

語篇的基頻構組與語流韻律體現*

鄭秋豫

中央研究院

本研究從語篇組織角度重新定義語篇韻律單位，依「階層式多短語語流韻律 HPG 架構」(Hierarchical Prosodic Phrase Grouping) 之層疊與覆蓋關係，解構不同韻律格式語料，以音高的構組為例，證明口語語流的基頻聲學體現，乃由字調、韻律詞調、句法調、語篇聯繫位置、邊界延長、邊界停頓等成分系統性共構，並以各級韻律單位的相鄰與跨單位之對比與反差，證明較大韻律單位與語篇邊界效應亦為韻律語境之構組成分。各級韻律單位對韻律生成均有貢獻，層級貢獻度可透過統計分析釐清，不同韻律格式之體現反映於層級貢獻度的分布差異。研究方法則凸顯語料庫語言學及計算語言學解析動態超音段聲學參數相對性的重要與意義。

關鍵詞：語流韻律組織，韻律單位，韻律語境，相對韻律資訊，基頻，邊界效應，線性相鄰，跨單位呼應，韻律階層

1. 引言

語流中各級語音及韻律單位的實際體現，充滿多元的變化：有時清晰完整，有時多變而片面。本文旨在分析基頻信號，解釋表面的多變其實是有所本且有系統規則可循的，單字調與短語句調構成語段時必須變化。本研究提供語篇韻律語境的成分及各級韻律單位的互動證據，從聲學層面證明語段的存在，進而論述語篇韻律的特徵及其語言學上的意義。

* 本研究為國科會研究計畫「新一代語音科學與技術——由基礎到應用」部分研究成果（計畫編號：NSC 96-2218-E-002-001，2005.08~2008.05）。

參與研究人員鄭雲卿（2000 迄今）、李瑋珊（2000 迄今）、蘇昭宇（2005~2009）對本研究貢獻良多，謹此一併致謝。

本文部分內容曾發表於以下二篇會議論文：(1)〈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」，2007 年 9 月 6-7 日），及 (2)〈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〉（「第二十屆自然語言與語音處理研討會」，2008 年 9 月 4-5 日）。

我們強調，連續說話時所體現的語流韻律實為語篇韻律，語流韻律反映的是語篇的心理語言規劃及結構，涉及因素除語言即時產製的單位與組織外，至少還必須包括生理語音的限制及二者必須的互動。從語篇的階層式結構切入研究語流韻律，重新界定語篇韻律的單位，分析每一個韻律單位本身構成與語篇結構的互動的階層意義，才可釐清並解釋為什麼音段、音節、字調以外還有上層的韻律單位，進而進一步找出造成超音段成分的不完整與多變的重要成因。本論文將以此前提，採語料庫語音學的研究方法，解析語流語音信號裡的基頻，再建構基頻體現音高變化的構組成分及意義，解釋語流韻律輸出的原因。

語篇 (spoken discourse) 可包括一個以上的語段，語段可包括一個以上的短語，因此語篇與語段也都是韻律單位。韻律階層裡每一韻律層級所構成的韻律語境，都同時包括線性串接平滑及上層覆蓋，也就是相鄰單位的連接和跨單位的呼應，若僅考慮相鄰單位線性的串接平滑，則既不足以解釋語流韻律的成因，也不足以表達語流韻律的體現。這些階層式的關係及連接覆蓋的規則都具有特定基型，在物理信號上造成部分單位充盈或縮減的調整，共構表意韻律語境，造成語流韻律的體現。我們認為，語者與聽者依據來自語篇的大範圍表意韻律訊息，以同樣的基型產製編碼及接收解碼，結合已產生變異的字調句調，共同達到溝通的目的。因此，從語篇組織切入語流韻律，提供了一個由上而下檢視語料的視角，使所有檢視的韻律單位與韻律現象，都必須考慮單位本身與來自上層單位的管轄制約以及對下轄單位可能造成的影響。由上而下的管轄制約，並不限於一對多的關係；既然有一對多的管轄，所有的單位都不能僅接受個別及相鄰的檢驗而不考慮跨單位的相關。因此，階層結構式的相關，必然同時包含了同層級單位的相鄰線性互動，及來自上層管轄的跨單位呼應互動。所以，我們認為語篇韻律研究首要之務，是重新界定語篇韻律單位及語流韻律現象涉及的結構來源。

九〇年代至今的韻律研究成為熱門研究議題，主要以語音學與音系學的介面為主，二者均採由下而上的角度，著重小單位語音現象的分析，不但鮮有從語篇組織的角度檢視短語成段的韻律或語篇韻律，也鮮有在語音單位、音系單位、與韻律單位的定義間重新詳加釐清。舉例來說，語音學的研究，最主要的單位是音段 (segment)，包括輔音與元音 (Scott 1941, Pike 1943)，而句調或語調的單位，則從一段語音在結尾時是否有音高下降拖延及能量減弱的現象來判定 (Crystal 1969)。韻律的研究以音段的高低長短變化為主，但一向不是語音學研究的重點。從音系學出發的韻律研究，在韻律單位方面可略分為純音韻律與兼顧句法（或語法）的韻律，前者以 Selkirk (1984) 為代表，後者以 Nespor & Vogel (1986) 為代表。專注於句調討論的則以 Pierrehumbert (1980) 為主。純音韻律單位有音節 (syllable)、

音步 (foot)、韻律詞 (prosodic word)、大小韻律短語 (major and minor prosodic phrase)、語調短語 IP (intonation phrase)、話語 (discourse) 等，討論的重心多為音步、韻律詞，少有比韻律詞大的討論；兼顧句法屬性的韻律單位則有莫拉 (mora)、音節、音步、韻律詞、黏附組 (clitic group)、韻律短語、語調短語 IP、語篇等等 (Nespor & Vogel 1986)，討論的重心也多以音步、韻律詞為主。在通行多年的語氣與停頓標註系統 ToBI (Tone and Break Indices) (Beckman & Hirschberg 1994, Beckman & Elam 1997) 中，也僅將短語 (phrase) 分為主要語調短語 (intonational/major phrase) 和次要語調短語 (intermediate/minor phrase)，以短語為分析句調 (sentence intonation) 的單位。也就是說，一般的韻律分析中經常是將語音單位與音系單位混用的 (Halliday 1967, Crystal 1969)，以致在使用語料從事聲學語音分析時，單位混淆十分常見。

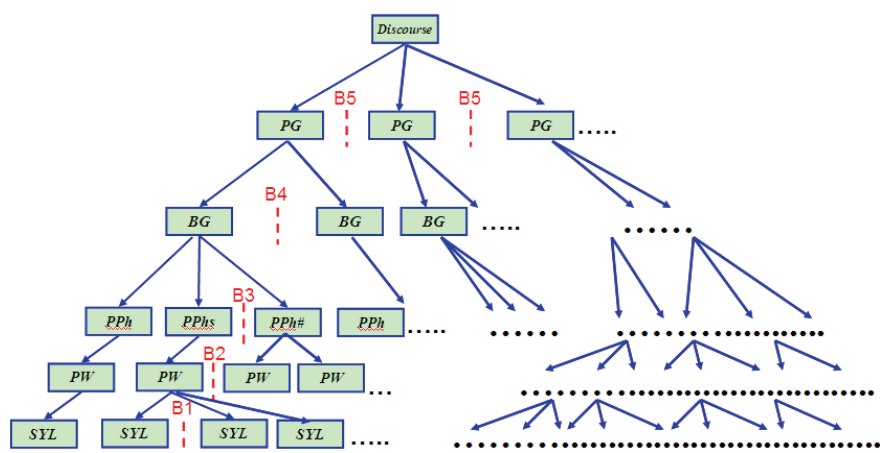
未釐清以上這些單位的結果，導致音系單位 IP 既是最大的語音單位，也是最大的韻律單位，但 IP 的身分不明，也沒有清楚界定的基本模版。試問 (1) 沒有句法訊息時（如 Selkirk 1984），如何界定主要短語及次要短語？它們是有從屬關係的語調短語？或是各不相關的語調短語？此時一個 IP 可以包括多少短語？是否有基型？(2) 有句法訊息時，語調 (intonation) 單位¹ 最大為何？複雜句 (complex sentence) 與複合句 (compound sentence) 的語調如何分析？一個 IP 又可以包括多少短語？句法分析節點與韻律節點不符時如何解釋及解決？此外，以句法訊息界定 IP 時，最大單位止於句子 (sentence)，篇章 (discourse)² 訊息是無法考慮的，未能將 IP 的單位擴大至篇章單位時，便無法也無力處理、解釋語篇裡相同的字調為什麼變化多端？結構類似的短語，為什麼會呈現多樣的韻律體現？而漢語的韻律研究受到上述研究的影響，一方面必須考慮漢字的字調，一方面又因複雜句與短語句段的界線不明，於是，韻律研究的取向變成字調、短語句調為主的研究 (Xu 1999, 2004, Liu 2009)。

可是，為什麼在連續說話的口語語料中，往往字調和句調的調型都有許多變異？一方面並不是每一個音節都有清楚可辨認的字調，另一方面句調又往往沒有清晰的下傾或上揚。本研究將分析國語語篇語料，整合聲學語音信號基頻 (fundamental frequency F0) 的分析，進一步在字調、韻律短語句調以外，加入表語段開始、延續與結束的三個表語篇關連的因素，提取並證明各層級的韻律單位

¹ 本文以句調 (sentence intonation) 一詞，表示根據句法界定為句子的單位所產生的調型；以語調 (intonation) 一詞，表示根據聲學語音信號具有可辨識的基頻走勢所產生的調型，二者的單位並非一定相等。

² 本文以語篇 (spoken discourse) 一詞，表示從聽感獲得的篇章單位；以篇章 (discourse) 一詞，表示與聽感無關，從文本與標點獲得的單位，二者的單位並非一定相等。

對最終的韻律輸出體現皆有貢獻，而非僅是字調連讀、及變調短語句調所能解釋。說明語流體現的基頻的構組成分，必須將語篇成分納入，才能理解字調、句調必變的原因。我們採用「階層式多短語語流韻律 HPG 架構」(Tseng et al. 2004, 2005, Tseng 2006, 鄭秋豫 2008, 見〈圖 1〉)，依該架構之韻律層級由下而上，將漢語的語篇韻律單位定義為音節 Syl (syllable)、韻律詞 PW (prosodic word)、韻律短語 PPh (prosodic phrase)、呼吸組 (breath-group) 及多短語韻律句群 PG (prosodic phrase group, 即語段) 共五級，由小而大的韻律單位；相對應的邊界停頓也為五級，分別為 B1, B2, B3, B4 及 B5。語篇韻律單位由音節 Syl 開始，它是一個漢語的字調單位也是漢語的節奏(時間)單位。韻律詞 PW 在聽感上不受音節個數限制，是韻律區塊的單位。同理韻律短語 PPh 是一個 IP，通常受句法結構規範，直接互動的下級單位不是 Syl 而是 PW。而呼吸組 BG 原本是生理單位，但產製口語篇章時，連續的說話必須配合呼吸與換氣，因此呼吸組 BG 既是一個生理單位又是一個語篇心理語言規劃的韻律單位，包含一個以上的 IP；而韻律句群 PG 則為語段的韻律單位，並上承語篇下轄韻律短語 PPh。因此語篇由語段組成，語段由韻律短語組成，以此向下類推。依階層式多短語語流韻律 HPG 架構，韻律句群 PG 可由三個以韻律短語 PPh 為單位的相對韻律位置：PG 首 (PG-Initial)、PG 中 (PG-Medial) 與 PG 尾 (PG-Final) 所構成，表示口語段落的起始、延續和結束；而 PG 首和 PG 尾都是 PPh 單句，而 PG 中可以是一個或多個短語。換句話說，在 HPG 架構中，韻律短語 PPh 是一語篇關連的韻律單位而非不相關且獨立的韻律單位，且語篇的關連性包含相鄰單位的連接與跨單位的呼應。



〈圖 1〉「階層式多短語語流韻律 HPG」韻律單位架構圖

本研究一方面採語料庫語言學使用大批語料，一方面採計算語言學研究方式量化語料驗證：構組語流韻律基頻參數最終體現的成分中，(1) 並非僅來自字調與短語句調的調型，而是所有語篇韻律單位調型的相加與抵消；(2) 各級調型的相加與抵消並非僅包括單位串接平滑（或平順），還需包含上層訊息的覆蓋；(3) 各級韻律邊界對韻律單位所造成的影響，也必須在分析時加以考慮。最後我們將討論所獲結果在語言學上的解釋性意義。

2. 研究方法

2.1 以 Fujisaki Model 為基礎發展自動擷取國語基頻曲線特徵參數系統

2.1.1 Fujisaki Model 簡介

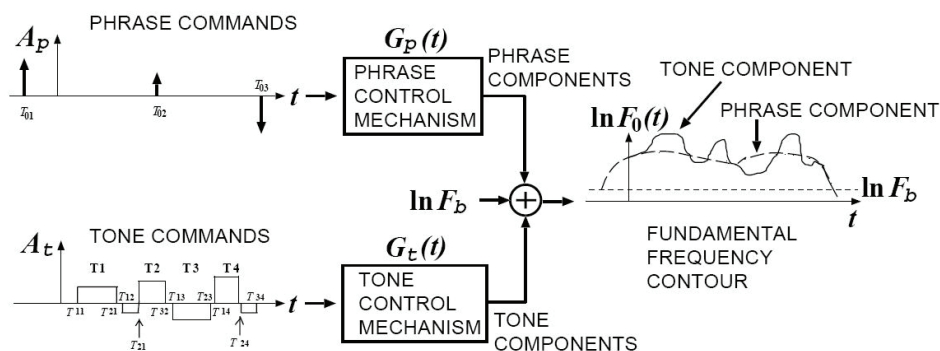
Fujisaki 等人 1984 年提出基於發聲機制的疊加式 command-response 基頻曲線模型，主張產生基頻主要器官為聲帶 (vocal tract) 及相鄰的隨意肌。控制隨意肌時，有一定物理特性的限制，對應於語言單位，在時間程上產生聲帶高低變化；所以基頻曲線並非任意直線或曲線，必須受制於一定的規則在時間程所產生的變化，而此限制顯現於對數尺度之中，此為基頻曲線的函數限制。此一 command-response 模型一般通稱為 Fujisaki Model，其特點在於將時間程上動態變化的不規則基頻曲線，拆解為三個單位大小不同的元件函數 (component) 疊加總合（如〈圖 2〉所示），並分別對應發聲器官的物理特性。此三元件函數分別為 (1) 短語句調元件 (Phrase Component Ap)，反映較大單位基頻曲線的控制與發聲限制；(2) 強調元件 (Accent Component Aa)，反映較小單位基頻曲線的控制發聲限制；與 (3) 基底頻率 (base frequency Fb) 代表基本音高。此模型定義如下：

$$\ln F_0(t) = \ln F_b + \sum_{i=1}^I A_{p_i} G_p(t - T_{0i}) + \sum_{j=1}^J A_{a_j} [G_a(t - T_{1j}) - G_a(t - T_{2j})]$$

i, j = Index of phrase command, Index of accent command
 F_b = Base frequency
 A_{p_i} = Phrase command magnitude
 A_{a_j} = Accent command magnitude
 $G_p(t - T_{0i})$ = Phrase command response function
 $[G_a(t - T_{1j}) - G_a(t - T_{2j})]$ = Accent command response function

原 Fujisaki Model 中 Phrase Component Ap 泛指短語句調或語調對全面性基頻

曲線造成的影響，其單位或範圍並未界定，一般應用是以短語句調為單位。Accent Command A_a 則泛指強調、加重語氣對局部基頻曲線造成的影響。此模型應用於非聲調語言如英語、德語時， A_p 用以描述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及下傾程度； A_a 則用來描述某些局部的加重或加強 (emphasis) (Mixdorff 2001, 2002)。此模型應用於聲調語言如國語時， A_p 不變，仍用來描述句調的高低走勢，即片語的語調或短語的句調；而 A_a 則用來描述每一個單音節的局部變化，以表現字調 (Mixdorff et al. 2003, Tseng & Pin 2004)。本研究分析國語連續語流的基頻走勢時，以 A_a 表示字調成分，以下稱作字調元件； A_p 表示短語句調成分，以下稱作短語調元件。

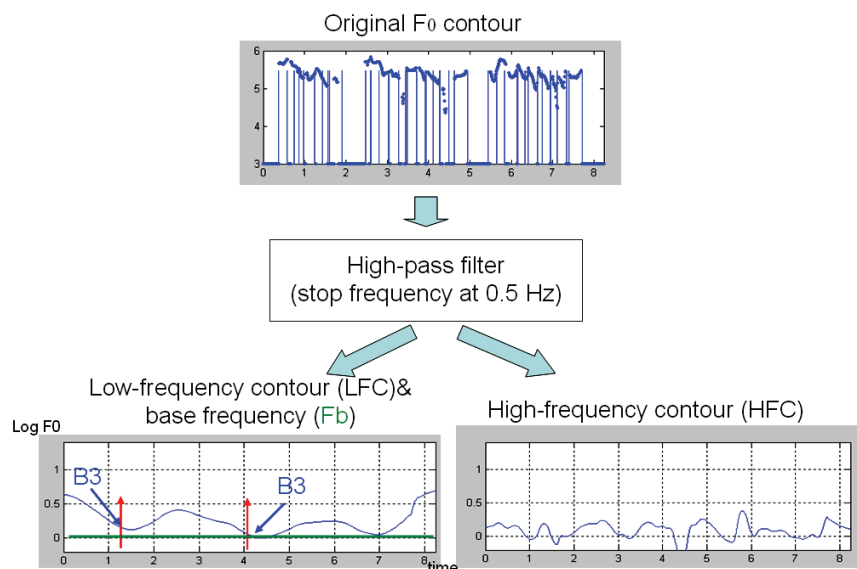


〈圖 2〉強調元件與基底頻率疊加後的基頻曲線 (Fujisaki & Hirose 1984)

2.1.2 自動擷取 Fujisaki 特徵參數之設計

若以 Fujisaki Model 的參數表示國語語流的基頻曲線呈現，陳述句句調呈現可視為開高後走低全面下傾、各局部字調起伏變化與基本頻率三元件疊加所成。根據 Fujisaki Model 之定義，相同的字調元件必須在語流中各位置保持大致相同的權重，因此文獻中應用 Fujisaki Model 解構聲調語的傳統的作法，多以目測原始基頻曲線，一次一句人工調整決定這三個元件的最佳組合，以達盡量趨近基頻的整體表現 (Gu et al. 2005)。但這樣的作法除相當耗費人力，也無法處理語音資料庫所收得的大批語料，因此我們採用了 Mixdorff 提出的方法 (Mixdorff 2000, Mixdorff et al. 2003) 解析原有的基頻曲線，並以程式抽取大量語料的 A_p 與 A_a 二參數值：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分即為語流的基本單位，對應 Fujisaki

Model 中的字調元件；而變化和緩的部分，則為語流中語調全面下傾的趨勢，對應 Fujisaki Model 中的短語調元件。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)，即為字調元件需逼近的目標曲線；(2) 扣掉高通部分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) 扣掉基底直線後的低通曲線 (LFC) 即為短語調元件需逼近的目標曲線。〈圖 3〉為 Mixdorff 自動擷取 Fujisaki 參數架構。



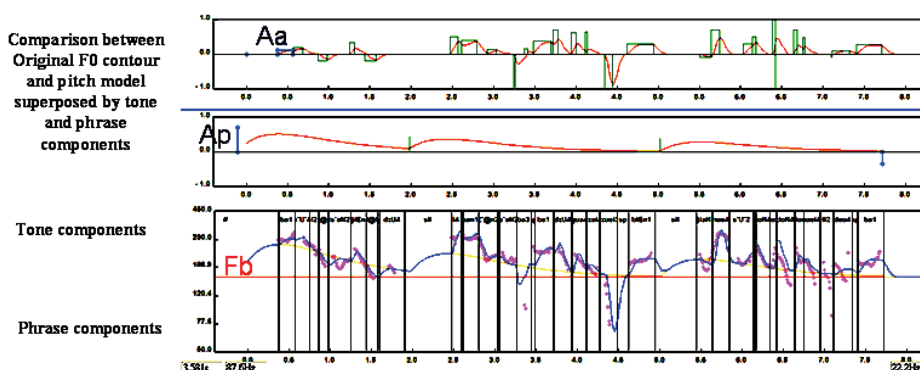
〈圖 3〉利用濾波器將基頻曲線分離為高通 (HFC) 及低通曲線 (LFC)，即分別為字調元件與短語調元件需逼近的目標曲線，進行自動擷取程序 (Mixdorff 2000)；圖中縱軸表時間，橫軸表 $\log F_0$ 。

傳統人工調整步驟 (1) 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趨近一音節內較複雜的基頻曲線 (Wang et al. 1999)。然而在本研究中，為了能夠自動擷取大量語料的參數，而非使用傳統人工逼近方式，我們在自動擷取 Fujisaki 特徵參數程式中，只採用一字調元件來趨近一音節內高頻的基頻曲線，抽取最概括性的字調變化。

2.2 加入聽感所得語篇韻律邊界的方法與自動擷取結果

在本研究中，我們除驗證 HPG 架構，也希望探討由聽覺感知所獲之語篇韻

律邊界與基頻曲線的關係，因此我們加入邊界資訊作為自動擷取 Fujisaki Model 的 A_p 參數的輸入參數。初步觀察顯示，局部基頻最低點通常就出現在韻律短語的聽感語篇韻律邊界上，因此加入聽感韻律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 4〉為加入聽感韻律邊界為輸入參數後，自動擷取 Fujisaki Model 的 A_p 參數結果。



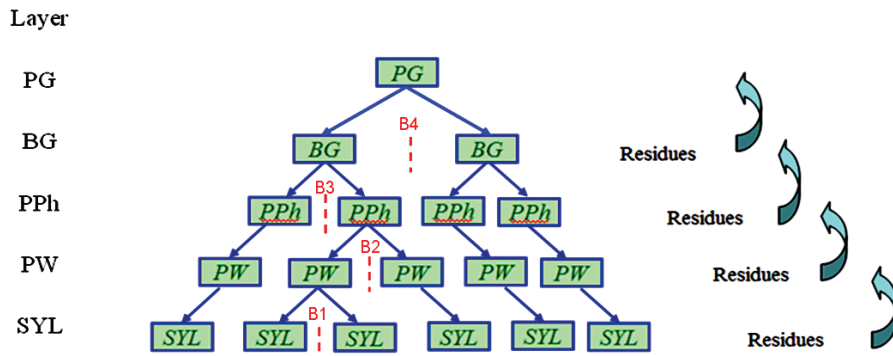
〈圖 4〉加入聽感的邊界資訊後自動擷取 Fujisaki 參數的結果

2.3 以階層式多短語語流韻律 HPG 架構為基礎的多元迴歸分析

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字調元件與短語調元件，分別對應 HPG 架構中字調與句調的成分。然而原始的 Fujisaki Model 僅針對短語句調及局部強調，並無對應韻律詞層、呼吸句群層與韻律短語層的元件函數。Nakai 等人在 1994、1995 年雖然以語音工程研究的慣例，提出將 A_p 歸為多類建立語流中的短語模型 (Simodaira & Nakai 1994, Nakai et al. 1995)，但歸類的方式並未有任何的理論基礎，將語流中的階層性關係列入考慮。因此我們參考結合階層性訊息的線性迴歸模型，找出各層的貢獻度與特徵的方法 (Keller & Zellner 1996, Zellner Keller & Keller 2001)，進行基於階層式多短語語流韻律 HPG 架構的多元迴歸分析，所獲致 HPG 階層的貢獻度，將有助釐清分層貢獻度的相對大小及對語篇韻律輸出的意義。

階層性線性模型是簡單線性迴歸的衍生，每個輸入都附有多層次的韻律邊界停頓標註，與聽覺感知所獲致語篇韻律邊界對應（見〈圖 5〉基於 HPG 架構之階層性多元迴歸分析示意圖）。每一標註分別代表每一韻律層的參數與特性，逐層分析步驟如下：(1) 分析及預測某一韻律層時，只利用該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) 將原始值與預測模型的差距（或殘差）視為實驗來自更上層的效應而非誤差，將分析及預測某一韻律層的殘差作為分析及預測下一韻律層的輸入，根據上一層的標註進行線性迴歸分析，以獲得上一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及疊加後的累加貢獻度。換言之，基於 HPG 架構之階層性多元迴歸分析，是利用每一層的標註資訊，得到各層韻律單位的模型，並依此模型進行正規化。



〈圖 5〉基於 HPG 架構之階層性多元迴歸分析示意圖 (Tseng et al. 2004)

3. 實驗設計與分析

本研究以 (1) Aa 為分析字調以上韻律詞意之特徵。(2) Ap 為分析韻律短語句調以上語段或語篇語意之特徵。

3.1 Aa 分析

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，預測參數包括相鄰字調基頻走勢的線性串接及來自上一韻律層 PW 的管轄與覆蓋。即以每一字調 $X \text{ tone}$ 為基準，預測其與前後字調共構的線性字調語境 ($\text{Preceding Tone} + X \text{ Tone} + \text{Following Tone} = \text{Linear Tone Context}$)；根據 HPG 架構，除了前後字調形成的線性字調語境外，上一層的韻律詞層 (PW) 也影響了下一層音節，韻律詞層對字調的預測參數包括韻律詞邊界 (PW Boundary Info) 及每一音節在所屬韻律詞內的位置順序 (PW Position Sequence)。根據我們對較高層級邊界 (BG/PG) 效應的研究中發現（鄭秋豫、蘇昭宇 2007, 2008），下層的韻律單位，

是否位於高層級韻律邊界前後，對於聲學信號有很明顯的影響，因此我們也將邊界效應加入考慮，從音節層起，以是否位於邊界前為基準，分為二類分析。考慮的上層邊界包括：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，從 (1) 音節層 SYL，(2) 韻律詞層 PW，(3) 韻律短語層 PPh 以上對音節的邊界效應，進行獨立的 Aa 類別分析，其數學函數如下：

$$(1) Aa = f(\text{FollowingTone}, \text{PrecedingTone}, \text{CurrentTone}) + \text{Delta1}$$

$$(2) \text{Delta1} = f(\text{PW Boundary Info}, \text{PWSequence}) + \text{Delta2}$$

$$(3) \text{Delta2} = f(\text{PPh Boundary Info}, \text{PG Boundary Info}) + \text{Delta3}$$

3.2 Ap 分析

Ap 是對較大範圍韻律單位的基頻走勢的預估，雖然 Fujisaki 的原定義並未界定 Ap 的分析單位大小，我們以短語句調 phrase intonation 及 IU 的定義為參考，將 Ap 定義為一個 IU 基頻特徵的單位，亦即 HPG 系統中的韻律短語 PPh，因此分析 Ap 的步驟由階層式架構中韻律短語層 (PPh) 開始，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：(4) PPh 層：以目前 PPh 長度 (Current PPh Length)、其前一 PPh 長度 (Preceding PPh Length) 與後一 PPh 長度 (Following PPh Length) 的組合做為分析參數，預測其與前後短語調共構的線性短語調語境，經過線性迴歸後的殘差定義為 Delta1，輸入 BG 層進行下一步分析；(5) 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；(6) PG 層：與 BG 層輸入參數相同。其數學函數表示如下：

$$(4) Ap = f(\text{FollowingPPh_Length}, \text{PrecedingPPh_Length}, \text{CurrentPPh_Length}) + \text{Delta2}$$

$$(5) \text{Delta1} = f(\text{BGSequence}) + \text{Delta2}$$

$$(6) \text{Delta2} = f(\text{PGSequence}) + \text{Delta3}$$

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。

4. 實驗一：口語語流韻律中句調成分與字調成分對應於階層式多短語語流韻律 HPG 架構的意義

4.1 實驗目的及假設

本實驗以台灣地區國語語篇口語的字調及韻律詞基頻曲線模型，根據「階層式多短語語流韻律 HPG」理論架構，由下層到上層，將基頻曲線模型參數與 HPG 韻律階層結合，以驗證 (1) 句調成分與字調成分對應各韻律階層的貢獻度；(2) 字調成分在各階層韻律單位管轄下，基頻模型參數如何變化。

4.2 實驗語料

文本部分，採用 (1) 古典文體（以下簡稱 CL，大約 1,600 個音節），與 (2) 長篇敘事段落文本（以下簡稱 CNA，大約 6,700 個音節），共計 (1) 26 篇古典文體段落（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。

語料來自四位發音人，二類文本各由一男一女發音人朗讀。古典文體文本由一男一女 (m056 & f054) 未受特別訓練之母語者朗讀，在隔音室使用 Sony ECM-77B 迷你型麥克風及 Cool Edit 2000 錄音軟體錄製；白話敘事段落由一男一女播音員 (m051 & f051) 朗讀。〈表 1〉為 HPG 架構下古典文體 CL 與長篇敘事段落文本 CNA 語料的韻律邊界以及相對應韻律單位次數分配表。

〈表 1〉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次數分配

語料	發音人	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

平均語速分析結果顯示：CL 男性發音人 m056 為 202m/syl，女性發音人 f054 為 265ms/syl；CNA 男性發音人為 199ms/syl，女性發音人為 189m/syl。二組語速顯示：文本的格式與語速的快慢有直接的關係，與發音人的性別則無。語速與文體的正相關，亦可作為語篇規劃與語篇組織正相關的佐證。

4.3 實驗結果

4.3.1 句調成分與字調成分對應各韻律階層的貢獻度

依據第二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，所有韻律層的迴歸係數的總和代表累積正確率之百分比，百分比越高表示預測的正確性越大。

4.3.1.1 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 (Tseng & Su 2008)，包括此音節前後是否有韻律短語邊界及韻律句群層邊界。結果顯示，音節層的字調貢獻度的範圍是 38%-46%；加入字調的前後相鄰字調共構的線性字調語境後，累積貢獻度提升為 45%-55%。加入 PW 層的韻律邊界訊息後，累積貢獻度提升至 48%-60%；加入 PW 層在所屬韻律詞的位置順序後，累積貢獻度提升至 51%-67%（見〈表 2〉）。從 PPh 層開始，再加入 PPh 層以上來自 BG 與 PG 層的邊界效應對字調的貢獻，分別計算，結果見〈表 3〉。PPh 層的字調預估加上 BG 層的邊界訊息後，累積貢獻度提升至 54%-73%；最後加上 PG 層的邊界訊息後，累積貢獻度即 Aa 的正確率提升至 56%-74% 不等。結果顯示，大部分的字調貢獻度雖來自音節層與韻律詞層，但在 Syl 層不及 50%；值得注意的是，邊界效應的貢獻度介於 5~7%，不容忽視。

〈表 2〉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 3〉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%

4.3.1.2 短語調元件 Ap 分層貢獻度

我們以 HPG 架構，分別在韻律短語層、呼吸句群層與韻律句群層進行階層性多元迴歸分析，求出 Ap 的累積貢獻度。結果顯示（見〈表 4〉）各層語篇資訊對韻律輸出的貢獻度因語料而異，二者的階層貢獻度相反。韻律短語 PPh 在古文語料 CL 及白話長篇敘事語料 CNA 貢獻度大不相同。CNA 的正確率多來自與前後短語調共構的線性短語調語境，在 80% 以上，表示韻律短語句調較為分明，而且在二發音人間並無差異（80% 與 81%）。而 CL 的 PPh 貢獻度不但較低，且有相當的個別差異（59% 及 38%），男發音人的線性短語調語境甚至只占了階層貢獻度 38%；由〈表 4〉中也可發現，CL 與 CNA 最大的不同是來自上層語篇貢獻度，表示發音人在朗讀內在韻律架構分明的文本時，上層韻律資訊的規劃及表現較白話明顯。

〈表 4〉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 5〉。結果顯示音節字調的調值對 Aa 的貢獻度較短語調的調值對 Ap 的貢獻度為大。不過，前文已提到，語流韻律輸出是層層範圍不同的調值堆疊而成，因此自 4.3.2 節起，我們將討論如何從音節 Syl 層起，建立每一韻律層的韻律模型，及各韻律層如何疊加，共構韻律輸出。

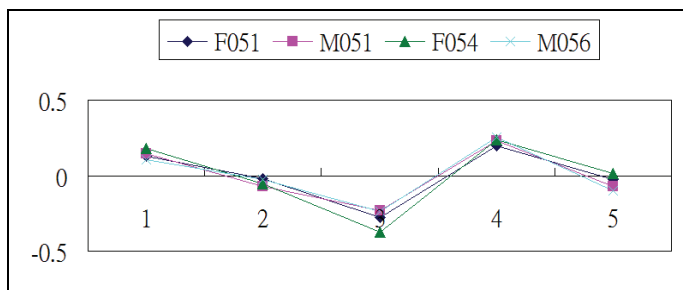
〈表 5〉套用 HPG 架構 Ap 與 Aa 對總體基頻曲線模型預測正確率

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%

4.3.2 Aa 對應音節層與韻律詞層模型

4.3.2.1 字調（音節層）模型

相對於每個字調的 Aa 模型如下圖（〈圖 6〉）所示，結果顯示對應五個國語字調（輕聲以第 5 調表示）的 Aa 模型，在不同語料間相當一致。

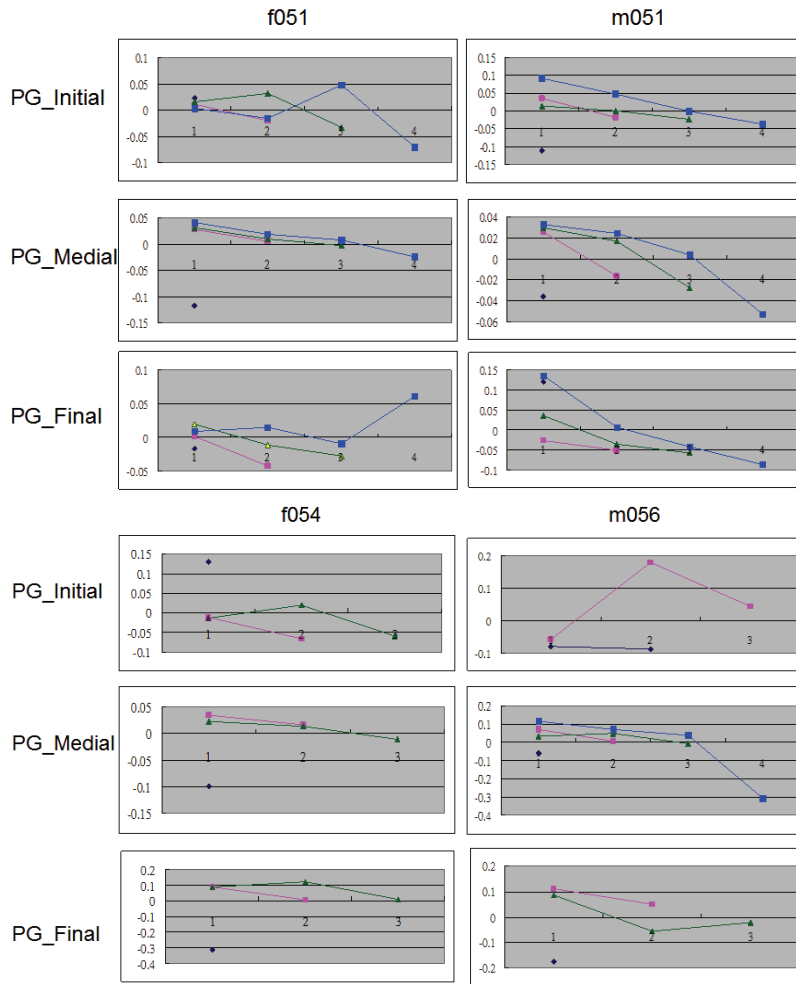


〈圖 6〉自動擷取出的字調元件對應各個字調之模型，縱軸為 Aa 值，橫軸為國語五個字調類別。

4.3.2.2 韻律詞模型

1. 韻律詞 PW 受韻律句群 PG 管轄之特徵

將除去字調效應後的韻律詞 PW 提取 Aa，獲得基頻模型依照 PG 的首中尾 (PG-Initial, PG-Medial, PG-Final) 三個位置分類後，我們發現 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。

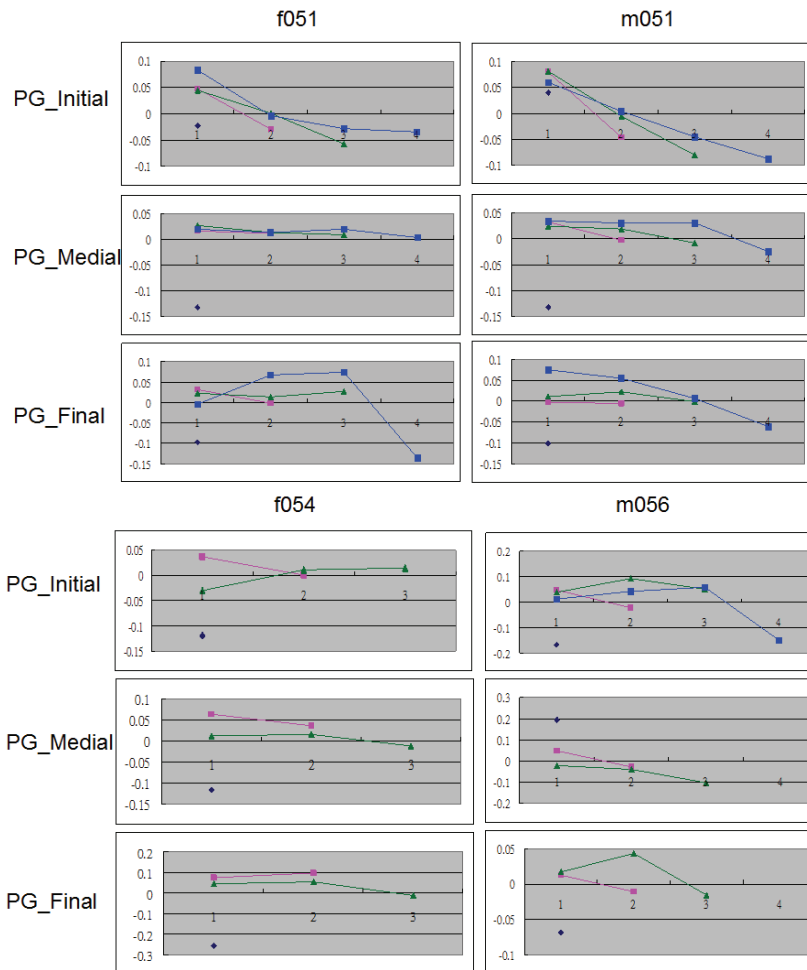


〈圖 7〉扣除字調成分，韻律詞模型在不同 PG 位置的特徵，每條曲線表示特定長度的韻律詞模型，橫軸表示在此韻律詞內音節順序，縱軸表示扣除字調成分後的殘差。

2. 韻律詞 PW 受韻律短語 PPh 管轄之特徵

同理我們將扣除字調成分的韻律詞 PW 模型依照韻律短語 PPh 位置分類後，得到各韻律詞的特徵（見〈圖 8〉），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 最為一致，而 PPh-Initial、PPh-Medial 最重要的特徵也大致發生在詞尾 Aa。從 PPh-Medial 位置的韻律詞表現出一致性可以看出，PPh-

Medial 最不受語體及相對位置的影響。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 在基頻曲線上的邊界效應，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 (Tseng et al. 2005, Zellner Keller & Keller 2001)。即 PPh 的邊界效應，主要表現於邊界前短語的末尾音節，在時長上有明顯的增加。



〈圖 8〉扣除字調成分後，韵律词模型在不同 PPh 位置的特徵，每條曲線表示特定長度的韵律词模型，橫軸表示在此韵律词內音節順序，縱軸表扣除字調成分後的殘差。

以上結果顯示：字調元件 Aa 分析中，音節層貢獻度的範圍是 38%-46%；加入字調的前後相鄰字調共構的線性字調語境後，累積貢獻度提升為 45%-55%。加入 PW 層的韻律邊界訊息後，累積貢獻度提升至 48%-60%；加入 PW 層在所屬韻律詞的位置順序後，累積貢獻度提升至 51%-67%（見〈表 2〉）。PPh 層的字調預估加上 BG 層的邊界訊息後，累積貢獻度提升至 54%-73%；最後加上 PG 層的邊界訊息後，累積貢獻度即 Aa 的正確率提升至 56%-74% 不等。由此可知，儘管大部分的字調貢獻度主要來自音節層與韻律詞層，但在 Syl 層不及 50%，仍需考慮邊界效應對字調的貢獻度。

短語調元件 Ap 分層貢獻度分析結果，對照先前研究的分析結果顯示，當每一個 PPh 被獨立分析時，結果顯示：正確預測只有 46.49%，53.51% 是殘差。在考慮韻律階層中上層 PG 效應後，預測值改善了 16.03%，累積預測值正確率為 62.52%（鄭秋豫 2008）。本實驗分析結果更進一步指出，各層語篇資訊對韻律輸出的貢獻度因語料而異。從〈表 4〉中可發現：韻律短語 PPh 在古文語料 CL 及白話長篇敘事語料 CNA 貢獻度大不相同。CNA 的正確率多來自與前後短語調共構的線性短語調語境，在 80% 以上，表示韻律短語句調較為分明，而且在發音人間並無差異（80% 與 81%）。而 CL 的 PPh 貢獻度不但較低，且有相當的個別差異（59% 及 38%），男性發音人 m056 線性短語調語境甚至只占了階層貢獻度 38%。因此，語流韻律的體現，並非僅由字調的連接與短語句調的串接所表現，還包括了上層韻律資訊的規劃。

5. 實驗二：不同韻律格式驗證階層式韻律架構

5.1 實驗目的與假設

本文以韻律短語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 (1) 文體工整性，及 (2) 內建的韻律性來檢視、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，驗證不同韻律格式實為 HPG 階層式貢獻度的不同分布的反映，韻律格式越規則，上層韻律單位的貢獻度越大。

5.2 實驗語料

除實驗一所使用的語料 CL 外，我們在實驗二加入 34 則氣象播報的語篇段落，並依文本結構的工整規則性，分為四類韻律格式：1. 規則 (R)、2. 半規則

(SMR)、3. 不規則 (IR) 與 4. 氣象播報 (WIR)，分類範例見附錄。古典文體文本包括三種韻律格式 (R, SMR, IR)，而氣象播報文本亦屬不規則類，但另以一類計 (WIR)，以示與古典文體有別，並可與古典文體比較。前三類的規律性區隔在內建的韻律格式，氣象播報文本雖無特定韻律格式，卻因其內容的變異較一般白話為少而略具規則性。我們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，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。氣象語料亦含男女發音人各一人，女性發音人同實驗一，男性不同 (m054)。各朗讀語料 34 篇，錄音環境、使用麥克風及錄音軟體同實驗一。〈表 6〉為古典文體語料 CL 與氣象播報語料 WIR 的音節數、韻律短語數、語篇數及平均語速統計。

〈表 6〉古典文體 CL 與氣象播報 WIR 語料的音節數、韻律短語數、語篇數及平均語速

韻律單元數/平均語速 語料庫/發音人		音節數	韻律短語數	語篇數	平均語速 (毫秒/音節)
CL	f054	3502	710	26	271
	m056	3510	711	26	202
WIR	f054	7054	720	34	193
	m054	7096	747	34	165

從〈表 6〉可看出，本研究語料設計的主要控制變項是韻律短語和語篇的數目，而非音節的數目。古典文體語料只有三千五百餘音節，氣象播報語料則有七千餘音節，二者差異頗大；但若以韻律短語和語篇的數目來看，古典文體語料包含韻律短語七百餘、語篇二十六，氣象播報語料則亦含韻律短語七百餘、語篇三十四，數目差異並不大。

5.3 實驗結果

5.3.1 聽感標註的重疊性

我們分析古典文體語料中二位發音人 (f054 與 m056) 聽感標註韻律邊界的重疊性後，將結果列於〈表 7〉。〈表 7〉中 B1、B2、B3、B4、B5 分別代表音節

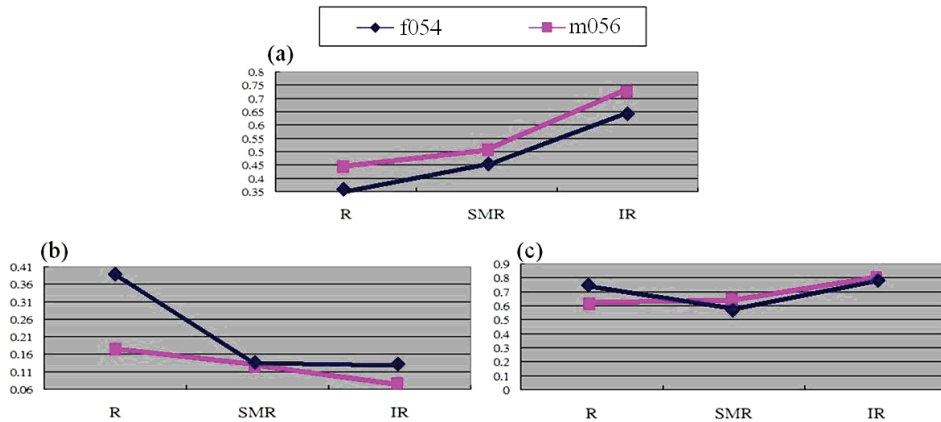
SYL、韻律詞 PW、韻律短語 PPh、呼吸句群 BG、語段（韻律句群 PG）之邊界。

〈表 7〉比較二語者朗讀古典文體韻律邊界的一致性結果

發音人 \ 韻律邊界	SYL/B1	PW/B2	PPh/B3	BG/B4	PG/B5
f054	97.98%	86.46%	82.79%	76.76%	62.30%
m056	96.71%	88.69%	80.19%	76.76%	80.85%

5.3.2 Ap 分析

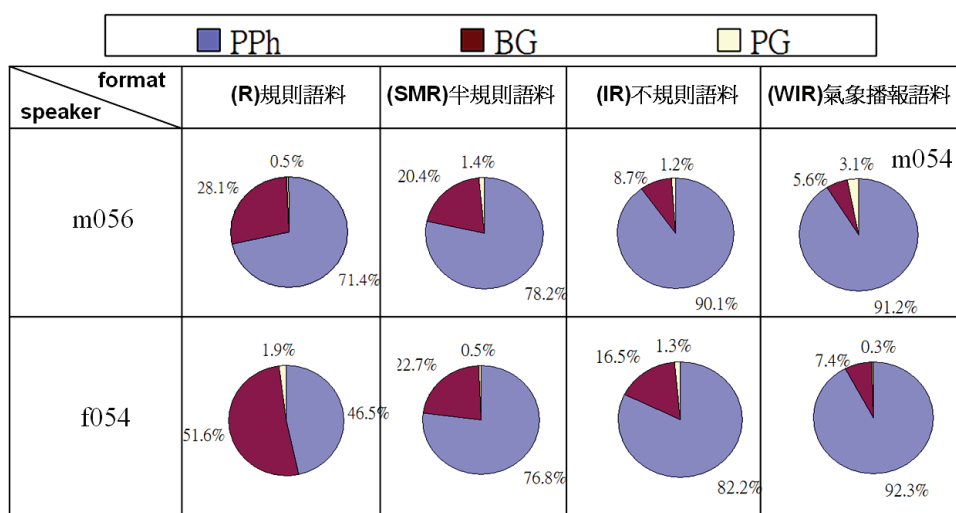
〈圖 9〉比較二發音人的三種古典文體語料 Ap 預測正確率，在特定的韻律層隨三種不同韻律格式 (R, SMR, IR) 的變化曲線，(a) PPh 層 (b) BG 層 (c) PPh 層與 BG 層疊加後的結果，顯示不同韻律階層的差異（見 (a) (b)），加總後因相抵，最後呈現相當一致的趨勢 (c)。



〈圖 9〉古典文體語料分別為 (a) PPh 層 (b) BG 層 (c) PPh 層加 BG 層的 Ap 正確率，橫軸表三種不同韻律格式 (R, SMR, IR)，縱軸表 Ap 的預測正確率。

各層的貢獻度在總預測率的比例分布，我們以〈圖 10〉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。上列表男性發音人，下列表女性發音人，由左至右的分類分別為規則 (R)，

半規則 (SMR)，不規則 (IR)，最後則為對照比較的氣象語料 (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 PPh 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%；氣象播報最低，分別是 5.6% 和 7.4%。



〈圖 10〉Ap 預測正確率在不同語者與不同韻律格式的分布比例整體比較

以上結果顯示：以 HPG 架構分析不同語料，韻律格式的不同，可由 HPG 階層式貢獻度的分布不同直接反映，韻律格式越規則，語篇韻律的規劃範圍越大，上層韻律單位的貢獻度也越大。HPG 架構可以視為一個語篇韻律的基型模版，因韻律格式而異的體現差異，其實只是韻律層級貢獻度的不同而已。

6. 實驗三：韻律短語的整體基頻高低與語篇組織的對應

6.1 實驗目的與假設

我們假設：口語連續語流中，每一韻律層都需將相鄰單位平滑與跨單位的呼應調整以形成並呈現必要的韻律語境。本實驗將檢視較大的韻律單位：韻律短語

PPh 的相鄰與跨單位關係，以 Fujisaki 的 command-response model 擷取韻律短語 PPh 整體基頻，分析語篇組織中的「韻律短語 PPh 整體基頻高度組型與調整」，呈現整體基頻組型變化，驗證此單位的整體基頻的相對高低是否與語篇組織對應，並討論其語言學意義。

6.2 實驗語料

同實驗一。

6.3 實驗過程

6.3.1 韻律短語 PPh 作為韻律句群 PG 特定的語篇單位

在前文已提及，不同位置的韻律短語 PPh 可表示口語段落的起始、延續和結束，是一語篇關連的韻律單位，而語篇的關連性可由相鄰韻律單位的連接與跨單位的呼應所構成，因此，韻律短語 PPh 在語段的位置、身分與功能亦可用來檢驗語篇組織。

6.3.2 韻律短語 PPh 與韻律句群 PG 位置標記

將語料中已標記的韻律短語 PPh 按以下所列的韻律句群 PG 序列位置：PG 首、PG 中與 PG 尾，分成三個相關的組別：

$$(7) \text{ PG-position} = \begin{cases} \text{PG-Initial} & \text{when sequence index}=1 \\ \text{PG-Final} & \text{when sequence index}=M \\ \text{PG-Medial} & \text{otherwise} \end{cases}$$

$$M = \text{Number of PPh in PG}$$

6.3.3 韻律特徵：基頻高度的擷取

擷取韻律短語 PPh 基頻高度的超音段聲學特徵。擷取特徵按韻律句群 PG 序列位置加以比較，以瞭解跨語料—語者間是否表現出不同但一致的整體組型。接著，進行統計分析，以瞭解在韻律句群 PG 不同位置的韻律短語聲學特徵差異是否顯著。

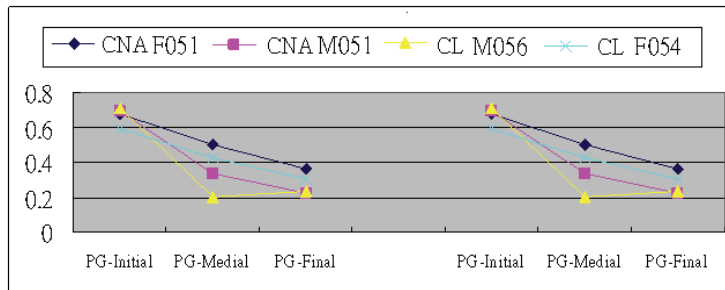
以 command-response model 擷取韻律短語 PPh 整體基頻

因為本實驗焦點放在整體基頻組型，並不涉及局部字調曲線，因此只擷取 Ap command 值。Ap command 的自動擷取方法參酌 Mixdorff 等人 (2003) 分析方法另行修改以適合本實驗。

6.4 實驗結果

6.4.1 韻律句群 PG 內與韻律句群 PG 間相對基頻高度

〈圖 11〉為跨語料—語者相鄰 PG 序列位置 Ap 平均值的例子，顯示兩韻律句群 PG 間相對基頻高度的調整。我們注意到：韻律句群 PG-Initial 的基頻重設，緊鄰著前一韻律句群 PG-Final 基頻高度的逐漸降低。類似的漸降組型也可以在韻律句群 PG 內觀察到，是漸降效應 (down-stepping effect) 的特徵。韻律句群 PG 內三個序列位置的韻律短語 PPh 整體基頻高度漸降程度為：PG-Initial > -Medial > -Final，表示語段內關連性的特徵。因此，PG 內短語關連性的聲學表現，可用基頻高度按 PG 序列位置由高到中、到低的形式調整加以說明。同一 PG 內線性序列所顯示的高—中或中—低的對比，均不及跨越中段的高—低對比的反差顯著。我們同時也注意到：相鄰二韻律句群 PG 線性序列所顯示的基頻組型特徵，則為以韻律邊界區隔呈現低—高對比，此一對比恰和同一語段內的高—低對比相反，而反差幅度類似。由此可見，相同幅度的跨單位對比反差，在同一 PG 裡是高—低的跨短語對比反差，表示語段的首尾呼應；在相鄰 PG 間則是低—高的線性對比，表示主題轉換。二者均為語篇關連的表達，但對應於語篇組織的層級所顯示的意義卻不同，PG 內 PPh 的跨單位高—低對比，反映的是 PG 內部的語意連貫；而 PG 間 PPh 的跨韻律邊界線性低—高對比，反映的是 PG 間的語意或主題轉換。以上特徵說明了語流段落的構成、主題轉換與上層語篇關連性。



〈圖 11〉跨語料-語者相鄰 PG 序列位置 Ap 平均值範例。橫軸代表韻律邊界標記 B3, B4 和 B5；縱軸代表 Ap 平均值。

6.4.2 統計分析結果

我們針對韻律句群 PG 序列位置間韻律短語 Ap 參數值進行統計分析，爲了比較兩兩相對位置間的區辨性，採用 t-test 比較韻律句群 PG 序列位置 Ap 參數值的韻律特性鑑別力是否顯著。其統計分析結果如〈表 8〉所示，星號 * 表示達 0.01 顯著水準。

〈表 8〉韻律句群 PG 序列位置間 Ap 參數值 t-test 摘要表

Speaker	Initial-Medial	Initial-Final	Medial-Final
f051	0.1764 *	0.313 *	0.1366 *
m051	0.35757 *	0.4674 *	0.1098 *
f054	0.162 *	0.286 *	0.124 *
m056	0.509 *	0.4739 *	-0.0352

Ap 參數值的統計結果顯示（見〈表 8〉）：韻律句群 PG 三序列位置，不管是相鄰單位或者跨單位，跨語料-語者間 Ap 參數值皆達顯著性差異，只有 m056 的 PG-Medial 與 PG-Final 的 Ap 參數值未達顯著水準。但對照〈圖 11〉我們可以看出，跨語料-語者間的整體基頻組型是一致的。

綜合來說，以韻律短語 PPh 爲單位所呈現的整體基頻高低，有以下三種關係：(1) 韻律句群 PG 內的韻律短語 PPh 基頻高度逐漸降低的漸降效應 (down-stepping effect)，表現句段內各相鄰短語的關連性；(2) 韻律句群 PG 內首尾二韻律短語 PPh 間的對比較任二位置短語的相對差距爲大，表現句段內非相鄰的跨短語呼應關連性；(3) 相鄰二韻律句群 PG 間，由於前一 PG 基頻降至最低與後接

PG 的基頻重設造成相鄰線性反差，則表現了句段與句段間語篇的關連性。PPh 整體基頻高低的相鄰對比與跨句關係呼應語流段落的構成、主題轉換與上層語篇關連性。

7. 綜合討論

7.1 實驗一

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，大約占字調元件 Aa 預測正確率的 38%-46%，層層疊加，包括加入字調前後線性字調語境、PW 層的韻律邊界訊息、PW 層在所屬韻律詞的序列位置、PPh 層的字調預估、BG 層的邊界訊息與 PG 層的邊界訊息，累積貢獻度則逐層提升，從原先字調本身的 38%-46% 提升至 45%-55%、48%-60%、51%-67%、54%-73% 至 56%-74% 不等，顯示每一階層皆有一定的貢獻度，解釋了在基頻曲線中，並非只以字調為基本單位進行音節串接與短語句調的覆蓋而已。由此亦可知，儘管大部分的字調貢獻度主要來自音節層與韻律詞層，但在 Syl 層不及 50%，且邊界效應的貢獻度介於 5~7%，因此，每一韻律階層亦皆須考慮邊界效應的貢獻度，才能更精確的掌握基頻變化的成因。

分析國語口語連續語流或語篇韻律基頻時，一般多以單音節為切分單位，將四種字調的調型個別與孤立調相比，再以短語句調的下傾或上揚，作為來自句調的考慮。這種分析方式，其實受到文獻中分析非聲調語以短語為韻律單位的影響，但因不得不考慮字調的調型，因而演變出以單字孤立調型為主、短語句調為輔的分析方式。這類的分析若以單音節為單位，通常以凸顯單字孤立調的連讀變調與相連時的平滑銜接為主，視基頻的變化為線性的串接，但也因此無法描述語篇連續長串的語音信號中，連讀變調及各字調線性平滑銜接之外的成分。若以句調為單位 (intonation unit IU)，則通常以凸顯短語的整體基頻走勢為主，結果一方面無法解釋短語句調的變化多端亦非陳述、疑問二調型及其間的銜接平順所能描述；另一方面又因需解決字調的存在，而無法在信號上釐清二者的成分與互動。此外，句調間是否可視為不相關的韻律單位此一問題，則無法觸及，而時程面還需處理各級韻律單位整體的高低對比，及各單位間的停頓、邊界停延等現象，才能全面性的解釋語流韻律。總括而言，語流韻律的基頻走勢的產生，非字調與句調的銜接所能概括；分析語流韻律與基頻變化，必須從時程面入手同時考慮韻律單位及構組基頻變化的來源，才能解釋。

7.2 實驗二

實驗二「聽感標註的重疊性」實驗結果顯示兩語者的差異性並不大，而且從〈表 7〉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。

而 Ap 分析中，我們比較了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 層貢獻度低；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不規則間。兩者互補的結果，總預測率曲線則趨於平緩。

本實驗結果顯示：語流最終的韻律輸出，各韻律層都有貢獻，其中又以 BG 層與 PG 層對短調元件 Ap 的貢獻度差異最大，約占 7~35%。以上結果一方面與實驗一的階層式貢獻度呼應，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；一方面顯示因語體而產生的韻律輸出差異，只是階層式貢獻度的分布差異而已，其韻律成分完全可由 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 (Tseng & Su 2007)。

相對於實驗二檢視不同語體的 Ap 階層貢獻度之結果，實驗一中經由「韻律詞受韻律句群 PG 管轄之特徵」實驗結果（〈圖 7〉）指出：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。此結果亦顯示，線性的串接平滑不及跨單位的呼應更具對比性，超音段訊息的相對意義比相鄰串接更具語篇組織的意義。

實驗二則更進一步證實，同一規劃的基型，即可生成不同的韻律格式，亦可作為處理不同韻律格式的模版，無論從語言產製、語言習得、語言處理、認知資源需求的角度，都符合語言學上的深層意義。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布具體表示。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，下層韻律短語的貢獻度越高，表示韻律規劃的範圍越小。以上結果證明，我們所提出的階層式多短語語流韻律 HPG 架構，可系統性的解釋並產製不同的語篇韻律格式，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、句調對應於短語相同。總括而言，HPG 是對應於多短語語流韻律的基型，也是語篇韻律的規劃處理基型。從語音科技應用的角度，對 TTS (Text-to-Speech) 而言，只需要一個韻律模型，將各韻律層級的貢獻度調整，便可產製不同的韻律體現。對語音辨識而言，使用上層韻律資訊，可提高語篇單位辨識與擷取的正確率。

7.3 實驗三

實驗三研究結果證明：以韻律短語 PPh 為單位呈現韻律語境確實存在於基頻信號中，體現語篇規劃閾限與語篇組織的整體韻律語境，可由解構基頻信號獲得。從不同 PG 序列位置韻律短語 PPh 相對基頻高度的分析結果顯示，PG 內基頻高度隨著韻律短語 PPh 所在的 PG 序列位置由起始經延續漸降至終結，表示語段內關連性是由相鄰單位的對比性與跨單位關連性所共構而成（如〈圖 11〉所示），在韻律短語 PPh 層亦不例外。倘若將韻律短語 PPh 視為相互獨立、不相干的單位，則極有可能將基頻走勢不同的下傾程度視為變異，而無法整理出短語共構語段、語篇的系統性關係。

除此以外，PG 整體漸降效應 PG-Initial > -Medial > -Final 為次級韻律短語 PPh 語調往上疊加的結果，觸發全面性基頻相對調節，形成必須的語段韻律語境。因此，此實驗結果也說明了：語音段落語境需要至少三個相對序列位置的全面性調節，而非直接相鄰單位的局部連接，才能完整說明。我們強調，短語關連性與變異須考慮語調單位 IU 之上的上層語篇韻律語境，而且語調的變異絕不是隨機、無規則可循的。再者，PG 相鄰單位的分析研究也闡明了：語段間的語篇關連性可由相鄰的對比組型加以具體說明。也就是說，基頻重設與基頻相對高低有關，當 PG-Final 的低降緊接著下一 PG-Initial 的高起，相鄰基頻反差使得主題轉換在聽感上更加明顯與突出。如此基頻相鄰對比形式所形成的語段間語篇韻律語境，與 PG 內關連組型有別。因此，明顯的語篇韻律語境包含，如：整體基頻走勢的漸降 (PG-Initial > -Medial > -Final) 組型、PG 內跨單位的關連性 (PG-Initial vs. PG-Final) 組型與 PG 間的對比性 (PG-Final vs. PG-Initial) 組型。我們相信：如此資訊反映語者說話時的語篇規劃，為語者即時韻律向前規劃 (on-line prosody

look-ahead) 與整體韻律處理 (global prosody processing), 提供相對性線索, 共構成全面性的韻律語境。

更重要的是, 以階層式多短語語流韻律 HPG 架構闡釋此一實驗結果, 我們認為, 在同一 PG 內 PPh 跨越單位的高-低對比反差, 所呈現的正是來自上層單位 PG 的覆蓋效應, 對於同韻律層級姊妹組成成分 (sister constituent) 造成影響; 而同級相鄰 PG 間的 PPh 逆向低-高反差, 則是來自更上一層單位的同級管轄效應而非覆蓋效應, 此時的相鄰同韻律層級姊妹組成成分是 PG 而非 PPh。因此, 語篇韻律語境的構成必須同時考慮相鄰單位與跨越單位的對比。另外, 值得注意的是, 實驗三的研究結果亦證明了: 以 PPh 為單位的基頻高低對語流韻律的體現也有其貢獻, 不容忽視; 所共構出的語篇韻律語境, 是語流韻律的重要特徵。而在聲學語音研究方法上, 我們強調, 基頻參數不應以赫茲為測量單位, 必須經過非線性的對數轉換後, 以對數值 (logarithmic value) 呈現, 才能精確表達基頻高低在聽感上的對應。動態基頻現象的解析, 需將大小不同的單位明確界定, 強調其對比性, 才能回歸超音段聲學語音參數的本質意義。另外, 對比反差的觀察, 不應侷限於差異值的大小, 而需從相對性的意義探索, 才能將信號的特徵及處理, 與言語鍊裡口語的產製聽解結合, 進而提出具語言意義的解釋。

7.4 綜合討論

根據趙元任先生的說法, 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, 以小波浪比喻字調, 大波浪比喻短語句調, 大小波浪的相疊與覆蓋可以代數總和表示, 相位相等時互相加成, 相位相反時互相抵消 (Chao 1968:39-40)。我們根據這個邏輯推斷: 一個大波浪同時覆蓋於一個以上的小波浪之上, 被大波浪覆蓋的一串小波浪受制於大波浪; 大小波浪之間的關係, 是上下層不對等單位的階層式管轄制約; 被同一大波浪所覆蓋的多個小波浪, 彼此之間則是對等單位的線性串接。由於大小波浪擁有各自的波型, 經由階層式的管轄制約與對等的線性串接的互動後, 會形成一個全新的波浪形式, 因此這個全新波浪的波型體現, 既非原來的大波浪、亦非各個的小波浪的串接。同理, 在 HPG 架構裡音節 (SYL) 以上的每一個韻律層裡的同級韻律單位, 即韻律詞 (PW)、韻律短語 (PPh)、呼吸組 (BG) 和韻律句群 (PG), 都同時包含了平行線性串接和上層制約管轄的訊息, 從韻律詞到韻律句群的每一個單位都是層級不同的較大波浪, 必須同時處理與其下轄較小一波的互動。在 HPG 架構下, 語流韻律的成分, 不僅止於字調與短語句調; 亦即階層式的關係不只一層, 我們因此要問: 語流輸出

(output) 的體現，一共涉及了多少層級的大小波浪？除了字調與句調，層層大小波浪間的互動為何？又，若從聲學參數的角度切入語流輸出，如何釐清並證明最終的基頻體現是由層層韻律單位的波浪互動共構？

在所有基頻曲線模型中最能表達趙元任先生大小波浪疊加概念的，是日本學者 Fujisaki 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成分，二者單位大小不同，解釋了疊加後的體現，有些部分充盈有些部分縮減。而應用於聲調語言時，局部強調則被轉換成描述字調的成分 (Fujisaki & Hirose 1984, Mixdorff 2000, Mixdorff et al. 2003)，此模型一方面能正確的表示與字調無關的語調或句調成分，一方面證明且呼應了趙元任先生所指大波浪與小漣漪的關係。

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，可視為字調成分與句調成分疊加而成，亦即除了相鄰字調的连接外，還有來自上層句調的覆蓋，連續語流不僅是字調的平滑串接，還有較大範圍語調對字調的規範，構成韻律語境。而根據 HPG 架構，以上這種相鄰平滑與上層覆蓋應存在於每一個韻律階層。連續語流裡不僅只有字調成分與句調成分這兩種韻律層級與單位間的互動，向下應有音節字調與韻律詞間的關係，向上也應有短語句調成分與語段成分的互動關係。因此，我們也希望在對應 HPG 架構中的韻律詞層中，確認韻律詞在連續語流中作為基本韻律單位的聲學語音證據與韻律意義；也希望在語段中確認短語句調間的語境與關係。本研究因而在字調與短語調之外，分析了韻律詞與韻律短語的基頻曲線變化與相對關係，提出韻律詞與韻律短語確為語篇韻律單位的證據；進一步證明韻律短語也是連續語流的基本韻律單位而非終極韻律單位，所謂的語調單位不應僅限於短語，而應擴及複雜句、語段及語篇，也有基型。除了觀察 HPG 架構下各級韻律單位在連續語流的共通性外，我們也針對不同語體及韻律格式進行分析，找出 HPG 架構對應不同韻律格式的變化，發現：不同韻律輸出來自於不同的 HPG 分層貢獻度，連續語流所存在的跨短語韻律語境，其實只是韻律層間分布的變異造成了不同語體韻律格式的差異。

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，詞彙詞約占韻律詞的 67.5% (Peng et al. 2004)，構成韻律詞意範圍內的音節，同樣有著關聯性與連貫性，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，

語音信號處理必須兼顧 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，才能完整地模擬出連續語流的特性，並計算每一韻律層對 HPG 架構的分層貢獻度。

在 HPG 架構下，語段下的不同位置的韻律短語 PPh 共同組成表意韻律語境，用來表示語段的起始、持續與結束，三者的關係既包含由首至中到尾線性連接所表示的語段起始、延續到結束的相鄰接續，也包含 PG 首尾二短語跨越中段短語所表達的起始與終止呼應，建構語篇語意關連時，它們都是重要語篇成分，缺一不可。也因為如此，位於三個 PG 位置的韻律短語，即便是句法結構相同，如均為陳述短句，其體現則因位置不同而有所不同。語段首尾的韻律短語不但下傾的幅度不同，語段中的韻律短語因表語篇語意尚未結束，甚至沒有明顯的下傾現象，三者所表達的相對性關係必須一起考慮，非比較個別短語下傾斜率所能盡述。

因此，語流韻律語境同時包括多短語語段及語篇結構的語境，二者韻律階層不同。語段內的多個短語間，既有相鄰短語的連接，也有表語段連貫的連接與覆蓋。語段間的語篇關連也有同樣的關係，既有相連也有跨語段的覆蓋。二種語境均不可僅視為短語句調的線性串接結果。我們強調，語流韻律的研究單位，必須從單獨的短語擴及為一個短語以上的複雜句 (complex sentence) 或語段 (speech paragraph)，亦即 HPG 架構所提出的「多短語韻律句群 PG (Multi-phrase Prosodic Group)」。PG 是語篇韻律的單位 (discourse unit)，上承語篇語意的管轄，下含短語或簡單句法句的句調單位 IU 為其直轄的次級單位，主要功能是表示特定語意段落的開始、延續與結束。因此，單一句調單位 IU 並不是最終或最大的韻律單位，當短語形成語段時，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。從 HPG 架構的角度來看，單一 IU 走勢或串接平滑無法解釋的句調變異，其實可從韻律語境所建構的相鄰 IU 串接平滑及上層管轄所造成的「開始、延續與結束」三者間的相互呼應關係得到解釋。

8. 結語

本研究結合聲學語音學、語料庫語言學及計算語言學，解析及量化語料的聲學基頻證據，證明口語語流中的詞意韻律、句法韻律及表意韻律如何共構語篇韻律及語流韻律的韻律語境；呈現出較大範圍韻律單位的整體基頻走勢的具體證據；具體提出線性輸出的語音訊號中非線性貢獻的證據，增進對口語韻律的瞭解。本研究首先從語篇結構定義語流韻律單位，再以量化方式將由上而下的階層

式管轄制約，從最小韻律單位音節，由下而上的解構超音段聲學語音參數基頻，推導各層級如何共構最終輸出與體現，最後提出符合語言學理論的解釋，證明基頻信號中各韻律層級均有貢獻，而非僅由字調與短語句調組成。

由實驗一結果，我們得知，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，語者與聽者依據此產製與接收來自語篇的大範圍表意韻律訊息，達到溝通意圖。實驗二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布具體表示。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，下層韻律短語的貢獻度越高，表示韻律規劃的範圍越小。實驗三則證明了各韻律層的基頻語境均需解構相鄰與上層成分，才得以系統性的獲知不同韻律單位、韻律層次間的相鄰與跨單位關係；而表層語流中句調的變異實來自各句調單位間的相鄰連貫語境與上層語篇表意的首尾呼應所構成的跨短語語境。

從以上這些感知的語音單位解構語篇產製規劃的證據與特徵，以及所獲對韻律與語音輸出體現的瞭解，我們從韻律單位與組織結構角度，期望這些證據及特徵能提供開發語音科技的新思維：語流韻律是有架構的，因有系統性關係且有跡可循，所以，我們強調，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線性串接的完美平滑，必須同時包含多短語語段間的跨短語表意韻律語境。因此，語音合成的輸出語音之所以聽起來不流暢、不自然，我們認為大部分是由於在音段層次只注重單字的連讀變調，而未含韻律詞的上層資訊，以致小單位韻律語境不足；在超音段層次只個別處理短語句調，而未含短語調的連讀及跨短語呼應，以致完全未能體現大範圍語篇韻律語境。因此，各級韻律語境的缺乏，才是輸出不自然的主因，加入語篇韻律語境的模型，可有效的產製更自然的語音合成輸出 (Chen et al. 2006)。在語音識別方面，以上研究讓我們瞭解，聽者在即時處理語流時，即便每一個單音節的字調訊息不完整，但只要有相當成分的韻律語境，便能彌補並提供有效判定即時韻律單位處理及向前預估的訊息。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訊息。如果一段連續語流內的單位無法同時體現語篇規範的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，延誤即時處理。由此亦可見，語音識別並不同音段識別、音節識別、字調識別而已，而應考慮各種口語語流中高低、快慢、強弱的對比訊息，以

及各級停頓節點與節點前的停延等語音信號上的對比差異特性，獲致語篇單位的訊息。近來已有研究證明：以 HPG 架構為基礎，有效的利用語篇韻律訊息，結合有限的句法訊息，自動將連續語流切分為小句、複雜句、語段等不同層級的單位，可得出包括正確的語篇標點符號的識別結果 (Xie 2008, Chiang 2009)，提升語音識別的層級。

未來研究方向包括將 HPG 架構套用至授課等溝通意圖明顯、語篇及語段訊息分明的自發性語料，進行量化分析研究，一方面希望進一步探討語言產製、語篇組織與語言認知的互動，以驗證此規劃基型也存在於看似不規則的自發性語料中，另一方面希望更全面的解釋語流中字調的變異，解構字調即韻律的迷思。

附錄：古典文體分類範例

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，斗酒十千恣歡譴。
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。

引用文獻

- Beckman, Mary E., and Julia Hirschberg. 1994. *The ToBI Annotation Conventions*, Appendix A. Columbus: The Ohio State University Research Foundation.
- Beckman, Mary E., and Gayle Ayers Elam. 1997. *Guidelines for ToBI Labelling*, Version 3.0. Columbus: The Ohio State University Research Foundation.
- Chao, Yuen Ren. 1968. *A Grammar of Spoken Chinese*. Berkeley: University of California Press.
- Chen, Sin-Horng, Chiu-yu Tseng, and Hsin-min Wang. 2006. Tone modeling for speech synthesis. *Advances in Chinese Spoken Language Processing*, ed. by Chin-Hui Lee, Haizhou Li, Lin-shan Lee, Ren-Hua Wang & Qiang Huo, 77-98. Singapore: World Scientific.
- Chiang, Chen-Yu. 2009. *Unsupervised Joint Prosody Labeling and Modeling for Mandarin Speech*. Hsinchu: National Chiao Tung University dissertation.
- Crystal, David. 1969. *Prosodic Systems and Intonation in English*. Cambridge: Cambridge University Press.
- Fujisaki, Hiroya, and Keikichi Hirose. 1984. Analysis of voice fundamental frequency contours for declarative sentences of Japanese. *Journal of the Acoustical Society of Japan (E)* 5.4:233-242.
- Gu, Wentao, Keikichi Hirose, and Hiroya Fujisaki. 2005. Identification and synthesis of Cantonese tones based on the command-response model for F0 contour generation. *Proceedings of 2005 IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP 2005)*, Vol. 1, I289-I292. Philadelphia, Pennsylvania, USA.
- Halliday, Michael A. K. 1967. *Intonation and Grammar in British English*. The Hague: Mouton.
- Keller, Eric, and Brigitte Zellner. 1996. A timing model for fast French. *York Papers in Linguistics* 17:53-75.
- Liu, Fang. 2009. *Intonation Systems of Mandarin and English: A Functional Approach*. Chicago: University of Chicago dissertation.
- Mixdorff, Hansjörg. 2000. A novel approach to the fully automatic extraction of Fujisaki Model parameters. *Proceedings of 2000 IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP 2000)*, Vol. 3, 1281-1284. Istanbul, Turkey.
- Mixdorff, Hansjörg. 2001. MFGI, a linguistically motivated quantitative model of German prosody. *Improvements in Speech Synthesis*, ed. by Eric Keller, Gérard Bailly, Alex Monaghan, Jacques Terken & Mark Huckvale, 134-143. Chichester: John Wiley.
- Mixdorff, Hansjörg. 2002. *An Integrated Approach to Modeling German Prosody*. Dresden: Technische Universität Dresden dissertation.

- Mixdorff, Hansjörg, Hiroya Fujisaki, Gao Peng Chen, and Yu Hu. 2003. Towards the automatic extraction of Fujisaki Model parameters for Mandarin. *Proceedings of the 8th European Conference on Speech Communication and Technology (EUROSPEECH 2003)*, 873-876. Geneva, Switzerland.
- Nakai, Mitsuru, Harald Singer, Yoshinori Sagisaka, and Hiroshi Shimodaira. 1995. Automatic prosodic segmentation by F0 clustering using superpositional modeling. *Proceedings of 1995 IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP 1995)*, Vol. 1, 624-627. Detroit, Michigan, USA.
- Nespor, Marina, and Irene Vogel. 1986. *Prosodic Phonology*. Dordrecht: Foris.
- Peng, Hua-jui, Chi-ching Chen, Chiu-yu Tseng, and Keh-jiann Chen. 2004. Predicting prosodic words from lexical words: a first step towards predicting prosody from text. *Proceedings of the 4th International Symposium on Chinese Spoken Language Processing (ISCSLP 2004)*, 173-176. Hong Kong SAR, China.
- Pike, Kenneth L. 1943. *Phonetics: A Critical Analysis of Phonetic Theory and a Technic for the Practical Description of Sounds*. Ann Arbor: University of Michigan Press.
- Pierrehumbert, Janet B. 1980. *The Phonology and Phonetics of English Intonation*. Cambridge: MIT dissertation.
- Scott, N. C. 1941. Broad transcriptions. *Le Maître Phonétique* 76:48-51.
- Selkirk, Elisabeth O. 1984. *Phonology and Syntax: The Relation between Sound and Structure*. Cambridge: MIT Press.
- Shimodaira, Hiroshi, and Mitsuru Nakai. 1994. Prosodic phrase segmentation by pitch pattern clustering. *Proceedings of 1994 IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP 1994)*, Vol. 2, 185-188. Adelaide, South Australia, Australia.
- Tseng, Chiu-yu. 2006. Prosody analysis. *Advances in Chinese Spoken Language Processing*, ed. by Chin-Hui Lee, Haizhou Li, Lin-shan Lee, Ren-Hua Wang & Qiang Huo, 57-75. Singapore: World Scientific.
- Tseng, Chiu-yu, and Yeh-lin Lee. 2004. Speech rate and prosody units: evidence of interaction from Mandarin Chinese. *Proceedings of the International Conference on Speech Prosody 2004*, 251-254. Nara, Japan.
- Tseng, Chiu-yu, and Shao-huang Pin. 2004. Modeling prosody of Mandarin Chinese fluent speech via phrase grouping. *Proceedings of Speech and Language Systems for Human Communication (SPLASH-2004/Oriental-COCOSDA2004)*, 53-57. New Delhi, India.
- Tseng, Chiu-yu, Shao-huang Pin, and Yeh-lin Lee. 2004. Speech prosody: issues, approaches and implications. *From Traditional Phonology to Modern Speech Processing*, ed. by Gunnar Fant, Hiroya Fujisaki, Jianfen Cao & Yi Xu, 417-437. Beijing: Foreign Language Teaching and Research Press.
- Tseng, Chiu-yu, Shao-huang Pin, Yeh-lin Lee, Hsin-min Wang, and Yong-cheng Chen.

2005. Fluent speech prosody: framework and modeling. *Speech Communication* 46.3-4:284-309.
- Tseng, Chiu-yu, and Zhao-yu Su. 2007. From one base form to multiple output styles—Predicting stylistic dynamics of discourse prosody. *Proceedings of the 8th Annual Conference of the International Speech Communication Association (INTERSPEECH 2007)*, 110-113. Antwerp, Belgium.
- Tseng, Chiu-yu, and Zhao-yu Su. 2008. Boundary and lengthening—On relative phonetic information. Paper presented at the 8th Phonetics Conference of China and the International Symposium on Phonetic Frontiers. Beijing, China.
- Wang, Changfu, Hiroya Fujisaki, Sumio Ohno, and Tomohiro Kodama. 1999. Analysis and synthesis of the four tones in connected speech of the standard Chinese based on a command-response model. *Proceedings of the 6th European Conference on Speech Communication and Technology (EUROSPEECH '99)*, 1655-1658. Budapest, Hungary.
- Xie, Lei. 2008. Discovering salient prosodic cues and their interactions for automatic story segmentation in Mandarin broadcast news. *Multimedia Systems* 14.4:237-253.
- Xu, Yi. 1999. Effects of tone and focus on the formation and alignment of F0 contours. *Journal of Phonetics* 27.1:55-105.
- Xu, Yi. 2004. Understanding tone from the perspective of production and perception. *Language and Linguistics* 5.4:757-797.
- Zellner Keller, Brigitte, and Eric Keller. 2001. Representing speech rhythm. *Improvements in Speech Synthesis*, ed. by Eric Keller, Gérard Bailly, Alex Monaghan, Jacques Terken & Mark Huckvale, 154-164. Chichester: John Wiley.
- 鄭秋豫. 2008. 〈語篇韻律與上層訊息——兼論語音學研究方法與發現〉,《語言暨語言學》9.3:659-719。
- 鄭秋豫, 蘇昭宇. 2007. 〈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論文。台北：中華民國計算語言學會。
- 鄭秋豫, 蘇昭宇. 2008. 〈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〉,第二十屆自然語言與語音處理研討會論文。台北：中華民國計算語言學會。

[Received 16 April 2009; revised 18 November 2009; accepted 18 December 2009]

Institute of Linguistics
Academia Sinica
130, Sec. 2, Academia Road
Nankang, Taipei 115, Taiwan
cytling@sinica.edu.tw

An F0 Analysis of Discourse Construction and Global Information in Realized Narrative Prosody

Chiu-yu Tseng

Academia Sinica

The aim of the present study is to show how the time varying F0 signal in paragraph prosody output is composed of contributions from multiple ranks of local between-unit concatenation and global higher-level layering. Discourse specified positioning, pre-boundary lengthening, and post-boundary pauses are also contributing variables. Both adjacency smoothing and higher level layering collectively make up the output F0, as supported by quantitative results. Results also reveal distinct patterns of phrase-level F0 height modulations within and between paragraphs, confirming the significance of phrase-level global prosody context in addition to tone-intonation interaction. Cross-speaker analysis of speech data varying in prosodic style show systematic patterns of contribution distribution by style, thus proving one base form is sufficient for both the planning and processing of surface variations. The study also shows why both methodology and interpretation must reflect the relative nature of suprasegmental cues, and how corpus analysis sheds new light on acoustic analysis.

Key words: discourse prosody organization, prosody hierarchy, global F0 contrasts, relative acoustic cues, higher level information