

# 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

鄭秋豫 蘇昭宇

中央研究院語言所語音實驗室

[cytling@sinica.edu.tw](mailto:cytling@sinica.edu.tw), [morison@gate.sinica.edu.tw](mailto:morison@gate.sinica.edu.tw)

## 摘要

本論文以四種韻律格式語料的基頻變化，驗證鄭秋豫[1][2][3]所提出的「階層式韻律句群 HPG 架構」，結果證明一種基型即可解釋並產製不同的韻律格式。HPG 架構解釋：口語語流的韻律輸出，並非僅止於字調與句調的線性串接，而同時含有跨短句語篇上層訊息，各韻律階層對總體韻律輸出的貢獻度可驗證。所用語料為朗讀古典文體三種（詩、詞二種韻文及古典散文）、氣象撥報語料一種。我們採 Fujisaki model 檢驗基頻（F0）參數，用 HPG 階層式模型進行分層分析與預測，結果發現：每類語體的韻律輸出都含有上層語篇訊息（higher level information）效應，唯韻律貢獻度依韻律格式呈不同分佈。三種古典文體語料的上層貢獻度均高於非氣象撥報語料，且文體結構越工整、語料的樂律性越高時，上層貢獻度所佔比例越重。因此韻律格式的差異，具體表現於階層貢獻度的分佈，而 HPG 架構可解釋韻律輸出的差異。我們以此結果更進一步證明：無論語體或韻律格式為何，上層語篇效應所造成的跨短句韻律語境，都是語流韻律不可或缺的因素。我們也因此推論，語音合成語流韻律輸出時，只需一個韻律基型，配以不同的韻律層貢獻度，便可產生不同韻律輸出。

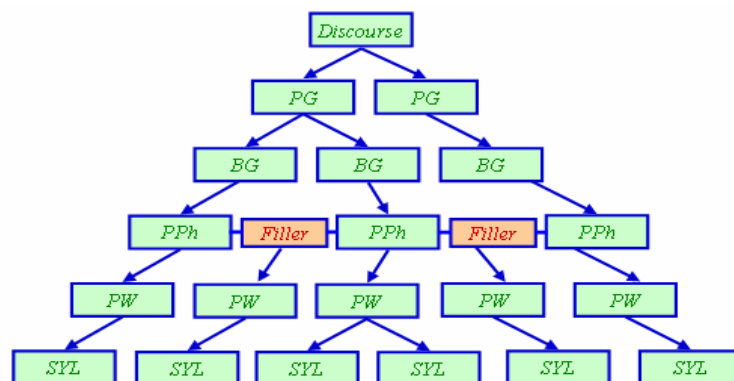
關鍵詞：階層式韻律句群架構，語流韻律，古典文體，Fujisaki model，上層效應，語體

## 一、緒論

與韻律訊息相關的聲學語音參數有四：基頻曲線(F0 contour)、音節時長(syllable duration)、能量分佈(Intensity distribution)、和停頓時長(pause duration)。與韻律相關的文獻中，最受重視的聲學參數一向是基頻，研究聲調語言時，字調基頻曲線的變化也一直是重點，本論文將以韻律短句的基頻曲線變化為主要分析參數，聽感的韻律邊界停頓為輔，探討不同韻律格式的韻律輸出，究竟如何不同？是否可以透過量化語料，加以驗

證。

傳統韻律研究一向著重於孤立音節的字調及孤立短句句調的分析，然而在長篇口語敘述段落的語料中，相同字調的音節或同類的句調，卻有相當大的變異，語音訊號的變化非常複雜。鄭秋豫從聽感出發，發現這些變異對聽者完全沒有影響，聽者輕易的同時處理不同韻律範圍的資訊，辨識音段、音節、字調、句調甚至更大單位的韻律訊息，更有效的應用短句以上的語篇語意訊息，以上層語篇訊息（higher level information）的效應，辨識出字調與句調以上的語段單位，作為口語分析的基礎。她經由量化分析大批口語料，發現語流中的音節長短在時程上的長短分佈、能量分佈、停頓時長，都含有上層語篇訊息，語流韻律輸出不但具有系統性，而且是層層韻律效應疊加的總和[3]，也因此證明語段中語篇語意的連接與連貫性確實存在於語流語音信號中，若只考慮字調與語調效應，不但不足以說明多短語語流的韻律特性，且無法系統性的呈現語段中各短語間的關聯性（association），她以此基礎，提出「階層式多短語韻律架構 Hierarchical Prosodic Phrase Grouping Framework (HPG)」及數學模型[3]。2006 年她進一步得到基頻走勢的證據[Tseng 2006]，至此結合基頻、音長、能量、停頓的證據，將 HPG 架構（如下圖所示）擴及語篇。



圖一、階層式多短語韻律句群架構圖。其中 SYL、PW、PPh、BG、PG 分別代表口語語流中的音節、韻律詞、韻律短句、呼吸句群，韻律句群。（Tseng, 2006）

在 HPG 架構中，下層韻律單位（如音節、韻律詞、韻律短句）接受上層韻律單位（如呼吸群、韻律句群、語篇）的管轄，因此音節與短句一旦進入多短語語段，除了字調（tone）與短語調（intonation）外，因為來自上層語篇的語意訊息造成語流的連接性（cohesion）與連貫性（coherence），以至韻律短句間產生韻律關連性（association），各下級韻律單位必須依照更語篇韻律（Discourse Prosody）的上層訊息進行系統性的調整，提供語段從

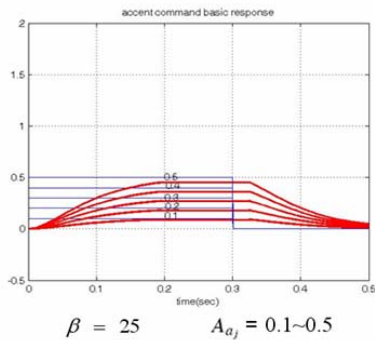
何處開始、維持到結束的韻律語境訊息，各語段下轄之次次級韻律單位需依此規範，做系統性的調整。由於此架構的提出，前文所提語流中，字調及孤立短句效應外的變異性，其實皆可由階層式架構韻律短句上層的資訊加以解釋，而不再屬於無法預測的變異。由此得知，進行語音合成時，語音信號必須兼顧架構中各層的韻律效應，而不僅只是做字調與句調的線性串接及平滑，才能將完整地模擬出連續語流的特性。

## 二、Fujisaki 特徵參數自動擷取系統

### (一) Fujisaki model 簡介

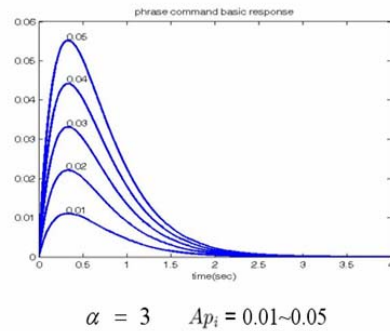
與韻律相關的文獻中，前文所提四項聲學語音參數（基頻曲線、音節時長、能量分佈和停頓時長）又以基頻曲線最受重視，基頻曲線的變化也是研究聲調語言字調的重點，本論文即以韻律短句的基頻曲線變化為主要分析參數。傳統研究中，是以直線、二次方或三次方函數曲線趨近基頻曲線，但 Fujisaki 等 1984 年[4]提出基於發聲器官的模型：產生基頻主要器官為聲道（vocal tract）及相鄰的隨意肌，而應於語言單位在時間程上產生高低的變化而控制隨意肌時，有一定物理特性的限制，所以基頻曲線並非任意直線或曲線，而是必須受制於一定的規則在時間程所產生的變化，且此限制出現於對數尺度上，此為基頻曲線的函數限制，此一模型一般通稱為 Fujisaki model。Fujisaki model 由三個不同的元件（component）所構成，分別為(1.)短語元件（phrase component  $A_p$ ），反應發聲器官產製較大單位基頻曲線的控制與發聲限制；(2.)語氣元件(tone component  $A_a$ ），反應發聲器官產製較小單位基頻曲線的控制發聲限制；與(3.)基底直線(base frequency  $F_b$ ) 代表基本音高。原 Fujisaki model 中 tone component  $A_a$  中的 tone，泛指語氣變化（tone of voice），如強調、加重語氣對基頻曲線造成的影響。此模型本身對於大、小單位為何並未規範，又因為基於發聲器官，所以也與語言無關，迄今已被用來成功的模擬過多種語言。應用到輕重音語言與英語、德語時，大單位指的是片語或短語的語調、小單位則用來表示加重（emphasis）；應用到聲調語言模擬中文（國語）時，大單位表示短語語調的下傾、小單位被用來表示字調變化[5]。圖二分別列出此二元件的函數及輸出特徵，其中短語元件和語氣元件函數分別由一個函數能量  $A_p$  及  $A_a$  所控制，各自控制輸出的突起程度以及上升／下降斜率， $A_p$  及  $A_a$  越大，元件函數曲線凸起程度越明顯，下降斜率也更陡峭。

$$A_{a_j} G_t(t) = \begin{cases} \min[1 - (1 + \beta t) \exp(-\beta t), \gamma], & t \geq 0 \\ 0, & t < 0 \end{cases}$$



(a)

$$A_{p_i} G_p(t) = \begin{cases} \alpha^2 t \exp(-\alpha t), & t \geq 0 \\ 0, & t < 0 \end{cases}$$

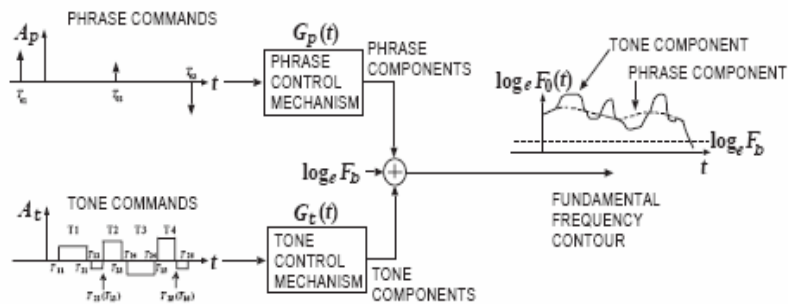


(b)

圖二、(a)Aa (b)Ap 隨函數能量變化的時間響應圖

## (二) 自動擷取 Fujisaki 特徵參數的方法

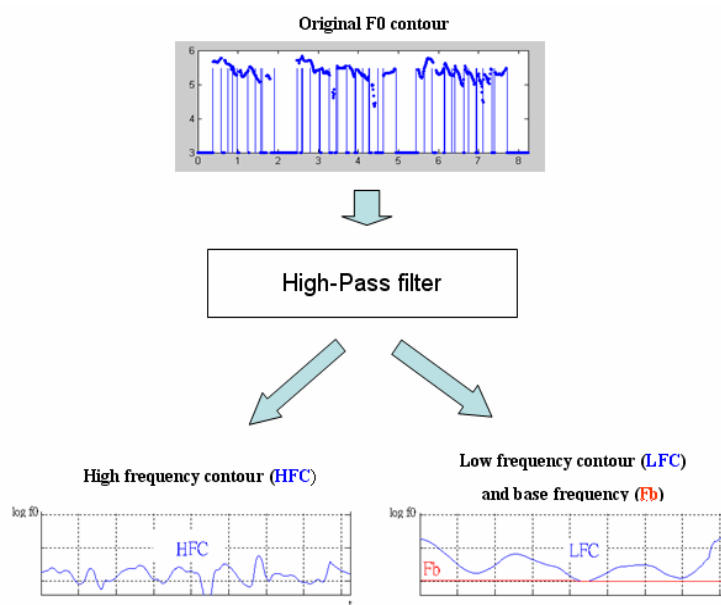
Fujisaki model 假設，基頻曲線可由短語元件、語氣元件和基底直線疊加而趨近，圖三表示 Fujisaki model 三個基本元件的疊加輸出。



圖三、短語元件、語氣元件與基底直線疊加後的基頻曲線 (Fujisaki, 1984)

換句話說，基於 Fujisaki Model 的國語基頻曲線，可視為語調下傾、字調變化與基本音高三元件疊加所成，而該模型泛指語氣變化的語氣元件 (tone component)，很巧合的剛好可用來表示字調 (tone) 的變化。雖然有了 Fujisaki model 可以來趨近原始語音，但由於 Fujisaki model 係由三個元件構成，所以趨近原始基頻曲線必須以三個元件的最佳組合為參考指標。此外，根據 Fujisaki model 之定義，相同的語氣元件必須在語流中各位置保持大致相同的權重，因此傳統的作法是以目測原始基頻曲線，手動調整決定這三個元件的最佳組合，這樣的作法相當耗費人力。我們則採用了 Mixdorff 2000、2003 年

[6][7]提出的方法並加以改善：以高通濾波器(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 中的語氣元件；而變化和緩的部份，則為語流中語調的下傾趨勢，可對應 Fujisaki model 中的短語元件，因此採用一組截止頻率為 0.5Hz 的高通濾波器分離原有基頻曲線為三個部份(1)高通濾波器的輸出定義為高通曲線(HFC)，為語氣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。圖四是 Mixdorff 自動擷取 Fujisaki 參數架構。



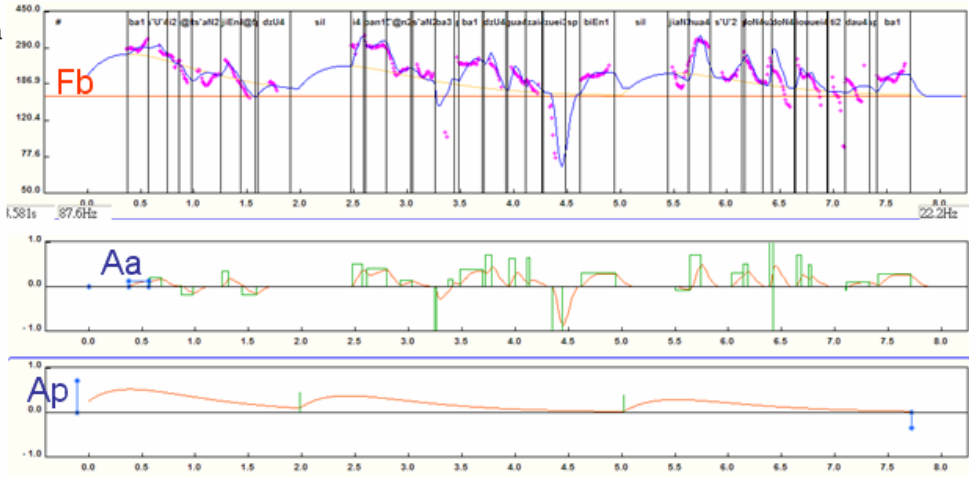
圖四、利用濾波器將基頻曲線分離為高通及低通曲線，再分別以字調元件與短語元件趨近高通與低通曲線，進行自動擷取程序 (Mixdorff, 2000)。

至於自動提取 Fujisaki 參數的方法，近來又有 Bu 等[8]提出以一階微分曲線來，不過，Mixdorff 處理的語言是國語，而 Bu 等處理的是日語，所以我們仍維持 Mixdorff 的方法。

### (三) 加入聽感邊界的方法與自動擷取結果

在本研究中，我們除驗證 HPG 韻律架構，也希望探討聽感獲得的韻律邊界與基頻曲線的關係，因此我們加入聽感的邊界資訊作為自動擷取 Fujisaki model 的  $A_p$  參數。初步觀察顯示，局部最低點通常就出現韻律短語的聽感邊界上，因此加入聽感邊界的自動擷取方法，可減少 Mixdorff 方法在搜尋整條低通曲線的範圍的運算時間。圖五為加入聽感邊界為輸入參數後，自動擷取 Fujisaki model 的  $A_p$  參數結果。

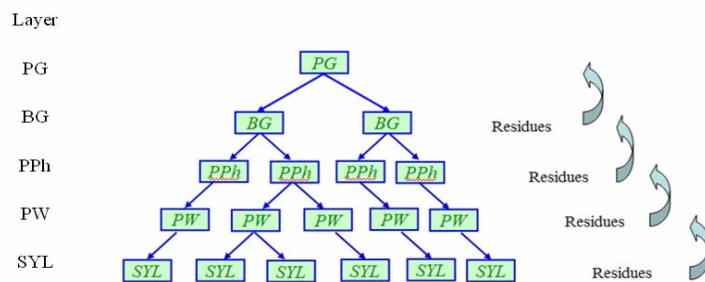
Comparison between Original F0 contour and pitch model superposed by tone and phrase components



圖五、加入聽感的邊界資訊後自動擷取 Fujisaki 參數的結果

### 三、階層式多元線性迴歸模型與計算

Nakai 等於 1994, 1995 [9][10]提出歸類 Ap 的方式來建立語流中的短語模型，然而並未將語流中的階層性關係列入考慮，我們採用了 Keller 等[11]與 Zellner 等[12]用於法語的階層性線性迴歸分析，並考慮國語的音韻結構加以改良，用此適用於國語的模型來分析語音信號。階層性線性模型是簡單線性迴歸的衍生，最大的不同點在於：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，可進行逐層分析，分析步驟如下：(1) 分析及預測某一韻律層時，只利用當層的參數做為線性迴歸的預測變項並得到此層的預測模型，(2)由於有更上層的標註，我們假設，原始值與預測模型的差距並不被視為實驗誤差，而是視為來自更上層的效應，因此將分析及預測某一韻律層的殘差（原始值與預測模型的差距）作為分析及預測下一韻律層的輸入，以更上層的標註進行線性迴歸分析，因此可得到更上層的預測模型與貢獻度，(3) 逐層分析、預測後計算出各韻律層的貢獻度。圖六以圖式表示階層式線性迴歸逐層分析。



圖六、以階層式線性迴歸逐層分析示意圖 (Tseng et al, 2004)

本研究以  $A_p$  為輸入參數，由於  $A_p$  是對應韻律短語的參數，因此，分析的步驟由階層式架構中韻律短語層(PPh)開始、之後對上層的呼吸句群層(BG)及更上層韻律短語句群(PG)進行逐層的線性迴歸，分析參數如下：(1)PPh 層：以目前 PPh 長度(Current PPh Length)與前一 PPh 長度(Preceding 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  $\Delta_1$ ，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：跟 BG 層輸入參數相同，其數學函數表示如下：

PPh:

$$A_p = f(\text{precedingPPhLength}, \text{currentPPhLength}) + \Delta_1$$

BG:

$$\Delta_1 = f(\text{BGSequence}) + \Delta_2$$

PG:

$$\Delta_2 = f(\text{PGSequence}) + \Delta_3$$

#### 四、實驗語料

文本部份，採用古典文體（詩、詞等韻文及古典散文）與氣象播報文本，共計 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 34 則氣象播報的語篇段落——以此分別代表四種韻律格式。古典文體各自承載不同文體的韻律變化特性，氣象播報文本則當作白話文本。我們依文本結構的工整規則性，定義四類韻律格式：規則(R)、半規則(SMR)、不規則(IR)與氣象播報(WIR)。其中詩詞等韻文主要為規則及半規則的文體，而沒有固定重複特徵的古典散文則被歸於不規則類；而氣象播報文本亦屬不規則類，但另以一類計，以示與古典文體有別，並可與古典文體比較。分類範例見附註一。古典文體文本包括的三種韻律格式(R、SMR、IR)，採人工方式判別詩歌的工整性，依據準則為：(1)標點符號。採用較大的段落標點符號，如：句號，將文本分成包含多短語的段落區塊；(2)短語長度。挑出相同文本中重複的多短語段落區塊，人工檢視重複區塊的工整性比例，加以分類。

語料部分，由三位發音人朗讀以上文本，使用 Sony ECM-77B 迷你麥克風、以及 Cool Edit 2000 在隔音室進行錄音。發音人為(1)女性發音人一位，f054，朗讀所有文本，(2)男性

發音人兩位，一位 m056 朗讀三種風格的古典文體語料（26 篇），另一位 m054 朗讀氣象語料（34 篇）。表一統計三種古典韻律格式加總後的音節、韻律短句數目及篇數，並以音節平均時長代表平均語速；表二統計氣象撥報語料的音節、韻律短句、語篇數目語料及平均語速。

表一、古典文體語料的音節、韻律短句、語篇數目語料及平均語速統計

韻律單元數/平均語速 發音人	音節數	韻律短句數	語篇數	平均語速 (毫秒/音節)
女 f054	3502	710	26	271
男 m056	3510	711	26	202

表二、氣象撥報語料的音節、韻律短句、語篇數目語料及平均語速統計

韻律單元數/平均語速 發音人	音節數	韻律短句數	語篇數	平均語速 (毫秒/音節)
女 f054	7054	720	34	193
男 m054	7096	747	34	165

從表一、表二可看出，本研究語料設計的主要控制組是韻律短句和語篇的數目，而非音節的數目。古典文體語料只有三千五百餘音節，氣象撥報語料則有七千餘音節，二者差異頗大；但古典文體語料含韻律短句七百餘、語篇二十六，氣象撥報語料則亦含韻律短句七百餘、語篇三十四，數目差異並不大。

## 五、實驗結果與分析

### (一) 聽感標註的重疊性

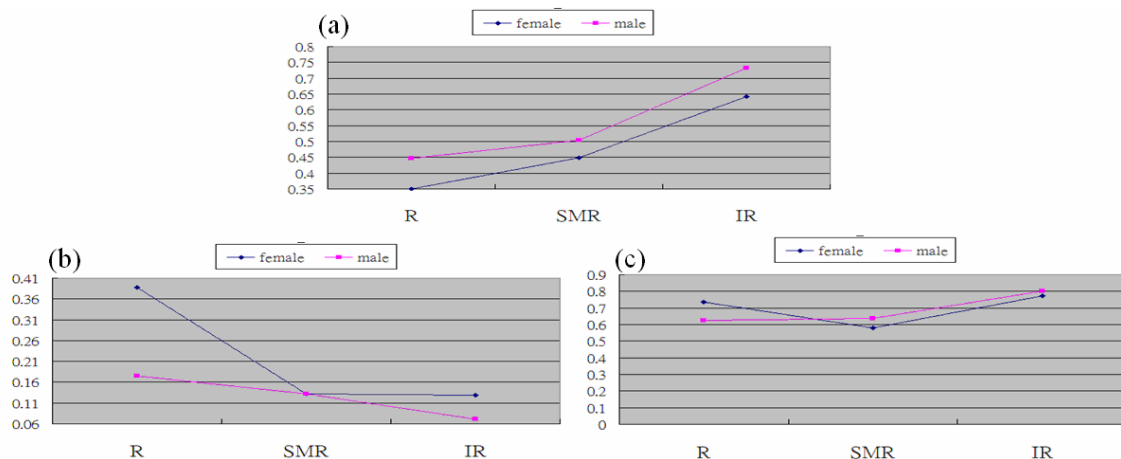
我們分析古典體文語料中二位發音人（f054 與 m056）聽感標註韻律邊界的重疊性後，將結果列於表三。表三中 B1、B2、B3、B4、B5 分別代表音節、韻律詞、韻律短語、呼吸句群，語段（韻律句群）之邊界。結果顯示二人的差異性並不大，而且從表三列出 HPG 韻律階層上層到下層的重疊比例，可發現上層標註的重疊率雖然不如下層，但在上層都能維持一定的重疊比例，可見在古典文體中，語篇的韻律效應仍存有相當的一致性。

表三、比較兩語者朗讀古典文體韻律邊界的一致性結果

韻律邊界 發音人	B1 音節邊界	B2 韻律詞邊 界	B3 韻律短句 邊界	呼吸群 B4	語段（韻律 句群）B5
f054	97.98%	86.46%	82.79%	76.76%	62.30%
m056	96.71%	88.69%	80.19%	76.76%	80.85%

## (二) Ap 分析

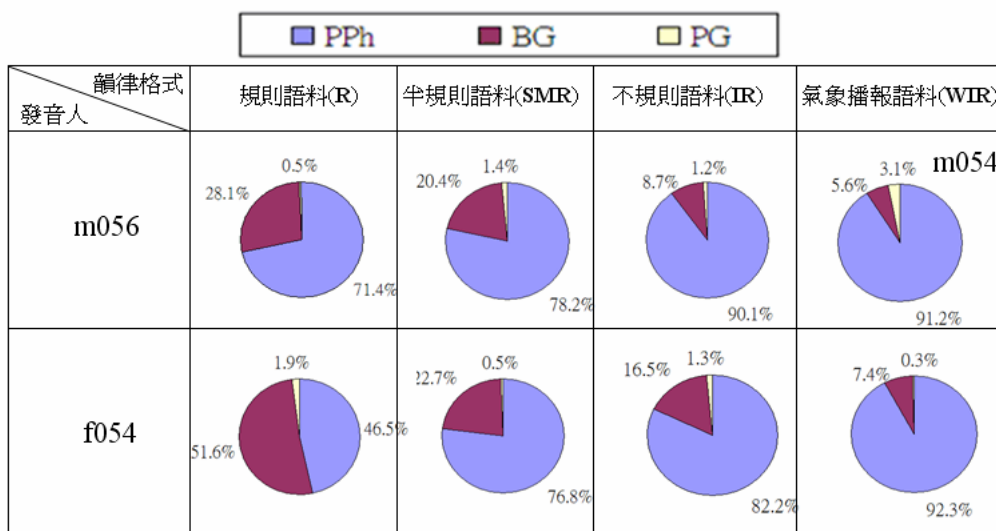
我們比較古典文體語料在單一韻律層中 Ap 正確率隨文體（規則、半規則到不規則）的變化趨勢後發現：在韻律短語層(PPh)的預測，隨著文體從規則到不規則的分類，兩語者 Ap 的正確率的預測皆呈現相同的趨勢。也就是說，預測的正確率隨文體的不規則程度上升，顯示出越不規則的短語結構變化，越是提供了較多的預測資訊，使不規則語料的 Ap 參數在 PPh 層獲得較準確的預測。然而，在呼吸句群層(BG)的預測，卻呈現相反的趨勢，即 Ap 預測正確率隨著文體工整性的提升，呈現上升的趨勢，顯示在文體越工整，上層韻律效應越明顯。因此在 BG 層中，預測度隨著文體工整性與 Ap 呈正相關。若將兩層的 Ap 預測正確率相加，可發現相加後的 Ap 預測率，對總預測率具有互補效應。在規則語料中，BG 貢獻度高，PPh 層貢獻低；不規則語料中，PPh 貢獻度高而 BG 貢獻度低；而半規則語料的 PPh 及 BG 貢獻度則介於規則跟半規則間。兩者互補的結果，總預測率曲線則趨於平緩。圖七比較二發音人的三種古典文體語料古文語料 Ap 預測正確率，在特定的韻律層隨三種不同韻律格式(R、SMR、IR)變化曲線，(a) PPh 層 (b) BG 層 (c) PPh 層與 BG 層疊加後的結果，顯示不同韻律階層的差異(b)，加總後因相抵，最後呈現相當一致的趨勢(c)。



圖七、古典文體語料 Ap 預測正確率在特定的韻律層隨三種不同韻律格式(R、SMR、IR)

變化曲線，(a) PPh 層 (b) BG 層 (c) PPh 層與 BG 層疊加後結果。

各層的貢獻度在總預測率的比例分佈，我們以圖八圓餅圖表示。靛色表示 PPh 層的貢獻比例，紫色表示 BG 層的貢獻比例，而黃色表示 PG 層的貢獻比例。上列表示男性發音人，下列表示女性發音人，由左至右的分類分別為規則(R)，半規則(SMR)，不規則(IR)，最後則為比較用的氣象語料(WIR)。相對於韻律格式的規則性，二位發音人的語料都顯示，最上層 PG 的貢獻度並不明顯，我們認為，這是因為此三類古典文體的語篇長度偏短。不過，在 BG 層則顯示，語料格式越規則，BG 層的貢獻度越高，且 BG 貢獻度的比重隨語料格式的規則度由規則向不規則遞減。BG 的貢獻度在規則語料(R)分別是 28.1% 和 51.6%；半規則語料(SMR)分別是 20.4% 和 22.7%；不規則語料(IR)分別是 8.7% 和 16.5%；氣象撥報最低，分別是 5.6% 和 7.4%。



圖八、Ap 預測正確率在不同語者與不同韻律格式的分佈比例總比較

分析四種語料後的總體趨勢顯示：韻律格式的差異，可以從不同韻律階層對整體韻律輸出的貢獻度分佈具體表示。韻律格式越規則、內含樂律性越高時，上層資訊的貢獻度越高、韻律規劃的範圍越大。韻律格式越不規則，下層韻律韻律短句的貢獻度越高，表示韻律規劃的範圍越小。基於的結果，我們認為：鄭所提出的階層式韻律句群 HPG 架構，可系統性的解釋並產製不同的韻律格式，因此可視為語流韻律的基型，其作用及意義，與字調對應於音節、句調對應於短句相同。總括而言，HPG 是對應於多短語語流韻律的基型。

## 六、結論與未來展望

本文以韻律短句的基頻曲線變化為主要分析參數，聽感的韻律邊界停頓為輔，從文體工整性及內建的韻律性來檢視語料、剖析朗讀古典文體語料的韻律格式，並從階層式韻律句群及架構分析語流韻律中的韻律成分，我們得到的結果顯示，上層語篇訊息提供跨短語韻律語境，是語流韻律不可或缺的成分。韻律格式越規則，上層訊息的貢獻度越高。隨著工整性變化造成不同的貢獻度分佈的研究結果可知，語者在朗讀大量重複結構的過程中，會更強調語意的銜接與語篇的轉換效應，因此大範圍的上層語意資訊及語段、語篇的效應更明顯。更重要的是，韻律格式的不同，從 HPG 韻律階層貢獻度而言，只不過是不同的分佈模式而已，完全可系統性的表示。而氣象撥報語料在四種語料中規則性最低，分析結果顯示，上層效應越不明顯。因此不論古典文體或氣象撥報語料，都提供階層式多短語韻律句群架構 HPG 階層韻律貢獻的證據。本研究中基頻曲線的實驗證據證明了多短語韻律句群架構具有一定的強健性，適用不同語體的語料，因此以 HPG 做為長篇語段的基本架構，但具有一定的自由度，可依照語體而變化調整，若能適當的運用 HPG，只需一種跨短語基型，藉由操弄 HPG 各層的貢獻度，便能達到韻律格式轉換的結果，不但能將連續語流中韻律基本特性表現出來，更能使語音合成系統的韻律輸出更具豐富的變化性。

本研究限於篇幅，僅以韻律短句的基頻曲線變化為主要分析參數，不過，我們強調，韻律的研究並非僅止基頻，其他的聲學參數也具有韻律效應，全面的韻律研究，應考慮語音信號中所有的聲學參數，以及它們的互動模式。鄭較早的研究結果[2, 3]顯示，語流韻律的節奏，也是重要的聲學參數，她也獲得音節時長在時間層的分佈的效應，得到對應於 HPG 架構系統性的節奏模版。未來我們將繼續分析不同韻律格式的語料中，進一步求取語流韻律的節奏特性，並與基頻研究的結果整合，以期對語流韻律有更全面性的瞭解。

## 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] 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.
- [5] 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.
- [6] Mixdorff, H. "A Novel Approach to the Fully Automatic Extraction of Fujisaki Model Parameters", *Proceedings of ICASSP 2000*, vol. 3, pp.1281-1284, 2000.
- [7] 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.
- [8] Bu, S., Yamamoto, M. and Itahashi, S. "An Automatic Extraction Method of F0 Generation Model Parameters", *Trans. IEICE, TRANS. INF. & SYST., VOL.E89-D, NO.1*, Jan 2006.
- [9] Nakai, M., Singer, H., Sagisaka, Y. and Shimodaira, H. "Automatic prosodic segmentation by F0 clustering using superpositional modeling" , *Proceedings of ICASSP95*, pp. 624-627, 1995.
- [10] Nakai, M., Shimodaira, H. and Sagayama, S. "Prosodic Phrase Segmentation Based on Pitch-Pattern Clustering", *Trans. IEICE, J77-A, 2*, pp. 206-214, Feb. 1994.
- [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.

## 附註一、古典文體三種分類範例

### 1. 規則(R)

古詩十九首之九（漢代五言古詩）

庭中有奇樹，綠葉發華滋。  
攀條折其榮，將以遺所思。  
馨香盈懷袖，路遠莫致之。  
此物何足貴，但感別經時。

### 2. 半規則(SMR)

將進酒（唐代樂府 李白）

君不見黃河之水天上來，奔流到海不復回；君不見高堂明鏡悲白髮，朝如青絲暮成雪。  
人生得意須盡歡，莫使金樽空對月。  
天生我材必有用，千金散盡還復來。  
烹羊宰牛且爲樂，會須一飲三百杯。  
岑夫子，丹丘生，將進酒，杯莫停。  
與君歌一曲，請君爲我側耳聽。  
鐘鼓饌玉不足貴，但願長醉不願醒。  
古來聖賢皆寂寞，惟有飲者留其名。  
陳王昔時宴平樂，鬥酒十千恣歡譔。  
主人何爲言少錢，徑須沽取對君酌。  
五花馬，千金裘，呼兒將出換美酒，與爾同消萬古愁。

### 3. 不規則(IR)

禮運大同篇（先秦散文）

大道之行也，天下爲公；選賢舉能講信修睦。  
故人不獨親其親，不獨子其子；使老有所終，壯有所用，幼有所長，矜、寡、孤、獨、廢疾者皆有所養。  
男有分，女有歸。  
貨惡其棄於地也，不必藏於己；力惡其不出於身也，不必爲己。  
是故謀閉而不興，盜竊亂賊而不作；故外戶而不閉。  
是謂大同。