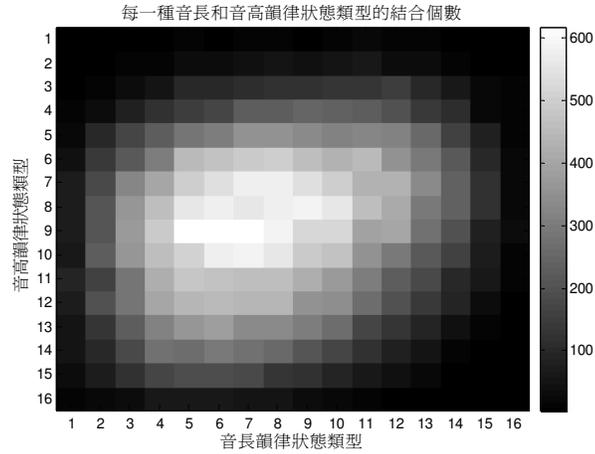


圖五.17： $C(pps)$ 。

再來我們將 $C(pps, dps)$, $pps \in \{1, 2, \dots, 16\}$, $dps \in \{1, 2, \dots, 16\}$ 列出如表五.15 和繪出如圖五.18 所示：

表五.15： $C(pps, dps)$ 。

$pps \backslash dps$	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
1	0	2	4	7	12	8	17	14	7	12	21	11	10	2	0	0
2	1	3	15	13	33	31	39	51	48	49	61	38	33	17	8	1
3	7	13	31	51	102	103	108	125	119	130	134	146	98	64	25	10
4	10	33	82	118	148	174	227	227	245	237	225	204	141	112	25	10
5	29	101	166	225	285	307	352	353	338	318	324	314	256	156	86	14
6	44	137	221	305	450	476	492	505	466	434	454	354	300	223	105	27
7	71	181	322	406	496	536	581	582	542	505	434	431	337	235	121	26
8	71	207	362	461	558	578	563	575	586	556	464	410	298	235	121	22
9	71	210	378	486	614	622	618	588	517	517	391	402	274	201	85	31
10	68	224	365	463	509	576	588	563	492	475	354	299	222	158	72	12
11	88	163	289	427	480	472	461	464	420	372	306	231	178	100	59	15
12	77	204	290	401	439	450	440	446	352	349	263	200	140	88	37	9
13	50	129	226	318	368	388	339	338	306	265	174	129	91	46	15	9
14	50	101	176	275	269	292	277	259	212	171	118	86	55	17	7	1
15	32	74	119	169	188	185	176	129	126	93	63	42	20	9	1	0
16	11	26	37	67	61	59	55	51	31	29	12	7	4	2	0	0



圖五.18： $C(pps, dps)$ 。

5.4.1 實驗結果分析

首先我們由表五.14 和圖五.17 可以觀察到其分佈與表五.1 和圖五.1 相似皆為高斯分佈，但整體而言圖五.17 為瘦高型而圖五.1 為矮胖型。再來我們由表五.15 和圖五.18 可以觀察到首先當一個音節的音長類型偏小也就是音長偏短的時候，音高類型偏大也就是音高偏高；再來當音長偏向適中的時候，音高除了絕大部份亦偏向適中以外偏低或偏高的機率也較前者為高；最後當音長偏長的時候，音高則會偏低。



第六章 高階音長模型的原理和數學說明

本章依據第五章所述的漢語語音階層韻律結構搭配著所觀察到的音長軌跡現象，建立一個高階語言層次階層音長模型(高階音長模型)，以數學的方式呈現出每一個階層之音長軌跡的 Pattern。本章內容首先對於高階語言層次的韻律訊息做介紹，而後對於模型的原理和數學做說明。

6.1 高階語言層次韻律訊息

本節將對於高階語言層次的韻律訊息做介紹，有別於在第二章所提及的低階音節層次之韻律訊息，高階語言層次的韻律訊息是屬於較大的韻律單元。我們將在以下的數個小節中以同於低階音節層次之韻律訊息的四個部份，其分別是音高、音量、音長和停頓，來對高階語言層次之韻律訊息做介紹。

6.1.1 音高訊息

一般而言，一句 Utterance 之音高軌跡的整體趨勢是類似浪潮般逐漸下降，並非是維持平穩，這種趨勢通常發生在以長停頓分開的 Utterance 之中，一般是伴隨著標點符號，我們將這種趨勢視為語調(Intonation)的變化。在逐漸下降的趨勢當中通常會伴隨著局部的浪潮般起伏，越到句尾起伏越趨緩和，這種局部的起伏是由於各個音節的聲調以及受其前後音節之影響所造成。而在以長停頓分開的語句當中通常會伴隨著 Pitch Reset 的現象，即下一個音節的 Pitch Level 突然增高，其通常是語者想要加重某個字或詞所造成，且其亦對應到停頓發生的地方。這種 Pitch Reset 的現象大部份是發生在功能詞(例如“出口及進口”的“及”字，其為連接詞，亦為功能詞的一種。)上和詞以及語法片語的邊界位置，在詞內則鮮少發生。

6.1.2 音量訊息

一般而言，一段語句的句首部份其音節通常含有比較高的音量，隨著發音到句尾其音量有逐漸下降的趨勢。這是一個很自然的現象，因為當語者在說一句話的時候一開始有比較足

夠的氣，隨著說到句尾氣漸漸地不足，所以才會有音量逐漸下降的趨勢。然而以上所述為一般的現象，語者有時候會在其想要加重的音節上以較大的音量來表達，因而造成局部的能量增強。如同其它韻律訊息般，這些加重音量的音節位置通常發生在語法片語的起始音節。

6.1.3 音長訊息

一般而言，一段語句的最後一個字或是詞的最後一個字會比句或詞中的字其音長要長一些，這就是所謂的 Lengthening Effect。通常這些拉長的音節所代表的是類似停頓的功能，以此方式來代表語意單位的結束。但是我們也可以觀察到並不是所有詞的最後一個音節皆會發生 Lengthening Effect，一般認為，相同於其它韻律訊息，Lengthening Effect 比較容易發生在語法片語和語句的最後一個字以及標點符號出現的地方。特別注意，因為不同的音節有不同的基本音節之 Phonetic Structure 與不同的聲調，故僅觀察音節是無法絕對的觀察到 Lengthening Effect。因此，我們在觀察 Lengthening Effect 的時候必須要先將基本音節和聲調對音長的影響去除掉，如此才能明確地觀察到高階語言層次韻律訊息中音長的 Lengthening Effect。



6.1.4 停頓訊息

一般而言，人們在朗讀文章時於標點符號出現的地方會產生明顯的停頓，以表示另一個語意段落的開始，然而在詞裡音節之間幾乎不會有停頓的發生，例如“行政院”的“行”和“政”與“政”和“院”之間。停頓的發生一般在詞和詞之間才有機會，其有兩點原因，第一是因為詞為最基本的語意單位；第二為語者為了能成功地傳遞訊息，因此會選擇在適當的詞結束位置處做停頓。再者我們可以觀察到雖然詞是最基本的語意單位，但是在韻律上數個詞可能連著一起唸，期間不會有停頓的發生，因而組成一個較大的韻律單元，我們稱之為韻律詞。經過眾多語料的觀察，除了標點符號所對應的明顯停頓以外，在功能詞或是詞以及語法片語的邊界處，也比較容易發生明顯的停頓。

6.2 高階音長模型的原理說明

本節首先將對於漢語的語法和韻律結構做說明，而後觀察高階語言層次的音長軌跡現

象，最後呈現出各個高階語言層次音長韻律單位的音長軌跡示意圖。

6.2.1 漢語語法結構

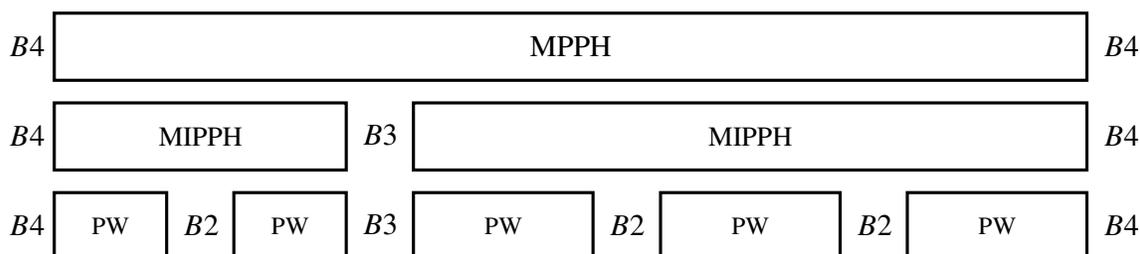
對於漢語文句的語法結構方面，漢語文句之組成的最基本單位為字(Character)，由一個字到數個字不等組成最小的語意單位詞(Lexical Word)，再由數個詞組成語法片語(Syntactic Phrase)，接著由數個語法片語組成句子(Sentence)，最後由數個句子組成文章。

6.2.2 漢語韻律結構

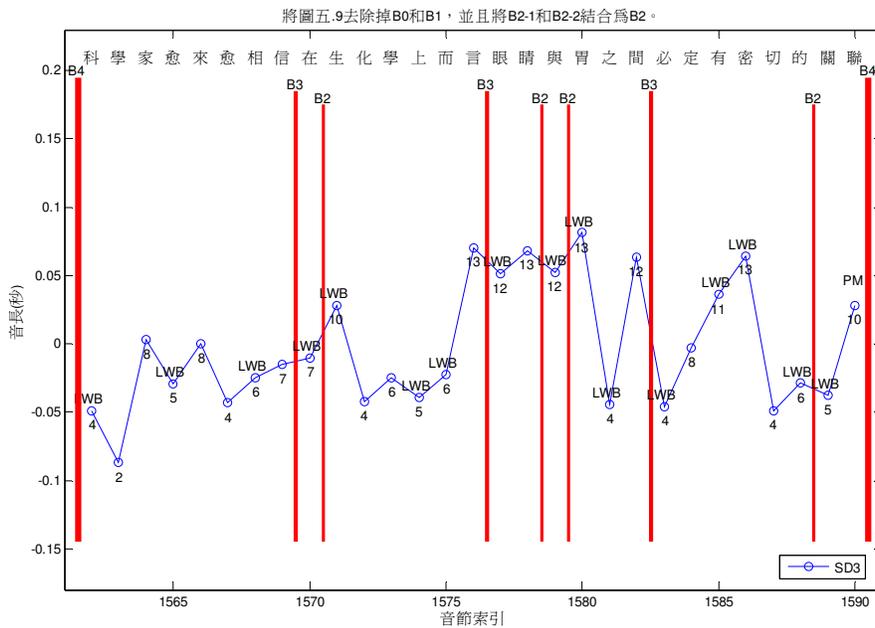
對於漢語語句的韻律結構方面，漢語相對於其它語言為一字、一音、一調的特色，其韻律架構是以音節(Syllable)為最基本的單位，由數個音節組成韻律詞(Prosodic Word)，再由數個韻律詞組成韻律片語(Prosodic Phrase)，最後由數個韻律片語組成 Discourse。

6.2.3 高階語言層次階層音長軌跡現象

我們由參考文獻[2、11、14]中所述與圖五.7、圖五.8和圖五.9的觀察可以知道，音長的漢語語音階層韻律結構是以韻律詞為最基本單位，因此我們將圖五.5與圖五.7、圖五.8和圖五.9去除掉B0和B1並且將B2-1和B2-2結合為B2，以分別呈現出音長的漢語語音階層韻律結構與對於高階語言層次的音長軌跡現象做進一步的觀察，如下的圖六.1與圖六.2、圖六.3和圖六.4所示。

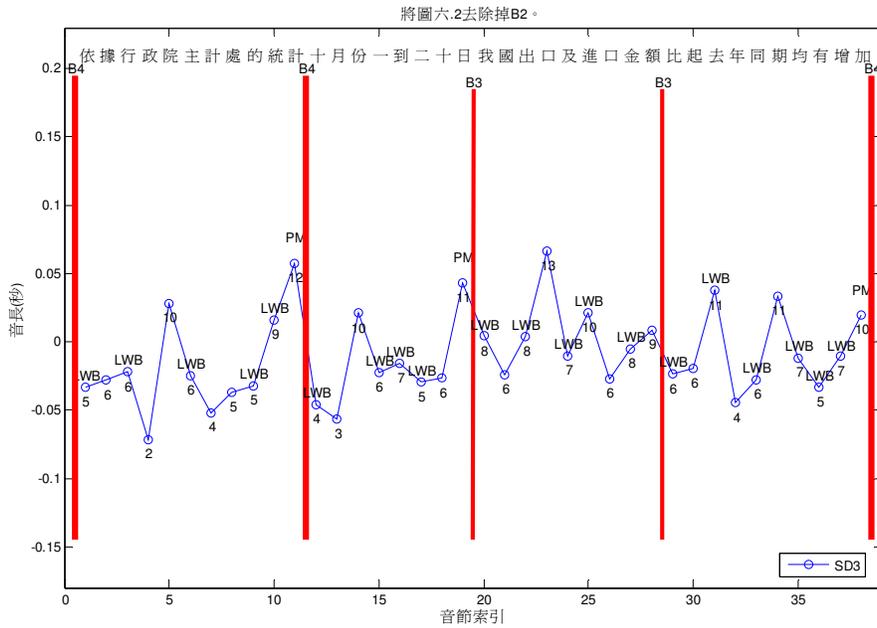


圖六.1：音長的漢語語音階層韻律結構。

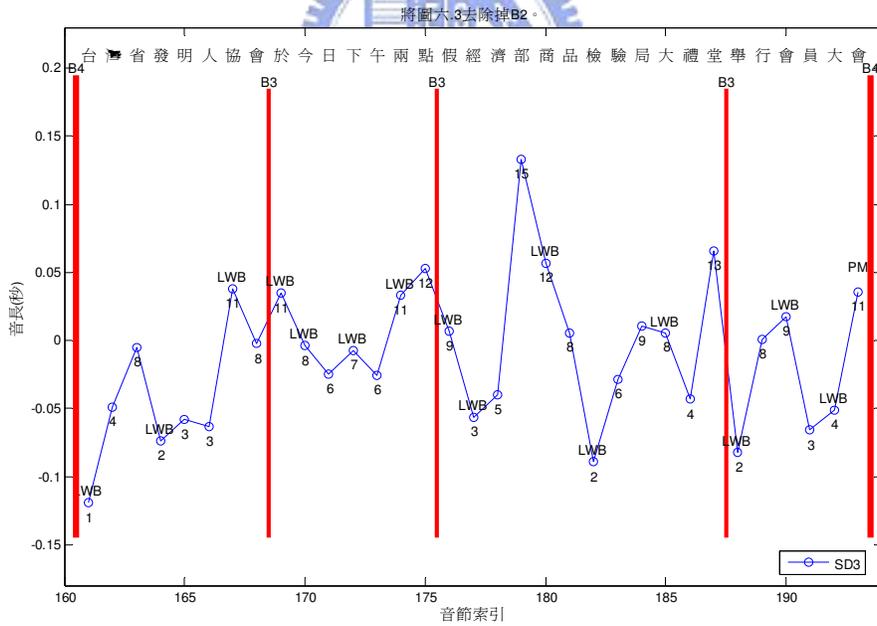


圖六.4：將圖五.9 去除掉 B0 和 B1，並且將 B2-1 和 B2-2 結合為 B2。

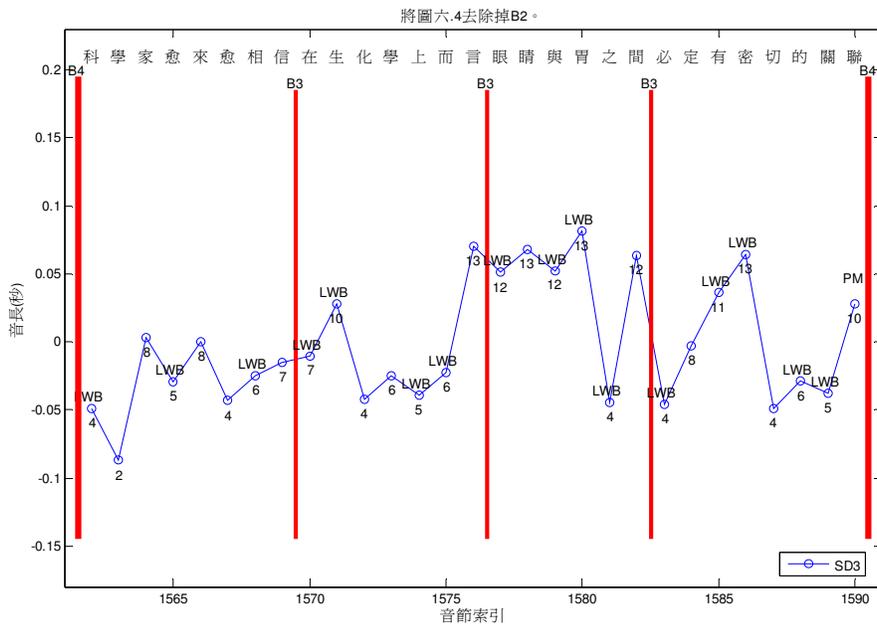
我們由圖六.2、圖六.3 和圖六.4 對應著圖六.1 可以觀察到其確實是存在著各個韻律階層之韻律單位的音長軌跡 Pattern，但是我們僅能夠明顯地觀察到 PW 的音長軌跡 Pattern，對於 MIPPH 和 MPPH 則較不易觀察出，因此我們將圖六.2、圖六.3 和圖六.4 依序去除掉 B2 與 B3，分別如下的圖六.5、圖六.6 和圖六.7 與圖六.8、圖六.9 和圖六.10 所示。



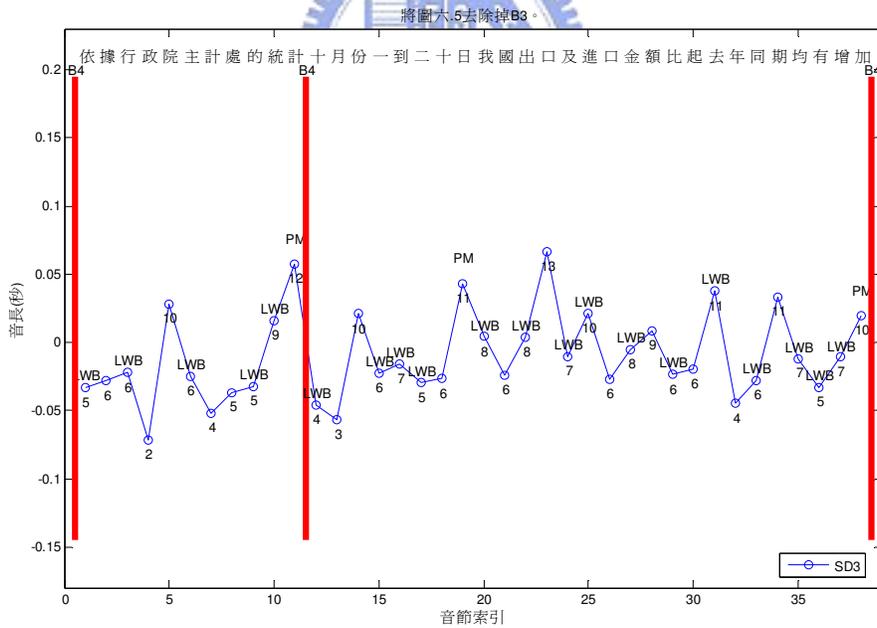
圖六.5：將圖六.2 去除掉 B2。



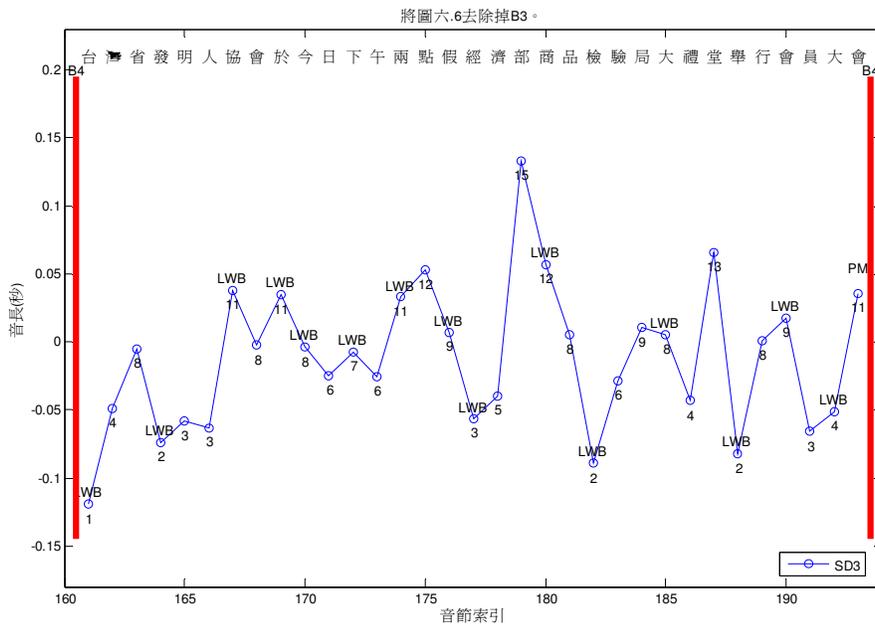
圖六.6：將圖六.3 去除掉 B2。



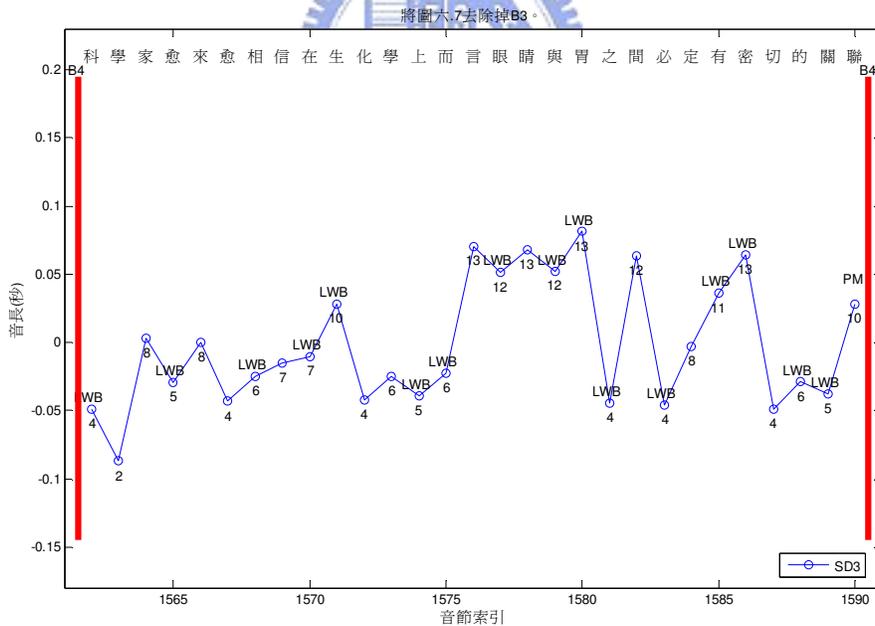
圖六.7：將圖六.4 去除掉 B2。



圖六.8：將圖六.5 去除掉 B3。



圖六.9：將圖六.6 去除掉 B3。

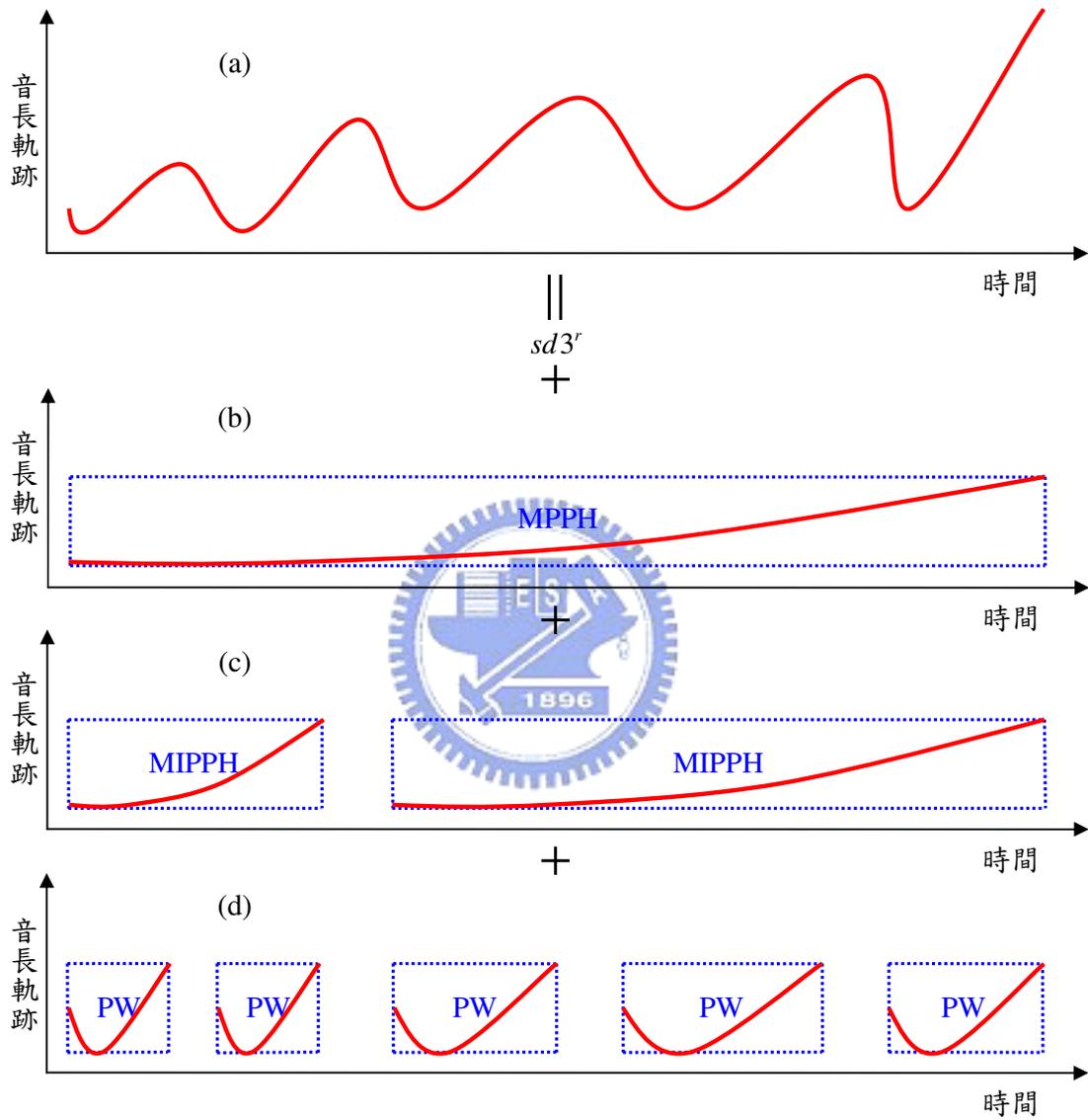


圖六.10：將圖六.7 去除掉 B3。

6.2.4 高階語言層次階層音長軌跡示意圖

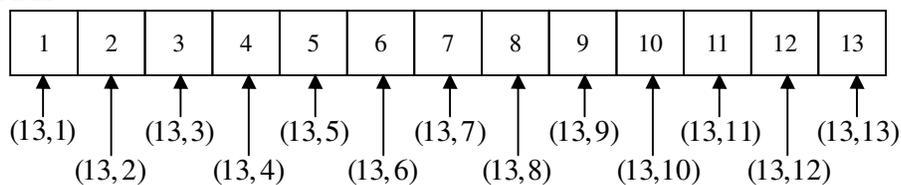
我們由 6.2.3 節可以觀察到各個韻律階層之韻律單位的音長軌跡 Pattern，而在下一節中

我們將以數學模型來呈現出這些 Pattern，因此我們將 MPPH、MIPPH 和 PW 的音長軌跡 Pattern 示意圖繪出如圖六.11 所示，並且以一句 13 個字的 MPPH 為例，繪出其和 MIPPH 與 PW 之間的關係，如圖六.12 所示。

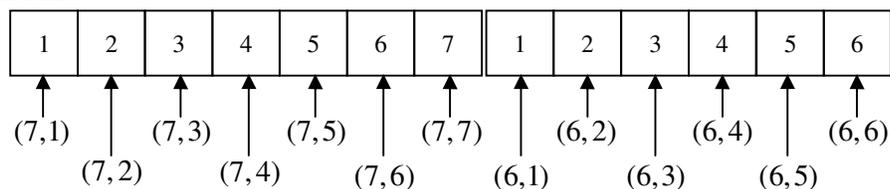


圖六.11：漢語語音階層韻律單位之音長軌跡 Pattern 示意圖。(a).原始音長軌跡，(b).MPPH 之音長軌跡 Pattern，(c). MIPPH 之音長軌跡 Pattern，(d). PW 之音長軌跡 Pattern。

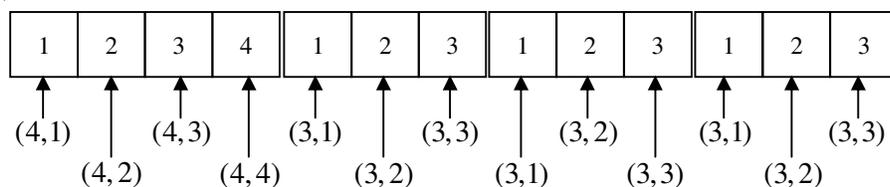
MPPH :



MIPPH :



PW :



圖六.12：一句 13 個字的 MPPH 與 MIPPH 和 PW 之間的關係圖。

6.3 高階音長模型的數學說明

在 4.4 節中我們定義了 SD 、 $SD1$ 、 $SD2$ 、 $SD3$ 和 $SD4$ ，而 $SD3$ 所代表的意義即為扣除掉低階音節層次參數後的高階語言層次之音長特性，我們將其重述如下：

$$sd3_{k,n} = sd_{k,n} - \gamma_{s_{k,n}} - \gamma_{t_{k,n}} - \gamma_{bs_{k,n}} \quad (6.1)$$

由 6.2.3 節我們可以觀察到高階語言層次的各個階層之音長特性，因此我們對 $sd3$ 建立一個高階語言層次階層音長模型，其之數學式如下所述：

$$sd3_{k,n} = sd3_{k,n}^r + \gamma_{PW_{k,n}} + \gamma_{MIPPH_{k,n}} + \gamma_{MPPH_{k,n}} \quad (6.2)$$

我們將各個數學符號條列說明如下：

- $\gamma_{PW_{k,n}}$ 代表語料庫中第 k 個 Utterance 裡第 n 個音節的 PW 類型影響因素， $PW \in \{(1,1), (2,1), (2,2), \dots, (MSC_{PW}, 1), (MSC_{PW}, 2), \dots, (MSC_{PW}, MSC_{PW})\}$ ， MSC_{PW} 代表語料庫中 PW 的最大音節個數 (Maximum Syllable Count, MSC)，其中我們撰寫了程式以找出每一個音節所屬的 PW 之長度和位置，並且得知 $MSC_{PW} = 15$ ；
- $\gamma_{MIPPH_{k,n}}$ 代表 MIPPH 類型影響因素， $MIPPH \in \{(1,1), (2,1), (2,2), \dots, (MSC_{MIPPH}, 1),$

$(MSC_{MIPPH}, 2), \dots, (MSC_{MIPPH}, MSC_{MIPPH})$ ， MSC_{MIPPH} 代表 MIPPH 的 MSC，其中我們撰寫了程式以找出每一個音節所屬的 MIPPH 之長度和位置，並且得知 $MSC_{MIPPH} = 28$ ；

- $\gamma_{MPPH_{k,n}}$ 代表 MPPH 類型影響因素， $MPPH \in \{(1,1), (2,1), (2,2), \dots, (MSC_{MPPH}, 1), (MSC_{MPPH}, 2), \dots, (MSC_{MPPH}, MSC_{MPPH})\}$ ， MSC_{MPPH} 代表 MPPH 的 MSC，其中我們撰寫了程式以找出每一個音節所屬的 MPPH 之長度和位置，並且得知 $MSC_{MPPH} = 63$ ；
- $sd3_{k,n}^r$ 代表 $sd3_{k,n}$ 扣除上述三個影響因素後的殘餘值。

6.3.1 目標函數的數學說明

我們定義了一個平方誤差(Squared Error)的目標函數 Q'' ，以對模型之收斂狀況做判斷，如下的數學式所述：

$$\begin{aligned}
 Q'' &= \sum_{k=1}^K \sum_{n=1}^{N_k} (sd3_{k,n}^r)^2 \\
 &= \sum_{k=1}^K \sum_{n=1}^{N_k} (sd3_{k,n} - \gamma_{PW_{k,n}} - \gamma_{MIPPH_{k,n}} - \gamma_{MPPH_{k,n}})^2
 \end{aligned} \tag{6.3}$$

因此我們使用了最小平方誤差法則(Minimum Squared Error Criterion, MSEC)，來推導出往後的數學式和建立模型。

6.3.2 PW 類型影響因素的數學說明

我們將 Q'' 對 γ_{PW} 做偏微分並令其等於零，以求得 Q'' 對 γ_{PW} 的極小值之所在的 γ_{PW} 值，數學式的推導如下所述：

$$\begin{aligned}
 \frac{\partial Q''}{\partial \gamma_{PW}} &= 0 \\
 \Rightarrow \sum_{k=1}^K \sum_{n=1}^{N_k} (2)(sd3_{k,n} - \gamma_{PW_{k,n}} - \gamma_{MIPPH_{k,n}} - \gamma_{MPPH_{k,n}})(-1)\delta(PW_{k,n} = PW) &= 0 \\
 \Rightarrow \sum_{k=1}^K \sum_{n=1}^{N_k} (sd3_{k,n} - \gamma_{PW_{k,n}} - \gamma_{MIPPH_{k,n}} - \gamma_{MPPH_{k,n}})\delta(PW_{k,n} = PW) &= 0 \\
 \Rightarrow \sum_{k=1}^K \sum_{n=1}^{N_k} \gamma_{PW_{k,n}} \delta(PW_{k,n} = PW) &= \sum_{k=1}^K \sum_{n=1}^{N_k} (sd3_{k,n} - \gamma_{MIPPH_{k,n}} - \gamma_{MPPH_{k,n}})
 \end{aligned}$$

$$\begin{aligned}
& \delta(PW_{k,n} = PW) \\
\Rightarrow \gamma_{PW} \sum_{k=1}^K \sum_{n=1}^{N_k} \delta(PW_{k,n} = PW) &= \sum_{k=1}^K \sum_{n=1}^{N_k} (sd3_{k,n} - \gamma_{MIPPH_{k,n}} - \gamma_{MPPH_{k,n}}) \\
& \delta(PW_{k,n} = PW) \\
\Rightarrow \gamma_{PW} &= \frac{\sum_{k=1}^K \sum_{n=1}^{N_k} (sd3_{k,n} - \gamma_{MIPPH_{k,n}} - \gamma_{MPPH_{k,n}}) \delta(PW_{k,n} = PW)}{\sum_{k=1}^K \sum_{n=1}^{N_k} \delta(PW_{k,n} = PW)}, PW \in \\
& \{(1,1), (2,1), (2,2), \dots, (MSC_{PW}, 1), (MSC_{PW}, 2), \dots, (MSC_{PW}, MSC_{PW})\} \quad (6.4)
\end{aligned}$$

其中

$$\delta(PW_{k,n} = PW) = \begin{cases} 1, & \text{if } PW_{k,n} = PW \\ 0, & \text{if } PW_{k,n} \neq PW \end{cases} \quad (6.5)$$

上式所代表的意義為將語料庫中所有的 $sd3$ 減去其所屬的 γ_{MIPPH} 和 γ_{MPPH} 後也就是 $sd3_{k,n} - \gamma_{MIPPH_{k,n}} - \gamma_{MPPH_{k,n}}$ ，再對每一種 PW 類型分別相加起來後除以其之個數。

6.3.3 MIPPH 類型影響因素的數學說明

我們將 Q 對 γ_{MIPPH} 做偏微分並令其等於零，以求得 Q 對 γ_{MIPPH} 的極小值之所在的 γ_{MIPPH} 值，數學式的推導如下所述：

$$\begin{aligned}
\frac{\partial Q}{\partial \gamma_{MIPPH}} &= 0 \\
\Rightarrow \sum_{k=1}^K \sum_{n=1}^{N_k} (2)(sd3_{k,n} - \gamma_{PW_{k,n}} - \gamma_{MIPPH_{k,n}} - \gamma_{MPPH_{k,n}})(-1) \delta(MIPPH_{k,n} = MIPPH) &= 0 \\
\Rightarrow \sum_{k=1}^K \sum_{n=1}^{N_k} (sd3_{k,n} - \gamma_{PW_{k,n}} - \gamma_{MIPPH_{k,n}} - \gamma_{MPPH_{k,n}}) \delta(MIPPH_{k,n} = MIPPH) &= 0 \\
\Rightarrow \sum_{k=1}^K \sum_{n=1}^{N_k} \gamma_{MIPPH_{k,n}} \delta(MIPPH_{k,n} = MIPPH) &= \sum_{k=1}^K \sum_{n=1}^{N_k} (sd3_{k,n} - \gamma_{PW_{k,n}} - \gamma_{MPPH_{k,n}}) \\
& \delta(MIPPH_{k,n} = MIPPH) \\
\Rightarrow \gamma_{MIPPH} \sum_{k=1}^K \sum_{n=1}^{N_k} \delta(MIPPH_{k,n} = MIPPH) &= \sum_{k=1}^K \sum_{n=1}^{N_k} (sd3_{k,n} - \gamma_{PW_{k,n}} - \gamma_{MPPH_{k,n}}) \\
& \delta(MIPPH_{k,n} = MIPPH)
\end{aligned}$$

$$\Rightarrow \gamma_{MIPPH} = \frac{\sum_{k=1}^K \sum_{n=1}^{N_k} (sd3_{k,n} - \gamma_{PW_{k,n}} - \gamma_{MPPH_{k,n}}) \delta(MIPPH_{k,n} = MIPPH)}{\sum_{k=1}^K \sum_{n=1}^{N_k} \delta(MIPPH_{k,n} = MIPPH)}, MIPPH \in \{(1,1), (2,1), (2,2), \dots, (MSC_{MIPPH}, 1), (MSC_{MIPPH}, 2), \dots, (MSC_{MIPPH}, MSC_{MIPPH})\} \quad (6.6)$$

其中

$$\delta(MIPPH_{k,n} = MIPPH) = \begin{cases} 1, & \text{if } MIPPH_{k,n} = MIPPH \\ 0, & \text{if } MIPPH_{k,n} \neq MIPPH \end{cases} \quad (6.7)$$

上式所代表的意義為將語料庫中所有的 $sd3$ 減去其所屬的 γ_{PW} 和 γ_{MPPH} 後也就是 $sd3_{k,n} - \gamma_{PW_{k,n}} - \gamma_{MPPH_{k,n}}$ ，再對每一種 MIPPH 類型分別相加起來後除以其之個數。

6.3.4 MPPH 類型影響因素的數學說明

我們將 Q'' 對 γ_{MPPH} 做偏微分並令其等於零，以求得 Q'' 對 γ_{MPPH} 的極小值之所在的 γ_{MPPH} 值，數學式的推導如下所述：

$$\begin{aligned} \frac{\partial Q''}{\partial \gamma_{MPPH}} &= 0 \\ \Rightarrow \sum_{k=1}^K \sum_{n=1}^{N_k} (2)(sd3_{k,n} - \gamma_{PW_{k,n}} - \gamma_{MPPH_{k,n}} - \gamma_{MPPH_{k,n}})(-1) \delta(MPPH_{k,n} = MPPH) &= 0 \\ \Rightarrow \sum_{k=1}^K \sum_{n=1}^{N_k} (sd3_{k,n} - \gamma_{PW_{k,n}} - \gamma_{MPPH_{k,n}} - \gamma_{MPPH_{k,n}}) \delta(MPPH_{k,n} = MPPH) &= 0 \\ \Rightarrow \sum_{k=1}^K \sum_{n=1}^{N_k} \gamma_{MPPH_{k,n}} \delta(MPPH_{k,n} = MPPH) &= \sum_{k=1}^K \sum_{n=1}^{N_k} (sd3_{k,n} - \gamma_{PW_{k,n}} - \gamma_{MPPH_{k,n}}) \delta(MPPH_{k,n} = MPPH) \\ \Rightarrow \gamma_{MPPH} \sum_{k=1}^K \sum_{n=1}^{N_k} \delta(MPPH_{k,n} = MPPH) &= \sum_{k=1}^K \sum_{n=1}^{N_k} (sd3_{k,n} - \gamma_{PW_{k,n}} - \gamma_{MPPH_{k,n}}) \delta(MPPH_{k,n} = MPPH) \\ \Rightarrow \gamma_{MPPH} &= \frac{\sum_{k=1}^K \sum_{n=1}^{N_k} (sd3_{k,n} - \gamma_{PW_{k,n}} - \gamma_{MPPH_{k,n}}) \delta(MPPH_{k,n} = MPPH)}{\sum_{k=1}^K \sum_{n=1}^{N_k} \delta(MPPH_{k,n} = MPPH)}, MPPH \in \{(1,1), (2,1), (2,2), \dots, (MSC_{MPPH}, 1), (MSC_{MPPH}, 2), \dots, (MSC_{MPPH}, MSC_{MPPH})\} \quad (6.8) \end{aligned}$$

其中

$$\delta(MPPH_{k,n} = MPPH) = \begin{cases} 1, & \text{if } MPPH_{k,n} = MPPH \\ 0, & \text{if } MPPH_{k,n} \neq MPPH \end{cases} \quad (6.9)$$

上式所代表的意義為將語料庫中所有的 $sd3$ 減去其所屬的 γ_{PW} 和 γ_{MIPPH} 後也就是 $sd3_{k,n} - \gamma_{PW_{k,n}} - \gamma_{MIPPH_{k,n}}$ ，再對每一種 MPPH 類型分別相加起來後除以其之個數。

6.4 高階音長模型之建立的流程說明

在完成了上述的數學說明之後接著要建立模型，我們將在以下的兩個小節分別對於模型參數的初始值設定和模型的訓練方式做說明。

6.4.1 模型參數的初始值設定

在訓練模型之前必須要先設定好各個模型參數的初始值，我們將在以下的數個小節中對於各個影響因素的初始值設定做說明。

6.4.1.1 PW 類型影響因素的初始值設定

γ_{PW} 的初始值設定如下的數學式所述：

$$\gamma_{PW} = \frac{\sum_{k=1}^K \sum_{n=1}^{N_k} (sd3_{k,n}) \delta(PW_{k,n} = PW)}{\sum_{k=1}^K \sum_{n=1}^{N_k} \delta(PW_{k,n} = PW)}, PW \in \{(1,1), (2,1), (2,2), \dots, (MSC_{PW}, 1), (MSC_{PW}, 2), \dots, (MSC_{PW}, MSC_{PW})\} \quad (6.10)$$

上式所代表的意義為將語料庫中所有的 $sd3$ 對每一種 PW 類型分別相加起來後除以其之個數。

6.4.1.2 MIPPH 類型影響因素的初始值設定

γ_{MIPPH} 的初始值設定如下的數學式所述：

$$\gamma_{MIPPH} = \frac{\sum_{k=1}^K \sum_{n=1}^{N_k} (sd3_{k,n} - \gamma_{PW_{k,n}}) \delta(MIPPH_{k,n} = MIPPH)}{\sum_{k=1}^K \sum_{n=1}^{N_k} \delta(MIPPH_{k,n} = MIPPH)}, MIPPH \in$$

$$\{(1,1), (2,1), (2,2), \dots, (\text{MSC}_{\text{MPPH}}, 1), (\text{MSC}_{\text{MPPH}}, 2), \dots, (\text{MSC}_{\text{MPPH}}, \text{MSC}_{\text{MPPH}})\} \quad (6.11)$$

上式所代表的意義為將語料庫中所有的 $sd3$ 減去其所屬的 γ_{PW} 後也就是 $sd3_{k,n} - \gamma_{PW_{k,n}}$ ，再對每一種 MPPH 類型分別相加起來後除以其之個數。

6.4.1.3 MPPH 類型影響因素的初始值設定

γ_{MPPH} 的初始值設定如下的數學式所述：

$$\gamma_{\text{MPPH}} = \frac{\sum_{k=1}^K \sum_{n=1}^{N_k} (sd3_{k,n} - \gamma_{PW_{k,n}} - \gamma_{\text{MPPH}_{k,n}}) \delta(\text{MPPH}_{k,n} = \text{MPPH})}{\sum_{k=1}^K \sum_{n=1}^{N_k} \delta(\text{MPPH}_{k,n} = \text{MPPH})}, \text{MPPH} \in \{(1,1), (2,1), (2,2), \dots, (\text{MSC}_{\text{MPPH}}, 1), (\text{MSC}_{\text{MPPH}}, 2), \dots, (\text{MSC}_{\text{MPPH}}, \text{MSC}_{\text{MPPH}})\} \quad (6.12)$$

上式所代表的意義為將語料庫中所有的 $sd3$ 減去其所屬的 γ_{PW} 和 γ_{MPPH} 後，也就是 $sd3_{k,n} - \gamma_{PW_{k,n}} - \gamma_{\text{MPPH}_{k,n}}$ ，再對每一種 MPPH 類型分別相加起來後除以其之個數。

6.4.1.4 目標函數的初始值計算

我們在設定好各個模型參數的初始值之後便將其代入(6.3)式，以計算 Q 的初始值，我們將 Q 的計算數學式重述如下：

$$Q = \sum_{k=1}^K \sum_{n=1}^{N_k} (sd3_{k,n} - \gamma_{PW_{k,n}} - \gamma_{\text{MPPH}_{k,n}} - \gamma_{\text{MPPH}_{k,n}})^2 \quad (6.13)$$

6.4.2 模型的訓練方式

在完成了上述模型參數的初始值設定之後接著便是要訓練模型，我們所採用的方式是以反覆循序最佳化的方式來不斷地訓練模型和更新模型參數，直到 Q 符合我們所訂下的收斂條件。我們將在以下的數個小節中呈現出 γ_{PW} 、 γ_{MPPH} 、 γ_{MPPH} 和 Q 的更新數學式，以及模型訓練的收斂條件。

6.4.2.1 PW 類型影響因素的更新數學式

γ_{PW} 的更新數學式為在之前所推導出來的(6.4)式，我們將其重述如下：

$$\gamma_{PW} = \frac{\sum_{k=1}^K \sum_{n=1}^{N_k} (sd3_{k,n} - \gamma_{MIPPH_{k,n}} - \gamma_{MPPH_{k,n}}) \delta(PW_{k,n} = PW)}{\sum_{k=1}^K \sum_{n=1}^{N_k} \delta(PW_{k,n} = PW)}, PW \in \{(1,1), (2,1), (2,2), \dots, (MSC_{PW}, 1), (MSC_{PW}, 2), \dots, (MSC_{PW}, MSC_{PW})\} \quad (6.14)$$

6.4.2.2 MIPPH 類型影響因素的更新數學式

γ_{MIPPH} 的更新數學式為在之前所推導出來的(6.6)式，我們將其重述如下所述：

$$\gamma_{MIPPH} = \frac{\sum_{k=1}^K \sum_{n=1}^{N_k} (sd3_{k,n} - \gamma_{PW_{k,n}} - \gamma_{MPPH_{k,n}}) \delta(MIPPH_{k,n} = MIPPH)}{\sum_{k=1}^K \sum_{n=1}^{N_k} \delta(MIPPH_{k,n} = MIPPH)}, MIPPH \in \{(1,1), (2,1), (2,2), \dots, (MSC_{MIPPH}, 1), (MSC_{MIPPH}, 2), \dots, (MSC_{MIPPH}, MSC_{MIPPH})\} \quad (6.15)$$

6.4.2.3 MPPH 類型影響因素的更新數學式

γ_{MPPH} 的更新數學式為在之前所推導出來的(6.8)式，我們將其重述如下所述：

$$\gamma_{MPPH} = \frac{\sum_{k=1}^K \sum_{n=1}^{N_k} (sd3_{k,n} - \gamma_{PW_{k,n}} - \gamma_{MIPPH_{k,n}}) \delta(MPPH_{k,n} = MPPH)}{\sum_{k=1}^K \sum_{n=1}^{N_k} \delta(MPPH_{k,n} = MPPH)}, MPPH \in \{(1,1), (2,1), (2,2), \dots, (MSC_{MPPH}, 1), (MSC_{MPPH}, 2), \dots, (MSC_{MPPH}, MSC_{MPPH})\} \quad (6.16)$$

6.4.2.4 目標函數的更新數學式

我們在更新好各個模型參數之後便將其代入上述的(6.3)式以更新 Q ，我們將其之計算數學式重述如下：

$$Q = \sum_{k=1}^K \sum_{n=1}^{N_k} (sd3_{k,n} - \gamma_{PW_{k,n}} - \gamma_{MIPPH_{k,n}} - \gamma_{MPPH_{k,n}})^2 \quad (6.17)$$

6.4.2.5 模型訓練的收斂條件

我們所訂下的模型訓練之收斂條件為“只要第 i 次更新之 Q 值和第 $i-1$ 次更新之 Q 值

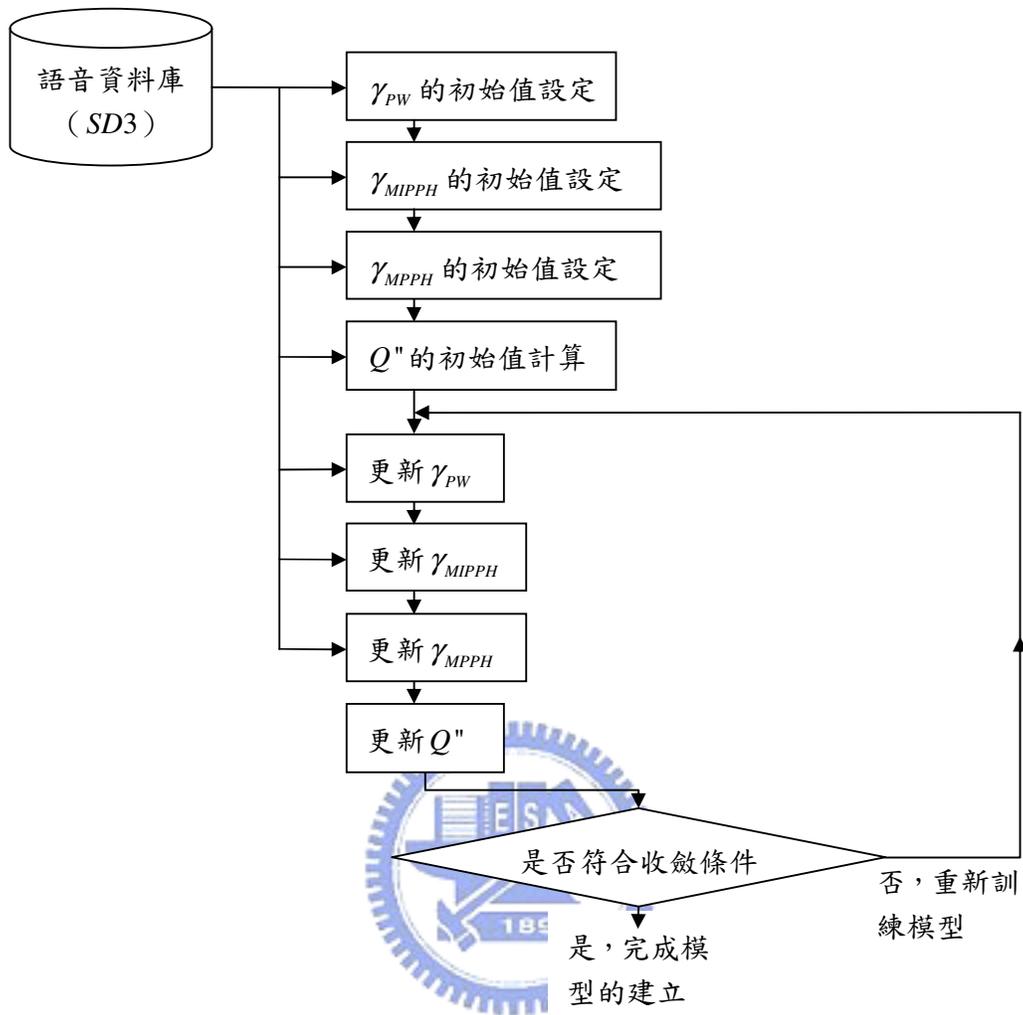
的差之負值小於第*i*次更新之 Q'' 值乘上一個極小的數，我們令這個極小的數為 10^{-7} ，如此便符合收斂條件。”，其數學表示如下式所示：

$$(-1) \times (Q''(i) - Q''(i-1)) < Q''(i) \times 10^{-7} \quad (6.18)$$

6.4.2.6 模型訓練的流程

我們所採用的模型訓練流程為先更新 γ_{PW} 再更新 γ_{MPPH} 最後更新 γ_{MPPH} ，而後再以上述的(6.17)式更新 Q'' 並且測試其是否符合上述的(6.18)式之收斂條件，如果符合便停止更新模型參數，否則繼續更新模型參數直到符合收斂條件為止，我們將模型訓練的流程方塊圖繪出如圖六.13所示。





圖六.13：高階音長模型的訓練流程方塊圖。

第七章 高階音長模型的實驗結果分析

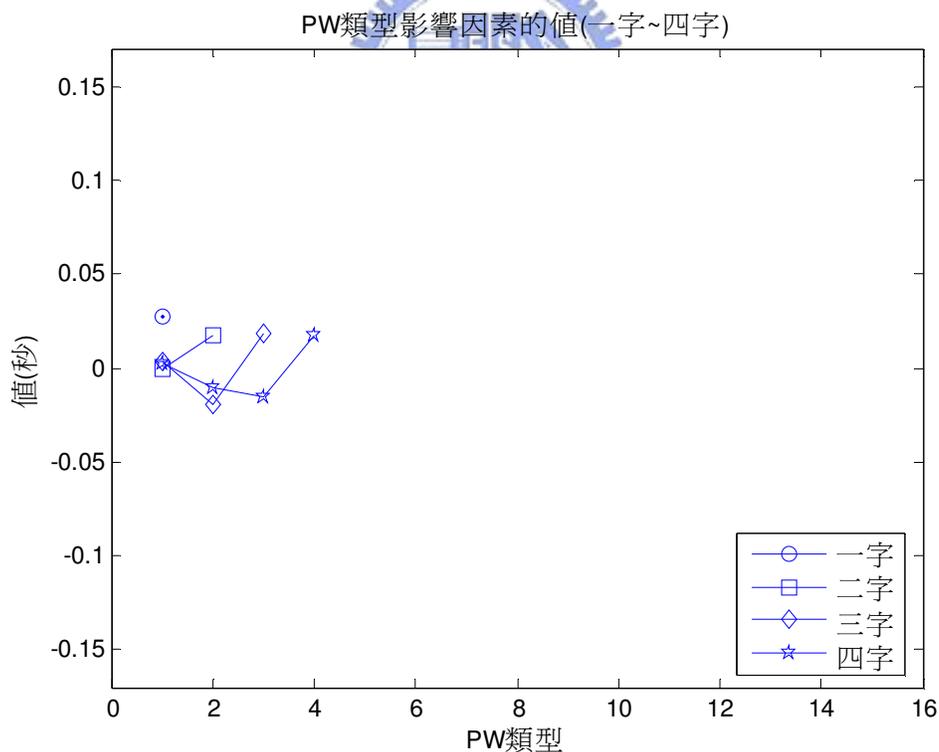
本章對於所建立的高階音長模型在符合收斂條件狀態下分析其參數，以觀察是否符合我們所認知的語音特性，藉此來判斷所建立的模型是否正確，並且呈現出各個漢語語音階層韻律結構的階層之高階音長軌跡 Pattern。

7.1 影響因素

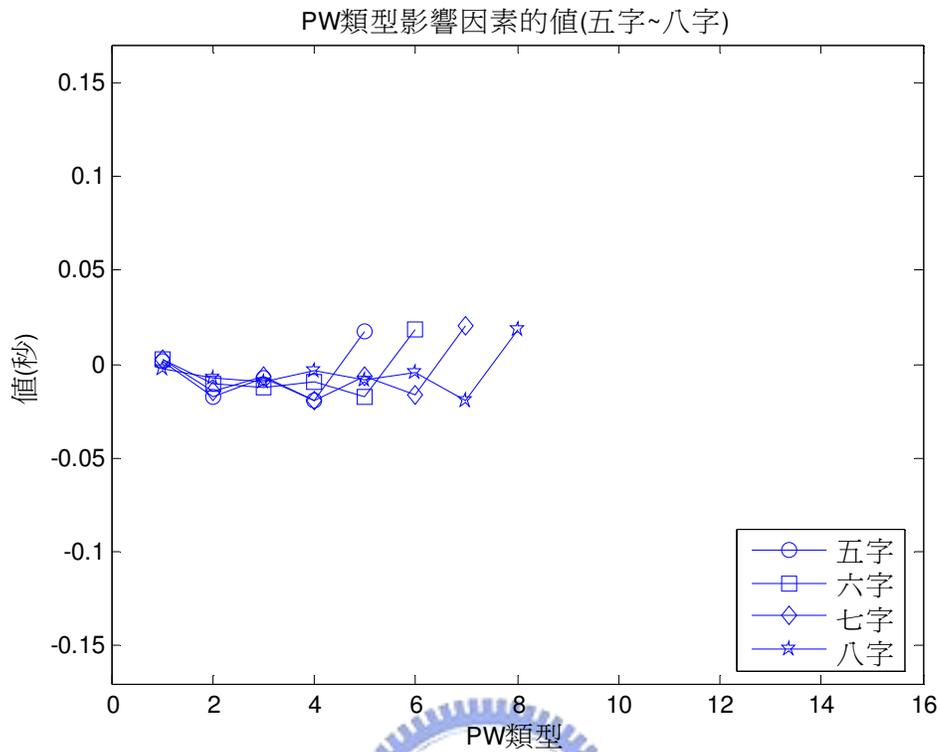
我們將在這一個小節中呈現出各個影響因素的值。

7.1.1 PW 類型影響因素

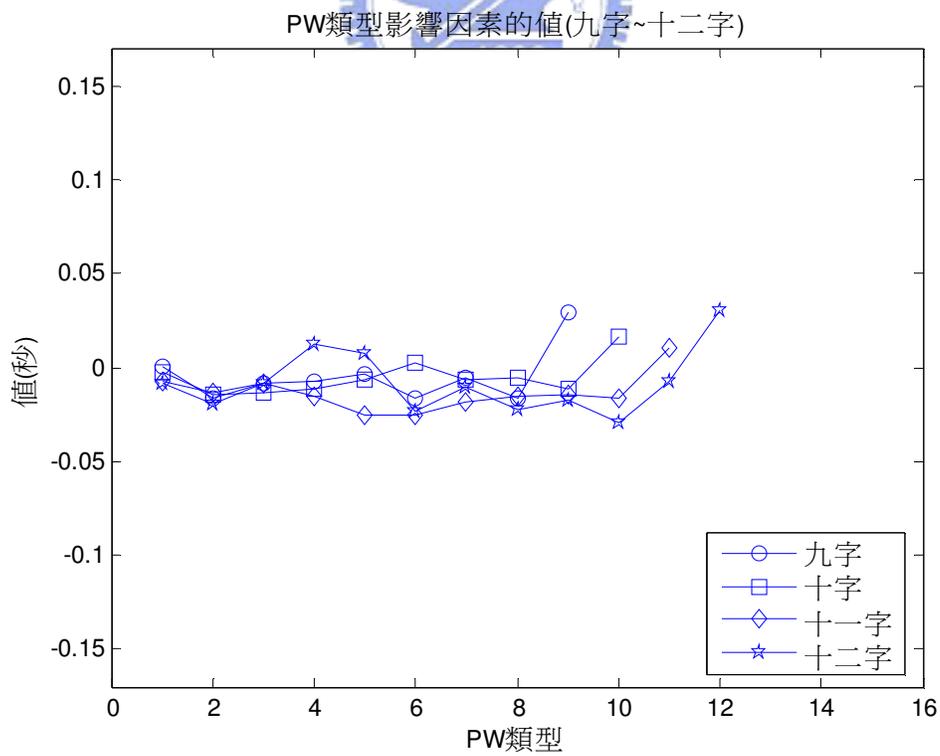
我們將 γ_{PW} 的值繪出如圖七.1 到圖七.4 所示。



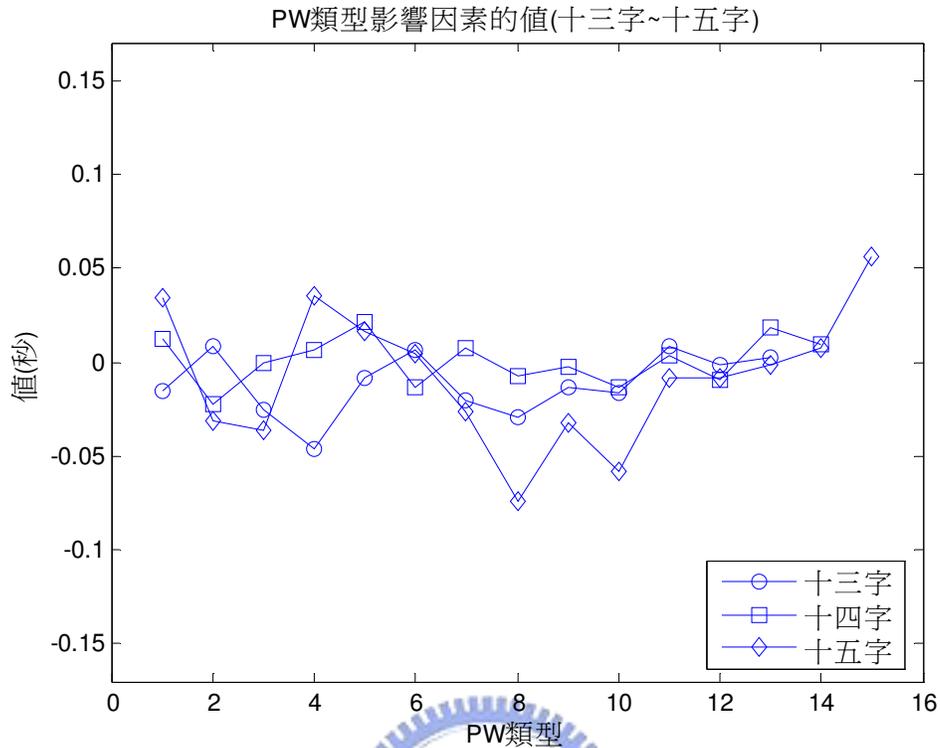
圖七.1： γ_{PW} 的值(一字~四字)。



圖七.2： γ_{PW} 的值(五字~八字)。



圖七.3： γ_{PW} 的值(九字~十二字)。



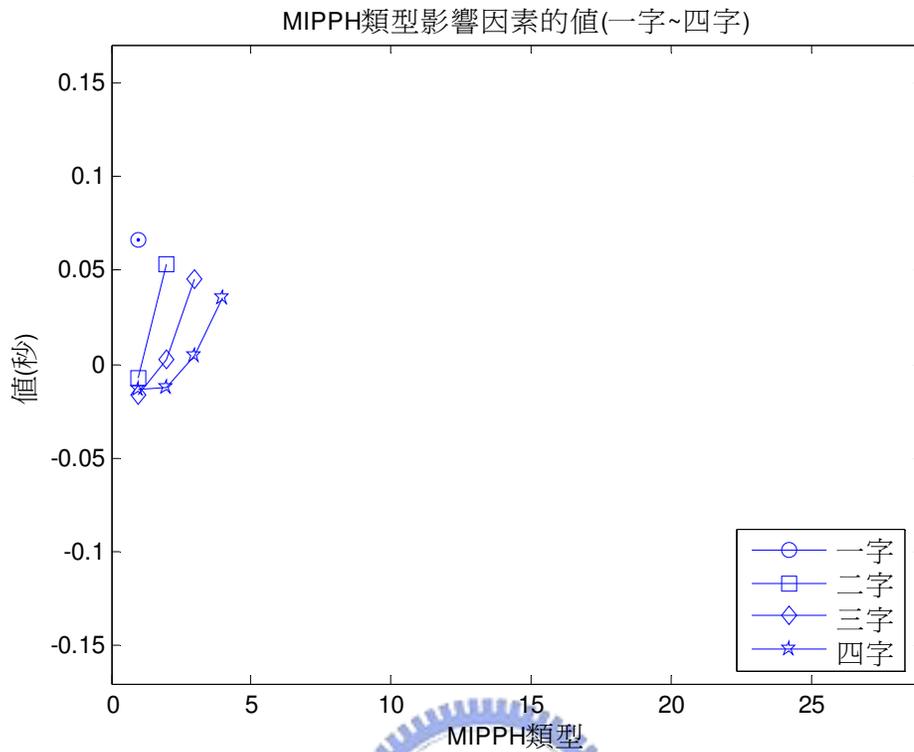
圖七.4： γ_{PW} 的值(十三字~十五字)。

7.1.1.1 實驗結果分析

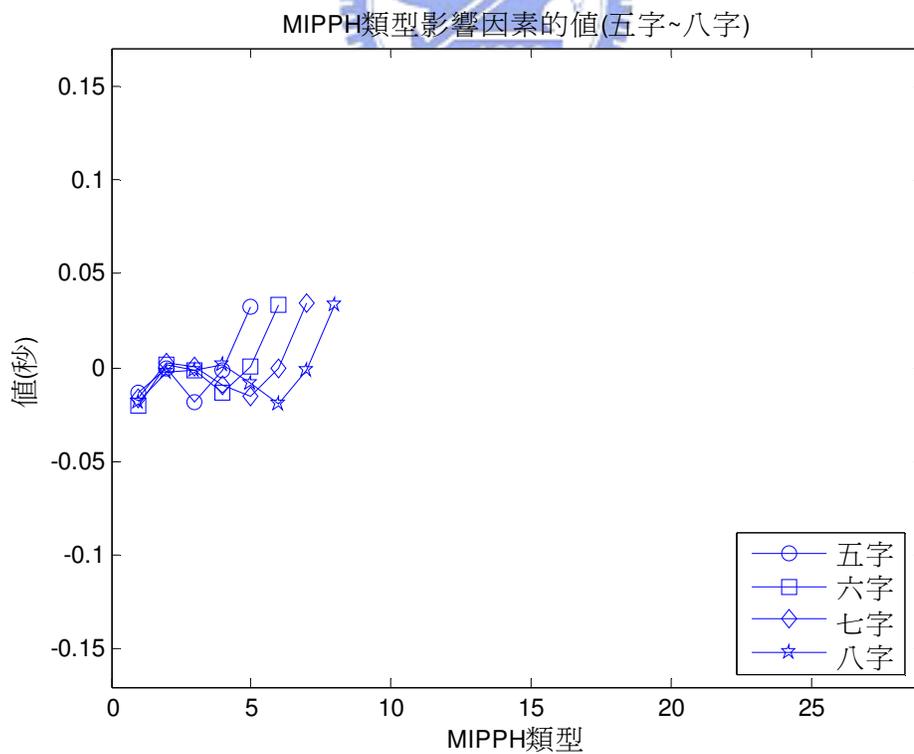
我們由圖七.1 到圖七.4 可以觀察到 PW 的音長軌跡 Pattern，特別注意二字的 Pattern 為“/”，三字和四字的 Pattern 為“v”，其確實如同我們於圖六.11 中所述的形狀，而二字、三字和四字的 Pattern 又可組成多字的 Pattern，因此模型所產生的 γ_{PW} 絕大部份是符合我們所認知的語音特性。

7.1.2 MIPPH 類型影響因素

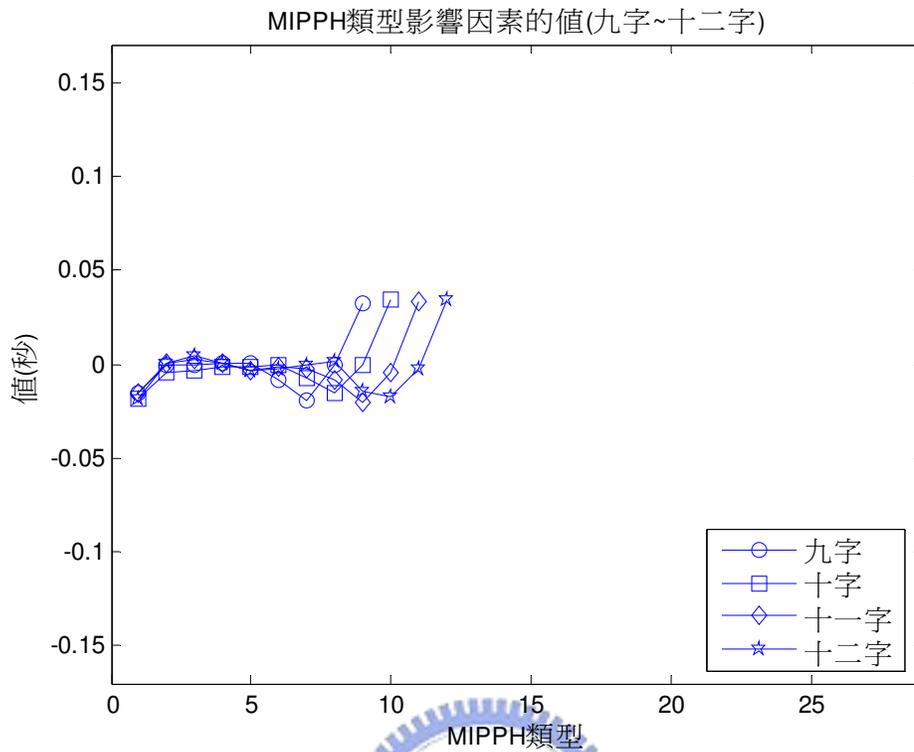
我們將 γ_{MIPPH} 的值繪出如圖七.5 到圖七.11 所示。



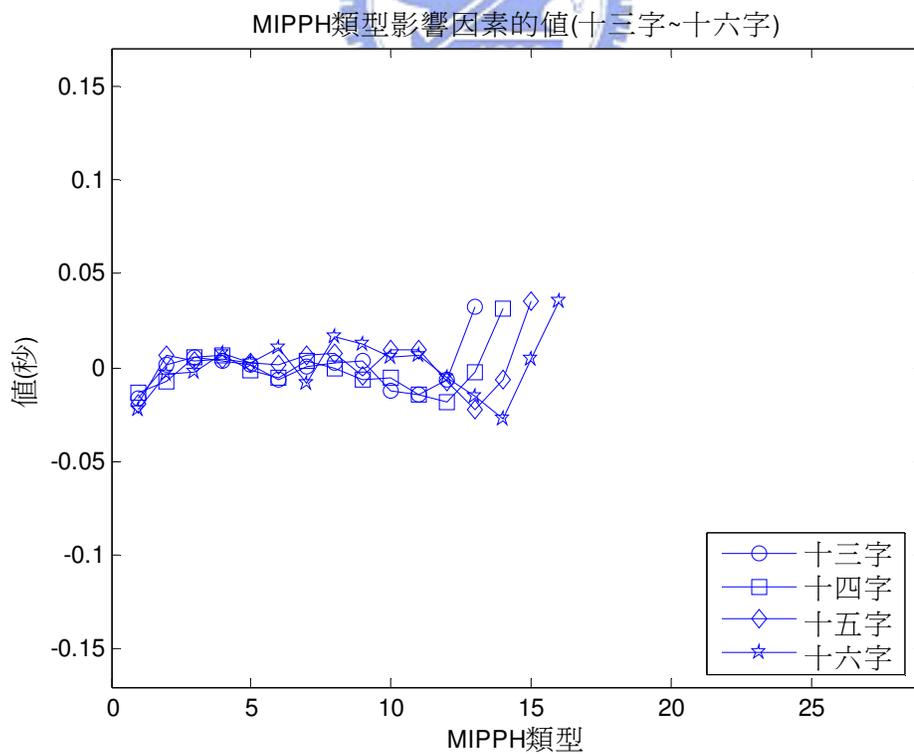
圖七.5： γ_{MIPPH} 的值(一字~四字)。



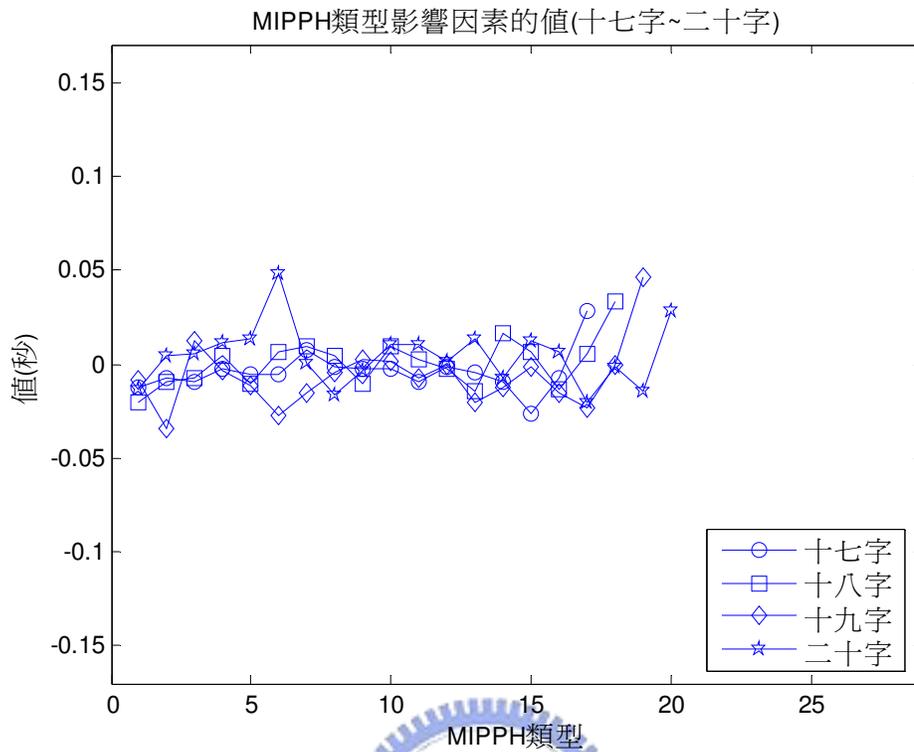
圖七.6： γ_{MIPPH} 的值(五字~八字)。



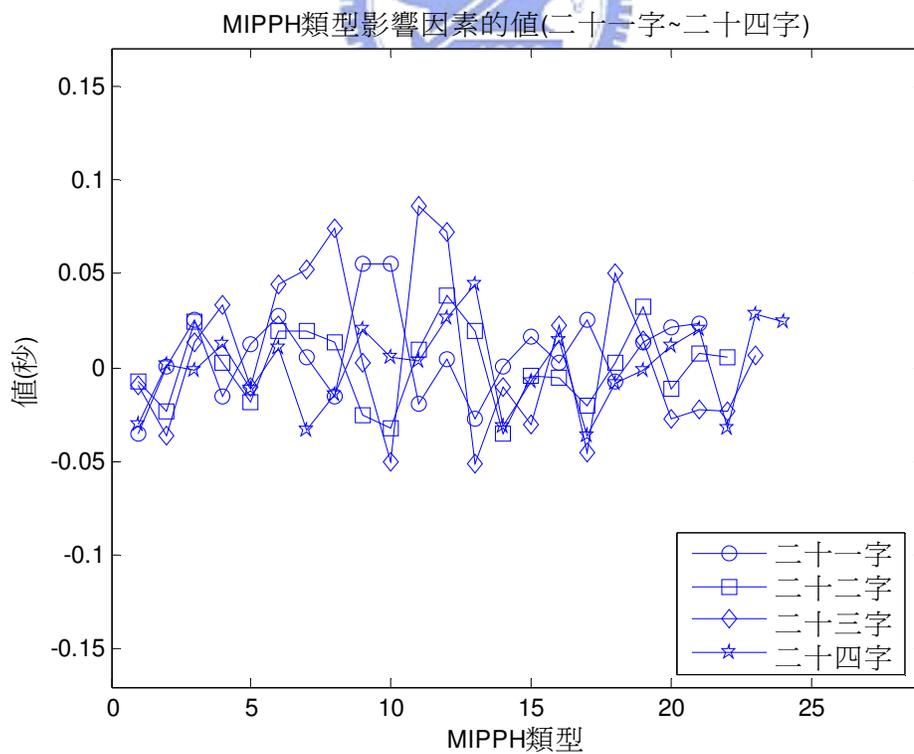
圖七.7： γ_{MIPPH} 的值(九字~十二字)。



圖七.8： γ_{MIPPH} 的值(十三字~十六字)。



圖七.9： γ_{MIPPH} 的值(十七字~二十字)。



圖七.10： γ_{MIPPH} 的值(二十一字~二十四字)。

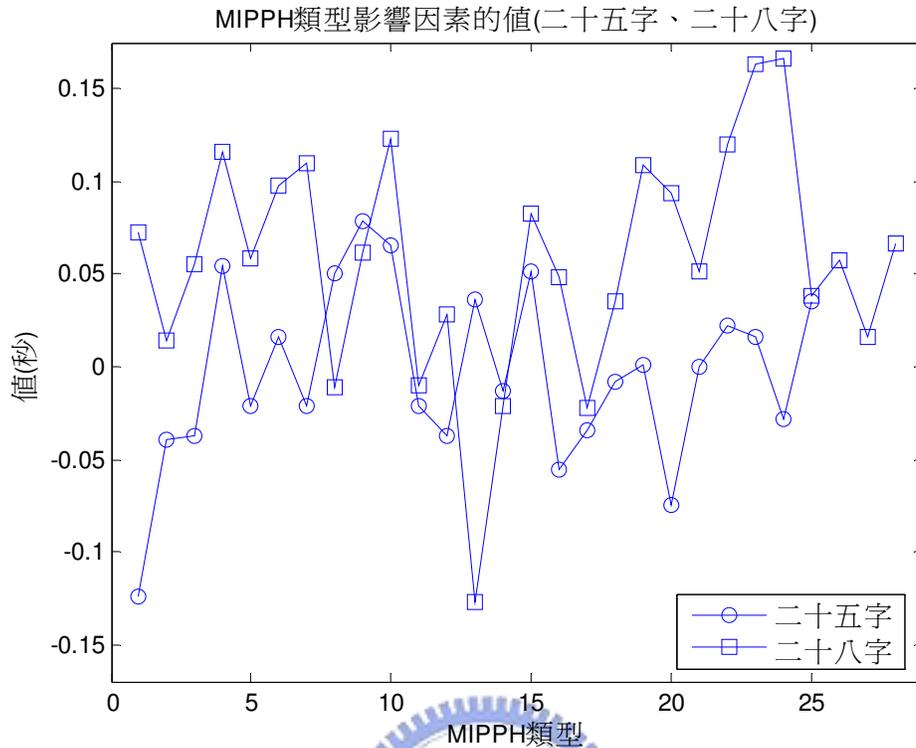


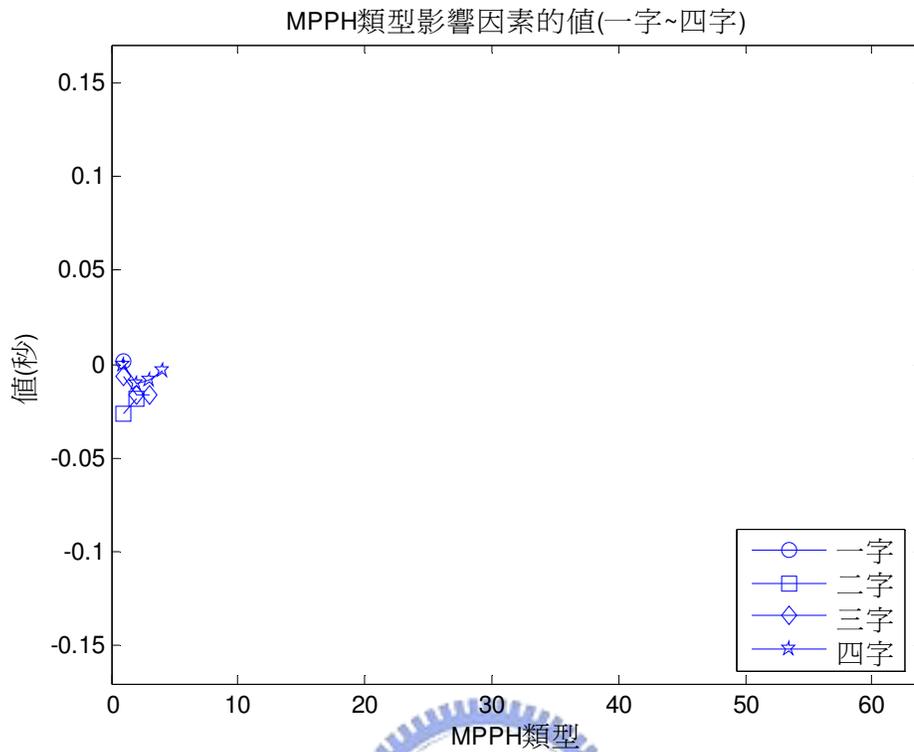
圖 7.11： γ_{MIPPH} 的值(二十五字、二十八字)。

7.1.2.1 實驗結果分析

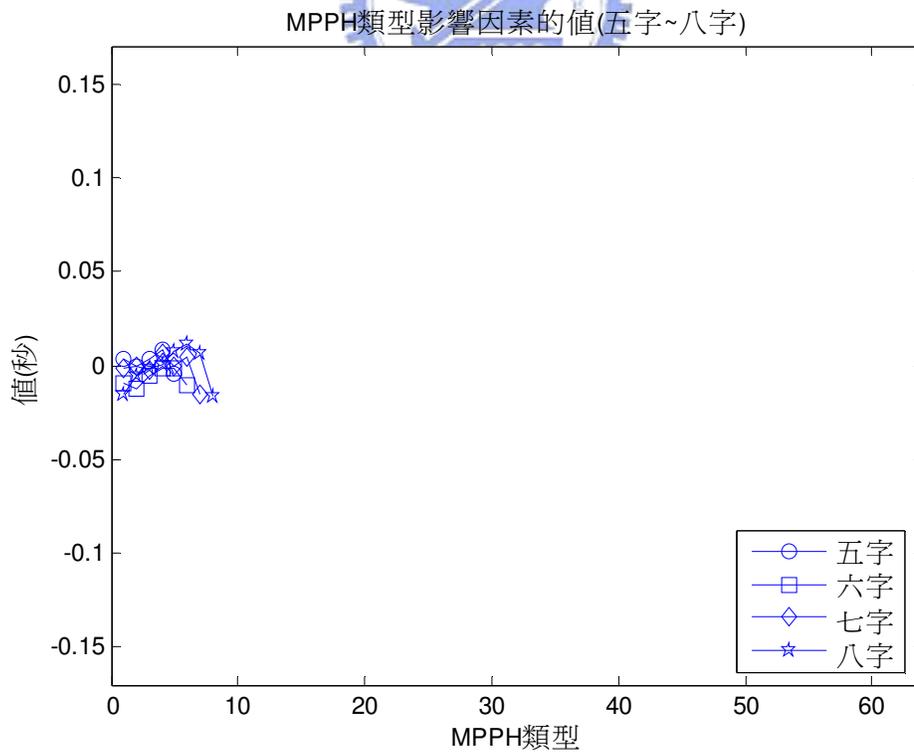
我們由圖 7.5 到圖 7.11 可以觀察到 MIPPH 的音長軌跡 Pattern，首先，我們的語料庫中沒有二十六字和二十七字的 MIPPH 存在；再者，一字到十八字的 Pattern 確實如同我們於圖 6.11 中所述的形狀，而十九字以上的 Pattern 則沒有固定的形狀，因此模型所產生的 γ_{MIPPH} 絕大部份是符合我們所認知的語音特性。

7.1.3 MPPH 類型影響因素

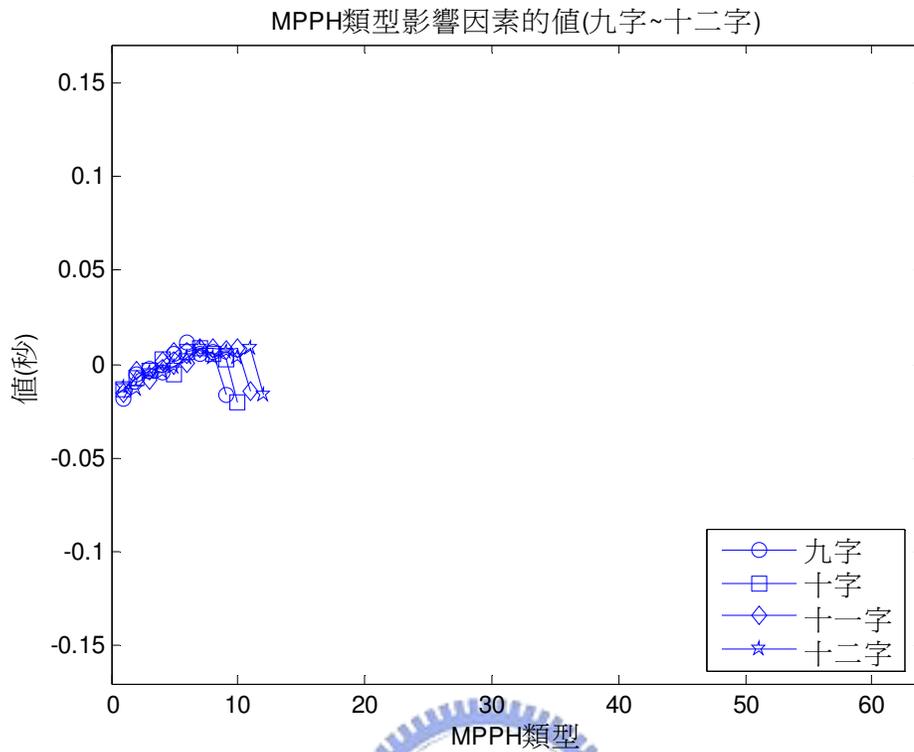
我們將 γ_{MPPH} 的值繪出如圖 7.12 到圖 7.26 所示。



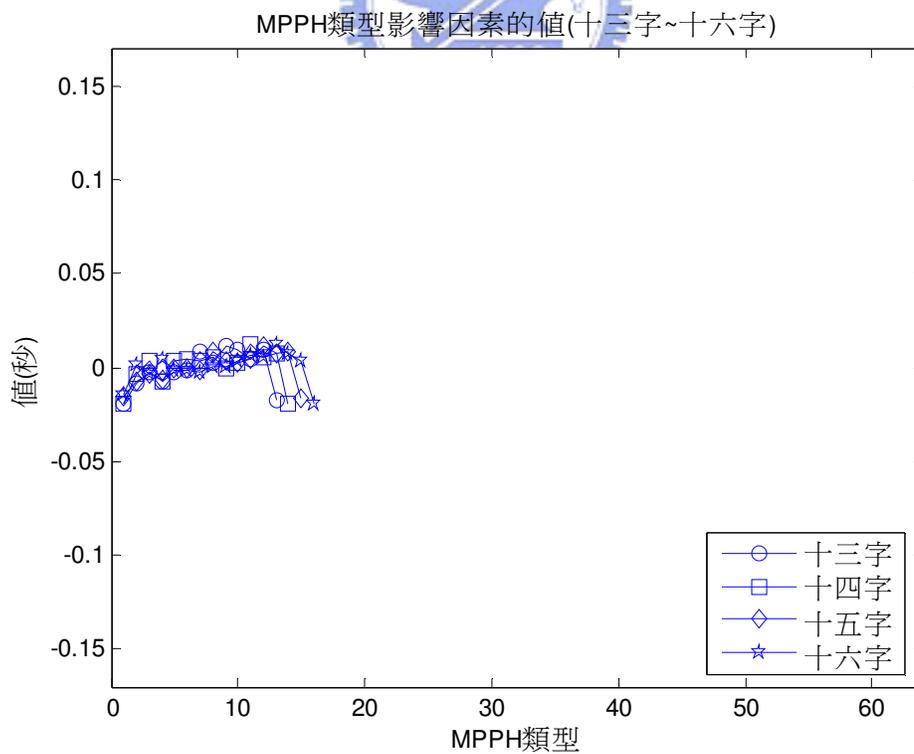
圖七.12： γ_{MPPH} 的值(一字~四字)。



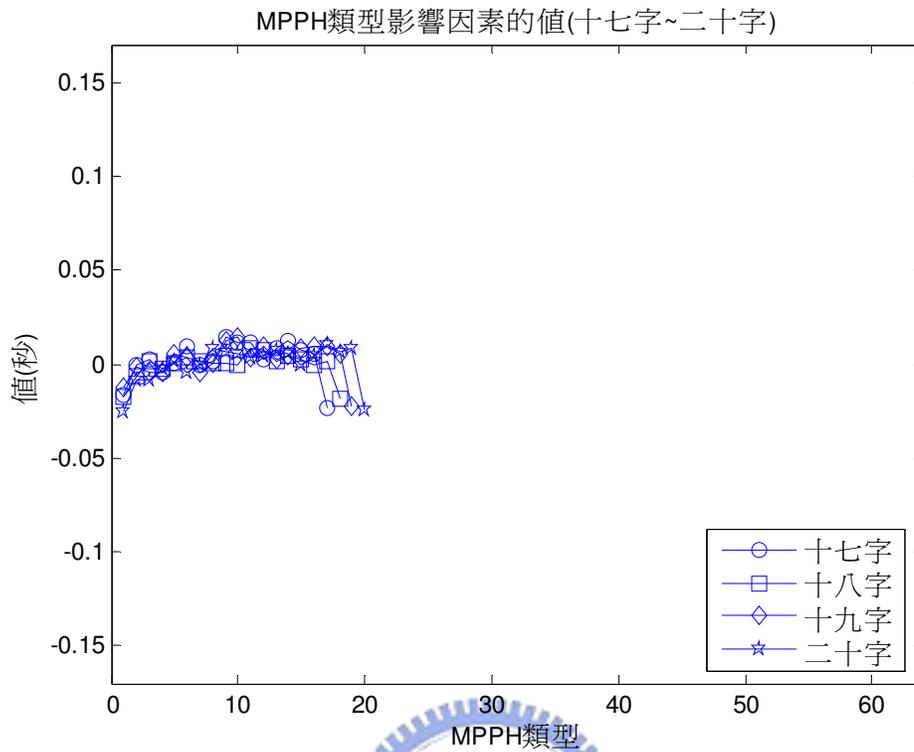
圖七.13： γ_{MPPH} 的值(五字~八字)。



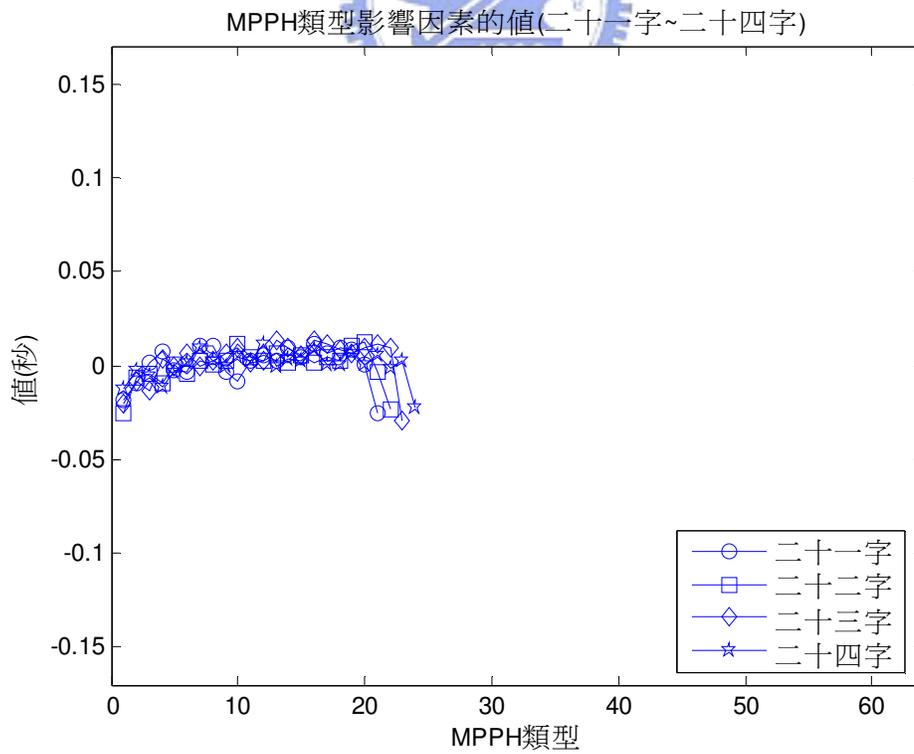
圖七.14： γ_{MPPH} 的值(九字~十二字)。



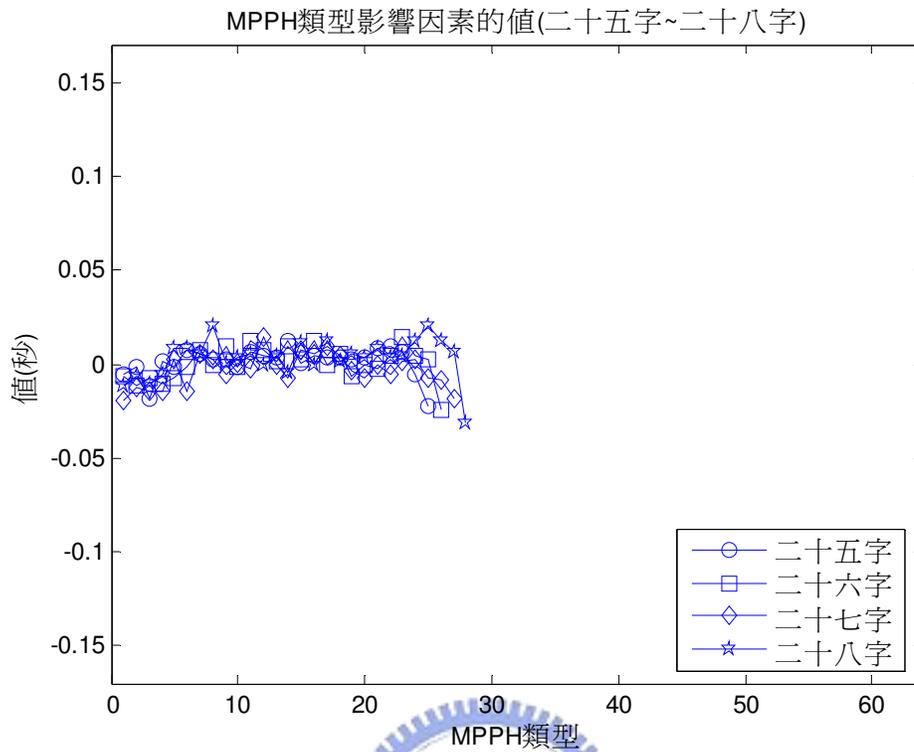
圖七.15： γ_{MPPH} 的值(十三字~十六字)。



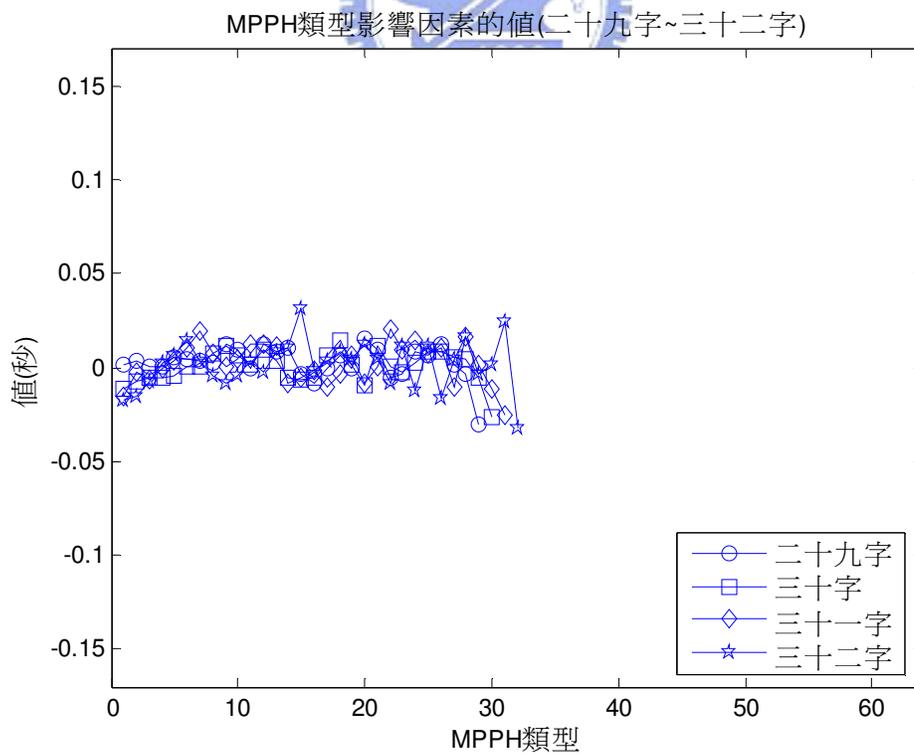
圖七.16： γ_{MPPH} 的值(十七字~二十字)。



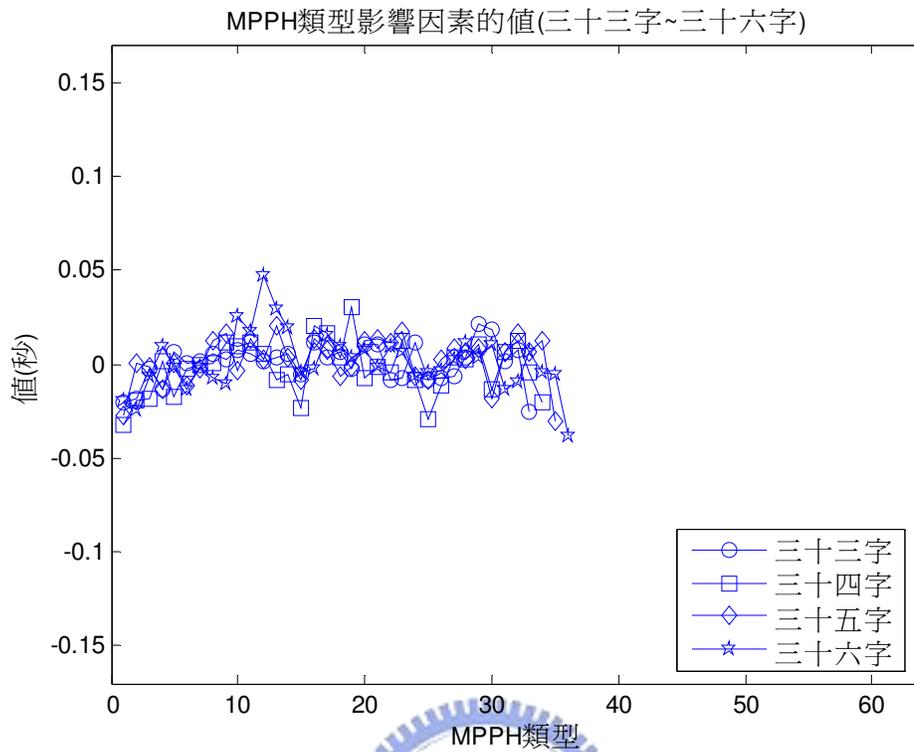
圖七.17： γ_{MPPH} 的值(二十一字~二十四字)。



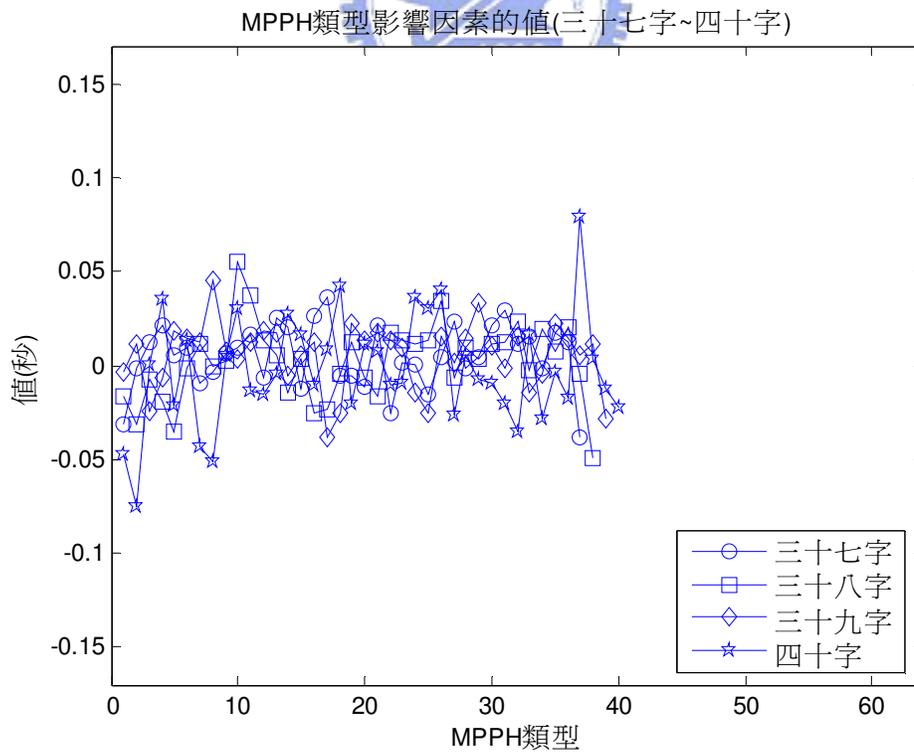
圖七.18： γ_{MPPH} 的值(二十五字~二十八字)。



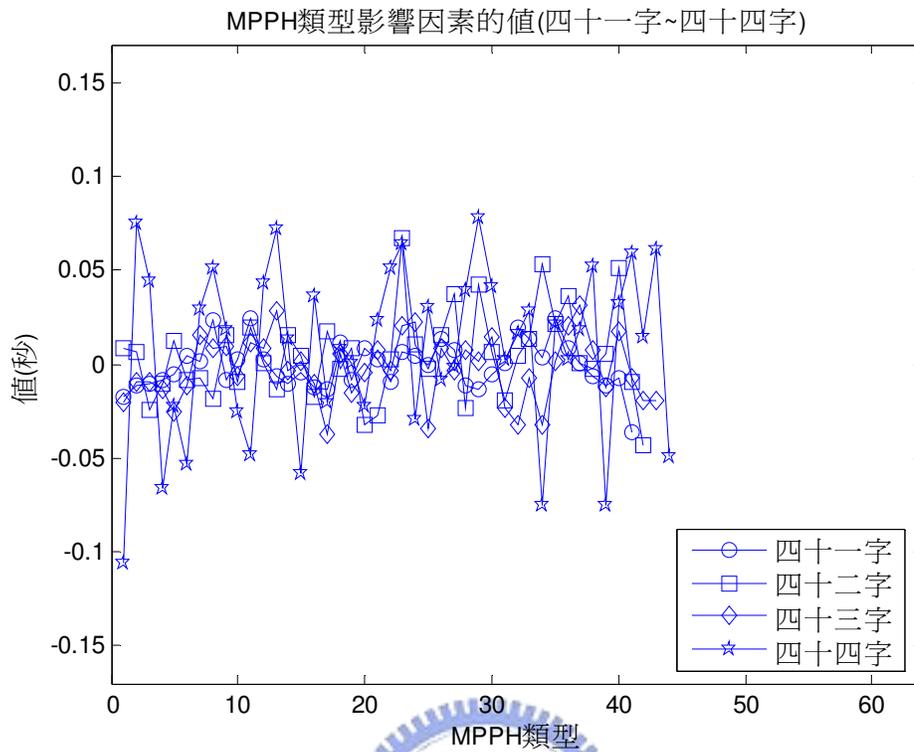
圖七.19： γ_{MPPH} 的值(二十九字~三十二字)。



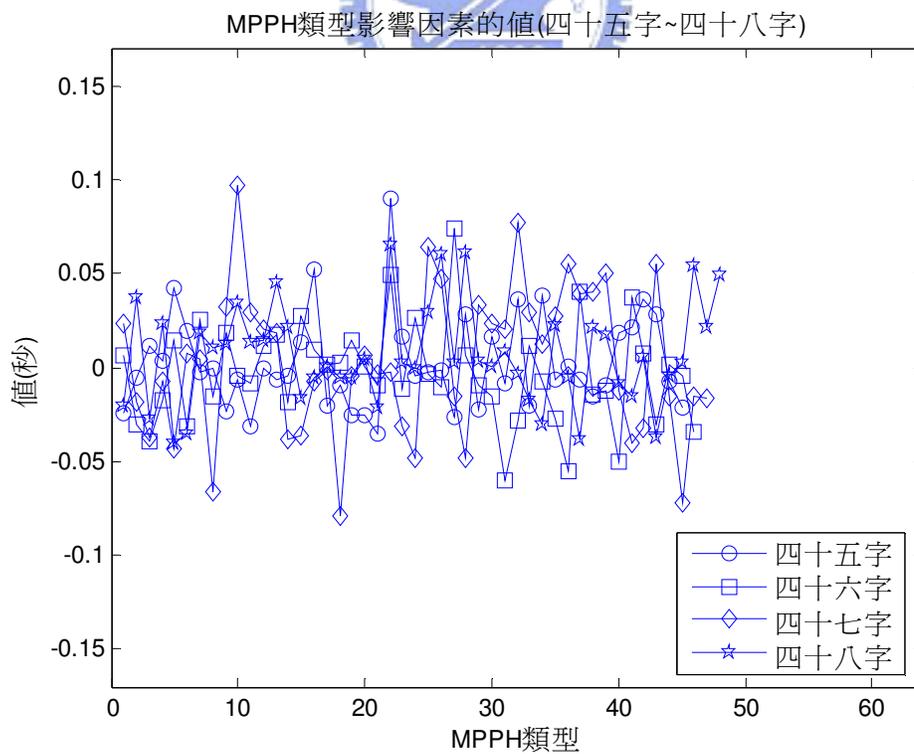
圖七.20： γ_{MPPH} 的值(三十三字~三十六字)。



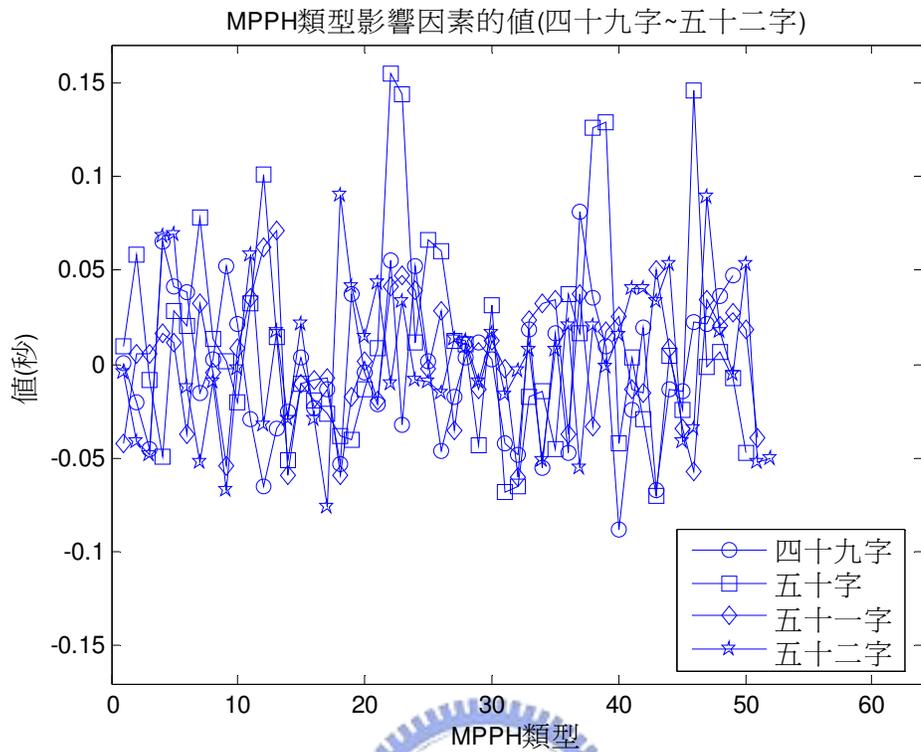
圖七.21： γ_{MPPH} 的值(三十七字~四十字)。



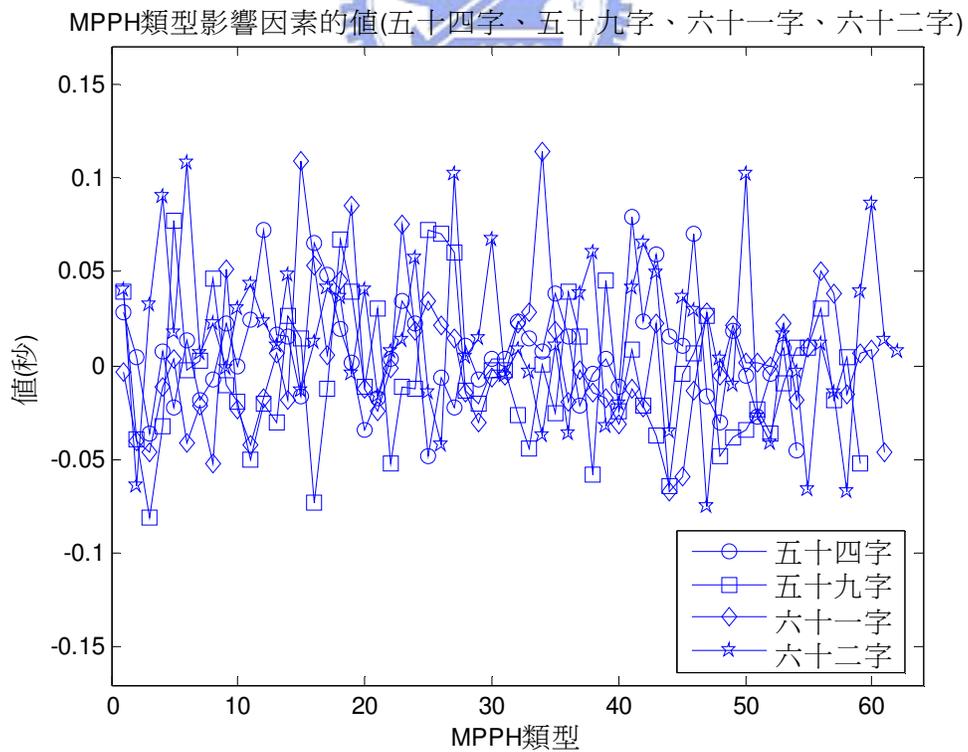
圖七.22： γ_{MPPH} 的值(四十一字~四十四字)。



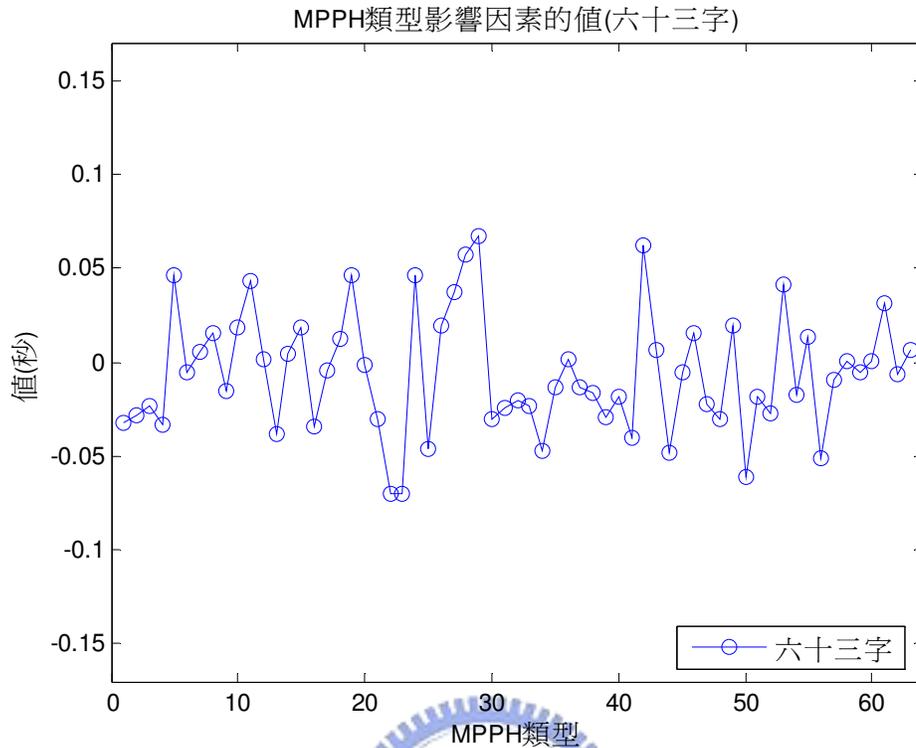
圖七.23： γ_{MPPH} 的值(四十五字~四十八字)。



圖七.24： γ_{MPPH} 的值(四十九字~五十二字)。



圖七.25： γ_{MPPH} 的值(五十四字、五十九字、六十一字、六十二字)。



圖七.26： γ_{MPPH} 的值(六十三字)。

7.1.3.1 實驗結果分析

我們由圖七.12 到圖七.26 可以觀察到 MPPH 的音長軌跡 Pattern，首先，我們的語料庫中沒有五十三字、五十五字到五十八字和六十字的 MPPH 存在；再者，一字到三十六字的 Pattern 有明顯固定的形狀，但是不同於我們在圖六.11 中所述在句尾會向上增長，其於句尾會向下減短，而三十七字以上的 Pattern 則沒有固定的形狀，因此模型所產生的 γ_{MPPH} 絕大部份是符合我們所認知的語音特性。

7.2 目標函數

對於 Q 我們將呈現出初始值、收斂值和更新曲線，以利於觀察對於模型之評估的平方誤差是否隨著模型的收斂而降低。

7.2.1 初始值

我們將 Q 的初始值呈現出如下：

$$Q'' = 87.639(\text{秒}^2) \quad (7.1)$$

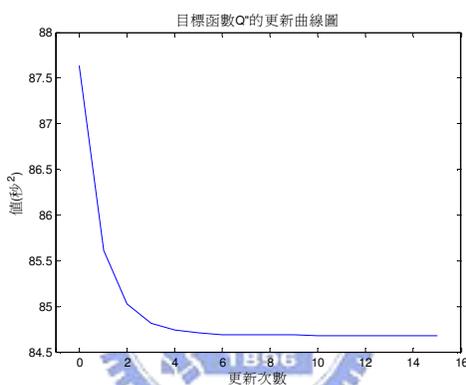
7.2.2 收斂值

我們將 Q'' 的收斂值呈現出如下：

$$Q'' = 84.684(\text{秒}^2) \quad (7.2)$$

7.2.3 更新曲線

我們將 Q'' 的更新曲線繪出如圖七.27 所示：



圖七.27：目標函數 Q'' 的更新曲線圖。

7.2.4 實驗結果分析

由以上所述，我們可以觀察到 Q'' 是隨著模型的收斂而降低，此即代表了模型是朝著合理的趨勢在收斂。

7.3 音長分佈和音長軌跡

我們將在這一個小節中呈現出語料庫中所有音節的 $SD3$ 分佈以及從語料庫中擷取一段句子的 $SD3$ 軌跡，而為了方便對所有圖形的顯示，我們定義了以下的數學符號：

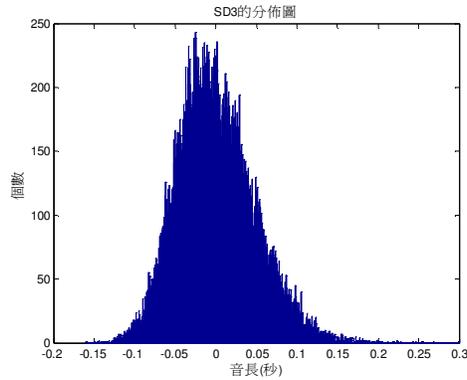
- $SD3$: $sd3$;
- $SD3_1$: $sd3 - \gamma_{PW}$;
- $SD3_2$: $sd3 - \gamma_{PW} - \gamma_{MIPPH}$;

- $SD3_3 : sd3 - \gamma_{PW} - \gamma_{MIPPH} - \gamma_{MPPH} ;$

7.3.1 音長分佈

我們將 $SD3$ 、 $SD3_1$ 、 $SD3_2$ 和 $SD3_3$ 的分佈條列呈述如下：

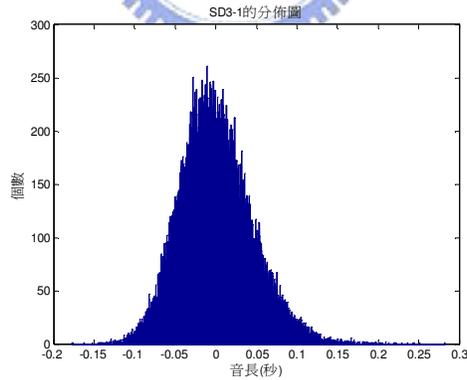
- 我們將 $SD3$ 的分佈繪出如圖七.28 所示：



圖七.28： $SD3$ 的分佈圖。

，其之變異數為 224.98×10^{-5} (秒²)。

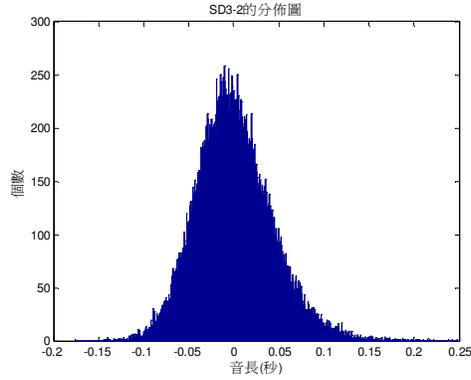
- 我們將 $SD3_1$ 的分佈繪出如圖七.29 所示：



圖七.29： $SD3_1$ 的分佈圖。

，其之變異數為 193.08×10^{-5} (秒²)。

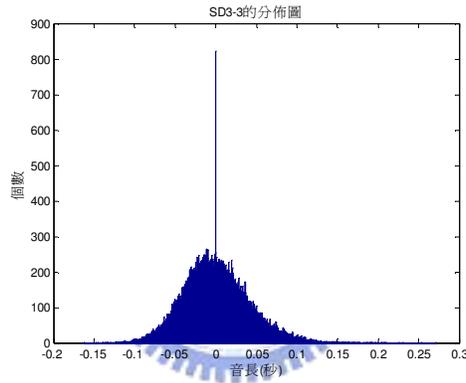
- 我們將 $SD3_2$ 的分佈繪出如圖七.30 所示：



圖七.30：SD3_2的分佈圖。

，其之變異數為 173.90×10^{-5} (秒²)。

- 我們將SD3_3的分佈繪出如圖七.31所示：



圖七.31：SD3_3的分佈圖。

，其之變異數為 162.25×10^{-5} (秒²)。

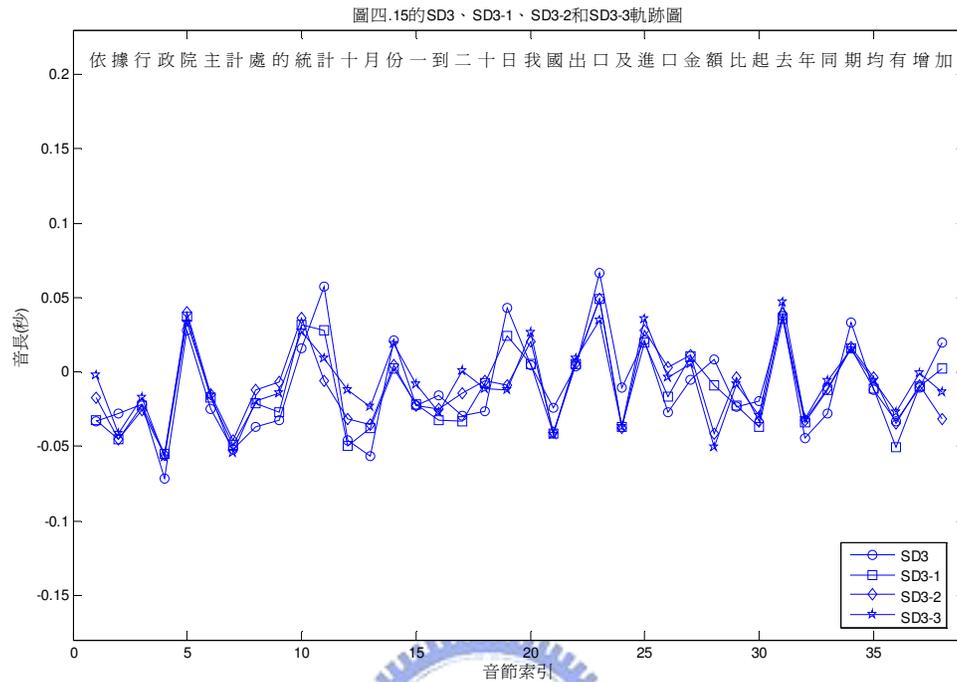
7.3.1.1 實驗結果分析

我們由上述的圖可以觀察到圖七.28到圖七.31的變異數逐漸變小，亦即SD3分佈漸趨集中，其所代表的意義為模型能夠有效地將造成每一個音節之SD3差異的因素分離並且抽取出來，使得SD3分佈之變異數相當程度地降低，由以上分析可以驗證我們的模型所造成的SD3分佈變化是合理的。

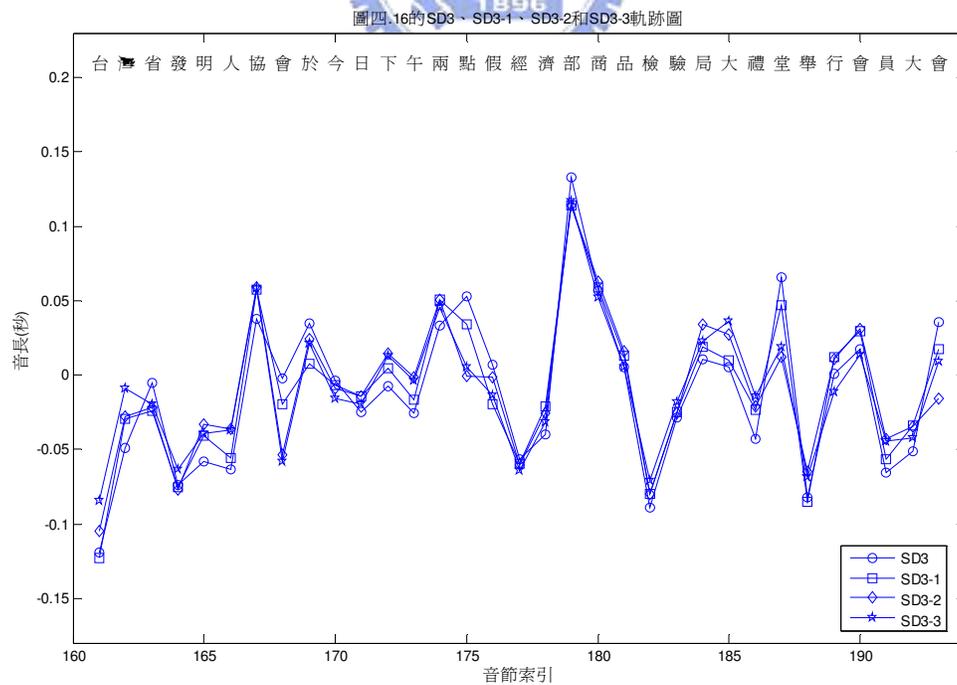
7.3.2 音長軌跡

我們呈現出圖四.15到圖四.17的SD3、SD3_1、SD3_2和SD3_3軌跡，如圖七.32到

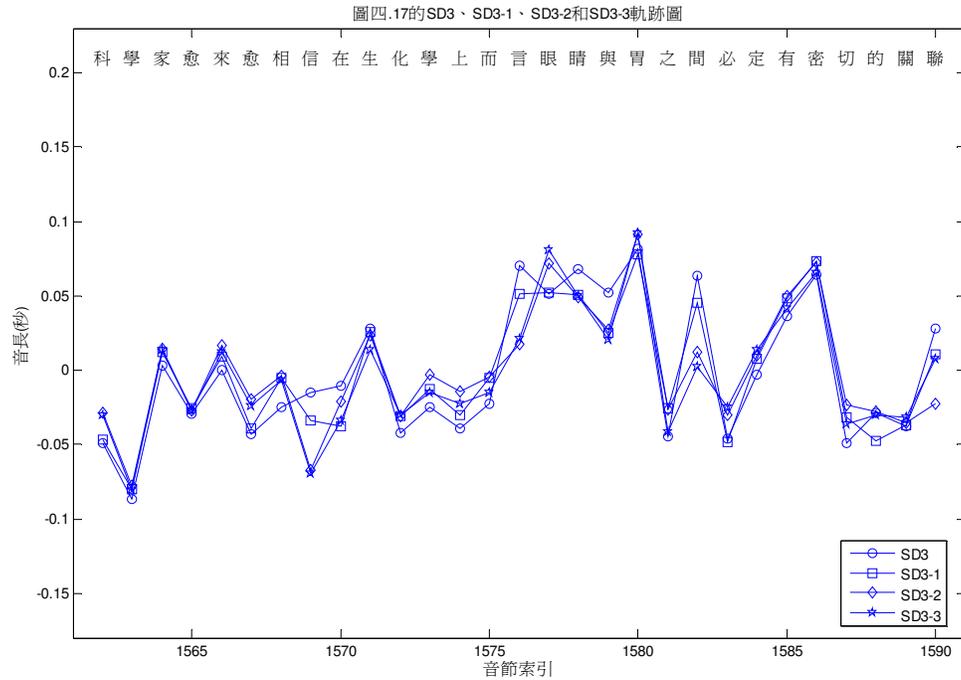
圖七.34 所示：



圖七.32：圖四.15 的SD3、SD3_1、SD3_2和SD3_3軌跡圖。



圖七.33：圖四.16 的SD3、SD3_1、SD3_2和SD3_3軌跡圖。



圖七.34：圖四.17 的 $SD3$ 、 $SD3_1$ 、 $SD3_2$ 和 $SD3_3$ 軌跡圖。

7.3.2.1 實驗結果分析

我們從圖七.32 到圖七.34 可以觀察到 $SD3$ 軌跡隨著減去各個影響因素後逐漸地平滑化，其之意義即為我們的模型能夠將造成每一個音節之 $SD3$ 差異的因素分離並且抽取出來，亦即驗證了我們的模型所造成的 $SD3$ 軌跡變化是合理的。

第八章 結論與未來展望

本章首先對於本論文所完成的工作做整體性的回顧和結論，而後提出幾個方向以在未來能夠做更進一步的研究。

8.1 結論

在本論文中我們提出了一個以統計為基礎的漢語語音音節音長模型，經由實驗證實我們的模型可以將影響音長的數個主要因素分離並且抽取出來，除了大幅地將音長分佈變異數降低和音長軌跡平滑化之外，各個影響因素也符合我們所認知的語音特性。而模型亦對語料庫中每一個音節標記上韻律狀態，其是扣除掉低階音節層次參數後代表著高階語言層次音長特性的標記，經過分析其亦符合我們所認知的語音特性。而後我們將韻律狀態和代表著漢語語音階層韻律結構標記的 Break Type 之關係做分析，以觀察高階語言層次音長特性的現象，也確實驗證了我們所認知的語音特性。而後我們也將音高和音長的韻律狀態做了相關性的分析，以觀察一個音節的音高和音長之間的相互現象，也確實符合我們所認知的語音特性。最後我們對代表著高階語言層次音長特性的音長建立了一個高階語言層次階層音長模型，經由實驗證實我們的模型可以將影響高階音長的三個階層影響因素分離並且抽取出來，除了將高階音長分佈變異數降低和高階音長軌跡平滑化之外，三個影響因素所代表的階層之高階音長軌跡 Pattern 也符合我們所認知的語音特性。

8.2 未來展望

對於本論文的研究方向有一些工作是值得未來進一步去研究的，首先本論文的模型可以再納入更多的影響因素，如此便可以更明確地掌握住各種影響音長的原因，進而去分析和了解它們，當然這將需要更精緻的文字分析工作。再者這套方法已經有效地被應用在韻律訊息中的音高訊息和停頓長度訊息上[11]，而韻律訊息中的音量訊息也是非常值得以這套方法去對其進行音量韻律現象的分析。



參考文獻

- [1] Sin-Horng Chen, and Yih-Ru Wang, “*Vector Quantization of Pitch Information in Mandarin Speech*,” IEEE Transactions on Communications, Vol. 38, No. 39, September, 1990.
- [2] Chiu-Yu Tseng, Shao-Huang Pin, Yehlin Lee, Hsin-Min Wang, and Yong-Cheng Chen, “*Fluent Speech Prosody: Framework and Modeling*,” Speech Communication 46, pp. 284-309, 2005.
- [3] 何鎮仲, “*混和統計與類神經網路法之台語韻律合成*”, 國立交通大學電信工程學系碩士班碩士論文, 中華民國八十九年六月。
- [4] C. C. Ho and S. H. Chen, “*A Maximum Likelihood Estimation of Duration Models for Taiwanese Speech*,” in Proceedings of ISAS/SCI’2000, Orlando, FL, July 2000.
- [5] C. C. Ho and S. H. Chen, “*A Hybrid Statistical/RNN Approach to Prosody Synthesis for Taiwanese TTS*,” in Proceedings of ICSLP’2000, Beijing, China, Oct. 2000.
- [6] Wen-Hsing Lai and Sin-Horng Chen, “*A Novel Syllable Duration Modeling Approach for Mandarin Speech*,” in Proceedings of ICASSP, Volume 1, 2001, pp. 93-96.
- [7] Wen-Hsing Lai and Sin-Horng Chen, “*Analysis of Syllable Duration Models for Mandarin Speech*,” in Proceedings of ICASSP, Volume 1, 2002, pp. 497-500.
- [8] Sin-Horng Chen, Wen-Hsing Lai, and Yih-Ru Wang, “*A New Duration Modeling Approach for Mandarin Speech*,” in IEEE Transactions on Speech and Audio Processing, Volume 11, NO. 4, July 2003, pp. 308-320.
- [9] Chen-Yu Chiang, Sin-Horng Chen, and Yih-Ru Wang, “*On the Inter-Syllable Coarticulation Effect of Pitch Modeling for Mandarin Speech*,” Interspeech, Lisbon, Portugal, November, 2006.
- [10] Chen-Yu Chiang, Xiao-Dong Wang, Yuan-Fu Liao, Yih-Ru Wang, Sin-Horng Chen, and Keikichi Hirose, “*Latent Prosody Model of Continuous Mandarin Speech*,” in Proceedings of ICASSP, 2007, pp. 625-628.
- [11] Chen-Yu Chiang, Hsiu-Min Yu, Yih-Ru Wang, and Sin-Horng Chen, “*An Automatic Prosody Labeling Method for Mandarin Speech*,” Interspeech, Antwerp, Belgium, August, 2007.

- [12] C. R. Huang, K. J. Chen, F. Y. Chen, Z. M. Gao, and K. Y. Chen, “*Sinica Treebank: Design Criteria, Annotation Guidelines, and On-Line Interface*,” in Proceedings of 2nd Chinese Language Processing Workshop 2000, Hong Kong, pp. 29-37.
- [13] 王小川, “**語音訊號處理**”, 全華科技圖書股份有限公司, 中華民國九十四年二月。
- [14] 江振宇, 蕭希群, 余秀敏, 廖元甫, “**語音韻律簡介**”, 中華民國計算機語言學會通訊, 中華民國九十六年。



附錄一

漢語 411 個基本音節碼與對應的注音符號、22 個聲母碼(見附錄二)、40 個韻母碼(見附錄三)的對照表。

音節碼	注音符號	聲母碼	韻母碼
1	ㄅ	2	1
2	ㄆ	3	1
3	ㄇ	4	1
4	ㄏ	5	1
5	ㄏ	6	40
6	ㄆ	7	40
7	ㄇ	8	40
8	ㄩ	1	2
9	ㄅㄩ	2	2
10	ㄆㄩ	3	2
11	ㄇㄩ	4	2
12	ㄏㄩ	6	2
13	ㄆㄩ	7	2
14	ㄇㄩ	8	2
15	ㄍㄩ	9	2
16	ㄎㄩ	10	2
17	ㄉㄩ	11	2
18	ㄎㄩ	15	2
19	ㄉㄩ	16	2
20	ㄎㄩ	17	2
21	ㄎㄩ	18	2
22	ㄎㄩ	19	2
23	ㄎㄩ	20	2
24	ㄇㄩ	21	2
25	ㄆㄩ	22	2
26	ㄅ	1	3
27	ㄆ	18	3
28	ㄆ	19	3
29	ㄆ	20	3

音節碼	注音符號	聲母碼	韻母碼
30	ㄇㄅ	21	3
31	ㄆㄅ	22	3
32	ㄅ	1	4
33	ㄅㄅ	2	4
34	ㄆㄅ	3	4
35	ㄇㄅ	4	4
36	ㄏㄅ	5	4
37	ㄏㄅ	6	4
38	ㄆㄅ	7	4
39	ㄇㄅ	8	4
40	ㄍㄅ	9	4
41	ㄎㄅ	10	4
42	ㄉㄅ	11	4
43	ㄎㄅ	15	4
44	ㄉㄅ	16	4
45	ㄎㄅ	17	4
46	ㄎㄅ	18	4
47	ㄅ	1	6
48	ㄅ	2	6
49	ㄆ	3	6
50	ㄇ	4	6
51	ㄏ	6	6
52	ㄆ	7	6
53	ㄇ	8	6
54	ㄍ	9	6
55	ㄎ	10	6
56	ㄉ	11	6
57	ㄎ	15	6
58	ㄉ	16	6

59	ㄋㄛ	17	6
60	ㄎㄛ	18	6
61	ㄏㄛ	19	6
62	ㄏㄛ	20	6
63	ㄏㄛ	21	6
64	せ	1	5
65	せゝ	2	7
66	せゝ	4	7
67	せゝ	6	7
68	せゝ	8	7
69	せゝ	9	7
70	せゝ	11	7
71	せゝ	15	7
72	せゝ	17	7
73	せゝ	18	7
74	せゝ	19	7
75	せゝ	20	7
76	せゝ	21	7
77	せゝ	22	7
78	せ	1	8
79	せゝ	2	8
80	せゝ	3	8
81	せゝ	4	8
82	せゝ	5	8
83	せゝ	6	8
84	せゝ	7	8
85	せゝ	8	8
86	せゝ	9	8
87	せゝ	10	8
88	せゝ	11	8
89	せゝ	15	8
90	せゝ	16	8
91	せゝ	17	8
92	せゝ	18	8
93	せゝ	19	8
94	せゝ	20	8
95	せゝ	21	8

96	又	1	9
97	又	2	9
98	又	3	9
99	又	4	9
100	又	5	9
101	又	6	9
102	又	7	9
103	又	8	9
104	又	9	9
105	又	10	9
106	又	11	9
107	又	15	9
108	又	16	9
109	又	17	9
110	又	18	9
111	又	20	9
112	又	21	9
113	又	22	9
114	又	1	10
115	又	2	10
116	又	3	10
117	又	4	10
118	又	5	10
119	又	6	10
120	又	7	10
121	又	8	10
122	又	9	10
123	又	10	10
124	又	11	10
125	又	15	10
126	又	16	10
127	又	17	10
128	又	18	10
129	又	19	10
130	又	20	10
131	又	21	10
132	又	22	10

133	ㇿ	1	11
134	𠂇	2	11
135	𠂈	3	11
136	𠂉	4	11
137	𠂊	5	11
138	𠂋	6	11
139	𠂌	7	11
140	𠂍	8	11
141	𠂎	9	11
142	𠂏	10	11
143	𠂐	11	11
144	𠂑	17	11
145	𠂒	19	11
146	𠂓	20	11
147	𠂔	21	11
148	𠂕	22	11
149	ㇿ	1	12
150	𠂇	2	12
151	𠂈	3	12
152	𠂉	4	12
153	𠂊	5	12
154	𠂋	6	12
155	𠂌	7	12
156	𠂍	8	12
157	𠂎	9	12
158	𠂏	10	12
159	𠂐	11	12
160	𠂑	15	12
161	𠂒	16	12
162	𠂓	17	12
163	𠂔	18	12
164	𠂕	19	12
165	𠂖	20	12
166	𠂗	21	12
167	𠂘	22	12
168	𠂙	2	13
169	𠂚	3	13

170	𠂛	4	13
171	𠂜	5	13
172	𠂝	6	13
173	𠂞	7	13
174	𠂟	8	13
175	𠂠	9	13
176	𠂡	10	13
177	𠂢	11	13
178	𠂣	15	13
179	𠂤	16	13
180	𠂥	17	13
181	𠂦	18	13
182	𠂧	19	13
183	𠂨	20	13
184	𠂩	21	13
185	𠂪	22	13
186	一	1	14
187	𠂫	12	14
188	𠂬	13	14
189	𠂭	14	14
190	𠂮	15	14
191	𠂯	16	14
192	𠂰	17	14
193	𠂱	18	14
194	𠂲	19	14
195	𠂳	20	14
196	𠂴	21	14
197	𠂵	1	15
198	𠂶	2	15
199	𠂷	3	15
200	𠂸	4	15
201	𠂹	5	15
202	𠂺	6	15
203	𠂻	7	15
204	𠂼	8	15
205	𠂽	9	15
206	𠂾	10	15

207	厂メ	11	15
208	夕メ	15	15
209	去メ	16	15
210	ろメ	17	15
211	カメ	18	15
212	ケメ	19	15
213	クメ	20	15
214	ケメ	21	15
215	クメ	22	15
216	口	1	16
217	口	12	16
218	く口	13	16
219	一口	14	16
220	ろ口	17	16
221	カ口	18	16
222	一ヤ	1	17
223	口一ヤ	12	17
224	く一ヤ	13	17
225	一口一ヤ	14	17
226	カ一ヤ	18	17
227	一セ	1	18
228	口一セ	12	18
229	く一セ	13	18
230	一口一セ	14	18
231	カ一セ	15	18
232	去一セ	16	18
233	ろ一セ	17	18
234	カ一セ	18	18
235	ケ一セ	19	18
236	ク一セ	20	18
237	ケ一セ	21	18
238	一カ	1	19
239	一カ	1	20
240	口一カ	12	20
241	く一カ	13	20
242	一口一カ	14	20
243	カ一カ	15	20

244	去一カ	16	20
245	ろ一カ	17	20
246	カ一カ	18	20
247	ケ一カ	19	20
248	ク一カ	20	20
249	ケ一カ	21	20
250	一カ	1	21
251	口一カ	12	21
252	く一カ	13	21
253	一口一カ	14	21
254	カ一カ	15	21
255	ろ一カ	17	21
256	カ一カ	18	21
257	ケ一カ	21	21
258	一カ	1	22
259	口一カ	12	22
260	く一カ	13	22
261	一口一カ	14	22
262	カ一カ	15	22
263	去一カ	16	22
264	ろ一カ	17	22
265	カ一カ	18	22
266	ケ一カ	19	22
267	ク一カ	20	22
268	ケ一カ	21	22
269	一カ	1	23
270	口一カ	12	23
271	く一カ	13	23
272	一口一カ	14	23
273	カ一カ	17	23
274	カ一カ	18	23
275	ケ一カ	19	23
276	ク一カ	20	23
277	ケ一カ	21	23
278	一カ	1	24
279	口一カ	12	24
280	く一カ	13	24

281	ㄒ一ㄅ	14	24
282	ㄟ一ㄅ	17	24
283	ㄎ一ㄅ	18	24
284	一ㄇ	1	25
285	ㄩ一ㄇ	12	25
286	ㄎ一ㄇ	13	25
287	ㄒ一ㄇ	14	25
288	ㄎ一ㄇ	15	25
289	ㄍ一ㄇ	16	25
290	ㄟ一ㄇ	17	25
291	ㄎ一ㄇ	18	25
292	ㄎ一ㄇ	19	25
293	ㄎ一ㄇ	20	25
294	ㄇ一ㄇ	21	25
295	ㄨㄚ	1	26
296	ㄗㄨㄚ	2	26
297	ㄨㄨㄚ	3	26
298	ㄨㄨㄚ	4	26
299	ㄨㄨㄚ	9	26
300	ㄨㄨㄚ	10	26
301	ㄨㄨㄚ	11	26
302	ㄨㄨㄚ	1	27
303	ㄗㄨㄨㄚ	2	27
304	ㄨㄨㄨㄚ	3	27
305	ㄨㄨㄨㄚ	4	27
306	ㄨㄨㄨㄚ	5	27
307	ㄨㄨㄨㄚ	6	27
308	ㄨㄨㄨㄚ	7	27
309	ㄨㄨㄨㄚ	8	27
310	ㄨㄨㄨㄚ	9	27
311	ㄨㄨㄨㄚ	10	27
312	ㄨㄨㄨㄚ	11	27
313	ㄨㄨㄨㄚ	15	27
314	ㄨㄨㄨㄚ	16	27
315	ㄨㄨㄨㄚ	17	27
316	ㄨㄨㄨㄚ	18	27
317	ㄨㄨㄨㄚ	1	28

318	ㄗㄨㄨㄚ	2	28
319	ㄨㄨㄨㄚ	3	28
320	ㄨㄨㄨㄚ	4	28
321	ㄨㄨㄨㄚ	9	28
322	ㄨㄨㄨㄚ	10	28
323	ㄨㄨㄨㄚ	11	28
324	ㄨㄨㄨㄚ	1	29
325	ㄗㄨㄨㄚ	2	29
326	ㄨㄨㄨㄚ	3	29
327	ㄨㄨㄨㄚ	4	29
328	ㄨㄨㄨㄚ	5	29
329	ㄨㄨㄨㄚ	6	29
330	ㄨㄨㄨㄚ	7	29
331	ㄨㄨㄨㄚ	8	29
332	ㄨㄨㄨㄚ	9	29
333	ㄨㄨㄨㄚ	10	29
334	ㄨㄨㄨㄚ	11	29
335	ㄨㄨㄨㄚ	15	29
336	ㄨㄨㄨㄚ	16	29
337	ㄨㄨㄨㄚ	1	30
338	ㄗㄨㄨㄚ	2	30
339	ㄨㄨㄨㄚ	3	30
340	ㄨㄨㄨㄚ	4	30
341	ㄨㄨㄨㄚ	5	30
342	ㄨㄨㄨㄚ	6	30
343	ㄨㄨㄨㄚ	7	30
344	ㄨㄨㄨㄚ	8	30
345	ㄨㄨㄨㄚ	9	30
346	ㄨㄨㄨㄚ	10	30
347	ㄨㄨㄨㄚ	11	30
348	ㄨㄨㄨㄚ	15	30
349	ㄨㄨㄨㄚ	16	30
350	ㄨㄨㄨㄚ	17	30
351	ㄨㄨㄨㄚ	18	30
352	ㄨㄨㄨㄚ	1	31
353	ㄗㄨㄨㄚ	2	31
354	ㄨㄨㄨㄚ	3	31

355	尸メㄥ	4	31
356	回メㄥ	5	31
357	尸メㄥ	6	31
358	ㄥメㄥ	7	31
359	厶メㄥ	8	31
360	ㄥメㄥ	9	31
361	ㄥメㄥ	10	31
362	厶メㄥ	11	31
363	ㄥメㄥ	15	31
364	厶メㄥ	16	31
365	ㄥメㄥ	18	31
366	メㄥ	1	32
367	ㄥメㄥ	2	32
368	ㄥメㄥ	3	32
369	尸メㄥ	4	32
370	ㄥメㄥ	9	32
371	ㄥメㄥ	10	32
372	厶メㄥ	11	32
373	メㄥ	1	33
374	ㄥメㄥ	2	33
375	ㄥメㄥ	3	33
376	回メㄥ	5	33
377	尸メㄥ	6	33
378	ㄥメㄥ	7	33
379	厶メㄥ	8	33
380	ㄥメㄥ	9	33
381	ㄥメㄥ	10	33
382	厶メㄥ	11	33
383	ㄥメㄥ	15	33
384	厶メㄥ	16	33
385	ㄥメㄥ	17	33
386	ㄥメㄥ	18	33
387	ㄥせ	1	34
388	ㄥㄥせ	12	34
389	ㄥㄥせ	13	34
390	ㄥㄥせ	14	34
391	ㄥㄥせ	17	34

392	ㄥㄥせ	18	34
393	ㄥㄥ	1	35
394	ㄥㄥㄥ	12	35
395	ㄥㄥㄥ	13	35
396	ㄥㄥㄥ	14	35
397	ㄥㄥㄥ	18	35
398	ㄥㄥ	1	36
399	ㄥㄥㄥ	12	36
400	ㄥㄥㄥ	13	36
401	ㄥㄥㄥ	14	36
402	ㄥㄥㄥ	18	36
403	ㄥㄥ	1	37
404	ㄥㄥㄥ	12	37
405	ㄥㄥㄥ	13	37
406	ㄥㄥㄥ	14	37
407	ㄥ	1	38
408	一ㄥ	1	39
409	ㄥ	1	13
410	ㄥ	1	7
411	ㄥㄥ	21	4

附錄二

漢語 22 個聲母碼與對應的注音符號對照表，其中Φ1 所代表的意義為空聲母，亦即在音節中沒有聲母只有韻母。

聲母碼	注音符號
1	Φ1
2	ㄅ
3	ㄆ
4	ㄇ
5	ㄏ
6	ㄏ
7	ㄏ
8	ㄏ
9	ㄏ
10	ㄏ
11	ㄏ
12	ㄏ
13	ㄏ
14	ㄏ
15	ㄏ
16	ㄏ
17	ㄏ
18	ㄏ
19	ㄏ
20	ㄏ
21	ㄏ
22	ㄏ



附錄三

漢語 40 個韻母碼與對應的注音符號對照表，其中Φ1 和Φ2 所代表的意義為兩種類型的空韻母，亦即在音節中只有聲母沒有韻母。

韻母碼	注音符號	韻母碼	注音符號
1	Φ1	30	ㄨㄛ
2	ㄚ	31	ㄨㄣ
3	ㄛ	32	ㄨㄨ
4	ㄜ	33	ㄨㄥ
5	ㄝ	34	ㄛㄝ
6	ㄞ	35	ㄛㄨ
7	ㄟ	36	ㄛㄣ
8	ㄚ	37	ㄛㄥ
9	ㄨ	38	ㄨ
10	ㄛ	39	ㄛ
11	ㄣ	40	Φ2
12	ㄨ		
13	ㄥ		
14	一		
15	ㄨ		
16	ㄛ		
17	一ㄚ		
18	一ㄛ		
19	一ㄜ		
20	一ㄝ		
21	一ㄞ		
22	一ㄟ		
23	一ㄚ		
24	一ㄨ		
25	一ㄛ		
26	ㄨㄚ		
27	ㄨㄛ		
28	ㄨㄜ		
29	ㄨㄝ		



附錄四

漢語 411 個基本音節碼與對應的注音符號、22 個聲母碼、40 個韻母碼和語料庫中每一種基本音節類型之音節個數的對照表。

音節碼	注音符號	聲母碼	韻母碼	個數	音節碼	注音符號	聲母碼	韻母碼	個數
1	ㄅ	2	1	926	30	ㄇㄛ	21	3	148
2	ㄆ	3	1	109	31	ㄇㄛ	22	3	25
3	ㄆ	4	1	2079	32	ㄛ	1	4	29
4	ㄇ	5	1	195	33	ㄅㄛ	2	4	365
5	ㄆ	6	40	419	34	ㄆㄛ	3	4	22
6	ㄇ	7	40	259	35	ㄆㄛ	4	4	134
7	ㄇ	8	40	278	36	ㄇㄛ	5	4	15
8	ㄩ	1	2	15	37	ㄆㄛ	6	4	121
9	ㄅㄩ	2	2	9	38	ㄇㄛ	7	4	41
10	ㄆㄩ	3	2	50	39	ㄇㄛ	8	4	32
11	ㄆㄩ	4	2	16	40	ㄍㄛ	9	4	307
12	ㄆㄩ	6	2	17	41	ㄎㄛ	10	4	274
13	ㄇㄩ	7	2	3	42	ㄏㄛ	11	4	214
14	ㄇㄩ	8	2	9	43	ㄉㄛ	15	4	2018
15	ㄍㄩ	9	2	3	44	ㄊㄛ	16	4	26
16	ㄎㄩ	10	2	3	45	ㄋㄛ	17	4	19
17	ㄏㄩ	11	2	8	46	ㄌㄛ	18	4	272
18	ㄉㄩ	15	2	435	47	ㄴ	1	6	82
19	ㄊㄩ	16	2	462	48	ㄅㄴ	2	6	10
20	ㄋㄩ	17	2	117	49	ㄆㄴ	3	6	10
21	ㄌㄩ	18	2	22	50	ㄆㄴ	4	6	2
22	ㄌㄩ	19	2	170	51	ㄆㄴ	6	6	529
23	ㄴ	20	2	16	52	ㄇㄴ	7	6	154
24	ㄇㄴ	21	2	48	53	ㄇㄴ	8	6	7
25	ㄇㄴ	22	2	311	54	ㄍㄴ	9	6	95
26	ㄛ	1	3	1	55	ㄎㄴ	10	6	78
27	ㄌㄛ	18	3	0	56	ㄏㄴ	11	6	154
28	ㄌㄛ	19	3	100	57	ㄉㄴ	15	6	115
29	ㄴ	20	3	97	58	ㄊㄴ	16	6	410

59	ㄋㄝ	17	6	16
60	ㄉㄝ	18	6	232
61	ㄅㄝ	19	6	88
62	ㄆㄝ	20	6	83
63	ㄇㄝ	21	6	22
64	ㄝ	1	5	0
65	ㄝㄨ	2	7	0
66	ㄝㄨ	4	7	13
67	ㄝㄨ	6	7	4
68	ㄝㄨ	8	7	0
69	ㄝㄨ	9	7	34
70	ㄝㄨ	11	7	18
71	ㄝㄨ	15	7	7
72	ㄝㄨ	17	7	109
73	ㄝㄨ	18	7	40
74	ㄝㄨ	19	7	183
75	ㄝㄨ	20	7	34
76	ㄝㄨ	21	7	252
77	ㄝㄨ	22	7	108
78	ㄝ	1	8	22
79	ㄝㄨ	2	8	71
80	ㄝㄨ	3	8	57
81	ㄝㄨ	4	8	73
82	ㄝㄨ	5	8	3
83	ㄝㄨ	6	8	75
84	ㄝㄨ	7	8	16
85	ㄝㄨ	8	8	8
86	ㄝㄨ	9	8	156
87	ㄝㄨ	10	8	28
88	ㄝㄨ	11	8	209
89	ㄝㄨ	15	8	344
90	ㄝㄨ	16	8	38
91	ㄝㄨ	17	8	9
92	ㄝㄨ	18	8	93
93	ㄝㄨ	19	8	191
94	ㄝㄨ	20	8	9
95	ㄝㄨ	21	8	33

96	ㄨ	1	9	23
97	ㄨㄨ	2	9	61
98	ㄨㄨ	3	9	25
99	ㄨㄨ	4	9	174
100	ㄨㄨ	5	9	17
101	ㄨㄨ	6	9	32
102	ㄨㄨ	7	9	0
103	ㄨㄨ	8	9	4
104	ㄨㄨ	9	9	54
105	ㄨㄨ	10	9	50
106	ㄨㄨ	11	9	192
107	ㄨㄨ	15	9	100
108	ㄨㄨ	16	9	99
109	ㄨㄨ	17	9	0
110	ㄨㄨ	18	9	16
111	ㄨㄨ	20	9	0
112	ㄨㄨ	21	9	30
113	ㄨㄨ	22	9	16
114	ㄨ	1	10	124
115	ㄨㄨ	2	10	103
116	ㄨㄨ	3	10	45
117	ㄨㄨ	4	10	57
118	ㄨㄨ	5	10	166
119	ㄨㄨ	6	10	5
120	ㄨㄨ	7	10	63
121	ㄨㄨ	8	10	126
122	ㄨㄨ	9	10	151
123	ㄨㄨ	10	10	84
124	ㄨㄨ	11	10	155
125	ㄨㄨ	15	10	204
126	ㄨㄨ	16	10	71
127	ㄨㄨ	17	10	100
128	ㄨㄨ	18	10	38
129	ㄨㄨ	19	10	86
130	ㄨㄨ	20	10	56
131	ㄨㄨ	21	10	54
132	ㄨㄨ	22	10	172

133	ㇿ	1	11	9
134	𠂇	2	11	105
135	𠂈	3	11	54
136	𠂉	4	11	180
137	𠂊	5	11	752
138	𠂋	6	11	15
139	𠂌	7	11	0
140	𠂍	8	11	7
141	𠂎	9	11	60
142	𠂏	10	11	8
143	𠂐	11	11	82
144	𠂑	17	11	2
145	𠂒	19	11	144
146	𠂓	20	11	5
147	𠂔	21	11	243
148	𠂕	22	11	188
149	尢	1	12	1
150	𠂇	2	12	285
151	𠂈	3	12	209
152	𠂉	4	12	244
153	𠂊	5	12	38
154	𠂋	6	12	11
155	𠂌	7	12	9
156	𠂍	8	12	7
157	𠂎	9	12	36
158	𠂏	10	12	57
159	𠂐	11	12	52
160	𠂑	15	12	243
161	𠂒	16	12	22
162	𠂓	17	12	0
163	𠂔	18	12	13
164	𠂕	19	12	25
165	𠂖	20	12	10
166	𠂗	21	12	40
167	𠂘	22	12	250
168	𠂙	2	13	384
169	𠂚	3	13	249

170	𠂛	4	13	311
171	𠂜	5	13	36
172	𠂝	6	13	26
173	𠂞	7	13	65
174	𠂟	8	13	4
175	𠂠	9	13	77
176	𠂡	10	13	5
177	𠂢	11	13	8
178	𠂣	15	13	122
179	𠂤	16	13	4
180	𠂥	17	13	176
181	𠂦	18	13	23
182	𠂧	19	13	2
183	𠂨	20	13	28
184	𠂩	21	13	70
185	𠂪	22	13	57
186	一	1	14	1610
187	𠂫	12	14	665
188	𠂬	13	14	501
189	𠂭	14	14	284
190	𠂮	15	14	239
191	𠂯	16	14	238
192	𠂰	17	14	100
193	𠂱	18	14	603
194	𠂲	19	14	156
195	𠂳	20	14	37
196	𠂴	21	14	62
197	𠂵	1	15	373
198	𠂶	2	15	323
199	𠂷	3	15	347
200	𠂸	4	15	175
201	𠂹	5	15	232
202	𠂺	6	15	99
203	𠂻	7	15	18
204	𠂼	8	15	95
205	𠂽	9	15	144
206	𠂾	10	15	20

207	厂メ	11	15	92
208	夕メ	15	15	198
209	去メ	16	15	104
210	ろメ	17	15	16
211	カメ	18	15	104
212	クメ	19	15	811
213	クメ	20	15	33
214	冂メ	21	15	95
215	匕メ	22	15	396
216	口	1	16	669
217	㇀口	12	16	206
218	く口	13	16	199
219	冂口	14	16	139
220	ろ口	17	16	119
221	カ口	18	16	41
222	一㇀	1	17	59
223	㇀一㇀	12	17	307
224	く一㇀	13	17	3
225	冂一㇀	14	17	149
226	カ一㇀	18	17	0
227	一セ	1	18	278
228	㇀一セ	12	18	319
229	く一セ	13	18	96
230	冂一セ	14	18	177
231	夕一セ	15	18	4
232	去一セ	16	18	8
233	ろ一セ	17	18	1
234	カ一セ	18	18	45
235	ク一セ	19	18	33
236	ク一セ	20	18	0
237	冂一セ	21	18	3
238	一𠂔	1	19	0
239	一𠂔	1	20	220
240	㇀一𠂔	12	20	242
241	く一𠂔	13	20	16
242	冂一𠂔	14	20	203
243	夕一𠂔	15	20	106

244	去一𠂔	16	20	52
245	ろ一𠂔	17	20	6
246	カ一𠂔	18	20	56
247	ク一𠂔	19	20	119
248	ク一𠂔	20	20	20
249	冂一𠂔	21	20	26
250	一又	1	21	754
251	㇀一又	12	21	302
252	く一又	13	21	74
253	冂一又	14	21	25
254	夕一又	15	21	3
255	ろ一又	17	21	16
256	カ一又	18	21	138
257	冂一又	21	21	1
258	一𠂔	1	22	223
259	㇀一𠂔	12	22	376
260	く一𠂔	13	22	179
261	冂一𠂔	14	22	294
262	夕一𠂔	15	22	128
263	去一𠂔	16	22	109
264	ろ一𠂔	17	22	259
265	カ一𠂔	18	22	106
266	ク一𠂔	19	22	112
267	ク一𠂔	20	22	32
268	冂一𠂔	21	22	132
269	一𠂔	1	23	271
270	㇀一𠂔	12	23	284
271	く一𠂔	13	23	63
272	冂一𠂔	14	23	273
273	ろ一𠂔	17	23	14
274	カ一𠂔	18	23	71
275	ク一𠂔	19	23	6
276	ク一𠂔	20	23	29
277	冂一𠂔	21	23	270
278	一尢	1	24	135
279	㇀一尢	12	24	182
280	く一尢	13	24	77

281	ㄒ一ㄅ	14	24	395
282	ㄗ一ㄅ	17	24	4
283	ㄎ一ㄅ	18	24	185
284	一ㄥ	1	25	137
285	ㄐ一ㄥ	12	25	326
286	ㄑ一ㄥ	13	25	237
287	ㄒ一ㄥ	14	25	368
288	ㄓ一ㄥ	15	25	121
289	ㄔ一ㄥ	16	25	82
290	ㄗ一ㄥ	17	25	20
291	ㄎ一ㄥ	18	25	130
292	ㄎ一ㄥ	19	25	163
293	ㄎ一ㄥ	20	25	114
294	ㄏ一ㄥ	21	25	210
295	ㄨㄚˊ	1	26	10
296	ㄨㄚˊ	2	26	4
297	ㄨㄚˊ	3	26	0
298	ㄨㄚˊ	4	26	0
299	ㄨㄚˊ	9	26	10
300	ㄨㄚˊ	10	26	2
301	ㄨㄚˊ	11	26	234
302	ㄨㄚˊ	1	27	379
303	ㄨㄚˊ	2	27	16
304	ㄨㄚˊ	3	27	0
305	ㄨㄚˊ	4	27	179
306	ㄨㄚˊ	5	27	36
307	ㄨㄚˊ	6	27	228
308	ㄨㄚˊ	7	27	18
309	ㄨㄚˊ	8	27	143
310	ㄨㄚˊ	9	27	726
311	ㄨㄚˊ	10	27	18
312	ㄨㄚˊ	11	27	241
313	ㄨㄚˊ	15	27	146
314	ㄨㄚˊ	16	27	31
315	ㄨㄚˊ	17	27	6
316	ㄨㄚˊ	18	27	50
317	ㄨㄚˊ	1	28	144

318	ㄨㄚˊ	2	28	0
319	ㄨㄚˊ	3	28	0
320	ㄨㄚˊ	4	28	13
321	ㄨㄚˊ	9	28	23
322	ㄨㄚˊ	10	28	35
323	ㄨㄚˊ	11	28	28
324	ㄨㄚˊ	1	29	696
325	ㄨㄚˊ	2	29	21
326	ㄨㄚˊ	3	29	9
327	ㄨㄚˊ	4	29	57
328	ㄨㄚˊ	5	29	9
329	ㄨㄚˊ	6	29	122
330	ㄨㄚˊ	7	29	16
331	ㄨㄚˊ	8	29	105
332	ㄨㄚˊ	9	29	50
333	ㄨㄚˊ	10	29	7
334	ㄨㄚˊ	11	29	455
335	ㄨㄚˊ	15	29	323
336	ㄨㄚˊ	16	29	42
337	ㄨㄚˊ	1	30	279
338	ㄨㄚˊ	2	30	51
339	ㄨㄚˊ	3	30	58
340	ㄨㄚˊ	4	30	0
341	ㄨㄚˊ	5	30	1
342	ㄨㄚˊ	6	30	3
343	ㄨㄚˊ	7	30	1
344	ㄨㄚˊ	8	30	15
345	ㄨㄚˊ	9	30	214
346	ㄨㄚˊ	10	30	10
347	ㄨㄚˊ	11	30	62
348	ㄨㄚˊ	15	30	52
349	ㄨㄚˊ	16	30	28
350	ㄨㄚˊ	17	30	4
351	ㄨㄚˊ	18	30	40
352	ㄨㄚˊ	1	31	237
353	ㄨㄚˊ	2	31	17
354	ㄨㄚˊ	3	31	25

355	尸メㄥ	4	31	10
356	回メㄥ	5	31	1
357	尸メㄥ	6	31	19
358	ㄥメㄥ	7	31	45
359	厶メㄥ	8	31	21
360	ㄥメㄥ	9	31	2
361	ㄥメㄥ	10	31	14
362	厶メㄥ	11	31	75
363	ㄥメㄥ	15	31	24
364	去メㄥ	16	31	5
365	ㄥメㄥ	18	31	52
366	メㄥ	1	32	147
367	虫メㄥ	2	32	28
368	彳メㄥ	3	32	21
369	尸メㄥ	4	32	26
370	ㄥメㄥ	9	32	84
371	ㄥメㄥ	10	32	24
372	厶メㄥ	11	32	30
373	メㄥ	1	33	3
374	虫メㄥ	2	33	658
375	彳メㄥ	3	33	75
376	回メㄥ	5	33	49
377	尸メㄥ	6	33	93
378	ㄥメㄥ	7	33	79
379	厶メㄥ	8	33	52
380	ㄥメㄥ	9	33	318
381	ㄥメㄥ	10	33	55
382	厶メㄥ	11	33	43
383	ㄥメㄥ	15	33	180
384	去メㄥ	16	33	294
385	ㄥメㄥ	17	33	15
386	ㄥメㄥ	18	33	46
387	ㄥせ	1	34	150
388	ㄥㄥせ	12	34	148
389	ㄥㄥせ	13	34	76
390	ㄥㄥせ	14	34	122
391	ㄥㄥせ	17	34	2

392	ㄥㄥせ	18	34	16
393	ㄥㄥ	1	35	355
394	ㄥㄥㄥ	12	35	10
395	ㄥㄥㄥ	13	35	214
396	ㄥㄥㄥ	14	35	64
397	ㄥㄥㄥ	18	35	0
398	ㄥㄥ	1	36	106
399	ㄥㄥㄥ	12	36	100
400	ㄥㄥㄥ	13	36	14
401	ㄥㄥㄥ	14	36	45
402	ㄥㄥㄥ	18	36	0
403	ㄥㄥ	1	37	106
404	ㄥㄥㄥ	12	37	4
405	ㄥㄥㄥ	13	37	3
406	ㄥㄥㄥ	14	37	30
407	ㄥ	1	38	414
408	一ㄥ	1	39	0
409	ㄥ	1	13	0
410	ㄥ	1	7	0
411	ㄥㄥ	21	4	6

附錄五

漢語 411 個基本音節碼與對應的注音符號、22 個聲母碼、40 個韻母碼和語料庫中每一種基本音節類型之音節音長的平均值對照表，其中的 NaN(Not a Number)為 MATLAB 中分母為零的表示方式，其意義為在語料庫中沒有此音節的語料。

音節碼	注音符號	聲母碼	韻母碼	值(秒)	音節碼	注音符號	聲母碼	韻母碼	值(秒)
1	ㄅ	2	1	0.1676	30	ㄇㄛ	21	3	0.1786
2	ㄆ	3	1	0.2437	31	ㄏㄛ	22	3	0.2203
3	ㄇ	4	1	0.2265	32	ㄛ	1	4	0.2140
4	ㄏ	5	1	0.1718	33	ㄅㄛ	2	4	0.1490
5	ㄆ	6	40	0.1819	34	ㄆㄛ	3	4	0.2581
6	ㄑ	7	40	0.2328	35	ㄇㄛ	4	4	0.2288
7	ㄌ	8	40	0.2332	36	ㄏㄛ	5	4	0.2185
8	ㄩ	1	2	0.1684	37	ㄆㄛ	6	4	0.2027
9	ㄅㄩ	2	2	0.2307	38	ㄑㄛ	7	4	0.2499
10	ㄆㄩ	3	2	0.2574	39	ㄌㄛ	8	4	0.2325
11	ㄇㄩ	4	2	0.2580	40	ㄍㄛ	9	4	0.1464
12	ㄏㄩ	6	2	0.2198	41	ㄎㄛ	10	4	0.1860
13	ㄑㄩ	7	2	0.3131	42	ㄏㄛ	11	4	0.2327
14	ㄌㄩ	8	2	0.2436	43	ㄉㄛ	15	4	0.1154
15	ㄍㄩ	9	2	0.1742	44	ㄊㄛ	16	4	0.1950
16	ㄎㄩ	10	2	0.2120	45	ㄋㄛ	17	4	0.1954
17	ㄏㄩ	11	2	0.2201	46	ㄌㄛ	18	4	0.1454
18	ㄉㄩ	15	2	0.1733	47	ㄛ	1	6	0.1895
19	ㄊㄩ	16	2	0.1702	48	ㄅㄛ	2	6	0.2327
20	ㄋㄩ	17	2	0.1613	49	ㄆㄛ	3	6	0.2741
21	ㄌㄩ	18	2	0.1977	50	ㄇㄛ	4	6	0.2950
22	ㄍㄩ	19	2	0.1629	51	ㄏㄛ	6	6	0.2117
23	ㄎㄩ	20	2	0.2400	52	ㄑㄛ	7	6	0.2394
24	ㄏㄩ	21	2	0.2028	53	ㄌㄛ	8	6	0.2590
25	ㄏㄩ	22	2	0.2070	54	ㄍㄛ	9	6	0.1822
26	ㄛ	1	3	0.3875	55	ㄎㄛ	10	6	0.2152
27	ㄎㄛ	18	3	NaN	56	ㄏㄛ	11	6	0.2372
28	ㄍㄛ	19	3	0.1668	57	ㄉㄛ	15	6	0.1702
29	ㄊㄛ	20	3	0.2105	58	ㄊㄛ	16	6	0.2277

59	ㄋㄝ	17	6	0.1914
60	ㄉㄝ	18	6	0.2063
61	ㄅㄝ	19	6	0.1699
62	ㄆㄝ	20	6	0.2444
63	ㄇㄝ	21	6	0.1808
64	ㄝ	1	5	NaN
65	ㄞㄟ	2	7	NaN
66	ㄟㄟ	4	7	0.3036
67	ㄟㄟ	6	7	0.2507
68	ㄟㄟ	8	7	NaN
69	ㄟㄟ	9	7	0.1460
70	ㄟㄟ	11	7	0.2220
71	ㄟㄟ	15	7	0.1472
72	ㄟㄟ	17	7	0.2037
73	ㄟㄟ	18	7	0.2026
74	ㄟㄟ	19	7	0.1586
75	ㄟㄟ	20	7	0.1924
76	ㄟㄟ	21	7	0.1698
77	ㄟㄟ	22	7	0.2112
78	ㄟ	1	8	0.1871
79	ㄟㄟ	2	8	0.2019
80	ㄟㄟ	3	8	0.2784
81	ㄟㄟ	4	8	0.2418
82	ㄟㄟ	5	8	0.1998
83	ㄟㄟ	6	8	0.2068
84	ㄟㄟ	7	8	0.2657
85	ㄟㄟ	8	8	0.2479
86	ㄟㄟ	9	8	0.1885
87	ㄟㄟ	10	8	0.2308
88	ㄟㄟ	11	8	0.2470
89	ㄟㄟ	15	8	0.1814
90	ㄟㄟ	16	8	0.2385
91	ㄟㄟ	17	8	0.2221
92	ㄟㄟ	18	8	0.2026
93	ㄟㄟ	19	8	0.1722
94	ㄟㄟ	20	8	0.2323
95	ㄟㄟ	21	8	0.2403

96	ㄟ	1	9	0.1758
97	ㄟㄟ	2	9	0.1994
98	ㄟㄟ	3	9	0.2710
99	ㄟㄟ	4	9	0.2510
100	ㄟㄟ	5	9	0.2099
101	ㄟㄟ	6	9	0.2018
102	ㄟㄟ	7	9	NaN
103	ㄟㄟ	8	9	0.2281
104	ㄟㄟ	9	9	0.1745
105	ㄟㄟ	10	9	0.2085
106	ㄟㄟ	11	9	0.2435
107	ㄟㄟ	15	9	0.1439
108	ㄟㄟ	16	9	0.2144
109	ㄟㄟ	17	9	NaN
110	ㄟㄟ	18	9	0.1880
111	ㄟㄟ	20	9	NaN
112	ㄟㄟ	21	9	0.2103
113	ㄟㄟ	22	9	0.2024
114	ㄟ	1	10	0.1927
115	ㄟㄟ	2	10	0.2107
116	ㄟㄟ	3	10	0.2571
117	ㄟㄟ	4	10	0.2712
118	ㄟㄟ	5	10	0.2106
119	ㄟㄟ	6	10	0.2071
120	ㄟㄟ	7	10	0.2461
121	ㄟㄟ	8	10	0.2455
122	ㄟㄟ	9	10	0.1884
123	ㄟㄟ	10	10	0.2237
124	ㄟㄟ	11	10	0.2733
125	ㄟㄟ	15	10	0.1706
126	ㄟㄟ	16	10	0.2455
127	ㄟㄟ	17	10	0.2066
128	ㄟㄟ	18	10	0.2161
129	ㄟㄟ	19	10	0.1771
130	ㄟㄟ	20	10	0.2286
131	ㄟㄟ	21	10	0.2086
132	ㄟㄟ	22	10	0.2269

133	ㄥ	1	11	0.1340
134	ㄷㄥ	2	11	0.1934
135	ㄸㄥ	3	11	0.2618
136	ㄹㄥ	4	11	0.2568
137	ㄱㄥ	5	11	0.2036
138	ㅋㄥ	6	11	0.1395
139	ㅌㄥ	7	11	NaN
140	ㄴㄥ	8	11	0.2338
141	ㄹㄥ	9	11	0.1762
142	ㅍㄥ	10	11	0.2189
143	ㅊㄥ	11	11	0.2252
144	ㅊㄥ	17	11	0.2228
145	ㅍㄥ	19	11	0.1577
146	ㅊㄥ	20	11	0.2158
147	ㅍㄥ	21	11	0.1398
148	ㅊㄥ	22	11	0.2069
149	ㄷ	1	12	0.2668
150	ㄷㄷ	2	12	0.2223
151	ㄷㄸ	3	12	0.2717
152	ㄷㄹ	4	12	0.2593
153	ㄷㄱ	5	12	0.2028
154	ㄷㅋ	6	12	0.2222
155	ㄷㅌ	7	12	0.3001
156	ㄷㄴ	8	12	0.2335
157	ㄷㄹ	9	12	0.2129
158	ㄷㅍ	10	12	0.2518
159	ㄷㅊ	11	12	0.3060
160	ㄷㅊ	15	12	0.1816
161	ㄷㅌ	16	12	0.2650
162	ㄷㅍ	17	12	NaN
163	ㄷㅊ	18	12	0.2250
164	ㄷㅍ	19	12	0.1929
165	ㄷㅊ	20	12	0.2533
166	ㄷㅍ	21	12	0.1823
167	ㄷㅊ	22	12	0.2282
168	ㄷㄷ	2	13	0.1941
169	ㄷㄸ	3	13	0.2500

170	ㄷㄷ	4	13	0.2439
171	ㄷㄸ	5	13	0.2076
172	ㄷㄹ	6	13	0.1997
173	ㄷㄱ	7	13	0.2511
174	ㄷㅋ	8	13	0.3074
175	ㄷㄹ	9	13	0.1843
176	ㄷㅍ	10	13	0.2555
177	ㄷㅊ	11	13	0.2528
178	ㄷㅊ	15	13	0.1610
179	ㄷㅌ	16	13	0.2094
180	ㄷㅍ	17	13	0.1850
181	ㄷㅊ	18	13	0.2356
182	ㄷㅍ	19	13	0.1585
183	ㄷㅊ	20	13	0.2271
184	ㄷㅍ	21	13	0.2367
185	ㄷㅊ	22	13	0.2347
186	ㄷ	1	14	0.1447
187	ㄷㄷ	12	14	0.1860
188	ㄷㄸ	13	14	0.2300
189	ㄷㄹ	14	14	0.2359
190	ㄷㄱ	15	14	0.1498
191	ㄷㅋ	16	14	0.2104
192	ㄷㅌ	17	14	0.1463
193	ㄷㄴ	18	14	0.1585
194	ㄷㄹ	19	14	0.1394
195	ㄷㅍ	20	14	0.1881
196	ㄷㅊ	21	14	0.1955
197	ㄷㅊ	1	15	0.1617
198	ㄷㅌ	2	15	0.1793
199	ㄷㄸ	3	15	0.2223
200	ㄷㄹ	4	15	0.2423
201	ㄷㄱ	5	15	0.1832
202	ㄷㅋ	6	15	0.2265
203	ㄷㅌ	7	15	0.2572
204	ㄷㄴ	8	15	0.2466
205	ㄷㄹ	9	15	0.1641
206	ㄷㅍ	10	15	0.2010

207	厂メ	11	15	0.2096
208	夕メ	15	15	0.1603
209	去メ	16	15	0.2145
210	弓メ	17	15	0.1648
211	夕メ	18	15	0.1667
212	夕メ	19	15	0.1092
213	夕メ	20	15	0.1751
214	冂メ	21	15	0.1804
215	匚メ	22	15	0.2023
216	口	1	16	0.1627
217	凵口	12	16	0.2035
218	く口	13	16	0.2301
219	冂口	14	16	0.2152
220	弓口	17	16	0.1677
221	夕口	18	16	0.1617
222	一冫	1	17	0.2160
223	凵一冫	12	17	0.2341
224	く一冫	13	17	0.2890
225	冂一冫	14	17	0.2620
226	夕一冫	18	17	NaN
227	一卩	1	18	0.1556
228	凵一卩	12	18	0.2035
229	く一卩	13	18	0.2433
230	冂一卩	14	18	0.2478
231	夕一卩	15	18	0.2269
232	去一卩	16	18	0.2338
233	弓一卩	17	18	0.1689
234	夕一卩	18	18	0.1815
235	夕一卩	19	18	0.1743
236	夕一卩	20	18	NaN
237	冂一卩	21	18	0.1896
238	一冫	1	19	NaN
239	一纟	1	20	0.1839
240	凵一纟	12	20	0.2164
241	く一纟	13	20	0.2847
242	冂一纟	14	20	0.2433
243	夕一纟	15	20	0.1941

244	去一纟	16	20	0.2660
245	弓一纟	17	20	0.2184
246	夕一纟	18	20	0.2275
247	夕一纟	19	20	0.1707
248	夕一纟	20	20	0.2249
249	冂一纟	21	20	0.2423
250	一冫	1	21	0.1791
251	凵一冫	12	21	0.1833
252	く一冫	13	21	0.2512
253	冂一冫	14	21	0.2579
254	夕一冫	15	21	0.1647
255	弓一冫	17	21	0.2143
256	夕一冫	18	21	0.2002
257	冂一冫	21	21	0.2100
258	一冫	1	22	0.2264
259	凵一冫	12	22	0.2279
260	く一冫	13	22	0.2921
261	冂一冫	14	22	0.2534
262	夕一冫	15	22	0.1874
263	去一冫	16	22	0.2422
264	弓一冫	17	22	0.2272
265	夕一冫	18	22	0.2176
266	夕一冫	19	22	0.1890
267	夕一冫	20	22	0.2395
268	冂一冫	21	22	0.2096
269	一冫	1	23	0.1742
270	凵一冫	12	23	0.2058
271	く一冫	13	23	0.2595
272	冂一冫	14	23	0.2482
273	夕一冫	17	23	0.1863
274	夕一冫	18	23	0.1971
275	夕一冫	19	23	0.2109
276	夕一冫	20	23	0.2064
277	冂一冫	21	23	0.1996
278	一冫	1	24	0.2043
279	凵一冫	12	24	0.2317
280	く一冫	13	24	0.3000

281	下一九	14	24	0.2610
282	弓一九	17	24	0.2404
283	カ一九	18	24	0.2244
284	一ム	1	25	0.2050
285	リ一ム	12	25	0.2118
286	く一ム	13	25	0.2638
287	下一ム	14	25	0.2498
288	夕一ム	15	25	0.1761
289	去一ム	16	25	0.2406
290	弓一ム	17	25	0.1990
291	カ一ム	18	25	0.1983
292	ク一ム	19	25	0.1692
293	夕一ム	20	25	0.2466
294	門一ム	21	25	0.2030
295	メヱ	1	26	0.2063
296	虫メヱ	2	26	0.2384
297	イメヱ	3	26	NaN
298	尸メヱ	4	26	NaN
299	《メヱ	9	26	0.2742
300	弓メヱ	10	26	0.2791
301	厂メヱ	11	26	0.2411
302	メヱ	1	27	0.1230
303	虫メヱ	2	27	0.2097
304	イメヱ	3	27	NaN
305	尸メヱ	4	27	0.2919
306	回メヱ	5	27	0.1889
307	尸メヱ	6	27	0.2110
308	チメヱ	7	27	0.2721
309	ムメヱ	8	27	0.2307
310	《メヱ	9	27	0.1659
311	弓メヱ	10	27	0.2381
312	厂メヱ	11	27	0.2049
313	夕メヱ	15	27	0.1666
314	去メヱ	16	27	0.2211
315	弓メヱ	17	27	0.1937
316	カメヱ	18	27	0.1930
317	メヱ	1	28	0.1930

318	虫メヱ	2	28	NaN
319	イメヱ	3	28	NaN
320	尸メヱ	4	28	0.2985
321	《メヱ	9	28	0.1990
322	弓メヱ	10	28	0.2447
323	厂メヱ	11	28	0.2284
324	メヱ	1	29	0.1875
325	虫メヱ	2	29	0.1963
326	イメヱ	3	29	0.2479
327	尸メヱ	4	29	0.2512
328	回メヱ	5	29	0.1912
329	尸メヱ	6	29	0.2055
330	チメヱ	7	29	0.2363
331	ムメヱ	8	29	0.2419
332	《メヱ	9	29	0.1745
333	弓メヱ	10	29	0.2266
334	厂メヱ	11	29	0.2004
335	夕メヱ	15	29	0.1648
336	去メヱ	16	29	0.1927
337	メヱ	1	30	0.1932
338	虫メヱ	2	30	0.2156
339	イメヱ	3	30	0.2951
340	尸メヱ	4	30	NaN
341	回メヱ	5	30	0.2928
342	尸メヱ	6	30	0.2669
343	チメヱ	7	30	0.3634
344	ムメヱ	8	30	0.2831
345	《メヱ	9	30	0.2041
346	弓メヱ	10	30	0.2346
347	厂メヱ	11	30	0.2431
348	夕メヱ	15	30	0.2074
349	去メヱ	16	30	0.2804
350	弓メヱ	17	30	0.2350
351	カメヱ	18	30	0.2077
352	メヱ	1	31	0.1977
353	虫メヱ	2	31	0.1941
354	イメヱ	3	31	0.2728

355	尸メㄥ	4	31	0.2946
356	冂メㄥ	5	31	0.1600
357	尸メㄥ	6	31	0.2486
358	ㄥメㄥ	7	31	0.2723
359	厶メㄥ	8	31	0.2611
360	ㄥメㄥ	9	31	0.1490
361	ㄥメㄥ	10	31	0.2372
362	厂メㄥ	11	31	0.2121
363	ㄥメㄥ	15	31	0.1837
364	去メㄥ	16	31	0.2650
365	ㄥメㄥ	18	31	0.2133
366	メㄥ	1	32	0.1995
367	虫メㄥ	2	32	0.2382
368	彳メㄥ	3	32	0.2850
369	尸メㄥ	4	32	0.2694
370	ㄥメㄥ	9	32	0.2243
371	ㄥメㄥ	10	32	0.2568
372	厂メㄥ	11	32	0.2609
373	メㄥ	1	33	0.1810
374	虫メㄥ	2	33	0.2021
375	彳メㄥ	3	33	0.2402
376	冂メㄥ	5	33	0.2238
377	尸メㄥ	6	33	0.2104
378	ㄥメㄥ	7	33	0.2651
379	厶メㄥ	8	33	0.2646
380	ㄥメㄥ	9	33	0.1793
381	ㄥメㄥ	10	33	0.2157
382	厂メㄥ	11	33	0.2511
383	ㄥメㄥ	15	33	0.1716
384	去メㄥ	16	33	0.2219
385	彳メㄥ	17	33	0.2051
386	ㄥメㄥ	18	33	0.2475
387	ㄥせ	1	34	0.1895
388	ㄥㄥせ	12	34	0.2223
389	ㄥㄥせ	13	34	0.2183
390	ㄥㄥせ	14	34	0.2360
391	ㄥㄥせ	17	34	0.2259

392	ㄥㄥせ	18	34	0.1787
393	ㄥㄥ	1	35	0.2123
394	ㄥㄥㄥ	12	35	0.2437
395	ㄥㄥㄥ	13	35	0.2929
396	ㄥㄥㄥ	14	35	0.2672
397	ㄥㄥㄥ	18	35	NaN
398	ㄥㄥ	1	36	0.1939
399	ㄥㄥㄥ	12	36	0.2180
400	ㄥㄥㄥ	13	36	0.2752
401	ㄥㄥㄥ	14	36	0.2530
402	ㄥㄥㄥ	18	36	NaN
403	ㄥㄥ	1	37	0.2013
404	ㄥㄥㄥ	12	37	0.2292
405	ㄥㄥㄥ	13	37	0.3222
406	ㄥㄥㄥ	14	37	0.2859
407	ㄥ	1	38	0.1595
408	一ㄥ	1	39	NaN
409	ㄥ	1	13	NaN
410	ㄥ	1	7	NaN
411	ㄥㄥ	21	4	0.1057

附錄六

漢語 411 個基本音節碼與對應的注音符號、22 個聲母碼、40 個韻母碼和 γ_{bs} 的值對照表，其中的 NaN(Not a Number)為 MATLAB 中分母為零的表示方式，其意義為在語料庫中沒有此音節的語料。

音節碼	注音符號	聲母碼	韻母碼	值(秒)
1	ㄅ	2	1	-0.0342
2	ㄆ	3	1	0.0293
3	ㄇ	4	1	0.0220
4	ㄏ	5	1	-0.0259
5	ㄏ	6	40	0.0059
6	ㄏ	7	40	0.0426
7	ㄏ	8	40	0.0240
8	ㄩ	1	2	-0.0301
9	ㄅㄩ	2	2	0.0269
10	ㄆㄩ	3	2	0.0425
11	ㄇㄩ	4	2	0.0514
12	ㄏㄩ	6	2	0.0107
13	ㄏㄩ	7	2	0.1015
14	ㄏㄩ	8	2	0.0470
15	ㄏㄩ	9	2	-0.0250
16	ㄏㄩ	10	2	0.0145
17	ㄏㄩ	11	2	0.0097
18	ㄏㄩ	15	2	-0.0248
19	ㄏㄩ	16	2	-0.0405
20	ㄏㄩ	17	2	-0.0349
21	ㄏㄩ	18	2	-0.0117
22	ㄏㄩ	19	2	-0.0350
23	ㄏㄩ	20	2	0.0385
24	ㄏㄩ	21	2	0.0190
25	ㄏㄩ	22	2	0.0108
26	ㄅ	1	3	0.1762
27	ㄆ	18	3	NaN
28	ㄇ	19	3	-0.0486
29	ㄏ	20	3	0.0007

音節碼	注音符號	聲母碼	韻母碼	值(秒)
30	ㄏㄩ	21	3	0.0069
31	ㄏㄩ	22	3	0.0047
32	ㄅ	1	4	0.0095
33	ㄅ	2	4	-0.0361
34	ㄆ	3	4	0.0562
35	ㄇ	4	4	0.0346
36	ㄏ	5	4	0.0244
37	ㄏ	6	4	-0.0098
38	ㄏ	7	4	0.0511
39	ㄏ	8	4	0.0341
40	ㄏ	9	4	-0.0239
41	ㄏ	10	4	-0.0116
42	ㄏ	11	4	0.0236
43	ㄏ	15	4	-0.0111
44	ㄏ	16	4	-0.0034
45	ㄏ	17	4	0.0390
46	ㄏ	18	4	0.0111
47	ㄆ	1	6	-0.0091
48	ㄅ	2	6	0.0335
49	ㄆ	3	6	0.0608
50	ㄇ	4	6	0.0903
51	ㄏ	6	6	0.0163
52	ㄏ	7	6	0.0344
53	ㄏ	8	6	0.0538
54	ㄏ	9	6	-0.0210
55	ㄏ	10	6	0.0066
56	ㄏ	11	6	0.0274
57	ㄏ	15	6	-0.0261
58	ㄏ	16	6	0.0155

59	ㄋㄝ	17	6	-0.0032
60	ㄉㄝ	18	6	-0.0095
61	ㄅㄝ	19	6	-0.0254
62	ㄆㄝ	20	6	0.0427
63	ㄇㄝ	21	6	-0.0131
64	ㄟ	1	5	NaN
65	ㄗㄨ	2	7	NaN
66	ㄗㄨ	4	7	0.0931
67	ㄗㄨ	6	7	0.0342
68	ㄇㄨ	8	7	NaN
69	ㄍㄨ	9	7	-0.0444
70	ㄉㄨ	11	7	0.0119
71	ㄉㄨ	15	7	-0.0339
72	ㄋㄨ	17	7	0.0082
73	ㄉㄨ	18	7	0.0065
74	ㄅㄨ	19	7	-0.0393
75	ㄆㄨ	20	7	-0.0096
76	ㄇㄨ	21	7	-0.0280
77	ㄇㄨ	22	7	0.0033
78	ㄟ	1	8	-0.0123
79	ㄗㄨ	2	8	0.0028
80	ㄗㄨ	3	8	0.0684
81	ㄗㄨ	4	8	0.0468
82	ㄇㄨ	5	8	0.0002
83	ㄗㄨ	6	8	0.0087
84	ㄗㄨ	7	8	0.0663
85	ㄇㄨ	8	8	0.0508
86	ㄍㄨ	9	8	-0.0236
87	ㄗㄨ	10	8	0.0359
88	ㄉㄨ	11	8	0.0509
89	ㄉㄨ	15	8	-0.0154
90	ㄉㄨ	16	8	0.0352
91	ㄋㄨ	17	8	0.0396
92	ㄉㄨ	18	8	0.0099
93	ㄅㄨ	19	8	-0.0263
94	ㄆㄨ	20	8	0.0250
95	ㄇㄨ	21	8	0.0346

96	ㄨ	1	9	-0.0255
97	ㄗㄨ	2	9	-0.0112
98	ㄗㄨ	3	9	0.0613
99	ㄗㄨ	4	9	0.0539
100	ㄇㄨ	5	9	0.0017
101	ㄗㄨ	6	9	0.0114
102	ㄗㄨ	7	9	NaN
103	ㄇㄨ	8	9	0.0271
104	ㄍㄨ	9	9	-0.0249
105	ㄗㄨ	10	9	0.0158
106	ㄉㄨ	11	9	0.0459
107	ㄉㄨ	15	9	-0.0669
108	ㄉㄨ	16	9	0.0071
109	ㄋㄨ	17	9	NaN
110	ㄉㄨ	18	9	-0.0190
111	ㄆㄨ	20	9	NaN
112	ㄇㄨ	21	9	0.0038
113	ㄇㄨ	22	9	0.0122
114	ㄗ	1	10	-0.0124
115	ㄗㄨ	2	10	0.0143
116	ㄗㄨ	3	10	0.0854
117	ㄗㄨ	4	10	0.0658
118	ㄇㄨ	5	10	0.0013
119	ㄗㄨ	6	10	0.0020
120	ㄗㄨ	7	10	0.0357
121	ㄇㄨ	8	10	0.0353
122	ㄍㄨ	9	10	-0.0010
123	ㄗㄨ	10	10	0.0257
124	ㄉㄨ	11	10	0.0723
125	ㄉㄨ	15	10	-0.0272
126	ㄉㄨ	16	10	0.0382
127	ㄋㄨ	17	10	-0.0076
128	ㄉㄨ	18	10	0.0083
129	ㄅㄨ	19	10	-0.0256
130	ㄆㄨ	20	10	0.0310
131	ㄇㄨ	21	10	0.0138
132	ㄇㄨ	22	10	0.0316

133	ㇿ	1	11	-0.0764
134	𪗇	2	11	-0.0148
135	𪗈	3	11	0.0565
136	𪗉	4	11	0.0515
137	𪗊	5	11	-0.0075
138	𪗋	6	11	-0.0514
139	𪗌	7	11	NaN
140	𪗍	8	11	0.0228
141	𪗎	9	11	-0.0384
142	𪗏	10	11	0.0308
143	𪗐	11	11	0.0293
144	𪗑	17	11	0.0244
145	𪗒	19	11	-0.0319
146	𪗓	20	11	0.0031
147	𪗔	21	11	0.0061
148	𪗕	22	11	0.0013
149	𪗖	1	12	0.0686
150	𪗗	2	12	0.0196
151	𪗘	3	12	0.0634
152	𪗙	4	12	0.0572
153	𪗚	5	12	0.0032
154	𪗛	6	12	0.0137
155	𪗜	7	12	0.0864
156	𪗝	8	12	0.0335
157	𪗞	9	12	0.0141
158	𪗟	10	12	0.0470
159	𪗠	11	12	0.0845
160	𪗡	15	12	-0.0176
161	𪗢	16	12	0.0478
162	𪗣	17	12	NaN
163	𪗤	18	12	0.0256
164	𪗥	19	12	-0.0174
165	𪗦	20	12	0.0430
166	𪗧	21	12	-0.0322
167	𪗨	22	12	0.0224
168	𪗩	2	13	-0.0049
169	𪗪	3	13	0.0346

170	𪗫	4	13	0.0355
171	𪗬	5	13	-0.0029
172	𪗭	6	13	-0.0134
173	𪗮	7	13	0.0388
174	𪗯	8	13	0.0953
175	𪗰	9	13	-0.0116
176	𪗱	10	13	0.0451
177	𪗲	11	13	0.0404
178	𪗳	15	13	-0.0339
179	𪗴	16	13	0.0008
180	𪗵	17	13	-0.0290
181	𪗶	18	13	0.0442
182	𪗷	19	13	-0.0527
183	𪗸	20	13	0.0165
184	𪗹	21	13	0.0252
185	𪗺	22	13	0.0228
186	𪗻	1	14	-0.0518
187	𪗼	12	14	-0.0205
188	𪗽	13	14	0.0302
189	𪗾	14	14	0.0320
190	𪗿	15	14	-0.0445
191	𪘀	16	14	0.0114
192	𪘁	17	14	-0.0478
193	𪘂	18	14	-0.0364
194	𪘃	19	14	-0.0573
195	𪘄	20	14	-0.0198
196	𪘅	21	14	-0.0055
197	𪘆	1	15	-0.0429
198	𪘇	2	15	-0.0092
199	𪘈	3	15	0.0074
200	𪘉	4	15	0.0394
201	𪘊	5	15	-0.0264
202	𪘋	6	15	0.0222
203	𪘌	7	15	0.0571
204	𪘍	8	15	0.0409
205	𪘎	9	15	-0.0231
206	𪘏	10	15	0.0124

207	厂メ	11	15	0.0048
208	夕メ	15	15	-0.0493
209	去メ	16	15	0.0118
210	ろメ	17	15	-0.0288
211	カメ	18	15	-0.0350
212	ケメ	19	15	-0.0940
213	クメ	20	15	-0.0187
214	コメ	21	15	-0.0138
215	クメ	22	15	0.0064
216	口	1	16	-0.0404
217	口	12	16	0.0044
218	く口	13	16	0.0354
219	一口	14	16	0.0127
220	ろ口	17	16	-0.0250
221	カ口	18	16	-0.0340
222	一Y	1	17	0.0212
223	口一Y	12	17	0.0246
224	く一Y	13	17	0.0845
225	一口Y	14	17	0.0629
226	カ一Y	18	17	NaN
227	一セ	1	18	-0.0334
228	口一セ	12	18	0.0001
229	く一セ	13	18	0.0453
230	一口セ	14	18	0.0416
231	夕一セ	15	18	0.0120
232	去一セ	16	18	0.0426
233	ろ一セ	17	18	-0.0424
234	カ一セ	18	18	-0.0164
235	ケ一セ	19	18	-0.0362
236	ク一セ	20	18	NaN
237	コ一セ	21	18	-0.0057
238	一カ	1	19	NaN
239	一カ	1	20	-0.0137
240	口一カ	12	20	0.0169
241	く一カ	13	20	0.0777
242	一口カ	14	20	0.0465
243	夕一カ	15	20	-0.0046

244	去一カ	16	20	0.0538
245	ろ一カ	17	20	0.0202
246	カ一カ	18	20	0.0228
247	ケ一カ	19	20	-0.0251
248	ク一カ	20	20	0.0261
249	コ一カ	21	20	0.0424
250	一カ	1	21	-0.0142
251	口一カ	12	21	-0.0152
252	く一カ	13	21	0.0354
253	一口カ	14	21	0.0510
254	夕一カ	15	21	-0.0472
255	ろ一カ	17	21	0.0164
256	カ一カ	18	21	0.0043
257	コ一カ	21	21	0.0119
258	一カ	1	22	0.0273
259	口一カ	12	22	0.0203
260	く一カ	13	22	0.0790
261	一口カ	14	22	0.0540
262	夕一カ	15	22	-0.0089
263	去一カ	16	22	0.0311
264	ろ一カ	17	22	0.0156
265	カ一カ	18	22	0.0060
266	ケ一カ	19	22	-0.0135
267	ク一カ	20	22	0.0369
268	コ一カ	21	22	0.0104
269	一カ	1	23	-0.0350
270	口一カ	12	23	0.0058
271	く一カ	13	23	0.0439
272	一口カ	14	23	0.0378
273	ろ一カ	17	23	-0.0259
274	カ一カ	18	23	-0.0201
275	ケ一カ	19	23	-0.0006
276	ク一カ	20	23	0.0147
277	コ一カ	21	23	-0.0160
278	一カ	1	24	0.0046
279	口一カ	12	24	0.0178
280	く一カ	13	24	0.0856

281	丁一尢	14	24	0.0647
282	ㄋ一尢	17	24	0.0355
283	ㄎ一尢	18	24	0.0273
284	一厶	1	25	0.0042
285	ㄩ一厶	12	25	0.0104
286	く一厶	13	25	0.0549
287	丁一厶	14	25	0.0430
288	ㄎ一厶	15	25	-0.0252
289	去一厶	16	25	0.0313
290	ㄋ一厶	17	25	-0.0138
291	ㄎ一厶	18	25	-0.0047
292	ㄎ一厶	19	25	-0.0306
293	ㄎ一厶	20	25	0.0326
294	冂一厶	21	25	-0.0061
295	メヱ	1	26	0.0052
296	虫メヱ	2	26	0.0283
297	イメヱ	3	26	NaN
298	尸メヱ	4	26	NaN
299	ㄥメヱ	9	26	0.0650
300	ㄎメヱ	10	26	0.0848
301	厂メヱ	11	26	0.0320
302	メㄥ	1	27	-0.0613
303	虫メㄥ	2	27	-0.0025
304	イメㄥ	3	27	NaN
305	尸メㄥ	4	27	0.0771
306	冂メㄥ	5	27	-0.0102
307	冂メㄥ	6	27	0.0054
308	ㄎメㄥ	7	27	0.0727
309	厶メㄥ	8	27	0.0366
310	ㄥメㄥ	9	27	-0.0438
311	ㄎメㄥ	10	27	0.0404
312	厂メㄥ	11	27	0.0026
313	ㄎメㄥ	15	27	-0.0407
314	去メㄥ	16	27	0.0159
315	ㄋメㄥ	17	27	-0.0102
316	ㄎメㄥ	18	27	-0.0145
317	メㄥ	1	28	-0.0002

318	虫メㄥ	2	28	NaN
319	イメㄥ	3	28	NaN
320	尸メㄥ	4	28	0.0988
321	ㄥメㄥ	9	28	-0.0005
322	ㄎメㄥ	10	28	0.0440
323	厂メㄥ	11	28	0.0192
324	メㄥ	1	29	-0.0154
325	虫メㄥ	2	29	-0.0156
326	イメㄥ	3	29	0.0432
327	尸メㄥ	4	29	0.0595
328	冂メㄥ	5	29	-0.0069
329	冂メㄥ	6	29	0.0122
330	ㄎメㄥ	7	29	0.0287
331	厶メㄥ	8	29	0.0333
332	ㄥメㄥ	9	29	-0.0313
333	ㄎメㄥ	10	29	0.0219
334	厂メㄥ	11	29	0.0027
335	ㄎメㄥ	15	29	-0.0334
336	去メㄥ	16	29	-0.0175
337	メㄥ	1	30	-0.0159
338	虫メㄥ	2	30	0.0117
339	イメㄥ	3	30	0.0774
340	尸メㄥ	4	30	NaN
341	冂メㄥ	5	30	0.1028
342	冂メㄥ	6	30	0.0639
343	ㄎメㄥ	7	30	0.1653
344	厶メㄥ	8	30	0.0781
345	ㄥメㄥ	9	30	-0.0045
346	ㄎメㄥ	10	30	0.0359
347	厂メㄥ	11	30	0.0374
348	ㄎメㄥ	15	30	0.0081
349	去メㄥ	16	30	0.0646
350	ㄋメㄥ	17	30	0.0463
351	ㄎメㄥ	18	30	0.0098
352	メㄥ	1	31	-0.0003
353	虫メㄥ	2	31	-0.0010
354	イメㄥ	3	31	0.0665

355	尸メㄥ	4	31	0.0963
356	回メㄥ	5	31	-0.0381
357	尸メㄥ	6	31	0.0379
358	ㄥメㄥ	7	31	0.0583
359	厶メㄥ	8	31	0.0585
360	ㄥメㄥ	9	31	-0.0417
361	ㄥメㄥ	10	31	0.0337
362	厶メㄥ	11	31	0.0035
363	ㄥメㄥ	15	31	-0.0217
364	去メㄥ	16	31	0.0630
365	ㄥメㄥ	18	31	0.0022
366	メㄥ	1	32	-0.0001
367	虫メㄥ	2	32	0.0418
368	彳メㄥ	3	32	0.0868
369	尸メㄥ	4	32	0.0647
370	ㄥメㄥ	9	32	0.0193
371	ㄥメㄥ	10	32	0.0558
372	厶メㄥ	11	32	0.0438
373	メㄥ	1	33	-0.0308
374	虫メㄥ	2	33	-0.0019
375	彳メㄥ	3	33	0.0259
376	回メㄥ	5	33	0.0093
377	尸メㄥ	6	33	0.0119
378	ㄥメㄥ	7	33	0.0502
379	厶メㄥ	8	33	0.0658
380	ㄥメㄥ	9	33	-0.0257
381	ㄥメㄥ	10	33	0.0067
382	厶メㄥ	11	33	0.0350
383	ㄥメㄥ	15	33	-0.0284
384	去メㄥ	16	33	0.0210
385	ㄥメㄥ	17	33	-0.0064
386	ㄥメㄥ	18	33	0.0341
387	ㄥせ	1	34	-0.0094
388	ㄥㄥせ	12	34	0.0167
389	ㄥㄥせ	13	34	0.0190
390	ㄥㄥせ	14	34	0.0239
391	ㄥㄥせ	17	34	0.0276

392	ㄥㄥせ	18	34	-0.0195
393	ㄥㄥ	1	35	0.0073
394	ㄥㄥㄥ	12	35	0.0497
395	ㄥㄥㄥ	13	35	0.0846
396	ㄥㄥㄥ	14	35	0.0636
397	ㄥㄥㄥ	18	35	NaN
398	ㄥㄥ	1	36	-0.0068
399	ㄥㄥㄥ	12	36	0.0089
400	ㄥㄥㄥ	13	36	0.0605
401	ㄥㄥㄥ	14	36	0.0394
402	ㄥㄥㄥ	18	36	NaN
403	ㄥㄥ	1	37	0.0086
404	ㄥㄥㄥ	12	37	0.0385
405	ㄥㄥㄥ	13	37	0.1123
406	ㄥㄥㄥ	14	37	0.0797
407	ㄥ	1	38	-0.0463
408	一ㄥ	1	39	NaN
409	ㄥ	1	13	NaN
410	ㄥ	1	7	NaN
411	ㄥㄥ	21	4	-0.0154

附錄七

漢語 22 個聲母的注音符號與漢語拼音對照表，其中 INULL 為 Initial Null 的縮寫，其所代表的意義為空聲母亦即在音節中沒有聲母只有韻母。

注音符號	漢語拼音
ㄨ	INULL
ㄗ	zh
ㄘ	ch
ㄙ	sh
ㄨㄛ	r
ㄗ	z
ㄘ	c
ㄙ	s
ㄍ	g
ㄎ	k
ㄏ	h
ㄐ	j
ㄑ	q
ㄒ	x
ㄉ	d
ㄊ	t
ㄋ	n
ㄌ	l
ㄅ	b
ㄆ	p
ㄇ	m
ㄈ	f



附錄八

漢語 40 個韻母的注音符號與漢語拼音對照表，其中 FNULL 為 Final Null 的縮寫，其所代表的意義為空韻母亦即在音節中只有聲母沒有韻母。

注音符號	漢語拼音	注音符號	漢語拼音
Φ1	FNULL1	ㄨㄢ	wan
ㄚ	a	ㄨㄣ	wen
ㄛ	o	ㄨㄤ	wang
ㄜ	e	ㄨㄥ	weng
ㄝ	eh	ㄩㄝ	yue
ㄟ	ai	ㄩㄢ	yuan
ㄟ	ei	ㄩㄣ	yun
ㄠ	ao	ㄩㄥ	yung
ㄡ	ou	ㄦ	er
ㄢ	an	ㄩㄛ	yo
ㄣ	en	Φ2	FNULL2
ㄤ	ang		
ㄥ	eng		
ㄩ	yi		
ㄨ	wu		
ㄩ	yu		
ㄩㄚ	ya		
ㄩㄝ	ye		
ㄩㄞ	yai		
ㄩㄠ	yao		
ㄩㄡ	you		
ㄩㄢ	yan		
ㄩㄣ	yin		
ㄩㄤ	yang		
ㄩㄥ	ying		
ㄨㄚ	wa		
ㄨㄛ	wo		
ㄨㄞ	wai		
ㄨㄟ	wei		