

以Fujisaki模型驗證連續語流中字調及韻律詞對應於階層性韻律架構HPG的意義

鄭秋豫 蘇昭宇

中央研究院語言所語音實驗室

cytling@sinica.edu.tw, morison@gate.sinica.edu.tw

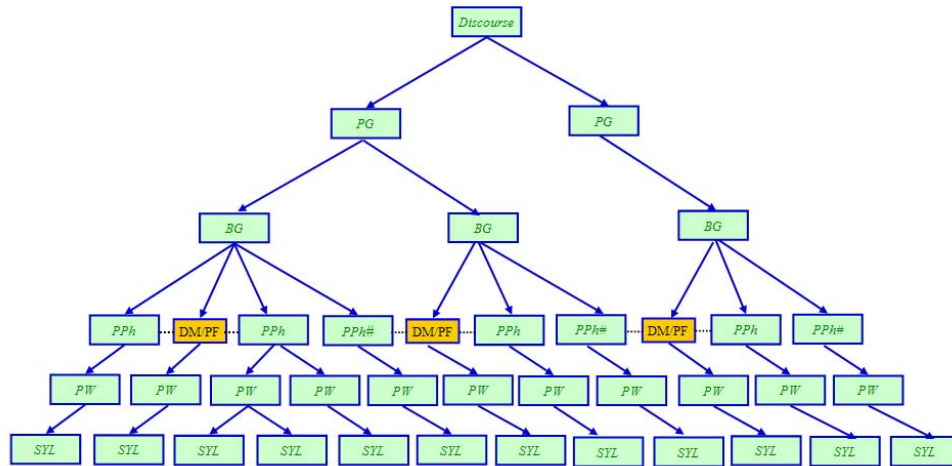
摘要

本文從台灣地區國語連續語流的字調及韻律詞基頻曲線模型，根據鄭秋豫所提出的「階層式韻律句群HPG架構」，由下層到上層，將基頻曲線模型參數與HPG韻律階層結合與驗證，探討（1.）句調成份與字調成份對應各韻律階層的貢獻度 與（2.）字調成份在各階層韻律單位管轄下，基頻模型參數如何變化。結果顯示（1.）HPG架構中各韻律層由上而下管轄制約，下層韻律單位必須承上層韻律訊息進行系統性的調整。（2.）字調接受韻律詞的上層制約，字調層及韻律詞層對基頻輸出均有貢獻，韻律詞與字調的基頻關係，不等於字調線性串接及平滑；表意語境所造成的制約，才是語流韻律的主要特徵。

關鍵詞：韻律詞，基頻曲線模型，階層式韻律句群HPG架構，句調成份，字調成份，表意韻律

一、緒論

國語連續語流韻律一向被視為充滿變異性且難以預測，本文主旨在討論韻律語境來自階層式管轄制約，韻律語境同時包括線性串接平滑及上層語篇的跨短語表意資訊，因此串接平滑不足以解釋語流中字調的變化。鄭秋豫由語段與語篇的角度切入，發現表面看似複雜的連續語流韻律，事實上有系統性規則可循，這些規則及階層式的關係，與串接共構表意韻律語境，並在物理信號上表現出特定基型，因此語者與聽者依據此基型產製與接收來自語篇的大範圍表意韻律訊息，結合區辨詞義的字調、區辨句法訊息的句調共同達到溝通的目的。鄭秋豫[1][2][3]於2004提出階層式語流韻律架HPG (Hierarchical Prosodic Phrase Grouping) 指出，從聲學語音訊息而言，口語連續語流韻律的多短語階層架構是以感知為基礎[2][3]，感知的最大成分是聽者預期，該架構主要精神在於將國語口語語流的韻律單位，定義為多短語韻律短語組PG (Multi-phrase Prosodic Group) 而非單一短語 (phrase)，表達上層語意訊息的連貫性，構成韻律短語組的相鄰及跨短語的語段韻律語境，從大範圍韻律單位表示特定語意段落的開始，延續與結束。因此此跨短語語段的韻律語境的基型，即為語者溝通時語言即時產製與接收處理的模版，當短語形成句段時，各短語必須受上層語篇語意資訊管轄制約而調整，呈現表意韻律語境，才能成為語意完整的語段。因此傳統語音信號分析或串接平滑皆無法解釋的句調變異，套用HPG架構後，其實可從韻律語境結構的角度得到解釋。我們先前也已提出基頻曲線、音節時長、能量分佈和停頓時長對應HPG架構的證據[1][2][3]。



圖一「階層式多短語韻律句群HPG」韻律單位架構圖。由下而上分別為音節(SYL)、韻律詞(PW)、韻律短語(PPh)、呼吸組(BG)、韻律句組(PG)及語篇(Discourse)。(Tseng 2006)。

以HPG韻律架構分析語流韻律的取向和傳統韻律分析最大的不同，在傳統韻律分析通常只分析單字字調和短語句調，卻忽略了連續語流裡，除了受詞義制約規範的字調和受句法結構制約規範的句調以外，還含有上層語篇語意造成的句調連貫訊息，表達語段訊息。亦即韻律語境構成句調以上語流全面性的連接性與連貫性，以致韻律短句間產生韻律關連性，各下級韻律單位必須依照更語篇韻律的上層訊息進行系統性的調整，提供語段從何處開始、維持到結束的韻律語境訊息；表意韻律語境不僅是相鄰韻律單位的關係，各語段下轄之各層次級韻律單位，均需依此規範做系統性的調整，同時表現相鄰單位及跨單位的韻律關係。鄭秋豫等稍早的韻律短語PPh基頻分析研究，已提出相鄰PPh及跨PPh調整的證據[4]。在HPG架構中，字調以上、韻律短語以下的韻律詞PW也是一個比詞彙詞略大的韻律單位，詞彙詞約佔韻律詞的80%，構成韻律詞意範圍內的音節，同樣有著關聯性與連結性，也必須依據多音節韻律詞的詞意範圍對音節字調進行相鄰音節及跨音節的調整，使聽者不受音節個數限制，輕易的將韻律詞正確的歸類，並據此判斷是否為另一個新韻律詞的開始。因此，由於HPG架構的提出，傳統語流韻律分析所視為字調及孤立短句的種種變異性，其實皆可由階層式架構韻律短句上層的資訊加以釐清並解釋，而且可以預測。由此得知，語音信號必須兼顧HPG架構中各層的韻律效應所造成的相鄰及跨單位韻律關係，而不僅只是做字調與句調的線性串接及平滑，才能將完整地模擬出連續語流的特性。此篇論文將以單字及韻律詞基頻曲線為物理特徵參數，由下至上提出對應HPG架構的證據。

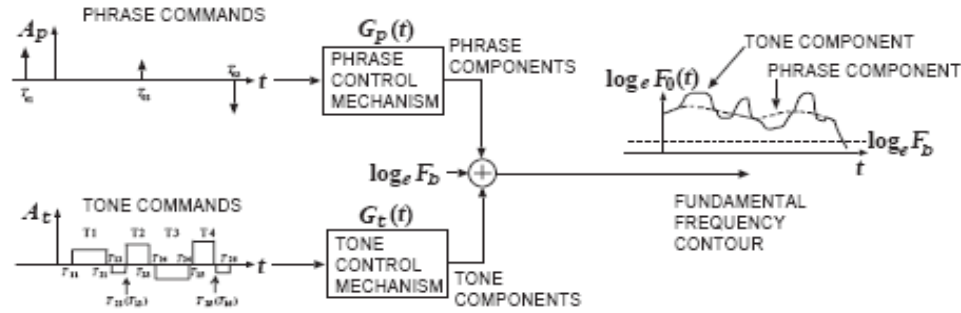
根據趙元任先生的說法[5]，短語句調與字調之間的關係，就好比波浪與連漪之間層層疊加關係，可以代數總和表示，相位相等時互相加成，相位相反時互相抵消。而在所有基頻曲線模型中最能表達這種疊加概念的，為日本學者Fujisaki[6]於1984年提

出的Fujisaki Model。此模型的精神是非聲調語的一個句調單位IU (intonation unit)的基頻曲線，必能拆解成全面句調成分與局部強調二個成分，單位大小不同，而套用於聲調語言時，局部強調則被轉換成描述字調的成分[6, 7, 8]，此即證明且呼應了趙元任所指大波浪與小漣漪的關係。因此，綜合趙元任先生與Fujisaki教授的看法，國語短語句調的基頻曲線，其實是字調成分與句調成分疊加而成，亦即除相鄰字調的連接外，還有來自上層的句調的覆蓋，因此不僅是字調的串接。若根據HPG架構，連續語流裡還更不僅只有字調與句調兩種韻律層級與單位。我們先前依據聽感標註得出的HPG架構進行分析，已找出句調彼此間的關連性與連貫性，即為所謂上層訊息，及對應HPG架構中的韻律句群關係[1][2][3]，並在稍早的研究中，特別探討如何使用Fujisaki Model提取短語句調的成分在語段中的體現[9]。同樣的，我們也希望在對應HPG架構中的韻律詞層，也找到字調間關係，並確認韻律詞在連續語流中作為基本韻律單位的韻律意義。我們知道，音段成分相同的音節，會因為字義不同而有不同的聲調，超音段成分基頻曲線的變化由詞義決定，因此嚴格說來，只能稱為詞義韻律(lexical prosody)，而非語流韻律的特徵或成分。因此本論文除了證明HPG架構外，也將以韻律詞的基頻曲線變化為主要分析參數，將強調字調以上韻律詞的在基頻曲線上的相對關係與特徵的確存在，且是連續語流中的基本韻律單位；短語也是連續語流的基本韻律單位而非終極韻律單位。

二、國語基頻曲線特徵參數自動擷取系統

(一) Fujisaki model 簡介

Fujisaki 等 1984 年[6]提出疊加式 command-response 基頻曲線模型，簡稱為 Fujisaki model，此模型的特點在於拆解看似不規則的基頻曲線為三個不同的元件函數 (component) 的疊加總合，並分別可找到相對應發聲器官的物理特性來解釋這些元件函數，此三元件函數分別為(1.)短語元件 (Phrase Component A_p)，反應較大單位基頻曲線的控制與發聲限制；(2.)強調元件(Accent Component A_a)，反應較小單位基頻曲線的控制發聲限制；與(3.)基底頻率(base frequency F_b) 代表基本音高。原 Fujisaki model 中的 Accent component A_a 泛指強調、加重語氣對局部基頻曲線造成的影響，此模型應用到非聲調語言與英語、德語時，大單位指的是片語的語調或短語的句調，即陳述句的由高走低的下傾趨勢，小單位則用來表示局部的加重或加強 (emphasis)；應用到聲調語言模擬國語時，大單位表示的成分與非聲調語相同，即片語的語調或短語的句調，而小單位被用來表示單音節的局部變化，亦即字調。本研究僅分析國語連續語流，所以 A_a 皆表示字調成分，因此下文均稱做字調元件。



圖二、短語元件、強調元件與基底頻率疊加後的基頻曲線 (Fujisaki, 1984)

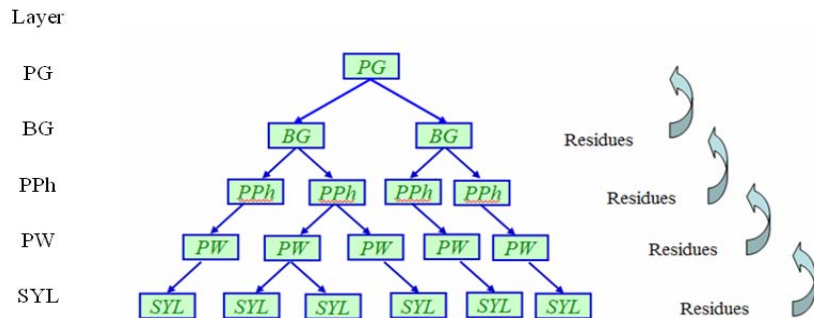
(二) Fujisaki 特徵參數的自動擷取

基於 Fujisaki Model 的國語基頻曲線，陳述句的句調可視為開高後走低全面下傾、局部字調起伏變化與基本頻率三元件疊加所成，我們則採用了 Mixdorff 2000、2003 年[7][8]提出的方法來解析原有的基頻曲線：以一組截止頻率為 0.5Hz 的高通濾波器 (high-pass filter) 來分離基頻曲線，自動提取出基頻曲線中變化劇烈的部份即為語流的基本單位，可對應 Fujisaki model 中的字調元件；而變化和緩的部份，則為語流中語調全面下傾的趨勢，可對應 Fujisaki model 中的短語元件。接著分別對三個元件進行逼近步驟：(1) 高通濾波器的輸出定義為高通曲線(HFC)，為字調元件逼近的目標曲線，(2) 扣掉高通部份剩餘平滑曲線則定義為低通曲線(LFC)，找出此低通曲線的最低點並定義為通過此最低點的直線為基底直線(Fb)，(3) 扣掉基底直線後的曲線視為短語元件的目標曲線，必須用短語元件函數來加以逼近。傳統人工逼近步驟(1)[10]中，遇到第三聲（字調三）跟第四聲（字調四）的音節，通常會指派一大一小的字調元件以逼近一音節內較複雜的基頻曲線，如圖二所示，然而為了自動擷取大量語料的參數，在自動擷取 Fujisaki 特徵參數程式中，我們忽略了較小的字調元件，換句話說，不論字調，我們皆只採用一字調元件來逼近一音節內高頻的基頻曲線。

三、基於HPG架構之階層性多元迴歸分析

從自動擷取Fujisaki特徵參數系統中，我們可將原始的基頻曲線拆解成短語元件與字調元件，分別對應HPG架構中字調與句調的成分。然而原始的Fujisaki Model並無對應韻律詞層、呼吸句群層與韻律短語層的元件函數，因此我們利用階層性線性迴歸找出各層的貢獻度與特徵。階層性線性模型是簡單線性迴歸的衍生[11][12]，每個輸入都附有多層次的停頓標註，每一標註分別代表著在每一韻律層的參數與特性，我們利用每一層的標註資訊，可得到各層韻律單位的模型。接著依據此模型進行正規化。由於有更上層的標註，正規化後數值的變異並不被視為實驗誤差，而是以更上

層的標註進行線性迴歸分析，分析及預測下一韻律層的輸入，因此可得到更上層的預測模型與貢獻度。逐層分析、預測後計算出各韻律層的貢獻度。圖三以圖式表示階層式線性迴歸逐層分析。



圖三、以階層式線性迴歸逐層分析示意圖 (Tseng et al, 2004)

本實驗以(1)Ap 為分析韻律短語句調以上語段或與篇語意之特徵 (2)Aa 為分析字調以上韻律詞意之特徵。Ap 是對應韻律短語的基頻特徵，因此分析的步驟由階層式架構中韻律短語層(PPh)開始、之後逐層向上，對上一層的呼吸句群層(BG)及更上一層韻律短語句群(PG)進行的線性迴歸，分析參數如下: (1)PPh 層: 以目前 PPh 長度(Current PPh Length)、前一 PPh 長度(Preceding PPh Length)與後一 PPh 長度(Following PPh Length)的組合做為分析參數，進行分析，經過線性迴歸後的殘差定義為 Delta1，並輸入 BG 層進行分析，(2)BG 層: 以目前 PPh 在 BG 中的位置(BG Sequence)做為分析參數，若 BG Sequence=1，表示目前 PPh 為此 BG 之起始 PPh，以此類推，進行線性迴歸，(3)PG 層: 與 BG 層輸入參數相同，其數學函數表示如下：

PPh

$$Ap=f(\text{FollowingPPh_Length}, \text{PrecedingPPh_Length}, \text{CurrentPPh_Length})+\text{Delta1}$$

BG

$$\text{Delta1}=f(\text{BGSequence})+\text{Delta2}$$

PG

$$\text{Delta2}=f(\text{PGSequence})+\text{Delta3}$$

其中 f 表示線性迴歸函數，迴歸係數與原始值 Ap 間的差值視為上層貢獻 Delta，並以上層的輸入參數對 Delta 再執行一次線性迴歸，此時得到的迴歸正確率視為上層的貢獻度，以此類推。

Aa 是對應字調的特徵，因此分析的步驟由階層式架構音節層(Syllable)開始，預測參數包括目前的字調(Current Tone)與前後字調的組合(Preceding Tone + Following Tone= Tone Context)，之後對上層的韻律詞層(PW)做分析，預測參數包括韻律詞邊界(PW Boundary Info)與此音節在韻律詞內的位置順序(PW Position Sequence)，由於在[12]中發現，在較高層級的韻律邊界常有邊界效應發生，因此我們也將邊界效應加入考慮，包括 1. 韻律短語邊界訊息(PPh Boundary Info)、2. 韻律句群層邊界訊息(PG Boundary Info)，將大範圍韻律單位首和尾音節的類別標記出來，進行獨立的

Aa 類別分析，其數學函數表示如下：

Syl

$$Aa=f(\text{FollowingTone}, \text{PrecedingTone}, \text{CurrentTone})+\text{Delta1}$$

PW

$$\text{Delta1}=f(\text{PW Boundary Infor}, \text{PWSequence})+\text{Delta2}$$

Boundary effect above PPh

$$\text{Delta2}=f(\text{PPh Boundary Infor}, \text{PG Boundary Infor})+\text{Delta3}$$

四、實驗語料

文本部份，採用 (1)古典文體 CL 與 (2)長篇敘事段落文本 CNA，共計(1)26 篇古典文體語篇段落（包含：4 篇古典散文，1 首賦，1 首民歌，6 首古詩，6 首唐代樂府詩和 8 首宋詞），以及(2) 26 則白話敘事段落。

語料部分，由一男一女發音員（M056 & F054）朗讀古典文體文本，平均語速分別為 202m/syl 和 265ms/syl；另外的一男一女發音員則（M051 & F051），負責朗讀白話敘事段落，平均語速分別為 189m/syl 和 199ms/syl，錄製過程使用 Sony ECM-77B 迷你麥克風、以及 Cool Edit 2000 在隔音室進行錄音。表一統計兩種韻律格式加總後的 HPG 架構下韻律邊界以及相對應韻律單位的個數。

表一、古典文體 CL 與長篇敘事段落文本 CNA 韻律邊界以及相對應韻律單位個數

語料	語者	SYL/B1	PW/B2	PPh/B3	BG/B4	PG/B5
CL	F054	1444	599	290	135	58
	M056	1551	619	318	142	47
CNA	F051	6583	3468	1092	297	151
	M051	6661	3332	1207	270	129

五、實驗結果與分析

(一)句調成份與字調成份對應各韻律階層的貢獻度

依據第三節的階層性多元迴歸分析方法，我們由下層到上層進行線性迴歸，每一層的迴歸正確率視為每一韻律層的分層貢獻度，而由最下層(音節層)累積到當層的迴歸正確率即被視為累積到當層的累積貢獻度，累積正確率 100% 表示最後所有韻

律層的迴歸係數的總和等於原始Fujisaki特徵參數值，Fujisaki特徵參數值完全可由階層性多元迴歸分析正確預測。

1. 字調元件Aa分層貢獻度

從階層性多元迴歸分析可求出字調元件Aa在音節層與韻律詞層的貢獻度，並加入韻律短語層以上邊界效應的考慮[13]，包括此音節前後是否有韻律短語邊界及韻律句群層邊界。最後結果顯示Aa的正確率可達73.80%到56.25%不等，其中大部份貢獻度來自音節層與韻律詞層，邊界效應的貢獻度則介於5~7%。

表二、字調元件在音節層與韻律詞層的累積正確率

語料	語者	Syl 層貢獻度		PW 層貢獻度	
		Tone	Tone Context	PW Boundary Info	PW Position Sequence
CL	F054	46.21%	54.74%	60.54%	66.61%
	M056	39.12%	47.86%	57.68%	61.45%
CNA	F051	38.40%	45.00%	48.43%	51.27%
	M051	41.61%	47.96%	51.33%	54.53%

表三、加上邊界效應後字調元件的累積正確率

語料	語者	PPh 層以上的邊界效應		邊界效應的貢獻度
		PPh Boundary Info	PG Boundary Info	
CL	F054	72.98%	73.80%	7.19%
	M056	64.13%	66.89%	5.43%
CNA	F051	54.41%	56.25%	4.98%
	M051	57.43%	59.32%	4.79%

2. 短語元件Ap分層貢獻度

從階層性多元迴歸分析可求出短語元件Ap在韻律短語層、呼吸句群層與韻律句群層的累積貢獻度，我們發現古文語料CL有許多貢獻度來自於韻律短語層以上的上層資訊，反之，白話長篇敘事語料CNA正確率則多來自目前PPh與前後PPh的長度資訊。

表四、短語元件Ap在韻律短語層、呼吸句群層與韻律詞層的累積正確率

語料	語者	PPh	BG	PG
CL	f054	58.79%	63.58%	76.66%
	m056	37.89%	48.99%	73.66%
CNA	F051	80.17%	81.46%	87.71%
	m051	81.53%	82.72%	88.20%

如前文所提，基頻曲線的變化由Ap與Aa構成，因此我們將Ap與Aa的預測率正確率平均作為套用HPG架構後Fujisaki model對總體基頻曲線模型預測正確率。結果如表五。

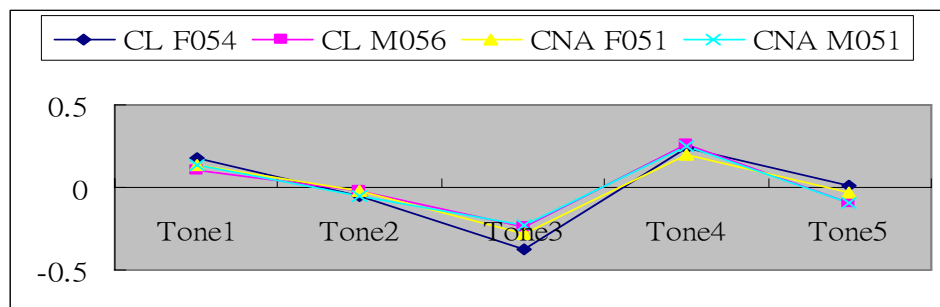
表五、套用HPG架構對總體基頻曲線模型預測正確率

語料	語者	Aa	Ap	Total
CL	f054	76.66%	73.80%	75.23%
	m056	73.66%	66.89%	70.28%
CNA	F051	87.71%	56.25%	71.98%
	m051	88.20%	59.32%	73.76%

(二) Aa對應音節層與韻律詞層模型

1. 字調(音節層)模型

在上表中我們得見字調對Aa的貢獻度為最大，此結果與先前研究Fujisaki Model的結果一致。相對於每個字調的Aa模型列於下圖，我們可看見Aa的能量模型在不同語料間相當一致。



圖四、自動擷取出的字調元件對應各個字調之模型

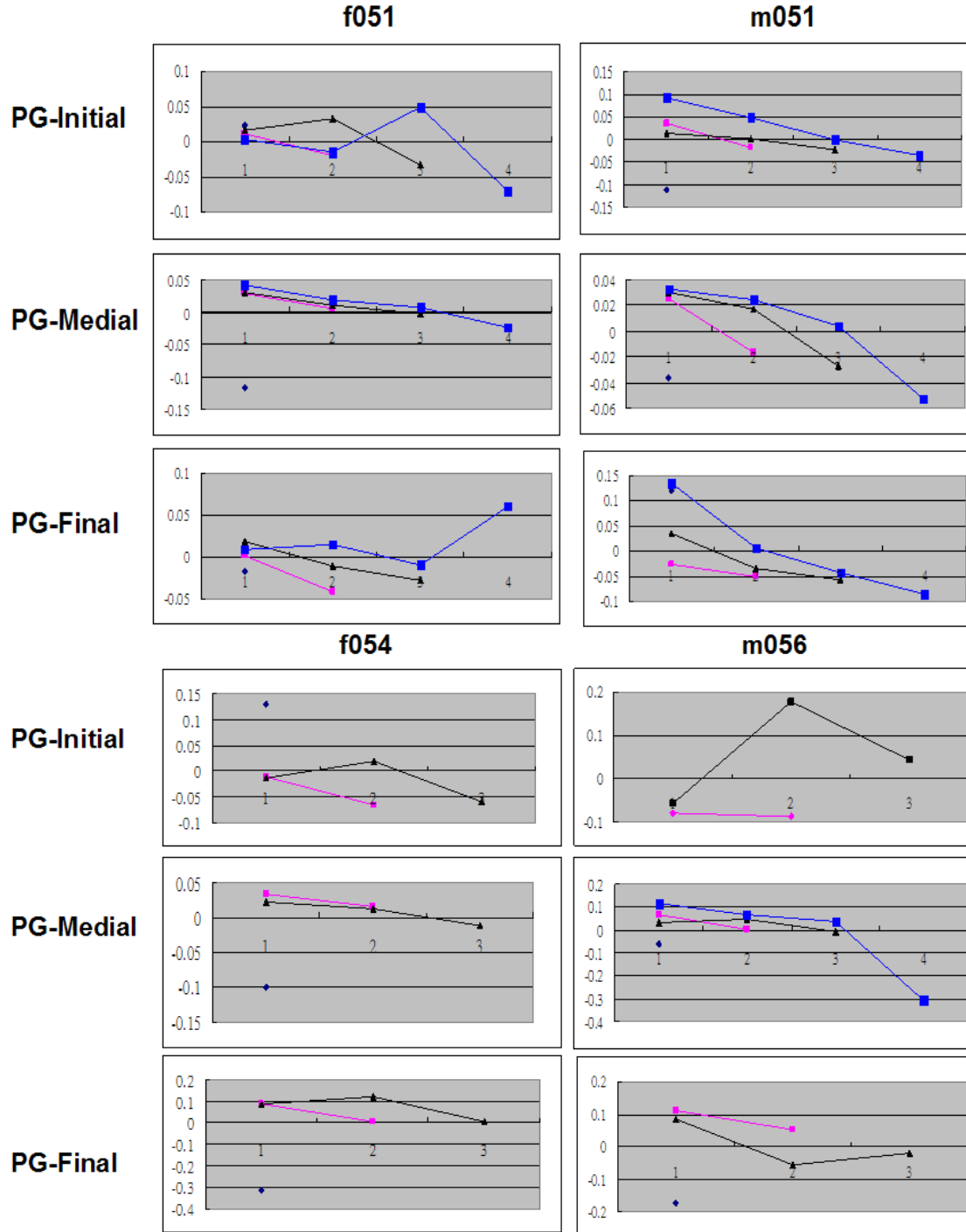
2. 韻律詞模型

2.1 韻律詞受韻律句群PG管轄之特徵

將消除字調效應後的韻律詞基頻模型依照PG位置分類後，我們發現在PG-Medial的韻律詞模型，在四筆語料間呈現一致性的特徵，而且韻律詞在基頻曲線上的邊界分隔，主要發生在韻律詞的詞尾，換言之，即便韻律詞間並沒有停頓，且韻律詞間受字調間平滑影響，並不易觀察到基頻重設，但經過消除字調效應的正規化處理後，可發現PW末的Aa強度不如其他PW音節，因此我們可從這詞尾的基頻能量衰減(decay)來辨別韻律詞單位邊界。在PG-Initial位置，我們也發現韻律詞在PW末存在相

同特徵，且PW末和倒數第二音節間相對於PG-Medial的對比也較大。

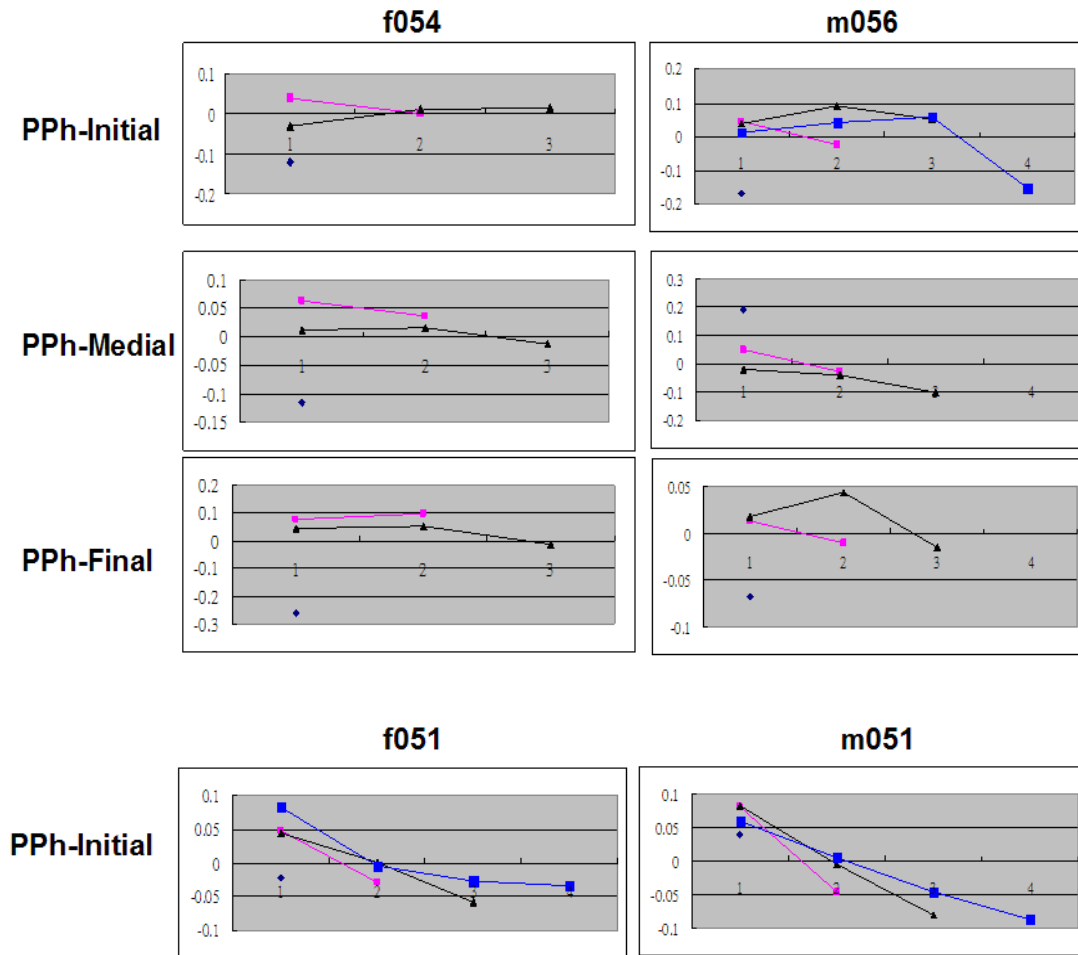
對照2007以Ap為實驗，獲不同語體的階層貢獻度不同之結論[9]，經由本實驗結果可進一步發現，由於跨語體間的韻律詞特徵，在PG-Medial最為穩定且最不明顯，因此大部分與語體變化相關的貢獻度差異，應來自PG-Initial與PG-Final而非PG-Medial。

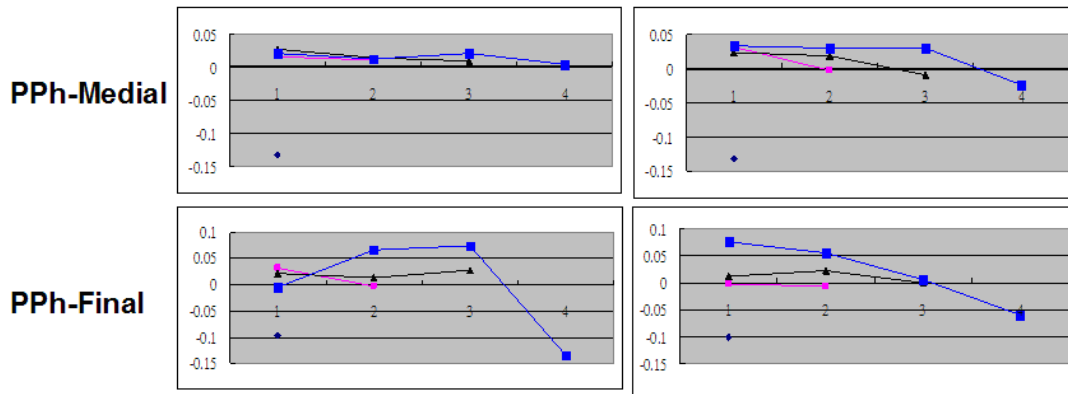


圖五、扣除字調成份後，韻律詞模型在不同 PG 位置的特徵，每條曲線表示特定長度的韻律詞模型，橫軸表示在此韻律詞內音節順序，縱軸表示扣除字調成份後的殘差。

2.2 韻律詞PW受韻律短句PPh管轄之特徵

同理我們將扣除字調成份的韻律詞PW模型依照韻律短句PPh位置分類後，得到各韻律詞的特徵，發現韻律詞詞尾的特徵與詞首、詞中最為不同，尤其又以PPh-Medial最為明顯，PPh-Initial、PPh-Medial最重要的特徵也大致發生在詞尾Aa。在PPh-Medial的位置的韻律詞的一致性表示，PPh-Medial為最不受語體影響的相對位置，如圖六所示。我們也可發現在PPh-Final，Aa強度韻律詞的最末音節存在衰減的特徵，雖然PW模型不如PPh-Medial規則，但最末音節的Aa衰減程度較PPh-Medial明顯，顯示PPh在基頻區線上的邊界效應，主要發生在邊界前的韻律詞。這個特性，與音節時長在PPh句尾延長效應吻合[2] [12]，顯示在物理信號上，主要的PPh邊界效應，主要表現在PPh邊界前，PPh句末的物理信號上。





圖六、扣除字調成份後，韻律詞模型在不同 PPh 位置的特徵，每條曲線表示特定長度的韻律詞模型，橫軸表示在此韻律詞內音節順序，縱軸表扣除字調成份後的殘差。

六、結論與展望

由 Aa 實驗結果可知，字調仍為影響音節內基頻曲線的主要因素，大約佔字調元件 Aa 預測正確率的 40~45%；韻律詞 PW 的階層性貢獻佔 15~20%。但以上數據亦顯示：從字調辨識的角度而言，字調的正確率不及一半，表示最終的語流韻律輸出中，字調的成分並非字字可辨。由扣除字調成份的韻律詞 PW 模型可發現，HPG 架構中的 PW 層存在一定特徵，也有一定的貢獻度，解釋了在基頻曲線中，並非只以字調為基本單位進行串接；PW 對基頻曲線也有一定程度的影響，並在最終的語流韻律輸出中，佔有一定的比例。以上結果和稍早 Ap 實驗的 PPh 表現的結果相符，即 PPh 的 Ap 必須考慮呼吸句群 BG 層以及韻律句群 PG 層的上層效應[9]，本研究結果顯示 BG 層與 PG 層對句調元件 Ap 的貢獻度約佔 7~35%，語流最終的韻律輸出，各韻律層都有貢獻。不同語體的朗讀語料，因語體而產生的韻律輸出差異，只是階層式貢獻度的分佈差異而已，其韻律成份完全可由 HPG 架構解釋，同一基型只需調整階層式的韻律貢獻度，便可產生不同的韻律輸出[9]。結合以上實驗結果，我們因而得知，不論字調或句調，皆為 HPG 的次級韻律單位，各自受到 HPG 架構的層層管轄，系統性的調整字調以及句調，以形成流暢的連續語流中的表意韻律語境，表達句調間的連續與跨句調的呼應。我們相信這些韻律語境的模型為理解與產製語音的重要單位，口語產製時，人們使用這些韻律基型，因此依據這些基型，可輕易快速的將複雜的連續語流的語境歸類成最適合的語段、語篇，以進行上層語意的組織與分析。

我們期望這些證據及特徵能提供語音合成系統一個新的思維：語流韻律是有架構的，因有系統性而有跡可循，所以連續語流並非充滿變異而難以預測。語流的流暢並非僅來自於字調與句調的線性串接的完美平滑，也必須同時包含多短語語段間的跨短語表意韻律語境。我們所提出的多短語階層性 HPG 架構，完整的解釋語流韻律規範的來源及如何互動。合成的語音輸出聽起來不流暢、不自然的感覺，則大部

分由於未含上層資訊，以致韻律語境不足或未能體現韻律語境。同理，在聽者即時處理語流時，即便每一個單音節的字調訊息不完全，只要有相當成分的韻律語境，便能彌補並提供聽者切分韻律單位及向前預估的訊息。只不過，所謂的相當成分，必須同時含有相鄰下層韻律單位的連接及上層韻律單位的呼應。反之，如一段連續語流內的單位無法同時體現階層性 HPG 架構組織所規範之全面表意及局部連接的韻律語境時，將會違背聽者預期，造成聽者切分單位的錯誤，需一再修正，而延誤即時處理，由此亦可見，語音識別並不同字調識別。未來研究方向包括將 HPG 架構套用至授課等自發性語料溝通意圖明顯、與篇及語段訊息分明的語料進行分析，一方面希望找出此規劃的基型也存在看似不規則的自發性語料中，一方面更全面的解釋語流中字調的變異，解構字調即韻律的迷思。

參考文獻

- [1] Tseng, C. "Prosody Analysis", *Advances in Chinese Spoken Language Processing*, World Scientific Publishing, Singapore, pp. 57-76, 2006.
- [2] Tseng, C., Pin, S., Lee, Y., Wang, H. and Chen, Y. "Fluent Speech Prosody: Framework and Modeling", *Speech Communication, Special Issue on Quantitative Prosody Modeling for Natural Speech Description and Generation*, Vol. 46:3-4, pp. 284-309, 2005.
- [3] Tseng, C. and Lee, Y. "Speech rate and prosody units: Evidence of interaction from Mandarin Chinese", *Proceedings of the International Conference on Speech Prosody 2004*, pp. 251-254, 2004.
- [4] Tseng, C., Pin, S., Lee, Y., "Speech prosody: Issues, approaches and implications". *From Traditional Phonology to Modern Speech Processing*. ed. by Fant, G., Fujisaki, H., Cao, J., and Y. Xu, 417-437. Beijing: Foreign Language Teaching and Research Press, 2004
- [5] Chao, Y., *A Grammar of Spoken Chinese*. University of California Press, Berkeley, 1968
- [6] Fujisaki, H. and Hirose, K. "Analysis of voice fundamental frequency contours for declarative sentences of Japanese", *J.Acoust. Soc. Jpn.(E)*, 1984; 5(4), pp. 233-242, 1984.
- [7] Mixdorff, H. "A Novel Approach to the Fully Automatic Extraction of Fujisaki Model Parameters", *Proceedings of ICASSP 2000*, vol. 3, pp.1281-1284, 2000.
- [8] Mixdorff, H., Hu, Y. and Chen, G. "Towards the Automatic Extraction of Fujisaki Model Parameters for Mandarin", *Proceedings of Eurospeech 2003*, pp. 873-876, 2003.
- [9] 鄭秋豫、蘇昭宇. "從不同韻律格式驗證階層式韻律架構並兼論對語音科技的應用" 第十九屆自然語言與語音處理研討會，中華民國計算語言學會，台北，2007
- [10] Wang, C., Fujisaki, H., Ohno, S. and Kodama, Tomohiro. "Analysis and synthesis of the four tones in connected speech of the standard Chinese based on a command-response model", *Proceedings of EUROSPEECH'99*, pp. 1655-1658,

1999.

- [11] Keller, E., and Zellner, K.,. “A Timing model for Fast French”, *York Papers in Linguistics*, 17, University of York, pp.53-75, 1996.
- [12]Zellner, K., and Keller, E., “Representing Speech Rhythm” *Improvements in Speech Synthesis*. Chichester: John Wiley, pp. 154-164, 2001.
- [13] Tseng, C., Su, Z., “Boundary and Lengthening—On Relative Phonetic Information.” *The 8th Phonetics Conference of China and the International Symposium on Phonetic Frontiers*, Beijing, China., 2008