

語篇韻律與上層訊息 ——兼論語音學研究方法與發現*

鄭秋豫

中央研究院

口語語流的超音段訊息中，除字調和句調外，還包含上層語篇訊息，形成韻律語境。我們提出「階層式多短語語流韻律 HPG (Hierarchical Prosodic Phrase Grouping)」架構，分析並提取各韻律層級的聲學語音參數，包括基頻、音長與響度的韻律貢獻度，證明語流韻律係由各韻律層級系統性的層層疊加所共構。我們因而論述，研究語篇韻律：一、應將上層語篇訊息納入考慮，釐清韻律語境如何致使聲調與句調產生變異，解釋語流韻律之來源。二、可創新研究方法，結合語料庫語音學如何建構量化證據，提升語音學研究的科學性。三、結合質與量的研究，進一步瞭解超音段聲學語音體現的相對性意義。

關鍵詞：韻律，語篇韻律，語流韻律，韻律產製，韻律處理，韻律語境，口語語流，「階層式多短語語流韻律 HPG (Hierarchical Prosodic Phrase Grouping)」架構，聲學語音參數，超音段，語篇訊息，上層訊息，語料庫語音學

1. 引言

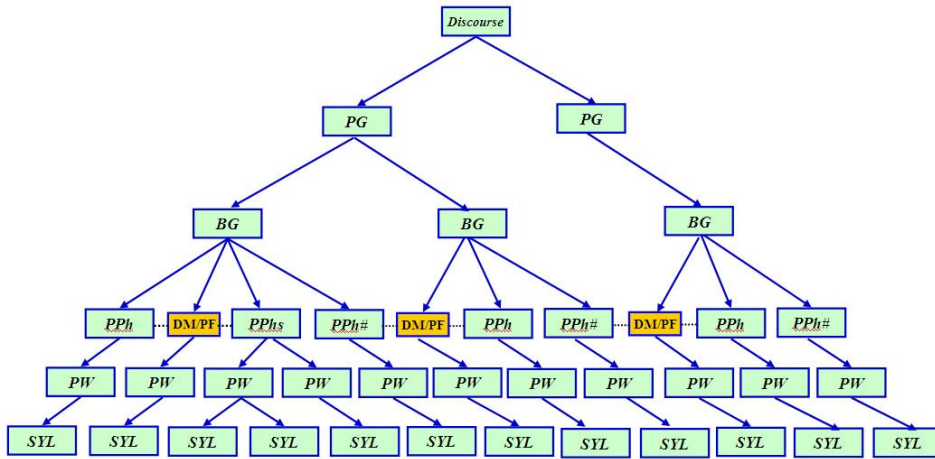
本文的研究課題是語流韻律的聲學語音結構，論述口語語流韻律最顯著的特徵與原因，及形成口語語篇的單位與意義。我們強調自由上而下的角度，探討韻律產製的規劃與人聽口語的處理，進而剖析語流韻律的結構。我們提出階層式語流韻律架構 HPG (Hierarchical Prosodic Phrase Grouping) 的說法，彰顯系統性規範

* 本研究為中央研究院主題計畫「國語語音合成的新方向——從韻律結構出發」部分研究成果(2003.01~2005.12)。參與研究人員：賓少煌(2000~2004)、鄭雲卿(2000迄今)、李瑋珊(2000迄今)、黃楓嵐(2002~2004)、李岳凌(2003~2004)、陳綺菁(2003~2004)、傅寶麟(2004~2005)、張俊祥(2004, 2007)、蘇昭宇(2005迄今)，謹此一併致謝。承蒙北京清華大學計算機與技術系蔡蓮紅教授提供實驗二與實驗三北京普通話女性播音員實驗語料 BFS，特此致謝。

語流或語篇韻律的上層制約，及此一說法所表示的語言學意義。我們認為：語篇裡的多短語韻律語段，是一個具有跨短語深層結構的韻律單位，表達跨短語的韻律語境。這種韻律語境可透過聲學語音學和語料庫研究方法的結合，採集大量語料後，佐以量化證據的方式來驗證。於是，語音學研究除現象描述外，亦可經由量化所得之證據，提出語言學上的解釋，語音學的研究方法也因此更實證、更科學。

首先我們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)，因為從聲學語音訊息而言，語流韻律是口語連續說話時，多短語語段及語篇的上層語意所導致的句調調整及韻律語境所造成，因句法結構所產生的單句句調 (phrase intonation) 並不包含這樣的資訊。跨短語語段的韻律語境，管轄制約的是如何將一個以上的短句組織成語意完整的語段，表達的是詞法與句法結構上層的語篇語意。由於語篇語意的聯繫，一個口語語段從開始、延續到結束，形成一個多短語的完整語意單位；而語段內的個別短語，因受到上層語意的制約，必須隨語段語意的開始、延續及結束這三個缺一不可的語段位置，各就各位，才能表達跨短語關係整體的成段意義。換言之，語流韻律所表達的，除了下層相鄰短語之間的線性銜接關係外，還必須包含上層語意的相鄰一致性與跨短語關聯性，因此語段的流暢 (paragraph flow)，自不可能僅由線性串接個別無關的短語句調來達成。從這個角度出發，語流韻律中個別短句的句調變化，就不再是無法解釋的變異；上述三個表示語段從開始經由延續最後到達終結的位置，所分別導致語段內各個短語句調的調整，自然也有其系統性。由上而下的角度，在自然語言分析方面，早有許多討論及證據（如 Jurafsky & Martin 2000），只不過不常見於一般語音學研究，尤其不見於語流韻律的研究而已。

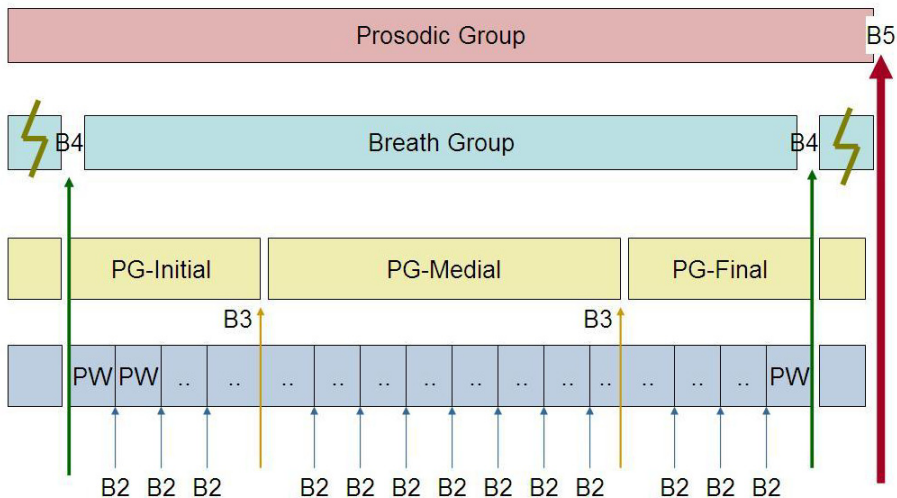
在「階層式多短語韻律句群 HPG」架構中（〈圖 1〉），由下而上的韻律單位分別為：音節 (SYL)、韻律詞 (PW)、韻律短語 (PPh)、呼吸組 (BG)、韻律句組 (PG) 及語篇 (Discourse)，每一層也對應了同名。相對應的韻律邊界停頓則分別為 SYL 後為 B1，PW 後為 B2，PPh 後為 B3，BG 後為 B4，PG 後為 B5。它們的關係是 $SYL < PW < PPh < BG < PG < Discourse$ ，而 $B1 < B2 < B3 < B4 < B5$ 。



〈圖 1〉「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)。

表示多短句韻律語段的三個相對韻律位置，則分別是以韻律短語 PPh 為單位的 PG 首 (PG-Initial)、PG 中 (PG-Medial) 與 PG 尾 (PG-Final)，其中 PG 首與 PG 尾是單一韻律短語，PG 中的複數韻律短語則可以是單數也可以是複數 (〈圖 2〉)。



〈圖 2〉「階層式多短語韻律句群 HPG」相對韻律位置與韻律邊界停頓示意圖

其中，韻律邊界與停頓分別對應為 B1, B2, B3, B4 與 B5。

HPG 階層式韻律語流架構，強調的是語流韻律中，必有韻律短語 PPh 層以上的語段訊息，以體現語段規劃及語意完整的單位。在一口氣即能表達完整語意的口語段落中，上達 BG (Breath Group) 層即已足夠，此時 BG 層併入 PG 層。但語段規劃單位可大於 BG，PG 可長可短，當一個呼吸組 BG 不足以表達完整的語意段落時，則有更上一層 PG。因此 HPG 可視為表現語段規劃的階層式模組，可伸可縮。

HPG 架構與其他韻律單位的主要差異有三：(1) 語篇中的多短語語段 PG 是大於短語（句調）小於語篇的韻律單位，換言之，PG 是短語的直屬上層韻律單位，同時也是語篇 (Discourse) 直屬的下層韻律單位，其規範制約的範圍與單位，遠大於句法所能管轄的孤立句。(2) 韻律短語 PPh 是一個短語句調單位 IU (Intonation Unit)，因此短語句調是一個 PG 的直屬下層韻律單位，接受三個相對 PG 位置 PG 首、PG 中與 PG 尾的規範與制約而調整。(3) 語流中的邊界停頓也是韻律單位，具有語篇韻律身分。

PG 強調語段中各短語間的韻律相關性及語意關連性，表現在聽感上最顯著的特徵是多短語語段如何起首、延續直到終了。各短語句調因受制於上層韻律訊息而做相對應的調整，亦即在小波浪、大波浪 (Chao 1968:39-40) 之上，找到範圍更大的浪潮，並證明它與所管轄的各級下層韻律層及韻律單位的關係，是由下往上層層相疊，「自漣由波、推波助瀾、逐浪到潮」，而各個層次分層貢獻不同範圍及比例的韻律成分，層層疊加後構成整體韻律輸出，解釋了語流中各短語多變的句調調型，其實是由於上層韻律單位系統性的規範制約所致。

我們以韻律短語 (Prosodic Phrase, PPh) 為單位，將語段韻律語境成分，分為：PG 首 (PG-Initial)、PG 中 (PG-Medial) 與 PG 尾 (PG-Final) 三個成分。PG 首和 PG 尾都是一個 PPh 單句，而 PG 中可以是一個或多個短句。這三個韻律成分系統性的導致各短語句調調整後，才能表達跨短語、大範圍語段語意的連貫及完整的韻律語境。各孤立短語必須先成語段，再通過語篇主旨將語段貫穿，才得以形成語篇。因此語篇不能視為不相關短句的線性串接，語篇韻律也不是由孤立的短語句調串接可構成。我們也因此強調：研究韻律語流的單位，必須取材完整語篇而非孤立短語或短句；檢視語流韻律的角度，必須包括短語句調以上的多短語語段及跨短語的韻律現象；研究語流韻律的目標，必須擴及尋找是否存有大範圍的韻律組型並提取韻律組型的深層結構；語流韻律及韻律語境才有可能浮現並得以釐清。要特別一提的是，即便是取材口語敘述或語篇，若於檢視時，或僅以單句調為單位，逐一挑出後個別檢視短語句調本身變化；或僅將檢視的重點放在字調的線性銜接所引起的連讀變調，及句調與字調的互動所導致的變化，則都將無法獲得體現語流的韻律組型。本研究也將證明，表達語段語意銜接與關聯的語篇

韻律具有深層結構的意義，因此呈現系統性變化，亦具語言學上的深層意義。

韻律短語組 PG 可以是多短語複雜句或多短句語段，而以口語篇章敘述裡最爲明顯（吳爲章、田小琳 2002）。HPG 一說強調並界定表面成短語串的韻律組型，非但不僅是相鄰孤立短語句調線性相連串接時的平滑，而是短語成組的跨短語整體韻律語境及其系統性，以及所反映的上層語意關連。本文將說明我們提出階層式多短語語流韻律架構 HPG (Hierarchical Prosodic phrase Group) 的假設，開創結合語料庫語言學研究方法的「語料庫語音學」，透過實驗設計，檢視大批口語資料，以數據逐步證明語流韻律其實是來自從字調、短語句調到語篇韻律各單位層層疊加 (super-positional and cumulative) 貢獻，及各層交互作用的加總 (Tseng et al. 2004, 2005, Tseng 2006a)。

我們所開創使用的語料庫語音學研究方法，從口語感知角度出發，分析大批口語語料的語流韻律，兼顧聽話人的語音感知、即時語音分析與說話人即時語音產製規劃間的對應，透過量化語料分析三個聲學語音參數，即跨短語的基頻曲勢、音節時長、邊界資訊，一一找到 HPG 各韻律階層貢獻的證據及聲學語音模版 (template)。這些通過量化所得的模版，有如樂曲的套譜，表達的正是跨短語 PG 的深層結構。對言語鏈裡語言產製和語音處理而言，這些模版是關鍵性的證據，意謂著人在連續講話時，說話人在即時語音產製，以及聽話人在即時語音訊號處理時，在語音層都涉及語流的規劃單元的上限及對應的聲學模版，就如同句法深層結構一樣，因此語流韻律也爲句法以上的深層語言、語意結構及認知閾限提供證據。

此外，我們更分析聽感與物理信號間的關連及落差，探討語音處理的認知意義，並進一步從語言學意義的角度論述：由語篇敘述得到的研究證據，一方面具有類似音位對立式的對應；另一方面也具有跨語言或語言共通性方面的意義。透過這些證據，我們不但釐清了字調、句調與語篇敘述韻律在詞義、句法和篇章訊息上系統性的階層對應、制約管轄及分層貢獻度，我們強調的疊加關係，更進一步解釋了由句法結構制約的句調與句調單位，只是篇章敘述語意裡的下屬單位；在語篇敘述中也必須同時經過上層語意連貫銜接的管轄而調整，因此若單一檢視語流中個別短語句調，所呈現的則僅是各種表象的不規則。

在同樣的語料中，我們進一步透過聲學語音訊息中的音域變化與基頻重設位置，進行對應句法結構與語意銜接的分析，檢視 HPG 如何組合成篇章敘述。我們的證據顯示：短語連接成韻律句組時，或可直接串接；或可插入篇章提語 (discourse marker) 和韻律填充 (prosodic filler)。這些插入的成分，致使語流韻律的表面樂律性更多變，但這些表面的不規則，其實亦可經過聲學檢測釐清。我們也

因此論述，口語語流韻律形似多變而實無損其共通，並因此提供共通性更具體的證據。

本文內容共分三部分：第一部分是實驗設計與結果，實驗一至實驗三分別從語流的基頻、音節時長和強度三個聲學參數檢驗、分析基頻走勢、時長配置及音強分布，來驗證是否有跨短句、表達上層語篇訊息的語段韻律訊息？是否可提取？第二部分是獲得以上三個聲學參數、得到語段韻律的訊息的證據後，透過實驗四：「口語韻律階層中的基頻重設與基頻範圍研究」進一步檢視：語段以上的語篇韻律結構，是否可透過語段與語段間基頻重設及頻域範圍的變化得到證據。第三部分則是討論這些證據在語流韻律結構方面的理論意義，及語音學的研究方法如何與語料庫語言學結合，以擴大分析單位的範圍及語料樣本的量，從語音學重質的描述推向語料量的處理。

2. 實驗一：基頻走勢分析及預測

2.1 實驗目的及假設

本實驗的目的是檢驗聲學語音參數基頻走勢，是否具有上層韻律效應，分二實驗階段，階段一以 HPG 架構檢驗上層效應的存在；階段二為 HPG 階層架構中基頻的分層貢獻度預測與原始語音的比對。由於基頻走勢是動態的，我們採用 Fujisaki model (Fujisaki & Hirose 1984) 透過該模型的兩個參數 A_p 和 A_a （前者檢測短語句調成分、後者檢測字調成分）來檢測基頻變化的整體走勢；並進一步檢驗是否可以釐清句法結構規範的短語句調和上層語意訊息所規範的語流韻律。

實驗階段一，我們假設：(1) 由語料中可擷取出 A_p （Fujisaki model 中的短句句調參數）組型；(2) 可找到 PPh 層的 A_p 預測值與更上層 A_p 預測值交互作用的佐證；(3) 可找出在連續語流中預測音高配置 (pitch allocation) 的佐證。以上的組型和佐證將證明：單句句調同為姐妹成分 (sister constituent) 時，各自獨立，彼此並沒有語流韻律上的關係；各單句受到上層資訊管轄制約成為語段 PG 的次級單位後，才產生跨單句的語段關係，因其在 PG 中的位置，而產生相對應的變化。換言之，句型相同的單句，在 PG 的不同位置，會導致句型相同的句調產生變化。

實驗階段二，我們假設：字調成分 (accent component)（Fujisaki model 中的 A_a 參數）也受 HPG 理論架構中的層層管轄的 PG 階層所管轄。因此，在 HPG 架構中音節層 (SYL)、PW 層與 PPh 層對於最終的輸出韻律 (output prosody) 應皆具貢獻度。

2.2 實驗語料

本實驗語料採用 COSPRO 08 承載語段 (carrier paragraph) 語料庫 (Tseng 2006a)。此承載語段一共有 30 個音節、包含 3 短句的複雜句。3 個短句表示 PG 首、中、尾三個 PG 位置，構成典型的短 PG 結構。每個承載語段句型為：「△是一個常見的字，一般人常把△字掛在嘴邊，講話時動不動就會提到△。」其中△表示目標音節（音節字調組合共 1300 音節），PG 首、中、尾短句長度分別為 8 音節、11 音節與 11 音節。此語料庫共 1300 句承載語段，由一男一女、未受專業播音訓練之發音人（M054&F054，語速分別為 289&308 ms/syllable；錄音時長分別為約 2 小時 21 分、3 小時 14 分；語料大小：650MB/898MB）在錄音室錄製而成。

此承載語段的控制包括：

- (1) 3 個短句長度差異不大，結構都是直述句，因此句調相同。
- (2) 盡可能移除詞義與語意焦點 (lexical and semantic focus)，使之成為一典型的 PG 整體組型。
- (3) 目標字分別嵌入 PG 首、中、尾短語中的第 1、第 6 和最後音節，如此一來，目標字正好位於此 PG 的首、中、尾位置，以便檢驗 PG 的位置對同一字調是否有基頻效應。

此承載語段的三個短句的句法結構均為直述句，單句句調均為下傾句調。觀察實驗語料承載語段中的基頻走勢（如〈圖 3〉所示），我們可以看到，三個短句個別來看雖都是下傾句調，但下傾的斜率並不相同，我們認為：這是受了上層訊息管轄的結果，為了表現口語段落的起始、延續和結束以及語流的連貫性與語意的關聯性，各短句句調在口語段落中就所在的語段位置，扮演了不同的語段角色，而非個別短句句調的隨機變異。換言之，起始句調為了突顯語篇新訊息的開始，通常開頭較高，相對的斜率也較為陡峭；中段句調則為了表示口語段落語篇主旨的延續，基頻走勢較為平緩，句調的下傾最不明顯；結尾句調則表示口語段落主旨的結束，通常開頭不高、基頻走勢緩緩降低，降至最低點以表示句調的結尾。為了表示同一主旨的關聯性，三個語段位置各有韻律特性。再者，即使是相同的字調（在本實驗中指的是承載語段中的目標字），因所處於口語段落中的位置不同，也會受制於上層句調、語段訊息，以致有音高的不同。以我們所蒐集的承載語段來說，目標字在 PG 首的音高最高，再來是 PG 中，最低的是 PG 尾，因此〈圖 3〉中的三個目標字「巴」雖均為第一聲，卻必須因應所在短句與語段的位置，呈現實際輸出的音高差異。以上證據再再顯示：不論是短句句調或音節字

調，皆會受到上層訊息的管轄，依口語段落中不同位置，系統性的產生不同的表面韻律形式。



〈圖 3〉承載語段基頻走勢圖（橢圓型小圈標出的是一聲目標字「巴」）

通常在一個口語段落中，一個多短語 PG 經常超過 3 個短語以上，意謂了 PG 中段短語通常不止 1 個短語。因此，相較於長篇的朗讀段落，雖然此承載語段只包含了相對少量的韻律訊息，但也不會超過 HPG 所包含在短語 PPh 以上 PG 以下的一個呼吸組 BG (breath group)，也就是一口氣能說完的短句；無疑的，此承載句型仍包含典型的語篇訊息，同時可提供相同脈絡下音節重複出現的字調研究。而我們的 HPG 階層式韻律語流架構，強調的是語流韻律中，必有韻律短語 PPh 層以上的訊息，但因 PG 可長可短，過長時一個 BG 不足以表達語意的完結，則有更上一層 PG；不過長時一個 BG 便表達完畢，因此 HPG 的結構本就是可伸可縮。相對應於本實驗的語料，BG 層與 PG 層合併，成為 PG 層。

考量到字調和句調可能存在的交互作用，因此，我們使用 COSPRO 08 中女性發音人 F054 tone 1 目標音節承載句 259 句做為樣本，用來說明 PG 效應；透過 Fujisaki model 中的 Ap 和 Aa 參數進行基頻數值的分析及預測。實驗語料特徵摘要如〈表 1〉所示：

〈表 1〉F054 tone 1 承載句語料摘要表

# of Paragraphs	# of Syls	# of Phrases
259	8008	891

2.3 語料處理與標記

以 HTK Toolkit 程式判定音段成分（聲母、韻母）再以 SAMPA-T (Tseng &

Chou 1999) 自動標記音段邊界，經人工檢查調整過後，以人工標記感知到的韻律邊界。不管是音段邊界標記人工檢查、抑或停延標記皆是由訓練有素的標記員所完成。

2.4 分析過程

我們使用 Fujisaki model，從 PG 階層中尋找同時包含額外上層訊息的短語成分 (phrase component Ap) 與字調成分 (accent component Aa)，以句調單位 (IU) 或韻律短語 (PPh) 做為操弄單位，擷取出音節及短語的基頻曲勢 (F0 curve tendency)。

以低通濾波器 (lowpass-filter) 分離、計算出語流中的短語成分 Ap 參數與字調成分 Aa 參數，藉以預測出隨著時程的整體基頻變化與局部基頻變化。由於是以韻律短語 (PPh) 做為操作單位，因此，此二參數可視為 PG 不同位置上層訊息的基頻曲勢代表值。

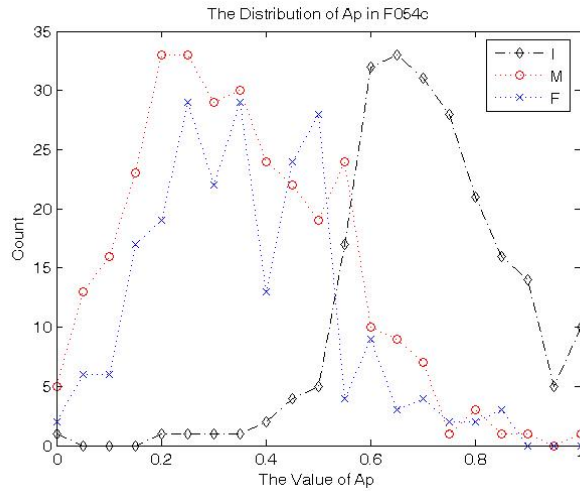
2.4.1 實驗階段一

在實驗階段一檢視 HPG 架構中的居上層的二韻律層。根據 HPG 階層的定義，在 PPh 層中的所有韻律短語 (PPh) 都是 PG 之下的相鄰次級附屬成分 (subjacent subordinate constituents)，彼此之間的關係是姐妹成分 (sister constituents)，沒有了上層 PG 訊息，各韻律短語間都將成為各自獨立、不相干的句調單位 (IU)。處在最接近 PG 層、且在 HPG 階層架構中屬於較高的韻律階層地位，則每一個韻律短語對應於三 PG 位置皆被分派不同的 PG 角色。因此，在實驗階段一第一步驟，我們首先假設在 PPh 層各短語間是獨立的韻律單位，忽略更上層的 PG 訊息，將 Ap 參數標準化後，以分派預測短語內的 Ap 大小；第二步驟，在 PG 層中，考慮到 PG 效應，根據三韻律短語所在的 PG 位置，不同的 Ap 數值以分派預測短語成分的大小；最後，比較有 PG 效應和無 PG 效應的韻律短語與原始語料的預測準確率 (prediction accuracy)。

在實際的分析運算中，我們首先提取語料中的 Ap 參數值並檢視其特徵，做為預測過程的參照。從語料中所提取出的 Ap 參數值在每個 PG 位置個別的音域範圍 (range) 及分布如〈表 2〉和〈圖 4〉所示。接著，將每一 PG 位置 Ap 參數值分布特徵標準化，做為隨後 Ap 預測之用。使用逐步回歸技術 (step-wise regression technique)，建立一適合國語的線性模型以預測 Ap 參數，利用階層式 HPG 組織架構將 Ap 值歸類到每一韻律層。分析過程所有的數據資料都是使用 Data Description 公司所開發的套裝軟體 DataDesk™ 加以計算。

〈表 2〉F054 個別 PG 位置之音域範圍 (range) 值

	音域範圍
PG 首	0.4-1.1
PG 中	0.01-0.8
PG 尾	0.025-0.8



〈圖 4〉F054 承載句 PG 首、中、尾短語之 Ap 參數分布圖

縱軸表示 Ap 發生的次數，橫軸表示 Ap 值。

在與原始語料的數值比較時，我們使用二項基準值以評量預測值與原始數值相近的程度。第一項是誤差平方和 (sum-squared error, SSE) 百分比，我們假設在 HPG 架構中下層誤差（不可解釋的變異量）係缺乏上層訊息所致；因此，殘差 (residual errors, RE) 可定義為在原始語料總平方和中誤差平方和所佔的比率，也就是預測值與原始值的差異；而無法解釋的殘差值則往更上層做進一步的解釋、預測。從 PPh 層向上到 PG 層，我們最終想檢驗 PG 層的貢獻度有多少。如果向上一層預測，預測值（可解釋的變異量）因而獲得改善的話，那麼緊鄰兩層預測值的差異可視做較上層的貢獻度。

2.4.2 實驗階段二

在實驗階段二，使用相似的回歸分析步驟，計算從音節層到 PPh 層每一韻律層 Aa 參數大小（局部基頻曲勢）對最終輸出的貢獻度。

在音節層，採用字調成分方法以趨近每一音節的基頻曲線，換句話說，每一個音節對應一個 Aa 參數。然而，在現階段這樣的方式無法精確地單由音節層擷取出 Aa 參數，於是我們假設了音節層、PW 層與 PPh 層的模型如下：

2.4.2.1 音節層模型

$$(1) \text{ Aa} = \text{constant} + \text{SYL} + \text{Delta } 1$$

以上的 SYL 代表音節類型，考慮因素包括 23 種音節類型（除了目標字之外）與 5 種字調（四聲及輕聲）。

2.4.2.2 PW 層模型

$$(2) \text{ Delta } 1 = f(\text{PW Length}, \text{PW Sequence}) + \text{Delta } 2$$

由於 Fujisaki model 所提出字調成分（Aa 參數），是指在沒有 PW 規範下的獨立字調成分，因此在 PW 層中，我們採用逐步回歸技術 (step-wise regression technique)，在每一個 Aa 成分加上此音節相對於 PW 的資訊，每一音節以一組向量值加以標記。例如：(3, 2) 表示考量的單位是三音節韻律詞中的第二音節，然後使用線性回歸技術計算出每一 PW 內音節所對應的 Aa 係數，將 PW 層的韻律資訊疊加至獨立字調成分。

2.4.2.3 PPh 層模型

$$(3) \text{ Delta } 2 = f(\text{PPh Length}, \text{PPh Sequence}) + \text{Delta } 3$$

每一音節以一組向量值加以標記，例如：(8, 4) 表示考量的單位是 8 音節韻律短語中的第四音節，然後，與前一層計算步驟相同，使用線性回歸技術計算出每一 PPh 內音節所對應的 Aa 係數。

2.5 實驗結果

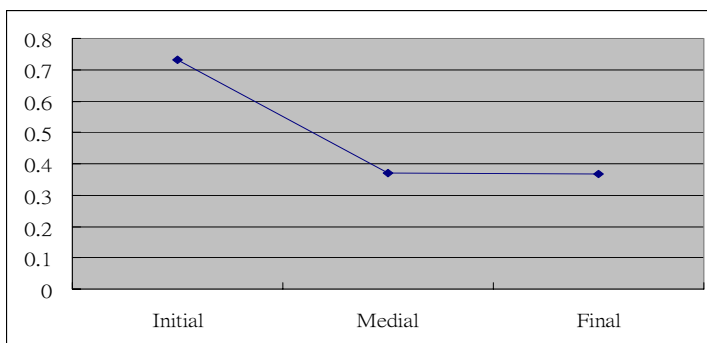
2.5.1 實驗階段一

〈表 3〉說明了從 PPh 到 PG 的 Ap 係數。在 PPh 層中，當每一個 PPh 被當作是獨立的韻律單位時，期望細格平均數 (expected cell mean) 為 0.45602；然而，

在 PG 層中，每個 PPh 依照所在的 PG 位置——PG 首、中、尾加以分類，PG 首、中、尾 PPh 伴隨 PG 效應的期望細格平均數分別是 0.7307、0.371 和 0.368。相對於 PG 首的 PPh，在 PG 尾 PPh 的 A_p 是變短的。PG 首與 PG 尾的係數反映出清楚的對比。

〈表 3〉有無 PG 效應 A_p 係數預測值之期望細格平均數

PPh 層無 PG 效應之期望細格平均數	0.45602		
PG 層有 PG 效應之期望細格平均數	PG 首	PG 中	PG 尾
	0.7307	0.371	0.368

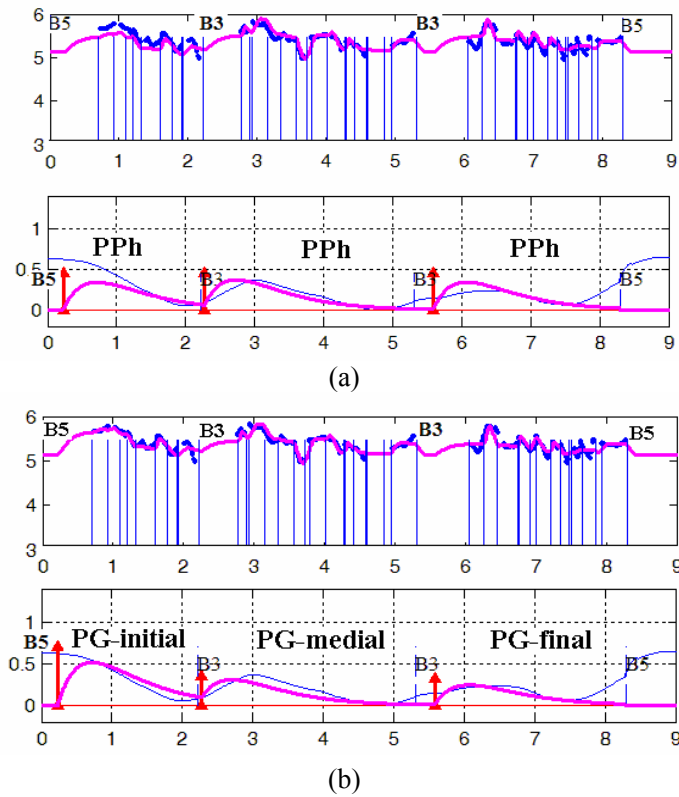


〈圖 5〉考慮 PG 效應後之短語組型圖

注意：PG 首和 PG 尾組前後呼應成為明顯的對比。

當每一個 IU（在我們的 HPG 架構中稱作 PPh）被獨立分析時，結果顯示：正確預測只有 46.49%，53.51% 是誤差。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%。依照 PG 相對位置而做的 A_p 調整，提供了韻律單位與層級如何對語流中的 A_p 起限制作用，以及上層韻律單位的限制因素可能不同於下層韻律單位限制因素的進一步證據。假使忽略了上層訊息，預測結果將不顯著。

最後，隨著 PG 層預測值的累加，我們可以得到 PG 三位置的基頻曲勢配置的預測值。〈圖 5〉和〈圖 6〉表示有無 PG 效應之間的預測與原始語料之比較，最後的累積預測值意謂著國語語流中的基頻配置組型 (patterns of F0 allocation) 無法單由 PPh 層所調整，必須包含 PG 層的結果。並且，此一實驗結果也證明了 PPh 是由更上層訊息管轄與限制。



〈圖 6〉無 PG 效應 (a) 與有 PG 效應 (b) 的預測值與原始語料之比較

上欄粗線顯示承載句 3 短句的基頻原始曲勢，而細線則表示基頻預測曲勢，垂直線代表音節邊界；下欄裡，細線顯示與低通濾波基頻曲勢 (lowpassed F0 curve) 的比較，而粗線則表示預測短語成分 (predicted phrase)，在下欄中每一個垂直箭頭表示 A_p 參數（其高度表示 A_p 參數的大小）。每一欄中，縱軸代表基頻曲勢的對數值，而橫軸表示時間。

總和來說，PPh 層僅構成韻律輸出近 47% 的貢獻度，而在 PG 層的上層訊息則能再多增加 16%。兩者加在一起，PPh 和 PG 層共組成韻律輸出的 63% 貢獻度。然而，值得注意的是，因為在 HPG 階層架構中，PG 層是屬於上層結構，管轄在它之下的所有短句，自然其效力不可忽視。沒有了 PG 層，就沒有語篇韻律。根據 HPG 階層架構的定義，其餘的 37% 應該來自較下層的音節（字調）及詞組（句法詞與韻律詞）層。在 HPG 階層架構下，向上層研究發現，字調訊息在連續語流中並非最顯著的貢獻因素。換言之，若個別檢視語流中的字調，會發

現大部分的字調調型不但不完整、且多數調型不明顯 (Tseng 1990, 2006b)。而短語中個別短語句調，亦即本實驗的 PPh 層，也僅構成韻律輸出近 40% 而已。因此同理亦可推論，句調訊息在連續語流中也並非最顯著的貢獻因素。

2.5.2 實驗階段二

〈表 4〉顯示 Aa 分析中各韻律層的貢獻度與累積預測正確率。

〈表 4〉韻律各層級對 Aa 預測分析的貢獻度與累積預測正確率

韻律層級	貢獻度	累積預測正確率
音節層	28.93%	28.93%
PW 層	2.36%	31.28%
PPh 層	2.00%	33.29%

如果只考慮五種字調而忽略音節種類時，Aa 預測正確率只有大約 25.71%，當包含音節種類，音節層的累積正確率提高到 28.93%，由音節層向上至 PW 層，累積預測值為 31.28%，最後 PPh 層，Aa 預測值的累積正確率為 33.29%。根據以上結果，我們得以釐清各韻律層的貢獻度，並導出 HPG 的基頻組型。

3. 實驗二：語流中音節時長分布預測的分析與比較

3.1 實驗目的與假設

本實驗的目的是證明國語語流韻律的節奏，並非由單一音節的時長串接所構成，不可視為音節時長的串接；也不止於詞組中音節數目的節律，如五字分為二加三（王洪君 1996）。我們的假設是：語流中各個音節的時長，需受制於單音節、詞組、句組及上層資訊各韻律層級的管轄制約，各層級必須互動，加總（含相抵）後構成語流的時長輸出，從音節時長在時程上的配置來表達。與一般研究節奏的論述最大的不同，是詞組以上，還存有對應於 HPG 韻律層的韻律時長深層結構。我們假設：(1) 語流的音節時長組型與 HPG 韻律層對應，可由語料中通過量化程式導出；(2) 語流韻律的音節時長組型及時長配置可由語流韻律組織的管轄效應解釋；(3) 台灣地區的國語與北京普通話的深層語流音節結構是一致的，二者雖有口音差異，但韻律時長的區分特徵另有因素。

3.2 實驗語料

在實驗二中，我們共分析了三批國語暨普通話語料庫，其中包括兩位台灣國語播音員之朗讀語料（COSPRO 05，一男一女，以下稱之為 TMS 和 TFS）以及一位北京女性播音員朗讀語料（以下稱之為 BFS）。其朗讀文本包含 15 段長段落，段落字數從 85 到 981 個音節不等，每一個段落平均字數 377 音節。這三批語料在性質上，與實驗一所設計的承載語段有很大的不同，它們除了包含典型的韻律句組的組型，每一 PG 也都有超過三個以上的短句，最長的 PG 有 12 個短語 PPh，一口氣說不完，必須中途換氣，所以在標註時 HPG 的 BG 層和 PG 層保留，我們也得到一個 PG 大約有 1.4 個 BG 的結果。不過，由於含 BG 的 PG 語料樣本有限，大部分的語料都在一個 BG 以內，也就是說，BG 層就是 PG 層，因此我們在本節時長預測分析及下一節（第 4 節）音強分析的實驗中，統計的計算只到 12 個 PPh、一個 BG 內，此類 BG 雖然根據定義也等於 PG 層，但為了忠於語料，我們仍以 BG 層表示，因而有 BG 首、中、尾的計算，而未有 PG 首、中、尾的計算，特在此先予說明。

本批語料在平均語速上，亦因語者有很大的不同。其中說的最快的是台灣男性播音員 TMS，平均音節長度為 182 ms，平均語速（包含停頓）每秒 4.3 個音節；最慢則為北京女性播音員 BFS，平均音節長度為 267ms，每秒 2.8 音節。而且 BFS 語料也顯示出：較長的音節時長與語流中較多停頓等慢語速的語流韻律特徵。三批語料的語料內容及語速特徵摘要如〈表 5〉所示：

〈表 5〉語料內容及語速特徵分析

語料 特徵	TFS	TMS	BFS
語料大小 (MB)	41.2	40.2	59.6
錄音時長	22m23s	22m50s	32m26s
音節總數	5655	5655	5483
音節平均時長 (ms) (僅考慮音段時長)	202.45	181.84	266.91
語速 (Syllable #/sec) (包含停頓)	4.23	4.33	2.816

3.3 語料處理與標記

首先以 HTK Toolkit 程式判定音段成分（聲母、韻母）再以 SAMPA-T (Tseng & Chou 1999) 自動標記音段邊界，而後以人工標記感知到的韻律邊界。不管是音段邊界標記人工檢查、抑或停延標記皆是由訓練有素的標記員所完成。

3.4 分析與量化

我們使用在 Tseng 等人 (2004) 研究中提出的研究步驟及方法，也就是線性回歸統計中的逐步回歸技術 (step-wise regression technique) 來分析估算語料。在本實驗中，我們建立一個四階層的線性模型以預測發音人的時長組型。此一階層式的回歸模型由下到上、一層一層地由下到上往上疊加，依序為：音節層、韻律詞層 (PW)、韻律短語層 (PPh) 以及韻律句組層 (BG)，最後把每一層的評估結果加總，得到最後的評估結果，再與實際語料比對，以檢視評估的比例是否有意義。全部過程都是要研究自變項 (IVs)：音段成分 (segment identities)，韻律架構 (prosodic organization) 與依變項 (DV)：音節時長 (syllable duration) 之間的關係，其分析步驟摘要如下：(1) DVs 直線轉換，(2) 決定音段成分的分群，(3) 建立線性回歸模型，(4) 刪減無效之 IVs，(5) 無法預測之殘差值 (residual) 由緊接著的更上層解釋之，重複步驟 (3) 與 (5)。

爲了要有更好的預測，DV 必須盡可能地呈現常態化分布。DV 與其轉換值 (log, square root, etc) 必須使用 N-plot 加以檢驗、比較何者接近常態化分布，最接近常態化分布者因而入選。於是，在接下來的回歸分析過程中，我們決定使用三批語料原始時長數值。

間斷資料之線性模型使用 DataDesk™ 的部分平方和 (partial sum of squares) (type3) 建立之。在音節層中，將語料中所有的音段分成若干的群組，以減少回歸分析的複雜度，是必要的作法。以時長回歸預測來說，決定分為 8 組子音群與母音群。分組的標準係依音節層中每個 IVs 項目的 DV 平均值而決定。換言之，具有相近的時長平均數之音段分在同一組，音節聲母類型、韻母類型以及字調類型則被當作自變項。同時，音節前後位置因素亦是考慮因素之一，另外，我們也考慮了其二因子交互作用，所以音節層之回歸分析可用下列公式表示：

$$\begin{aligned}
 (4) \quad \text{DV} \quad \text{IVs} \\
 \text{Dur (ms)} = & \text{constant} + CTy + VTy + Ton \\
 & + PCT + PVT + PrT + FCt + FVt + FlT \\
 & + 2\text{-way factors of each factor above} \\
 & + \text{Delta 1}
 \end{aligned}$$

回歸分析之後，將較小影響力的因素 ($p\text{-value} > 0.1$) 予以排除，無法由音段成分所預測之殘差 (Delta 1)，則移到再上一層分析。從 PW 層到 BG 層，我們以韻律結構當作自變項，所擷取出的係數表示韻律單元在特定音節位置的效應值（時長的增加或減少）。

在 PW 層中，我們的目的在於瞭解 DV 是否受其在 PW 的位置所影響，因此 PW 層模型之公式表示如下：

$$(5) \quad \text{Delta 1} = f(\text{PW Length, PW Sequence}) + \text{Delta 2}$$

無法由 PW 結構所預測之 PW 層的殘差：Delta 2，在緊接的上一層 PPh 層繼續分析。從我們最新的研究中發現，PPh 組織造成的效應，在 PPh 中的第一與最後 4 音節的位置最為顯著，因此我們標記少於 8 音節的 PPh 的音節韻律序列資訊為 [PPh 的長度，該音節在 PPh 中的序列位置]，對大於或等於 9 音節者，我們逐一標記前 4 與後 4 音節，而其餘中間音節為 [M]，也就是 { [I1], [I2], [I3], [I4], [M]...[M], [F1], [F2], [F3], [F4] } 的標記方式。如此一來，我們可略過語料中長 PPh 不足的問題。PW 層之公式表示如下：

$$(6) \quad \text{Delta 2} = f(\text{PP Length, PP Sequence}) + \text{Delta 3}$$

無法由 PW 與 PPh 所預測之 PPh 層的殘差：Delta 3，在緊接的上一層 BG 層繼續分析。我們標記 BG 中的第一與結尾 PPh，而其餘的 PPh 則一視同仁，當作中段的 PPh。在每一個 PPh 中，少於 7 音節的 PPh 的音節韻律序列資訊為 [PPh 長度，該音節在 PPh 中的序列位置]，對大於或等於 7 音節者，我們逐一標記前 3 與後 3 音節，而其餘中間音節則一視同仁，於是，BG 內的起首 PPh 的音節位置可標記為 { [II1], [II2], [II3], [IM]...[IM], [IF1], [IF2], [IF3] }。而 BG 層之公式表示如下：

$$(7) \quad \text{Delta 3} = f(\text{PPh I|M|F, PPh Length, PPh Sequence}) + \text{Delta 4}$$

採用兩項數值評估預測結果：(1) 相關係數 (Correlation Coefficient, r)，表示預測結果與原始資料的關聯程度。(2) 總殘差誤差 (T.R.E.) 為殘差平方和與原始資料平方和百分比。T.R.E. 意謂著由底層之音節層到目前這一層為止無法說明的殘差誤差比率。

3.5 實驗結果

3.5.1 音節層

在這一層中，每一批語料的回歸分析使用不同的音段分群，〈表 6〉到〈表 8〉描述了每一批語料中在時長回歸分析中的音段分群。其分群標準係根據每一音段成分的平均時長數值，將相近平均數的音段成分分在同一群。同時使用單因子和二因子之回歸分析，並刪減（忽略） p 值大於 0.1 之因素。〈表 9〉到〈表 11〉顯示出剩餘因素之 ANOVA 摘要表。在所列因素中，CTy/CT 代表子音因素，VTy/VT 代表母音因素，而 Ton/Tn 代表字調因素。以 P 和 F 為首碼者代表前/後音節的貢獻因素。

〈表 6〉TFS—時長音段分群

Group	Consonant	Group	Vowel
CDUR1	d, b, g	VDUR1	@, o, U', U
CDUR2	dz', l, f	VDUR2	i, u, a, ei
CDUR3	n, Z'	VDUR3	yE, y, @n, in
CDUR4	m, dz, dj	VDUR4	uo, iE, ai, ou, uei
CDUR5	t, p, k, h	VDUR5	@N, oN, iN, an, au
CDUR6	s', ts', sj, s	VDUR6	yn, iau, aN
CDUR7	ts, tj	VDUR7	ia, iou, u@n, @', iEn, ua
CDUR8	Zero	VDUR8	uan, yEn, iaN, uaN, uai, yoN

〈表 7〉TMS—時長音段分群

Group	Consonant	Group	Vowel
CDUR1	d, b, g	VDUR1	@, o
CDUR2	dz', l	VDUR2	o, U, ei
CDUR3	n, f, Z'	VDUR3	i, u
CDUR4	m, t, p, dz, dj, k	VDUR4	a, in, uo, y, @n, iE, yE, ou
CDUR5	h, ts'	VDUR5	uei, iN, ai, @n, yn
CDUR6	ts, sj, tj	VDUR6	oN, iou, au, aN, iau, an
CDUR7	s, s'	VDUR7	ia, iEn, u@n, uaN, uai, ua, uan
CDUR8	Zero	VDUR8	yEn, iaN, yoN

〈表 8〉BFS—時長音段分群

Group	Consonant	Group	Vowel
CDUR1	d, b	VDUR1	@, U, U'
CDUR2	g, l	VDUR2	u, o
CDUR3	n, dz', Z', m	VDUR3	a, I, yE, ou, ua
CDUR4	dz, dj	VDUR4	ei, ai, uo, au
CDUR5	t, p, f, k, h	VDUR5	oN, in, @n, u@n, uaN, @N, an, aN, iE, iN
CDUR6	ts'	VDUR6	uei, y, ia, yn, uan, iau
CDUR7	tj, s', sj, s, ts	VDUR7	iEn, @', iaN, iou
CDUR8	Zero	VDUR8	yEn, uai, yoN

〈表 9〉TFS—時長—音節層因素之 ANOVA 摘要表

Source	df	Sums of Squares	Mean Square	F-ratio	Prob	Source	df	Sums of Squares	Mean Square	F-ratio	Prob
CTy	7	615273	87896.1	42.767	<0.0001	CTy*VTy	49	199615	4073.77	1.9821	<0.0001
VTy	7	367437	52491.1	25.54	<0.0001	CTy*FTn	35	140762	4021.78	1.9568	0.0006
Ton	4	109460	27365	13.315	<0.0001	FCT	8	31730.5	3966.32	1.9299	0.0514
Ton*FCT	32	356105	11128.3	5.4146	<0.0001	PCT*PTn	25	95864.1	3834.56	1.8658	0.0056
CTy*Ton	25	183320	7332.81	3.5679	<0.0001	CTy*PTn	35	126427	3612.19	1.7576	0.0039
VTy*Ton	23	152705	6639.33	3.2305	<0.0001	PCT*PVT	49	173919	3549.36	1.727	0.0013
PVT	7	42277.2	6039.59	2.9386	0.0045	PVT*PTn	23	76195.2	3312.83	1.6119	0.0324
PTn	4	23319.3	5829.82	2.8366	0.023	FVT*FTn	23	74873	3255.35	1.5839	0.0377
PCT	8	37596.2	4699.52	2.2866	0.0193	Ton*PCT	31	96176.4	3102.46	1.5095	0.0348
FCT*FVT	56	262766	4692.25	2.2831	<0.0001	VTy*FCT	56	156598	2796.39	1.3606	0.0387
FCT*FTn	29	134237	4628.86	2.2522	0.0001	Const	1	2.31E+08	2.31E+08	1.12E+05	<0.0001
Error	5118	1.05E+07	2055.23			Total	5654	2.40E+07			

〈表 10〉TMS—時長—音節層因素之 ANOVA 摘要表

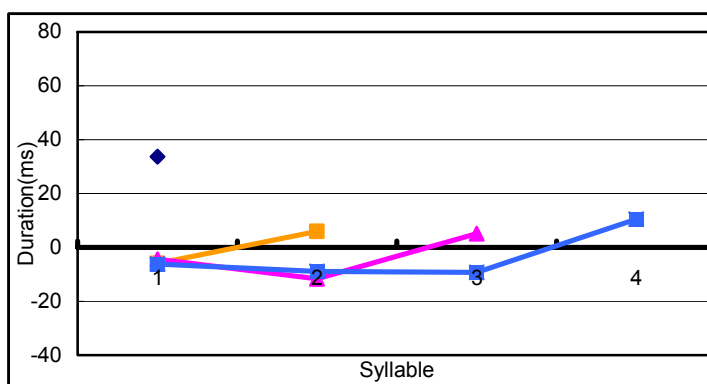
Source	df	Sums of Squares	Mean Square	F-ratio	Prob	Source	df	Sums of Squares	Mean Square	F-ratio	Prob
CTy	7	175279	25039.9	14.276	<0.0001	CTy*FTn	35	122844	3509.82	2.001	0.0004
PVT	9	74244.5	8249.38	4.7031	<0.0001	PCT*PTn	31	103806	3348.59	1.9091	0.0018
Ton*FTn	20	163881	8194.05	4.6716	<0.0001	PVT*PTn	26	80244.6	3086.33	1.7596	0.01
FCT	8	56309.5	7038.69	4.0129	<0.0001	VTy*FVT	49	148264	3025.8	1.7251	0.0013
FVT*FTn	26	178220	6854.63	3.9079	<0.0001	FCT*FVT	50	146530	2930.6	1.6708	0.0022
VTy	7	45882.4	6554.62	3.7369	0.0005	VTy*PCT	56	148779	2656.77	1.5147	0.0081
VTy*Ton	26	146987	5653.35	3.2231	<0.0001	VTy*FTn	28	72556.2	2591.29	1.4773	0.0504
CTy*VTy	43	226817	5274.82	3.0073	<0.0001	VTy*PVT	49	119575	2440.3	1.3913	0.0373
PTn	4	18084.9	4521.23	2.5776	0.0356	VTy*FCT	56	136265	2433.31	1.3873	0.0301
CTy*Ton	26	116595	4484.44	2.5567	<0.0001	PTn*FCT	39	92503.4	2371.88	1.3522	0.0709
FCT*FTn	29	128627	4435.42	2.5287	<0.0001	Const	1	2.08E+08	2.08E+08	1.19E+05	<0.0001
Error	5030	8.82E+06	1754.03			Total	5654	2.05E+07			

〈表 11〉BFS—時長—音節層因素之 ANOVA 摘要表

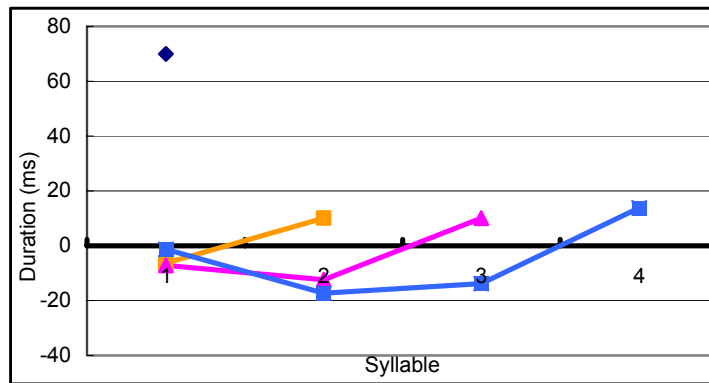
Source	df	Sums of Squares	Mean Square	F-ratio	Prob	Source	df	Sums of Squares	Mean Square	F-ratio	Prob
CTy	7	417148	59592.5	14.966	<0.0001	FTn	4	30036.5	7509.13	1.8858	0.11
Ton	4	96536.5	24134.1	6.0609	<0.0001	VTy*FTn	28	197720	7061.43	1.7734	0.0073
CTy*VTy	46	655573	14251.6	3.5791	<0.0001	VTy*PTn	33	223588	6775.41	1.7015	0.0075
FVT	7	85042.7	12149	3.051	0.0033	VTy*FCT	56	371720	6637.86	1.667	0.0014
Ton*FTn	16	184510	11531.9	2.896	<0.0001	FCT	8	53079.5	6634.94	1.6663	0.1013
VTy*Ton	26	294887	11341.8	2.8483	<0.0001	Ton*PCT	32	208780	6524.38	1.6385	0.0132
FCT*FTn	27	298044	11038.7	2.7722	<0.0001	VTy*PCT	56	351417	6275.3	1.5759	0.0041
Ton*FCT	32	312549	9767.16	2.4529	<0.0001	PVT*PTn	30	185741	6191.37	1.5549	0.0275
FVT*FTn	26	235986	9076.38	2.2794	0.0002	PCT*PVT	59	343430	5820.84	1.4618	0.0125
FCT*FVT	44	342737	7789.47	1.9562	0.0002	VTy*FVT	49	262980	5366.94	1.3478	0.0537
CTy*Ton	27	209032	7741.94	1.9443	0.0024	Const	1	3.91E+08	3.91E+08	98096	<0.0001
Error	4865	1.94E+07	3981.94			Total	5482	3.89E+07			

3.5.2 PW 層

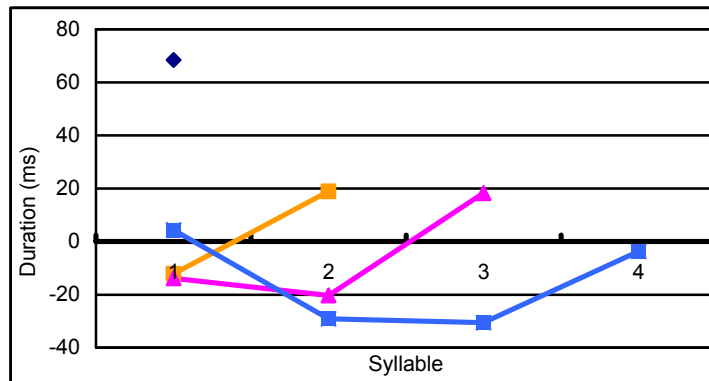
〈圖 7〉到〈圖 9〉顯示出在第三批語料中 PW 層的回歸係數。這些圖當中，每一條線代表不同長度的韻律詞單元，而 Y 軸中每一點代表在一個韻律詞中特定位置音節所得來的回歸係數，這些係數表示相較於低一層所獲得的預測，該音節時長延長/縮短的程度。結果顯示，PW 層的一般組型像一勺子狀曲線——PW 的中間音節縮短而最後音節拉長。在第三批語料中，TFS 的延長/縮短效應最小，在四音節的 PW 中最短和最長音節相差 19.68ms，是最為顯著的差異；BFS 在 3 音節 PW 則有 39.52ms 的差異。



〈圖 7〉TFS—時長—PW 層之回歸係數



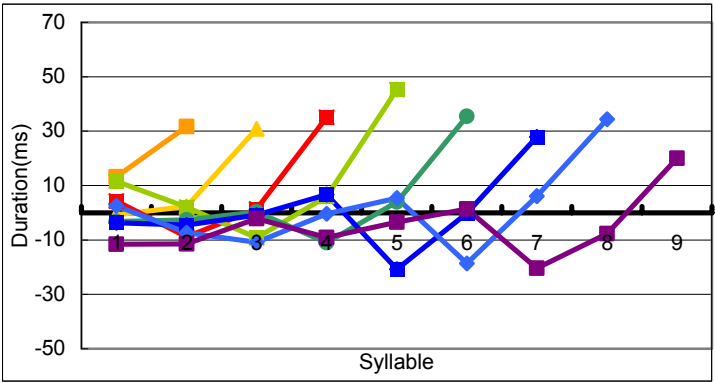
〈圖 8〉TMS－時長－PW 層之回歸係數



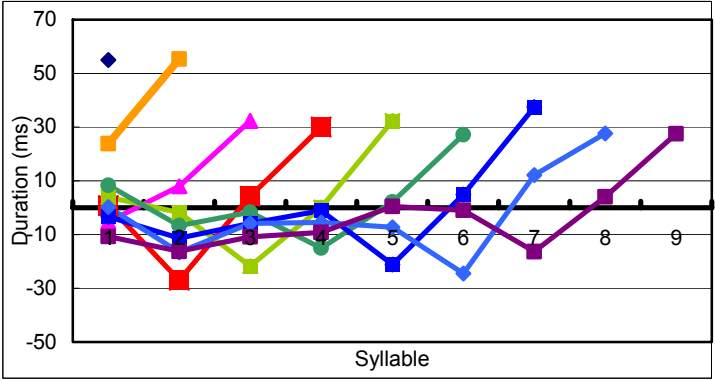
〈圖 9〉BFS－時長－PW 層之回歸係數

3.5.3 PPh 層

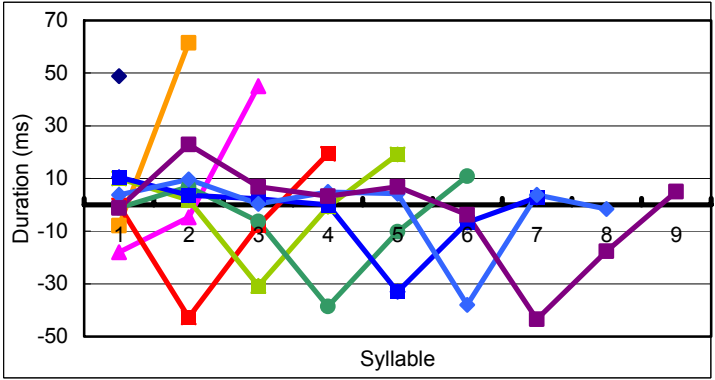
〈圖 10〉到〈圖 12〉顯示了三批語料中 PPh 層的回歸係數。PPh 層的一般組型可從 PPh 中最後 4 音節看出。不管什麼音節長度的韻律短語，從最後倒數第 4 個音節開始皆維持相同的組型，（倒數）第三音節明顯最短，而最後音節則拉長，這便是眾所熟知的「短語結尾拉長」(phrase final lengthening)，其實它就是不同韻律層級的邊界效應。TMS 和 TFS 在短語結尾表現出相同的組型，BFS 則有倒數第三音節最短，最後音節卻沒有拉長的傾向。



〈圖 10〉TFS－時長－PPh 層之回歸係數



〈圖 11〉TMS－時長－PPh 層之回歸係數



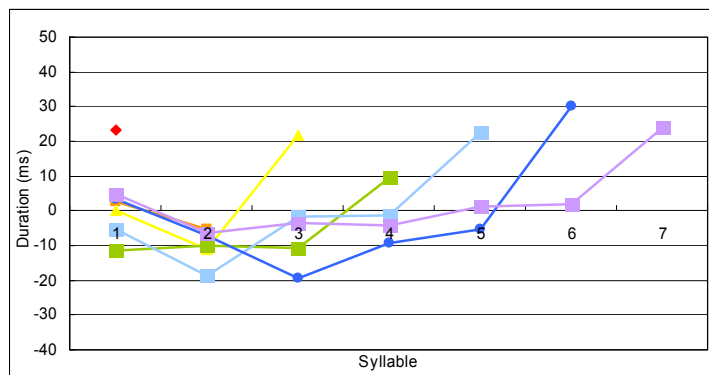
〈圖 12〉BFS－時長－PPh 層之回歸係數

3.5.4 BG 層

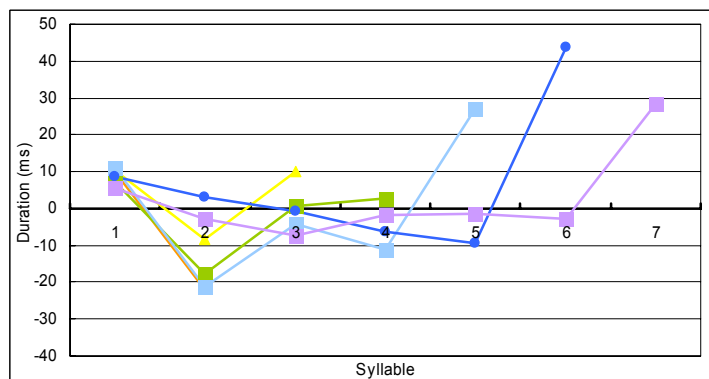
在 BG 中，我們假設在 BG 中不同位置的 PPh 的貢獻度並不相同，而進行個別分析。BG 中第一與最後的 PPh 視為起首與結尾 PPh，而其他的則為中段 PPh。

3.5.4.1 BG 中的起首 PPh (BG-initial PPh)

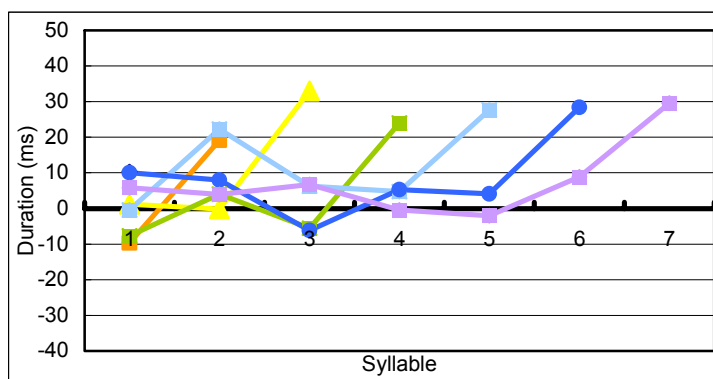
〈圖 13〉到〈圖 15〉顯示了三批語料中 BG 層起首 PPh 的回歸係數。三批語料中，由起首 PPh 所貢獻之一般時長組型十分相似。相較於 PPh 的預測，BG 層起首 PPh 短語末音節是更加地延長，起首 PPh 的最後音節均被延長了 23.85 (TFS) 到 29.55ms (BFS) 不等的長度。



〈圖 13〉 TFS－時長－BG 中起首 PPh 之回歸係數



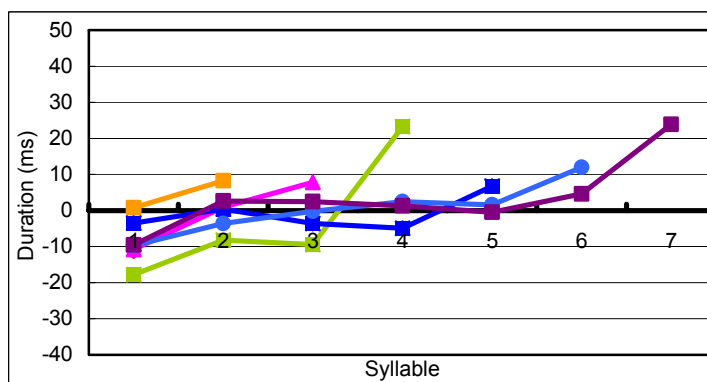
〈圖 14〉 TMS－時長－BG 中起首 PPh 之回歸係數



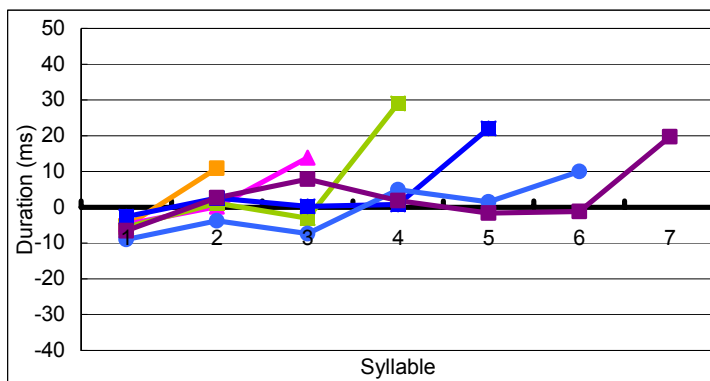
〈圖 15〉 BFS—時長—BG 中起首 PPh 之回歸係數

3.5.4.2 BG 中的中段 PPh (BG-medial PPh)

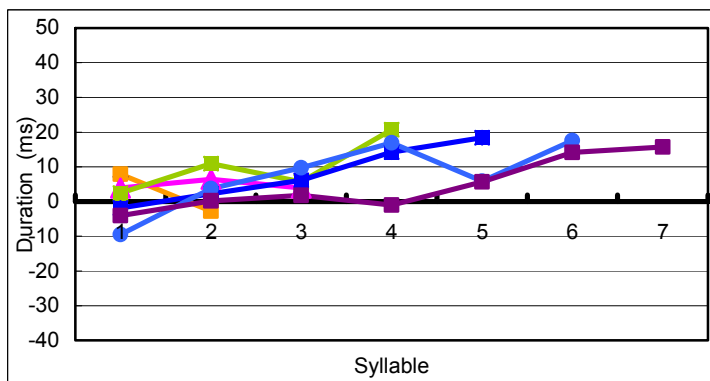
〈圖 16〉到〈圖 18〉顯示了三批語料中 BG 層中段 PPh 的回歸係數。除了已知的 PPh 結尾延長外，在這一層中，BG 的中段 PPh 更加地延長。在這裡，必須注意起首 PPh 與中段 PPh 之間的差異：(1) 起首 PPh 比中段 PPh 多了 10ms 左右，具有較大的延長效應。(2) 中段 PPh 的第一個音節時長縮短了 10 ms，這樣的現象在起首 PPh 並未發現。



〈圖 16〉 TFS—時長—BG 中段 PPh 之回歸係數



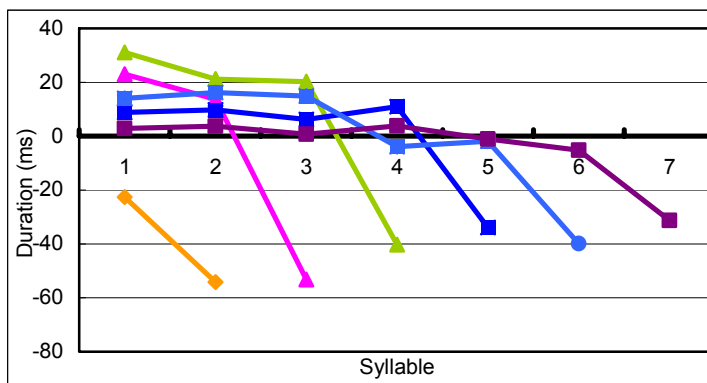
〈圖 17〉TMS－時長－BG 中段 PPh 之回歸係數



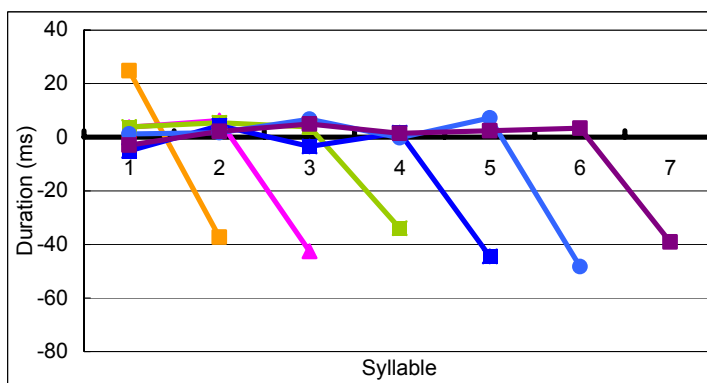
〈圖 18〉BFS－時長－BG 中段 PPh 之回歸係數

3.5.4.3 BG 中的結尾 PPh (BG-final PPh)

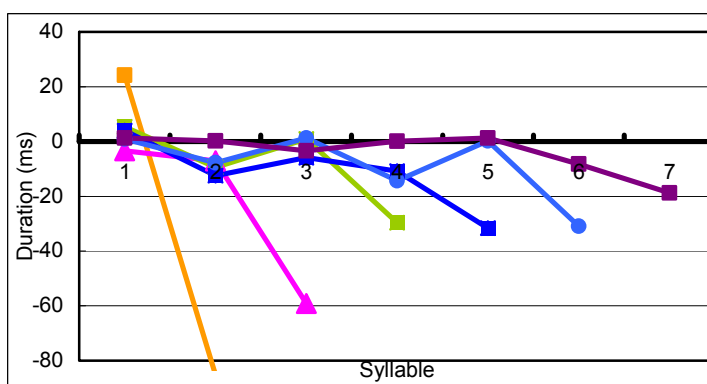
〈圖 19〉到〈圖 21〉顯示了第三批語料中 BG 層結尾 PPh 的回歸係數。BG 的結尾 PPh 的最後音節並沒有如起首與中段 PPh 一樣拉長，而是明顯地縮短。對超過 7 音節的結尾 PPh，TFS 縮短了 31.3ms，TMS 縮短 45.55ms，BFS 則為 18.8ms。然而，BG 最後音節的全面性效果應是預測結果的累積總和，其實際韻律表現依然呈現延長現象。



〈圖 19〉 TFS—時長—BG 結尾 PPh 之回歸係數



〈圖 20〉 TMS—時長—BG 結尾 PPh 之回歸係數



〈圖 21〉 BFS—時長—BG 結尾 PPh 之回歸係數

從以上結果，我們可以瞭解到：BG 層中不同位置的 PPhs 在已預測的 PPh 層之上貢獻了它本身的時長組型。〈表 12〉列出各層的全面性預測評估值，全面性評估值之相關係數 r ，TFS=0.806，TMS=0.842 以及 BFS=0.819。

〈表 12〉三批語料時長預測之全面性評估

Layer \ Corpus	TFS		TMS		BFS	
	T.R.E.	r	T.R.E.	r	T.R.E.	r
Syllable	43.88%	0.749	43.11%	0.754	49.97%	0.709
PW	42.43%	0.759	39.87%	0.776	43.22%	0.755
PP	38.07%	0.789	33.41%	0.822	36.13%	0.803
BG	35.13%	0.806	29.37%	0.842	33.56%	0.819

〈表 12〉中 T.R.E. 值仍有 30-35%，此數值仍偏高，可能變異來源為樣本數目仍可以增加，及未能考慮影響時長的其他參數，如因 PG 位置所造成的 PPh 短語總體時長的相對快慢調節等。

4. 實驗三：語流音強分布預測分析與比較

4.1 實驗目的與假設

我們先前已發現，語流中的時程分布及音節長短上可以由音段成分與韻律組織訊息所預測 (Tseng et al. 2004)。在本實驗中，我們的目的是要瞭解音強的分布模式，是否也可由 HPG 架構預測及解釋。我們因此假設：(1) 語流音強組型是否如時程分布一般，可由語料中通過量化程式導出；(2) 韻律音強組型可由韻律組織的管轄效應解釋；(3) 語流音強的單位在 HPG 的下層（或較小韻律單位）的意義不明，到短句層才有意義；(4) 台灣國語與北京普通話在韻律音強組型差異不大，跨短句音強分布組型不是口音區分的主要表現。

4.2 實驗語料

同實驗二。

4.3 語料處理與標記

首先以 HTK Toolkit 程式判定音段成分（聲母、韻母）再以 SAMPA-T (Tseng

& Chou 1999) 自動標記音段邊界，而後以人工標記感知到的韻律邊界。不管是音段邊界標記人工檢查、抑或停延標記皆是由訓練有素的標記員所完成。

在實驗數值截取方面，首先使用 ESPS Toolkit 取音段 RMS 值，每一音段以十個音框 (frame) 為一單位，計算 10 個音框的 RMS 平均值，少於 10 個音框的音段時長則直接取平均數。除此以外，為了排除因為每一段落落在錄音過程中的些微改變所造成的韻律層級差異，每一段落中的 RMS 值皆經過標準化過程，得到 NRMS 值。如此一來，我們即可將研究焦點放在韻律單元內的音強組型的比率差異，而不是每一音節所取得的絕對音強數值。

4.4 分析與量化

在音強預測部分，除了音節層的音段分群外，所有過程都與時長的回歸分析一樣。全部過程都是要研究自變項 (IVs)：音段成分 (segment identities)，韻律架構 (prosodic organization) 與依變項 (DV)：NRMS 之間的關係，其分析步驟摘要如下：(1) DVs 直線轉換，(2) 決定音段成分的分群，(3) 建立線性回歸模型，(4) 刪減無效之 IVs，(5) 無法預測之殘差值 (residual) 由緊接著的更上層解釋之，重複步驟 (3) 與 (5)。

爲了要有更好的預測，DV 必須盡可能地呈現常態化分布。DV 與其轉換值 (log, square root, etc) 必須使用 N-plot 加以檢驗、比較何者接近常態化分布，最接近常態化分布者因而入選。於是，在接下來的回歸分析過程中，我們決定使用三批語料的 NRMS 平方根。

間斷資料之線性模型使用 DataDesk™ 的部分平方和 (partial sum of squares) (type3) 建立之。在音節層中，將語料中所有的音段分成若干的群組，以減少回歸分析的複雜度，是必要的作法。以音強回歸預測來說，我們決定分成 10 組子音群與母音群。分組的標準係依音節層中每個 IVs 項目的 DV 平均值而決定。換言之，具有相近的 NRMS 平均數之音段分在同一組，音節聲母類型、韻母類型以及字調類型則被當作自變項。同時，音節前後位置因素亦是考慮因素之一，另外，我們也考慮了其因子交互作用，所以音節層之回歸分析可用下列公式表示：

$$\begin{array}{ll}
 \text{(8)} & \text{DV} \qquad \qquad \qquad \text{IVs} \\
 \text{Square root NRMS} & = \text{constant} + CT_y + VT_y + Ton \\
 & \quad + PCT + PVT + PrT + FCt + FVt + FIT \\
 & \quad + \text{2-way factors of each factor above} \\
 & \quad + \text{Delta 1}
 \end{array}$$

回歸分析之後，將較小影響力的因素 ($p\text{-value} > 0.1$) 予以排除，無法由音段成分所預測之殘差 (Delta 1)，則移到再上一層分析。從 PW 層到 BG 層，我們以韻律結構當作自變項，所擷取出的係數表示韻律單元在特定音節位置的效應值（時長或音強的增加或減少）。

在 PW 層中，我們的目的在於瞭解 DV 是否受其在 PW 的位置所影響，因此 PW 層模型之公式表示如下：

$$(9) \text{ Delta } 1 = f(\text{PW Length}, \text{PW Sequence}) + \text{Delta } 2$$

無法由 PW 結構所預測之 PW 層的殘差：Delta 2，在緊接的上一層 PPh 層繼續分析。從我們最新的研究中發現，PPh 組織造成的效應，在 PPh 中的第一與最後 4 音節的位置最為顯著，因此我們標記少於 8 音節的 PPh 的音節韻律序列資訊為 [PPh 的長度，該音節在 PPh 中的序列位置]，對大於或等於 9 音節者，我們逐一標記前 4 與後 4 音節，而其餘中間音節為 [M]，也就是 { [I1], [I2], [I3], [I4], [M]...[M], [F1], [F2], [F3], [F4] } 的標記方式。如此一來，我們略過語料中長 PPh 不足的問題。PW 層之公式表示如下：

$$(10) \text{ Delta } 2 = f(\text{PP Length}, \text{PP Sequence}) + \text{Delta } 3$$

無法由 PW 與 PPh 所預測之 PPh 層的殘差：Delta 3，在緊接的上一層 BG 層繼續分析。我們標記 BG 中的第一與結尾 PPh，而其餘的 PPh 則一視同仁，當作中段的 PPh。在每一個 PPh 中，少於 7 音節的 PPh 的音節韻律序列資訊為 [PPh 長度，該音節在 PPh 中的序列位置]，對大於或等於 7 音節者，我們逐一標記前 3 與後 3 音節，而其餘中間音節則一視同仁，於是，BG 內的起首 PPh 的音節位置可標記為 { [II1], [II2], [II3], [IM]...[IM], [IF1], [IF2], [IF3] }。而 BG 層之公式表示如下：

$$(11) \text{ Delta } 3 = f(\text{PPh I|M|F}, \text{PPh Length}, \text{PPh Sequence}) + \text{Delta } 4$$

採用兩項數值評估預測結果：(1) 相關係數 (Correlation Coefficient, r)，表示預測結果與原始資料的關聯程度。(2) 總殘差誤差 (T.R.E.) 為殘差平方和與原始資料平方和百分比。T.R.E. 意謂著由底層之音節層到目前這一層為止無法說明的殘差誤差比率。

4.5 實驗結果

4.5.1 音節層

在這一層中，每一批語料的回歸分析使用不同的音段分群，〈表 13〉到〈表 15〉描述了每一批語料在音強回歸分析中的音段分群。其分群標準係根據每一音段成分的 NRMS 的平方根數值的平均數，將相近平均數的音段成分分在同一群。語料間的子音群分群比起母音群分群更加地近似。同時使用單因數和二因數之回歸分析，並刪減（忽略）p 值大於 0.1 之因素。〈表 16〉到〈表 18〉顯示出剩餘因素之 ANOVA 摘要表。在所列因素中，CRM/CR 代表子音因素，VRM/VR 代表母音因素，而 Ton 代表字調因素。以 P 和 F 為首碼者代表前/後音節的貢獻因素。

〈表 13〉TFS—音強音段分群

Group	Consonant	Group	Vowel
CRMS1	h, s, k, p, s, ts', t	VRMS1	iau, @, y, ua, ia, U
CRMS2	tj, sj, dj, dz	VRMS2	i, U', ai, u, a, ou, iE
CRMS3	f, s', dz', g	VRMS3	au, yEn, an, iou, iaN
CRMS4	Z', b, d, l	VRMS4	yE, uai, @', yn, uan, iEn, aN, uei
CRMS5	m, n	VRMS5	uaN, @n, iN, @N, o
CRMS6	Zero	VRMS6	u@n, uo, oN, in, yoN, ei

〈表 14〉TMS—音強音段分群

Group	Consonant	Group	Vowel
CRMS1	ts, s, p, h, dz, tj, ts', k	VRMS1	y, U, i, u, U'
CRMS2	dj, t, sj, s'	VRMS2	iN, in, @N, iE, yE, @, yn, iEn
CRMS3	f, dz', g	VRMS3	lou, u@n, oN, @n
CRMS4	Z', l, b	VRMS4	iaN, ei, iau, yEn, o, yoN, ou, uei
CRMS5	d, m, n	VRMS5	uan, aN, an, ai, @', uo, uaN, uai
CRMS6	Zero	VRMS6	Ia, au, a, ua

〈表 15〉BFS—音強音段分群

Group	Consonant	Group	Vowel
CRMS1	ts, s, h, f, p, dz	VRMS1	@, yEn, U, u, y, iaN, iN
CRMS2	dj, t, ts', k	VRMS2	yoN, aN, u@n, iEn, i
CRMS3	sj, tj	VRMS3	@N, oN, an, @n, au
CRMS4	g, l, dz', Z'	VRMS4	U', iau, ai, uai, ia, in, uan, uaN, ua
CRMS5	m, s'n, d, b	VRMS5	yE, uei, iou, ou, iE, ei, yn
CRMS6	Zero	VRMS6	uo, o, a

〈表 16〉TFS—音強—音節層因素之 ANOVA 摘要表

Source	df	Sums of Squares	Mean Square	F-ratio	Prob	Source	df	Sums of Squares	Mean Square	F-ratio	Prob
CRM	5	7.19217	1.43843	107.64	<0.0001	VRM*FTn	21	0.577025	0.027477	2.0561	0.0031
PVR	7	2.42141	0.345916	25.885	<0.0001	PTn*Ton	19	0.51406	0.027056	2.0246	0.0053
FCR*FVR	37	4.13819	0.111843	8.3691	<0.0001	CRM*VRM	25	0.674099	0.026964	2.0177	0.002
VRM*Ton	20	1.04553	0.052276	3.9118	<0.0001	VRM*FCR	34	0.838438	0.02466	1.8453	0.002
FTn	4	0.196686	0.049171	3.6794	0.0054	VRM*PCR	35	0.796716	0.022763	1.7034	0.0061
FCR*Ton	27	1.28257	0.047503	3.5546	<0.0001	CRM*PTn	25	0.527438	0.021098	1.5787	0.0335
FTn*Ton	16	0.735772	0.045986	3.4411	<0.0001	FVR*Ton	20	0.402831	0.020142	1.5072	0.0681
PCR*PVR	33	1.12543	0.034104	2.5519	<0.0001						
FVR*FTn	19	0.631157	0.033219	2.4857	0.0003	Const	1	1662.33	1662.33	1.24E+05	<0.0001
Error	5308	70.668	0.013364			Total	5655	113.783			

〈表 17〉TMS—音強—音節層因素之 ANOVA 摘要表

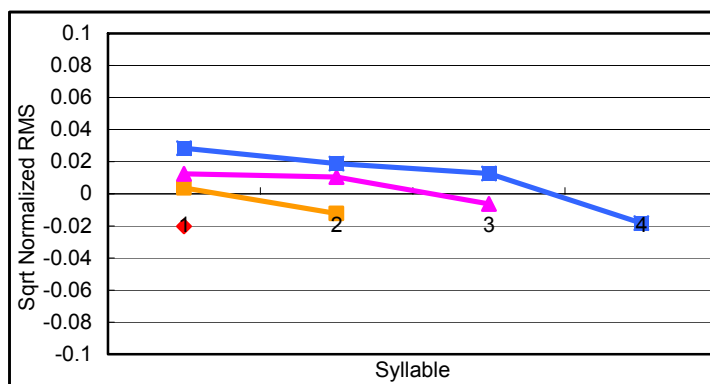
Source	df	Sums of Squares	Mean Square	F-ratio	Prob	Source	df	Sums of Squares	Mean Square	F-ratio	Prob
Ton	4	0.318537	0.079634	11.675	<0.0001	Ton*PTn	16	0.279218	0.017451	2.5585	0.0006
CRM*VRM	25	1.11392	0.044557	6.5324	<0.0001	CRM*PCR	30	0.466823	0.015561	2.2813	<0.0001
CRM	5	0.21493	0.042986	6.3021	<0.0001	CRM*FVR	25	0.350033	0.014001	2.0527	0.0015
VRM	5	0.189595	0.037919	5.5593	<0.0001	Ton*FTn	20	0.2714	0.01357	1.9895	0.0054
PTn	4	0.142174	0.035544	5.211	0.0003	PCR*PVR	30	0.37841	0.012614	1.8493	0.0033
PCR	7	0.234709	0.03353	4.9158	<0.0001	VRM*PVR	25	0.301156	0.012046	1.7661	0.0107
FCR*FVR	30	0.80378	0.026793	3.928	<0.0001	CRM*FCR	30	0.328262	0.010942	1.6042	0.0197
VRM*Ton	20	0.518252	0.025913	3.799	<0.0001	PVR*PTn	20	0.214018	0.010701	1.5688	0.0509
PCR*PTn	18	0.396491	0.022027	3.2294	<0.0001	VRM*PCR	30	0.307593	0.010253	1.5032	0.0384
CRM*FTn	19	0.400155	0.021061	3.0877	<0.0001	Ton*PCR	24	0.242049	0.010085	1.4786	0.0621
FCR*FTn	23	0.437687	0.01903	2.7899	<0.0001	FVR*FTn	20	0.197949	0.009897	1.4511	0.088
Ton*PVR	20	0.357034	0.017852	2.6172	0.0001	VRM*PTn	20	0.19212	0.009606	1.4083	0.1062
Ton*PVR	20	0.357034	0.017852	2.6172	0.0001						
FVR	6	0.105766	0.017628	2.5844	0.0168	Const	1	2286.83	2286.83	3.35E+05	<0.0001
Error	5178	35.3185	0.006821			Total	5654	77.7589			

〈表 18〉BFS—音強—音節層因素之 ANOVA 摘要表

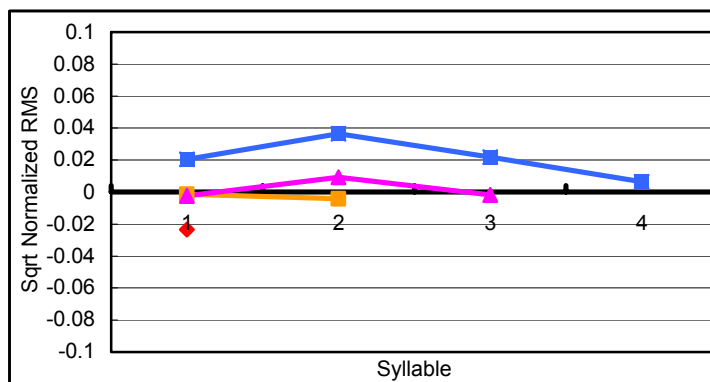
Source	df	Sums of Squares	Mean Square	F-ratio	Prob	Source	df	Sums of Squares	Mean Square	F-ratio	Prob
Ton	4	4.71969	1.17992	115.67	<0.0001	CRM*Ton	19	0.53875	0.028355	2.7796	<0.0001
CRM	5	0.808926	0.161785	15.86	<0.0001	Ton*PCR	23	0.576665	0.025072	2.4578	0.0001
FCR	6	0.520646	0.086774	8.5063	<0.0001	PCR*PTn	19	0.467188	0.024589	2.4104	0.0006
PTn	4	0.340328	0.085082	8.3404	<0.0001	FVR*FTn	30	0.69715	0.023238	2.278	<0.0001
Ton*PTn	16	1.25228	0.078268	7.6725	<0.0001	VRM*FCR	30	0.626958	0.020899	2.0487	0.0006
Ton*FTn	16	0.978621	0.061164	5.9958	<0.0001	CRM*FTn	20	0.374104	0.018705	1.8336	0.013
PCR	6	0.317725	0.052954	5.191	<0.0001	CRM*FCR	30	0.558874	0.018629	1.8262	0.0039
VRM	5	0.254008	0.050802	4.98	0.0001	CRM*PCR	30	0.558667	0.018622	1.8255	0.0039
FTn	4	0.157095	0.039274	3.8499	0.004	VRM*FVR	30	0.546392	0.018213	1.7854	0.0053
CRM*VRM	23	0.883748	0.038424	3.7666	<0.0001	Ton*FVR	28	0.482779	0.017242	1.6902	0.013
PVR	6	0.179831	0.029972	2.9381	0.0073	VRM*Ton	20	0.310572	0.015529	1.5222	0.0635
PVR*PTn	20	0.571958	0.028598	2.8034	<0.0001	Const	1	1716.77	1716.77	1.68E+05	<0.0001
Error	5088	51.9034	0.010201			Total	5482	112.8			

4.5.2 PW 層

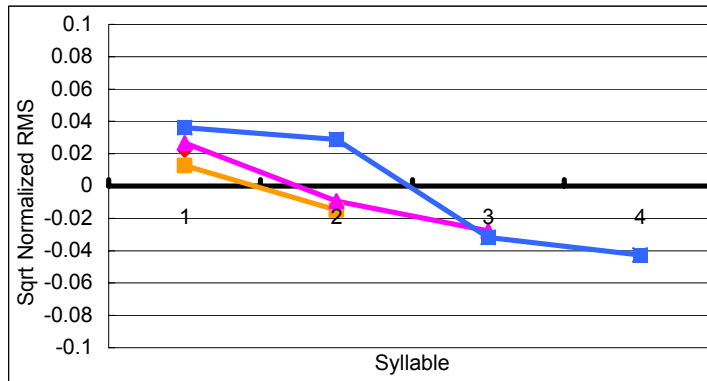
〈圖 22〉到〈圖 24〉顯示出在第三批語料中之 PW 層的音強組型。每一條線代表在一個韻律詞中特定位置音節所對應之回歸係數。Y 軸代表 NRMS 平方根的預測值，正值表示在特定位置音節的音強大於平均值，負值則表示音強小於平均值。PW 層的一般組型清楚地顯示出語料間的差異。越長的韻律詞，需要的能量越大，起始的音強越大。除此之外，以 2 音節與 3 音節的韻律詞來說，TFS 的回歸係數範圍為 0.016 和 0.019，TMS 的回歸係數範圍為 0.003 和 0.012，BFS 則是 0.028 和 0.054，顯示出兩批台灣國語語料在音強方面，較北京普通話低。



〈圖 22〉 TFS－音強－PW 層之回歸係數



〈圖 23〉 TMS－音強－PW 層之回歸係數

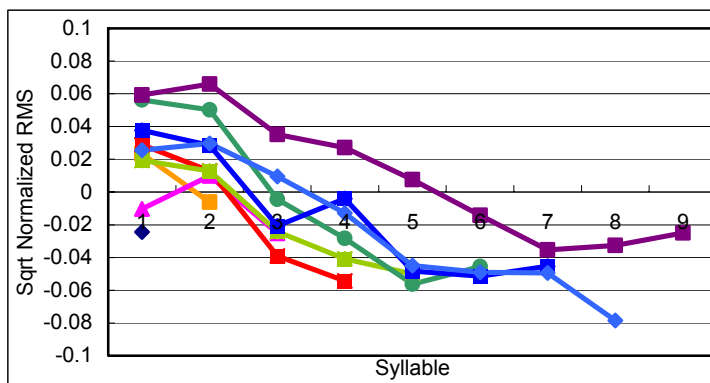


〈圖 24〉BFS—音強—PW 層之回歸係數

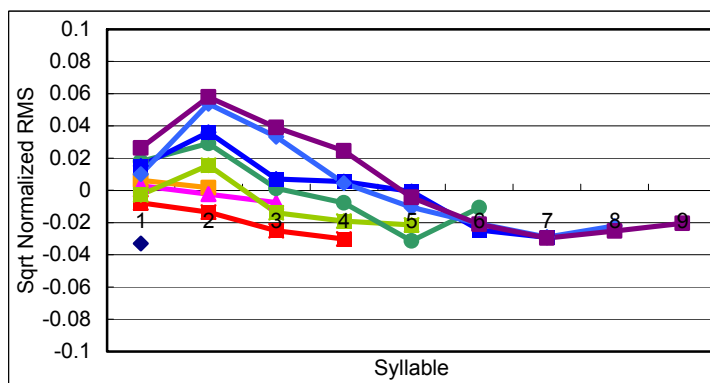
4.5.3 PPh 層

〈圖 25〉到〈圖 27〉顯示三批語料中的 PPh 音強比率組型。每一條線代表特定音節數之韻律短語的音強組型。超過 9 音節的 PPh 以紫色線條顯示，PPh 的中間部分以第四音節代表，而清楚地表現出前 4 與後 4 音節的音強組型。

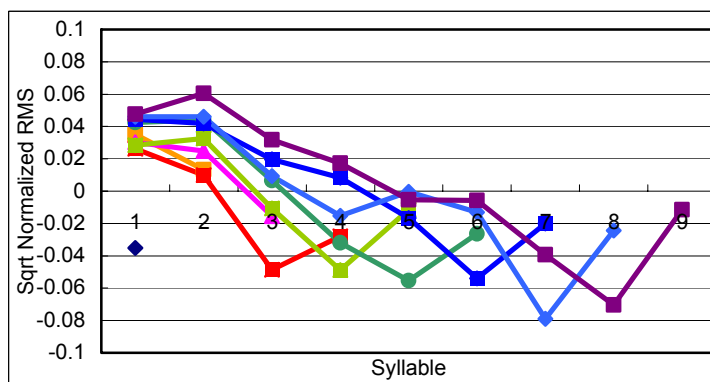
我們發現，越長的 PPh，需要的能量越多。對超過 9 音節的 PPh 而言，TFS 的係數範圍為 0.101，TMS 的係數範圍為 0.088，BFS 的係數範圍為 0.118。比起 TMS，TFS 和 BFS 皆具有較快的下降斜率。此一 PPh 音強係數下降斜率的差異現象可能導因於男性與女性在呼吸上有些差異，如肺活量、肌肉強度等等。除此之外，也可看出台灣國語與普通話的口音差異。以 PPh 尾來說，BFS 在 PPh 的最後一個音節表現出較強的音強，反觀 TFS 與 TMS 在 PPh 的倒數第二音節則表現出相當不清楚的突出 (prominence)。超過 7 音節的 PPhs 之回歸係數範圍，TFS 是 0.1，TMS 是 0.073，而 BFS 則為 0.096。超過 7 音節中段位置 PPhs 的回歸係數範圍，TFS 是 0.081，TMS 是 0.056，而 BFS 則為 0.094。



〈圖 25〉TFS—音強—PPh 層之回歸係數



〈圖 26〉TMS—音強—PPh 層之回歸係數



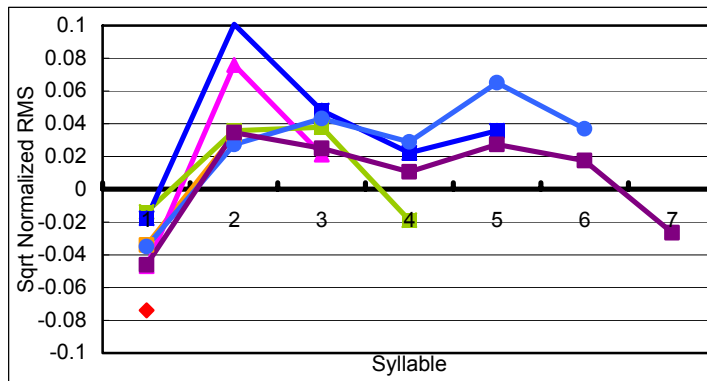
〈圖 27〉BFS—音強—PPh 層之回歸係數

4.5.4 BG 層

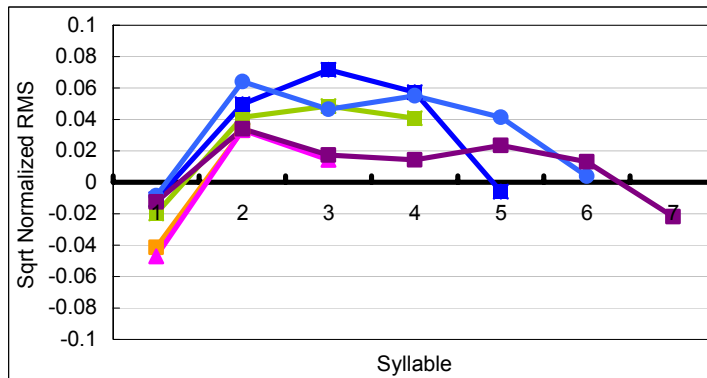
在 BG 中，我們假設在 BG 中不同位置的 PPh 的貢獻度並不相同，而進行個別分析。BG 中第一與最後的 PPh 視為起首與結尾 PPh，而其他的則為中段 PPh。

4.5.4.1 BG 中的起首 PPh (BG-initial PPh)

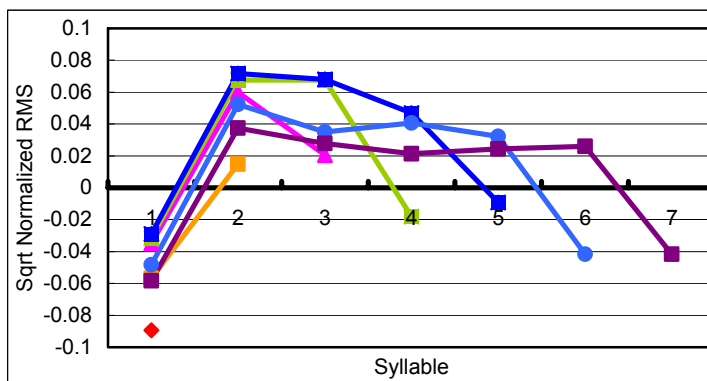
〈圖 28〉到〈圖 30〉顯示在 BG 中的起首 PPh 音強組型。每一條線代表不同長度 PPh 的組型。超過 7 音節的起首 PPh 以紫色線條顯示，我們發現：相對於前文中已出現過的 PPh 組型，當起首 PPh 比較長的時候，它反而沒有較強的音強。一般來說，除了第一與最後音節以外，BG 中的起首 PPh 音節和所導出的預測平均值比較起來，具有較強的音強。



〈圖 28〉TFS—音強—BG 中起首 PPh 之回歸係數



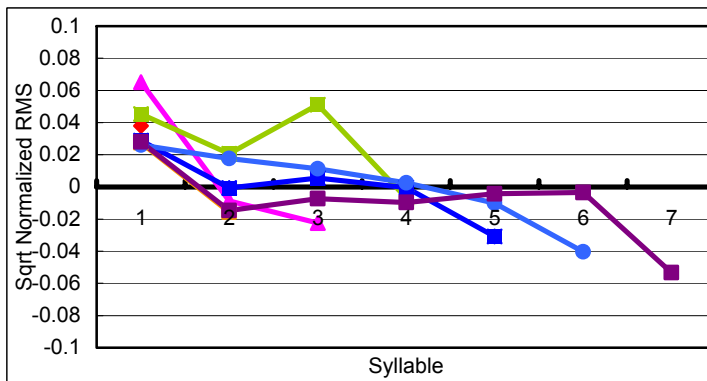
〈圖 29〉TMS—音強—BG 中起首 PPh 之回歸係數



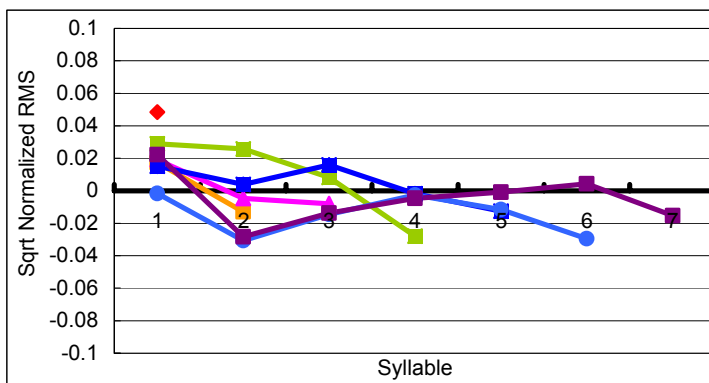
〈圖 30〉 BFS—音強—BG 中起首 PPh 之回歸係數

4.5.4.2 BG 中的中段 PPh (BG-medial PPh)

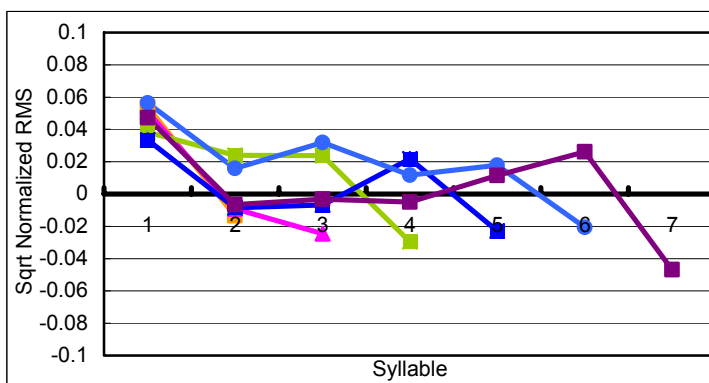
〈圖 31〉至〈圖 33〉顯示在 BG 中的中段 PPh 音強組型。與 PPh 層相類似，中段 PPh 有一個相似的下降組型。



〈圖 31〉 TFS—音強—BG 中段 PPh 之回歸係數



〈圖 32〉TMS－音強－BG 中段 PPh 之回歸係數

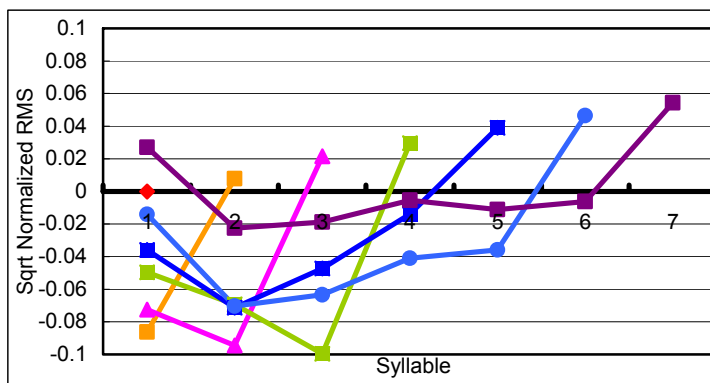


〈圖 33〉BFS－音強－BG 中段 PPh 之回歸係數

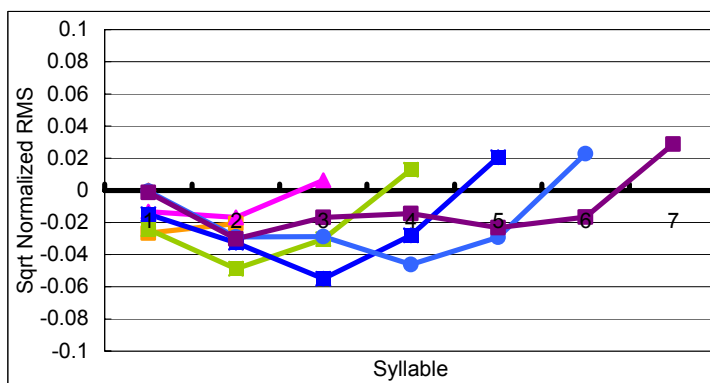
4.5.4.3 BG 中的結尾 PPh (BG-final PPh)

〈圖 34〉至〈圖 36〉顯示在 BG 中的結尾 PPh 音強組型。所有的語料均顯示出，結尾 PPh 並沒有表現出如前所見的 PPh 音強減弱，反而呈現一相反趨勢：在 BG 結尾伴隨著一個較強的音節。較短的結尾 PPh 具有一較寬的係數範圍。對超過 7 音節的結尾 PPh 而言，係數範圍 TFS 是 0.077，TMS 是 0.059，而 BFS 則為 0.094。

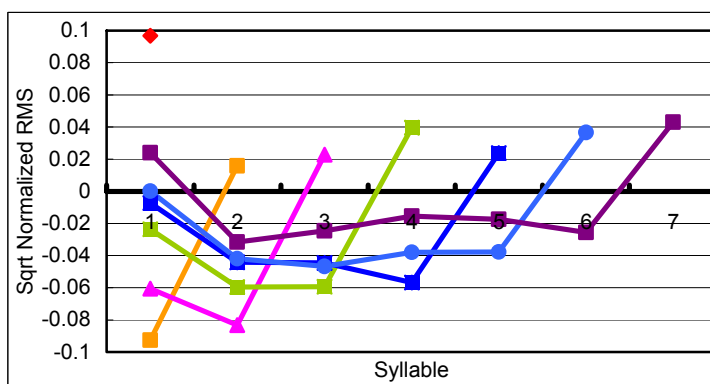
比起男性發音人語料，兩批女性語料皆顯示出在 BG 中的結尾 PPh 具有較寬的回歸係數範圍。雖然在這層中結尾 PPh 的最後音節是正值，但並不意謂著在 BG 結尾 PPh 具有一強勁的結尾音節，這僅能暗示著在這一層級中，這些音節比起在其他 BG 單元裡的音節來說還要來得強一些。若以整體來說，全面性預測考慮到所有層級的預測值，而 BG 結尾依然是變弱的。



〈圖 34〉 TFS－音強－BG 結尾 PPh 之回歸係數



〈圖 35〉 TMS－音強－BG 結尾 PPh 之回歸係數



〈圖 36〉 BFS－音強－BG 結尾 PPh 之回歸係數

全面性預測之評估如〈表 19〉摘要描述如下。值得注意的是：原始 DV（NRMS 值）與總和預測（預測總和之平方）轉換值之間的評估值，因此我們可以推論：殘差誤差的增加可能導因於此非線性轉換過程。另一方面，相關係數 r 意謂著當增加更高一層時，預測力獲得了改善。在第三批語料中，BFS 具有最好的預測結果（相關係數 $r=0.814$ ），而最差的是 TFS（相關係數 $r=0.693$ ）。

〈表 19〉第三批語料音強預測之全面性評估

Layer \ Corpus	TFS		TMS		BFS	
	T.R.E.	r	T.R.E.	r	T.R.E.	r
Syllable	63.80%	0.616	47.65%	0.732	48.00%	0.724
PW	62.10%	0.621	47.02%	0.736	40.53%	0.737
PP	48.19%	0.666	37.43%	0.766	23.42%	0.775
BG	49.04%	0.693	35.37%	0.787	27.88%	0.814

〈表 19〉中 T.R.E. 值與〈表 12〉相近，約有 27-49% 的誤差值，此數值偏高，可能來源為樣本數目仍可以增加，及未能考慮影響音強分布的其他參數，如因 PG 位置所造成的 PPh 短語總體時長的相對快慢調節等。

5. 實驗四：口語韻律階層的基頻重設 (F0 reset) 與基頻範圍 (F0 range) 研究

5.1 實驗目的與假設

本實驗以 PG 理論架構為基礎，進一步找尋語段與語篇結構的重要證據。我們分析了口語篇章中 PG 內和 PG 間韻律特徵的以下二項聲學特性：(1) PG 內與跨 PG 間基頻重設與重設前後停頓時長的關係。(2) 基頻範圍（或做基頻域）變化與基頻重設之間的關係。換言之，從基頻看進行中的朗讀語流的重設與音域範圍的大小變化，應可提供語篇結構的訊息。

5.2 實驗語料

本實驗所使用的語料為 COSPRO 05 中的朗讀段落，麥克風語料，男女發音人各一，其朗讀文本包含 26 段長段落，段落字數從 85 到 981 個音節不等，每一個段落平均字數 377 音節。

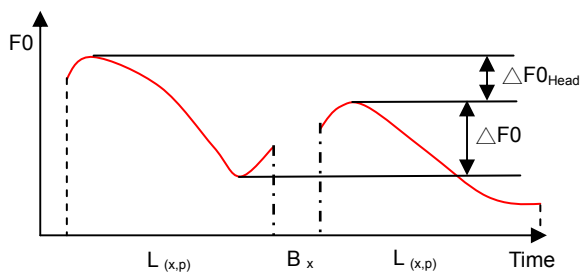
5.3 語料處理及標記

同實驗二、三。

5.4 分析過程

5.4.1 基頻重設分析

我們分析較大三級邊界停延（也就是 PPh、BG 和 PG 邊界停延，在標記上以 B3, B4 和 B5 表示之）前後基頻曲勢的變化。〈圖 37〉為基頻曲勢重設示意圖。圖中，我們使用兩項參數描述基頻曲勢的變化：(1) 以 $\Delta F0$ 表示較大邊界停延前後基頻曲勢的差異；(2) 以 $\Delta F0_{\text{Head}}$ 表示較大邊界停延前後基頻曲勢最大值的差異。



〈圖 37〉基頻曲勢重設分析示意圖

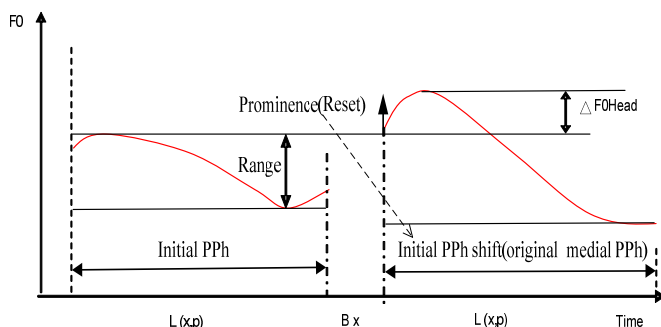
X=3 或 4 或 5

接著，分析較大三級邊界停延：PPh、BG 和 PG，基頻重設前後韻律單位的音節時長組型。在 BG 中，如果只包含 2 個 PPh，我們則假設關於 2 PPh 之間的位置分別為 BG 首和 BG 尾。為了推估男女發音人之間的基頻變化，每一組資料皆經過標準化程序。

5.4.2 基頻範圍分析

在我們理論架構的基頻基型裡，具體說明最突出的基頻重設發生在 BG/PG 首第一個 PPh 的位置。然而，仔細探究語料卻發現：少部分的句首短語的基頻範圍卻相對變窄了許多，且時長較短，同時，實際的基頻重設則向後位移到下一個短語位置，〈圖 38〉說明了較窄的起首基頻範圍後面伴隨著較寬基頻範圍的基頻

重設向後位移的現象。更進一步的觀察顯示：這些簡短的單位通常在本質上具有如“另一方面”、“換句話說”等過渡或者轉折短語的性質。我們將 PG 起首基頻範圍較窄且伴隨基頻重設向後位移的 PPh 稱作「韻律填充」(prosodic filler, 簡稱 PF), 意謂著在口語語流中, 它所扮演的只是「串場」或者「過門」的韻律角色, 訊息承載量不大, 有時甚至可略去不看; 與「篇章提語」(discourse marker) 相對, 係造成口語語流變化多端的來源之一。



〈圖 38〉較窄起首基頻範圍伴隨著基頻重設向後位移示意圖

獲得以上的初步語流觀察後, 我們僅以觀察到的聲學特徵為基礎, 實際進行「韻律填充」發生機率的統計分析; 同時, 也觀察基頻重設向後位移現象的基頻曲勢變化。

5.5 實驗結果

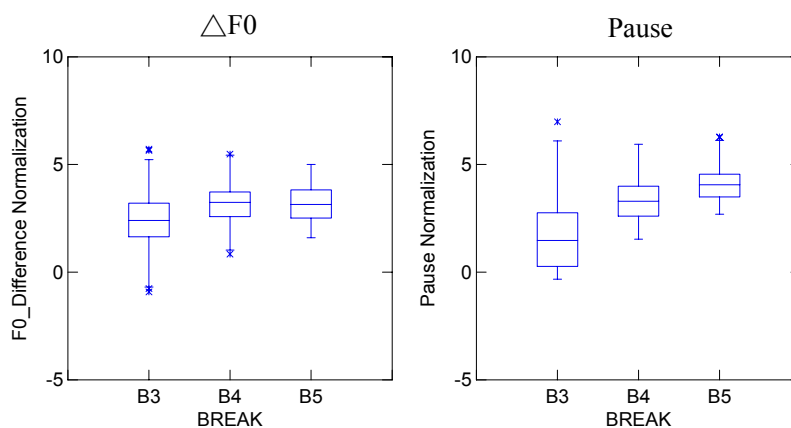
5.5.1 基頻重設分析

5.5.1.1 三級停延中的停頓與基頻統計資料

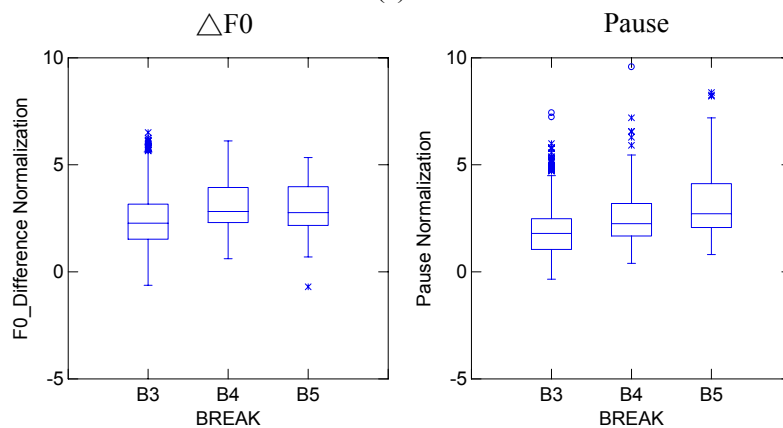
〈表 20〉列出三級邊界停延中的每一級邊界停延的次數, 〈圖 39〉則顯示了 M051 與 F051 在基頻與停頓的韻律特徵。二者的實驗結果相似, 皆顯示 BG 邊界 (B4) 和 PG 邊界 (B5) 比起 PPh 邊界 (B3) 存在較大的 $\Delta F0$ 值, 而 PG (B5) 比起 BG (B4) 和 PPh (B3) 有較長的停頓時長, 〈圖 39〉也顯示了 PG (B5) 和 BG (B4) 的韻律特徵在 $\Delta F0$ 值上只有極小的差異。

〈表 20〉M051 與 F051 各級邊界停延次數摘要表

M051	B3	B4	B5
Count	798	174	61
F051	B3	B4	B5
Count	724	168	68



(a)

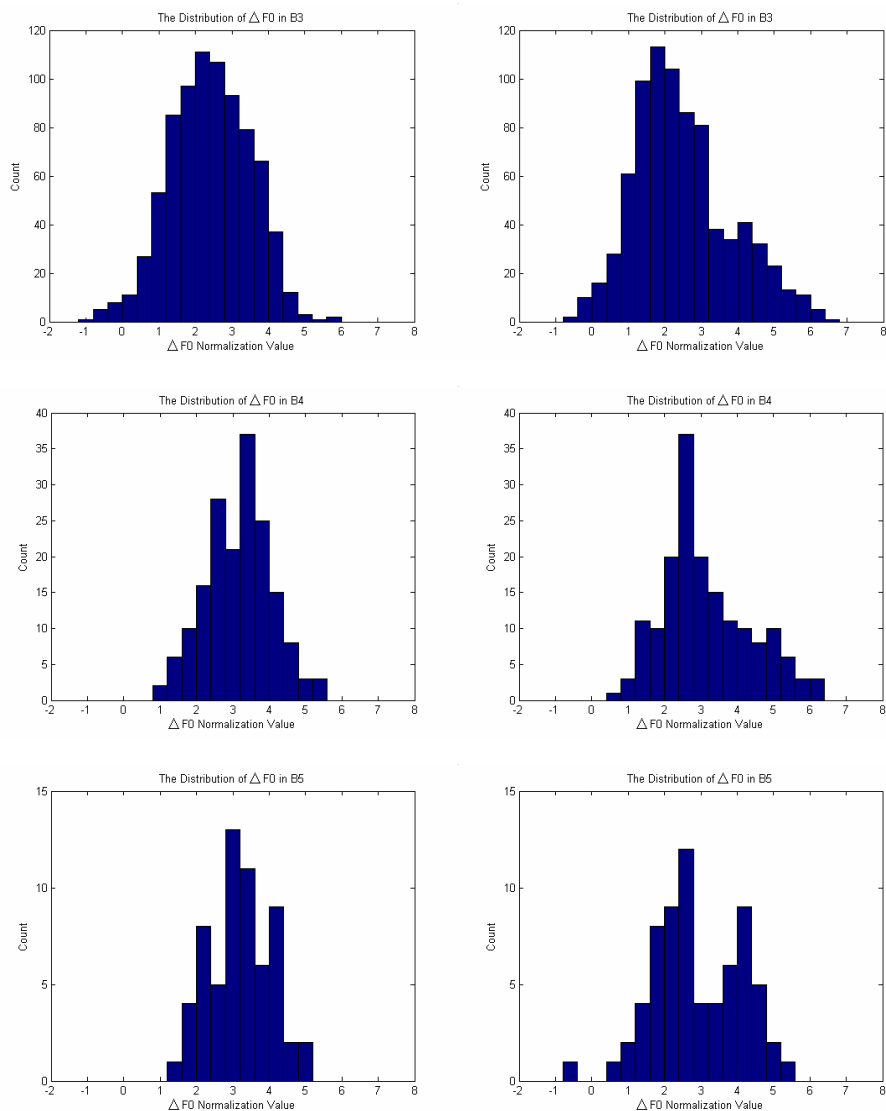


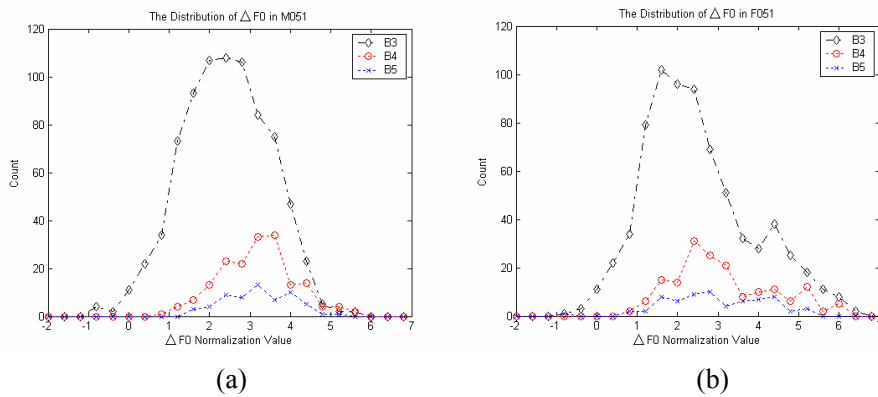
(b)

〈圖 39〉(a) M051 韻律特徵；(b) F051 韻律特徵

雖然 PG 和 BG 的 $\Delta F0$ 值並沒有顯著差異 (p 值 > 0.01)，在停頓時長上則有顯著性差異 ($z=6.58$, $p<0.01$ in M051; $z=2.958$, $p<0.01$ in F051)。因此，我們可以合理地假定停頓時長本身是 B4 與 B5 之間無關基頻重設的主要區別線索。

〈圖 40〉顯示 M051 與 F051 各級邊界停延間的 $\Delta F0$ 分布，我們可清楚地看出 M051 的 $\Delta F0$ 分布與 F051 的不同。M051 的 $\Delta F0$ 分布呈現對稱分布的傾向，反而 F051 的 $\Delta F0$ 分布呈現正偏態傾向。〈圖 40b〉中 B5 的 $\Delta F0$ 分布出現負值，除此之外，F051 的 $\Delta F0$ 分布範圍也比 M051 的分布還要來得大。

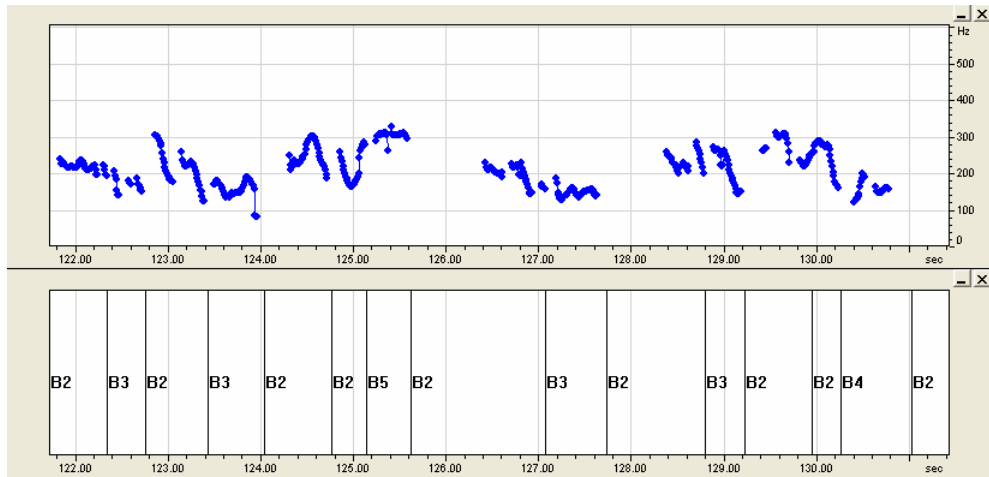




〈圖 40〉 (a) M051 各級邊界停延之 $\Delta F0$ 分布
(b) F051 各級邊界停延之 $\Delta F0$ 分布

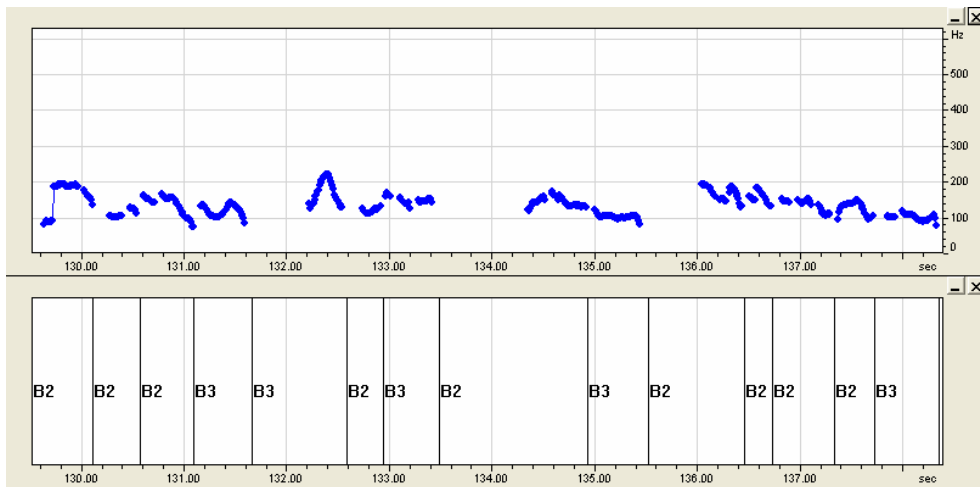
我們感興趣地注意到：朗讀風格的不同也會影響語流韻律的表現形式。在 Sinica COSPRO 語料庫中，我們發現，即使是相同的段落，男性與女性發音人在朗讀時使用不同的風格與形式，而且，清楚地透過邊界停延甚至是音色 (voice quality) 不同的選擇，在相同段落的斷句與詮釋上也不盡相同。

以〈圖 41〉中的例子為例，相同的朗讀段落的片段，但二者的斷句卻不相同。〈圖 41a〉中，F051 以音色上變換的技巧表示出發音人本身對段落上主題轉換的理解，即使在停頓邊界的兩側雖無明顯的基頻重設，但在聽感上卻聽得出明顯的音色差異，因此，標音員將邊界判定為 B5；但，〈圖 41b〉M051 則以一種形式上的不帶情感的方式在朗讀同一段落，並無音色上的差異，所以，標音員在相同位置上將其視為相同主題段落而標記為 B3。



媽拉著嗓子大聲往二樓問：玲兒，怎不開燈？ 媽知道她回來了，但不會知道她剛從醫院逃回來。

(a)



媽拉著嗓子大聲往二樓問：玲兒，怎不開燈？ 媽知道她回來了，但不會知道她剛從醫院逃回來。

(b)

〈圖 41〉(a) F051 例句；(b) M051 例句

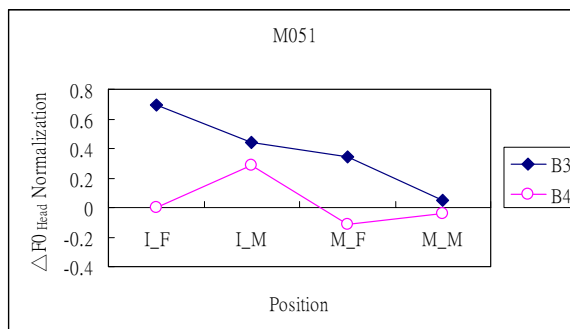
5.5.1.2 三級停延中不同位置的 $\Delta F0_{\text{Head}}$ 值統計

邊界停延依不同位置加以分組，如〈表 21〉所列。PG 相對位置標記 PT：Initial, Medial, Final 依韻律層級分成 4 組：Initial-Final (I_F), Initial-Medial (I_M), Medial-Final (M_F) 與 Medial-Medial (M_M) 等。由於 B5 的次數不如 B3 和 B4 那

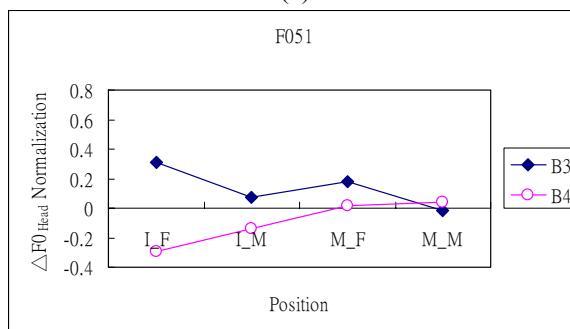
麼多，造成 B5 在 I_F 位置的次數不足，因此，在接下來的統計分析中 B5 將被略去不看：

〈表 21〉 M051 與 F051 不同位置的停延次數統計

M051					F051				
Break Position	B3	B4	B5	Sum	Break Position	B3	B4	B5	Sum
I_F	83	34	1	126	I_F	78	27	1	125
I_M	164	42	12	210	I_M	156	46	13	196
M_F	193	45	14	336	M_F	207	47	14	321
M_M	358	53	34	445	M_M	283	48	40	371
Sum	798	174	61	1033	Sum	724	168	68	960



(a)



(b)

〈圖 42〉 (a) M051 不同位置的 $\Delta F0_{Head}$ 平均數
(b) F051 不同位置的 $\Delta F0_{Head}$ 平均數

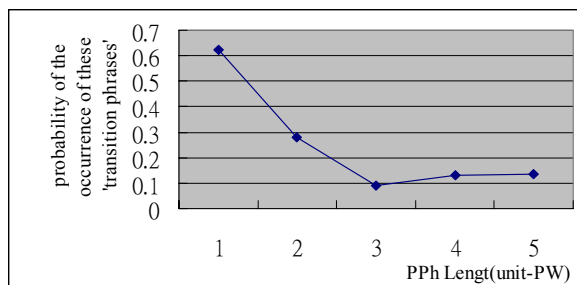
〈圖 42〉顯示 M051 與 F051 不同位置的 $\Delta F0_{Head}$ 平均數，從圖中我們可以

看到 B3 在 “I_F”, “I_M” 與 “M_F” 位置的 $\Delta F0_{\text{Head}}$ 平均數皆大於 B4 的 $\Delta F0_{\text{Head}}$ 平均, 其中以 B3 的 “I_F” 位置的 $\Delta F0_{\text{Head}}$ 的平均數為最大值; 除此之外, M051 與 F051 在 “M_M” 位置的 $\Delta F0_{\text{Head}}$ 值是近似的。

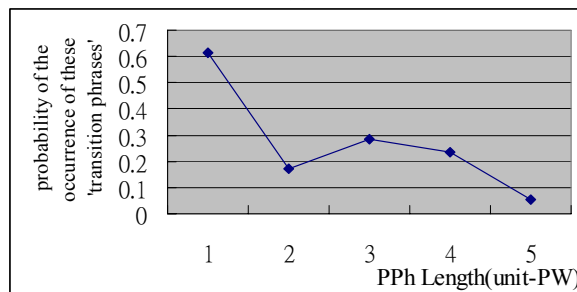
5.5.2 基頻範圍分析

5.5.2.1 基頻範圍變異

我們針對「韻律填充」(PF) 以觀察到的聲學特徵為基礎, 實際進行發生機率的統計分析, 結果發現: 當明顯的基頻重設 (中段短語 $F0_{\text{Head}} > 1.5*$ 起首短語 $F0_{\text{Head}}$) 發生在 BG 中段位置時 (如〈圖 43〉所示), 我們將這樣的情況定義為 “transition induced shift”, 統計分析結果如下, 發現: 「韻律填充」(PF) 過渡現象最常發生在短 PPh 中 (通常只包含一個 PW), 而出現在長 PPh 如 “過渡” 性質的機率明顯變小。我們可以看到, 在 BG 起首短 PPh 的情況下, 「韻律填充」(PF) 發生的機率比較大, 特別是只包含 1 個 PW 單位時。



(a)

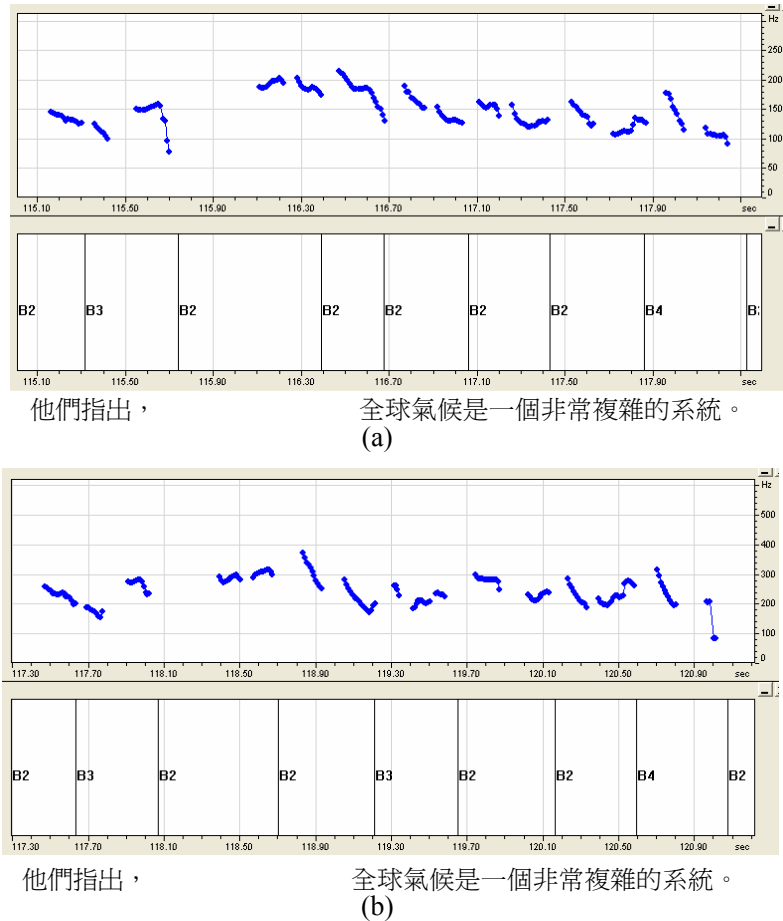


(b)

〈圖 43〉 (a) M051 不同 PPh 長度「韻律填充」(PF) 的出現機率
(b) F051 不同 PPh 長度「韻律填充」(PF) 的出現機率

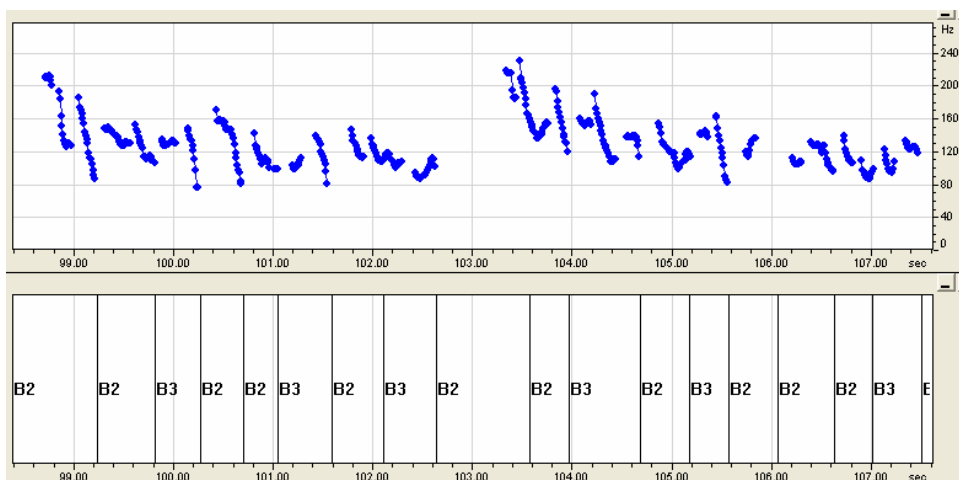
5.5.2.2 基頻重設向後位移現象的基頻曲勢變化

在 5.5.1.1 小節中第一部分，我們呈現了男女發音人不同的基頻重設組型與範圍，但卻表現出類似的 PG 起首過渡短語現象。在此一部分，〈圖 44〉進一步說明觀察整體基頻曲勢組型的結果，同樣地揭示了兩者間很少的發音人變異 (speaker variation)。從圖中可看出，M051 和 F051 語料呈現類似的整體基頻曲勢組型，二者在起首短語使用較窄的基頻範圍，同時將實際的基頻重設位移到下一個短語，這意謂著句法與語意的限制是凌駕發音人或說話風格 (speaking-style) 的變異，換言之，句法與語意的限制不會因發音人或說話風格的不同而有所變異，而且是可以預期的。



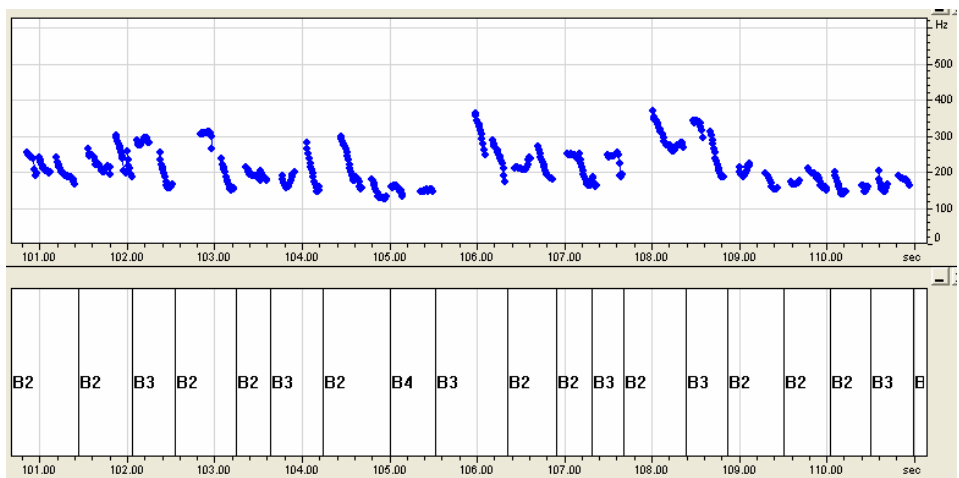
〈圖 44〉 (a) F051 例句；(b) M051 例句

〈圖 45〉顯示 M051 和 F051 相同的段落，卻選擇不同的焦點與 B3 邊界停頓，結果造成不同的整體輸出組型。男性發音人 M051 (〈圖 45a〉所示) 以明顯的整體下傾 (平鋪直述) 方式朗讀語句，然而，女性發音人 F051 (〈圖 45b〉所示) 在 B3 停頓後以較大的基頻重設手段表現強調語段中的特定焦點。



他總是以開玩笑的方式紓解他們的情緒，淡化問題。畢竟年少衝動的心較難接受苦口婆心的教條勸說，

(a)



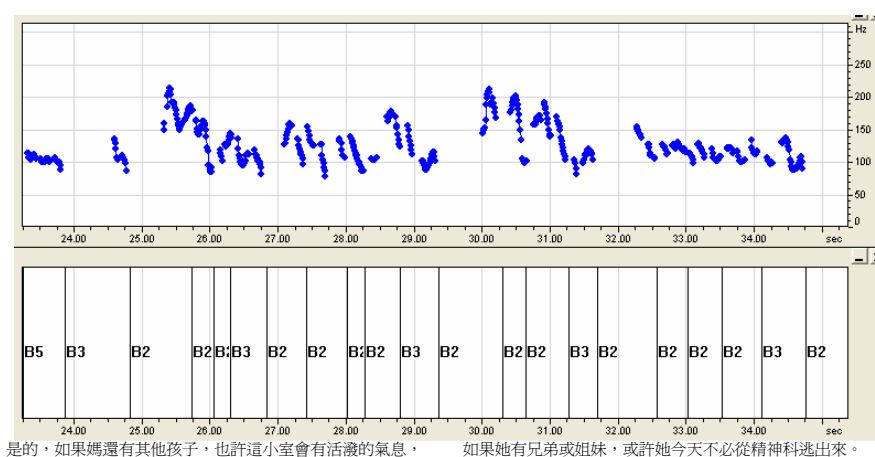
他總是以開玩笑的方式紓解他們的情緒，淡化問題。畢竟年少衝動的心較難接受苦口婆心的教條勸說，

(b)

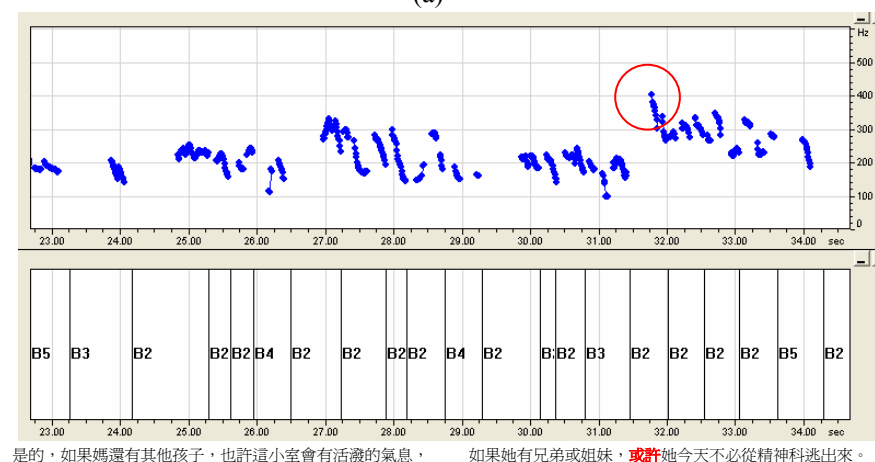
〈圖 45〉 M051 (a) 與 F051 (b) 相同文本朗讀語段

注意：二者在一開始不同的重設程度以及強調焦點，而這些因素對整體基頻走勢的影響。

〈圖 46〉中 M051 與 F051 在相同語段中則表現出相反的基頻走勢組型。我們觀察二者在朗讀風格上的差異，發現：男性發音人使用中性、不帶情感的朗讀方式；相對於男性發音人，女性發音人則減少起始基頻上揚程度，而最高的基頻重設向後位移許多。〈圖 46b〉中，B4 後本應伴隨基頻重設突顯的發生，然而，B4 邊界後的起始 PPh “如果她有兄弟姐妹”的基頻走勢卻顯得平緩，基頻重設高點落在接下來的 PPh 的起始 PW “或許”二字上面，以突顯語意焦點的所在。因此，語意焦點的不同選擇，因而整體輸出效應也就不同。女性發音人的選擇造成了更戲劇化的朗讀風格。



(a)



(b)

〈圖 46〉 M051 (a) 與 F051 (b) 相同文本朗讀語段
注意：不同位置的基頻重設突顯的發生對整體走勢的影響。

6. 綜合討論

6.1 實驗一

PG 的階層式架構最重要的特性是 (1) 說明了跨短語韻律的關聯性 (association) 不僅是緊鄰的線性關係，而且必須包括來自上層節點的規範制約所造成的跨短語關係如 PG 首與 PG 尾的呼應及關係，以及這些關係所造成的短語句調、音節時程分布調節的成因及系統性。(2) 解釋了為何字調和個別的句調走勢及二者的互動並不足以說明語流韻律。(3) 說明了語流韻律中的韻律單位必須包括韻律邊界與停頓，及語篇訊息成分的重要性。實驗一的結果，證明了雖然 PG 階層架構只強調 3 個 PG 相對位置，即 PG 首、中、尾位置，但階層式架構卻具體說明了次級個別 PPh 分別而相對的韻律角色，在多短語語段中產生必要的連貫性 (coherence)。在實驗一的結果中，我們注意到一旦 PPh 成為 PG 中的成分，它不再是一個獨立的句調單位而是需要調整其句調走勢，在韻律上傳達語段間各 PPh 的關聯性。PG 首與 PG 尾 PPh 的特徵說明了，因為 PG 效應，位於這兩個位置的 PPh 的句調走勢不管是在相對的起點、斜率、邊界效應以及邊界停延是如此地不同。即使 PG 首與 PG 尾 PPh 可能都表現出下傾的走勢 (declination)，但是二者在相對下傾的程度與斜率皆不相同，而且句尾延長以及弱化 (final lengthening and weakening) 以 PG 尾位置的 PPh 最為明顯；至於 PG 中段位置 PPh 則說明了在 PG 首尾之間其他 PPh 以較平緩的句調突顯語流的連續性 (continuation)，三者共同構成的關係，缺一不可。如此一來，PG 內韻律短語的相對位置彼此相依卻又互為一體。因此，連續語流中的短語應該整體看待，而不是各自不相干的獨立成分。以上的關係，都是以句法結構來構成的句調所無法解釋的，也就是如果從句調的角度來看語流中直述句的下傾短句句調，則無法解釋何以同類句型的眾多短句，其句調變化多端；也無法解釋簡單句與複雜句的句調是否有異、異於何處及原因為何。

同時，PG 架構也系統性地解釋了：為何在口語連續語流中句調變化不是隨機而是可預測的？為何語流韻律不可以當作是一個一個下傾句調的結果？為何在 PG 首與 PG 尾 PPh 之間成對呼應的脈絡是明顯可見的？大致上來說，整體韻律組型只有在所有 PPh 以特定的線性規則成群時才能呈現，如果把位於 PG 首與 PG 尾 PPh 的句調走勢互換，將無法成為可接受的韻律組型。同時注意到：實驗一所選擇的 3-PPh 複雜句因其在句法結構及字數上的控制，相對於標準韻律句組特徵的表現較不明顯，但仍提供一預設的 PG 原型，並可擴展至 COSPRO 語料庫中多達 12 PPh 的多短語句群；換言之，在 PG 首與 PG 尾 PPh 之間，依照語速的不同，可有多達 10 個相對平緩句調的 PPh，用來突顯語流的連續性。而且，正如

基頻走勢開高走低一路下傾是簡單直述句句調的基本調，具語言學上的深層結構意義，PG 架構也呈現的多短語語段的標準形式，亦具有類似的深層結構意義。而字調及輕重音實屬詞義，語句焦點屬語意，語者強調屬意圖，各有語言學意義的來源，都可以視作深層韻律結構以外或附加的成分。

實驗一的基頻走勢實驗結果證明了在口語語流中，上層訊息涉入語音產製規劃歷程中；語音單位不再是分離的句調單位。更大的多短語韻律單位反映出上層語篇組織在口語語流產製的作用。因此在方法論上，缺少語篇脈絡所產生的句調單位，勢必無法提供整體的韻律訊息。從口語連續語流中擷取出的片段以及分析音段或音節中微觀的語音或聲學細節，都只會指向語音變異的無法掌握，而無法系統性說明口語中整體語意連貫性結構。實驗一的結果，也解釋了語音合成時，同樣的聲調和句調出現在語流中不同的位置時，如不加上層語篇訊息的語流韻律的調整，則無法輸出自然的語流韻律的原因。

6.2 實驗二與實驗三

我們認為上層語篇韻律其實涉及的，是語意關聯性所導致的語流產製規劃範圍，任何描繪韻律組織與模型的意圖，除了基頻走勢之外，還必須包含特定語言的時長配置與音強分布（實驗二），以維持語篇的連貫性 (coherence) 與關聯性 (association)。實驗二結果顯示，語流韻律研究並非僅著眼於基頻走勢與音高的體現，也必須有系統的說明時長配置與語流韻律結構的關係；由此我們也得知，以音節為節奏單位的語言，並不表示音節的相對一致長度進入語流後會維持不變，語流韻律的節奏，也必須配合韻律架構與韻律單位從相對配置瞭解，才能得以釐清語流節奏快慢分布的組型及邊界效應。而音強的研究（實驗三）則顯示，語流中的音強配置至少是以韻律短語為單位，亦即較小的韻律單位的底層音強配置，不在組型內，也表示小範圍的韻律單位，在語流韻律上的意義不及韻律短語。

從實驗二和實驗三的結果，我們發現了在 PG 階層式結構中的各韻律層的典型韻律組型，如同時長組型一般，音強組型也可從音段成分與韻律結構中得到。除此以外，北京普通話與台灣國語間的差異也可由每一韻律單元層級中所看出。在本小節中，我們將分別討論每一層的實驗結果。

6.2.1 音節層

從音節層的實驗分析結果，我們可以看到，三位發音人 50% 左右的預測值可從音段成分獲得。預測結果解釋了發音人 TFS 的 56.12%，TMS 的 56.89% 以及 BFS 的 50.03%。也就是說，整體韻律輸出中的音節長度配置，只有將近 50% 來自音節本身的音長。以音強預測來說，音節層預測原始資料近 52% 的比率 (T.R.E.=47.65% in TMS and 48% in BFS)，其中最差的是 TFS，只解釋了 36.2%。至於時長預測部分，TFS 音節層解釋了 56.12%，TMS 解釋了 56.89%，以及 BFS 則為 50.03%。因此，音節層時長與音強的預測結果證明了語音合成若是只考慮到音節的串接與過渡，那將因缺少了更上層的韻律階層訊息而聽起來不自然，也提供了基頻走勢以外的聲學參數，對語流韻律的證據。

6.2.2 PW 層

在第三批語料中，PW 的時長與音強的典型組型是可預估獲得的（見〈圖 7〉至〈圖 9〉與〈圖 22〉至〈圖 24〉），音節時長配置與音強的逐漸下降像帶把勺子狀的組型，均顯示了韻律詞組型之原型 (prototype)：從一至四個音節的 PW 音節長度的長短配置及音強的漸弱配置都有固定的組型。其中以平均語速最慢的北京普通話發音人 BFS 語料的韻律組型是最為明顯，同時，在 PW 層的預估，韻律組型解釋的比率也以 BFS 較多。時長預測的結果顯示，三位發音人中，BFS 的 T.R.E. 減少了將近 6.75%，而 TFS 只有減少 1.45% 與 TMS 的減少 2.24%。而以音強預測來說，BFS 的 T.R.E. 減少了將近 8%，反而 TFS 只有減少 1.7% 與 TMS 的減少 0.63%。以上結果也顯示，在 PW 層的音節時長配置和音強二個聲學參數，台灣國語與北京普通話之間的差異並不大，除了音段成分的差異，一般聽感上北京普通話比起台灣國語具有較明顯的音節縮短與延長，多音節詞的輕重音型成分應大於 PW 成分。

6.2.3 PPh 層

從第三批語料中所取得之 PPh 層的音節時長配置的典型韻律組型（見〈圖 10〉至〈圖 12〉與〈圖 25〉至〈圖 27〉）清楚地顯示，PPh 的組型是由最後的三個音節的縮短與延長共同形成，再次證明此一韻律單位具有音段的固有時長以外的時長配置組型。PPh 層的音強分布也一樣，出現明顯的下降趨勢，可取得固定的組

型。此外，相對於在 PW 層台灣國語與北京普通話並不明顯的差異，往上一層到 PPh 後則顯示較明顯的差異。台灣國語的二位發音人的 PPh 層時長組型，都顯示出由句尾倒數第三個音節縮短後再遞增延長的組型，有明顯的延長效應。北京普通話發音人 BFS 的時長預測則顯示，相較於 PW 層的預測結果，其 PPh 結尾並沒有更進一步拉長，反而是倒數第二個音節縮短。在音強分布面，台灣國語的二位發音人有相對地結尾弱化現象產生，而北京普通話發音人卻顯示了短語最後音節的加強。在三層韻律層級中 (PW, PPh, BG layer)，PPh 層在時長與音強的預測上說明了絕大部分，在 BFS 的音強預測上，PPh 層解釋了將近 17% 的 T.R.E. 值。

6.2.4 BG 層

實驗結果顯示，在 BG 的起首 PPh、中段 PPh 與結尾 PPh 三個韻律位置的音節時長配置，不但都可發現清楚的韻律組型，而且起首 PPh 與結尾 PPh 組型類似，都有像 PPh 層一樣的組型，第一音節縮短，倒數第三個音節縮短後再遞增延長到結尾，但程度不同，而結尾 PPh 則呈現相反的組型，最後的二至三個音節是縮短的。在此特別值得一提，也是最重要的一點是：各韻律層雖各有組型，但總體的語流韻律輸出，是所有韻律層疊加後的總和，加總的過程中，正值相加、正負值則相抵，因此，結尾音節在疊加音節、PW、PPh 和 BG 各層後，結果仍是延長最終的二音節。強度分布方面，起首 PPh 的第一個音節強度最大，且中間的音節比 BG 其他位置的音節還強。中段 PPh 的第一個音節也被加重，而最後音節則弱化。至於結尾 PPh，也呈現與 PPh 相反的組型，最後兩音節呈現類似出重音上的輕重組型，但加總的方式及結果與音節時長完全相同。BG 層的結果，直接證明了上層語意對 PG 語段的效應，也就是 HPG 架構所強調的上層語篇語意存在於語流韻律的證據。此外，每一個韻律層都顯示，語流的邊界停延在口語篇章中也是語言學上顯著的成分同時在任何韻律架構中具合法地位。

在實驗二與實驗三中，我們通過配合「階層式多短語語流韻律架構」的回歸統計模型，檢驗大量語料的聲學語音特性，在語流中音節時長的時程配置和音強分布二方面做預測，發現當上層的韻律單元增加時，預測力也增加。進一步的證明語流的韻律最終輸出現象，無論是時程配置或音強，都是各韻律層級分層貢獻的加總所共同構成。結合實驗一的結果，我們在基頻、音長和強度三個聲學參數，都因為量化大批語料，獲得了階層式多短語韻律架構 HPG 的證據與組型模版，證明跨短語韻律的深層結構確實存在。

經由實驗比較兩岸口語在語流韻律，我們得到以下的結論，也強化了 HPG 的跨口音意義：

- (1) 由各層分析顯示，語流由多層次的架構所組合而成，每一層皆有典型的特徵，證明口語語流韻律存在著一定的規劃性及規則性。
- (2) 韻律結構中，時長和音強都具有規則性，而以音節長短在時程配置上規則性最是明顯，在各階層間顯現類似的邊界效應及節奏組型。層層疊加的結果，是口語語流韻律與孤立句的句調在聽感上顯現極大差異的最主要原因。
- (3) 兩岸雖有口音的差異，在語流韻律的特性上大致還是相當一致的，亦即基本的韻律差異並不大。我們認為，必須進行更進一步的實驗，以釐清兩岸口音差異還有哪些語音層面，如音段及字調系統、甚或是否還包括了語速差別。
- (4) 我們經由量化所獲得的韻律模型，可當做語音合成的基礎，作為語音科技開發方面的應用。

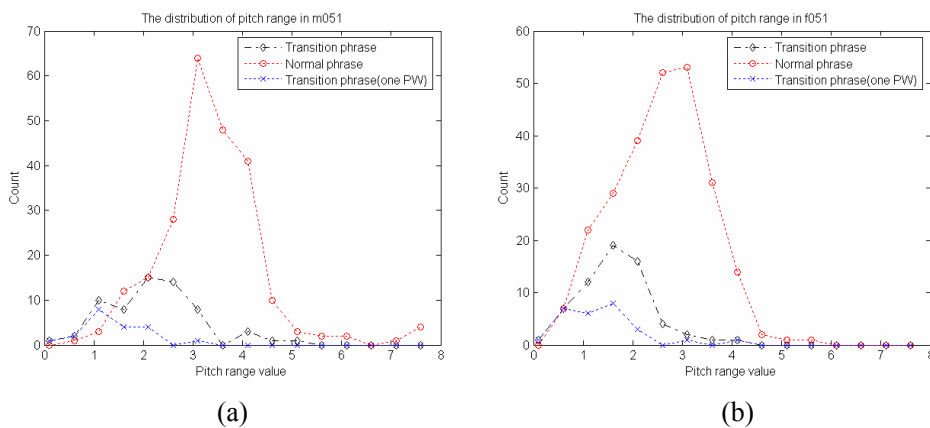
在研究方法方面，我們也提出結合實驗語音學與語料庫語言學後的語料庫語音學研究方法。此一方法，不但解決傳統語音學觀察個別語音樣本、著重音質描述，一旦面對樣本增多、語音變異便無法著力的窘境；也提供處理大量的語料，除必須量化，且將量化結合理論架構，對假設提出實證驗證，一方面可以複現，一方面提出解釋。此外，除了解釋語流韻律的構成及特性，我們也驗證經過量化所導出的韻律組型，實具有語言學上的深層結構意義。

6.3 實驗四

在實驗四中，透過基頻重設與範圍的分析，我們證明了口語語流韻律模型與理論架構可進一步藉由一不相應的「韻律填充」短語加以擴展、整合。這些起首且相對短（最多 5 音節）的「韻律填充」短語應可以較窄的基頻範圍與不伴隨基頻重設等特徵加以說明。一旦起首基頻範圍窄化現象發生時，PG 尚未真正開始，PG 起首的基頻重設現象則位移至下一個短語首。這樣的窄化與重設位移現象的發現，使得口語韻律架構在敘事 (narratives) 或者口語篇章中可調整必要的特徵，而使得口語語流輸出結果更顯自然與樂律性。實驗四主要證明了語流韻律的表面多變，其實可以透過「韻律填充」的銜接功能，回歸其系統性，實驗一至三所提出的語流韻律深層結構是存在的基型；多語段如何形成語篇，也可清楚顯

示。我們也推論，「韻律填充」也具有類似節律效應分明的詩詞中的領字領詞的虛化功能與意義，我們後續的研究，也將繼續進一步驗證語流韻律中的虛化與關鍵成分的對應。

我們也在男女發音人的基頻重設與基頻範圍參數上發現了相似性與相異性。在相似性方面，我們皆在男女發音人語料中觀察到預期中的因結構所造成的窄化現象與重設的發生。在相異性方面，我們觀察到基頻重設不同的位置與組型。男性發音人以中性、跨短語間無語意焦點的方式唸出大部分的文本，產生了無抑揚頓挫的朗讀風格；另一方面，女性發音人則採用了較寬的基頻範圍變化且大幅度的基頻重設，因此對於相同文本產生了更具戲劇效果的朗讀風格。然而，比起女性發音人，我們仍可以更容易地在男性發音人語料中區分出「韻律填充」短語與一般短語在基頻範圍上的不同，藉由計算「韻律填充」短語與一般短語在基頻範圍上的變化（如〈圖 47〉所示）；換言之，從〈圖 47a〉中，我們可清楚地看出：男性發音人的「韻律填充」短語基頻範圍分布與一般短語分布的差別，但女性發音人則否。



〈圖 47〉 M051 (a) 與 F051 (b) 基頻範圍分布曲線圖

注意：「韻律填充」短語基頻範圍分布往往小於一般短語分布。

7. 結語

口語語篇韻律的具體聲學語音參數表現為何？超音段訊息中，語流韻律的相對貢獻度為何？何謂韻律語境？為什麼要從由上而下、相對而非絕對、對立而非

孤立、關連不止相鄰、大範圍與小範圍等這些角度來切入聲學語音學的研究？這些都是語篇韻律研究亟欲解答的問題。我們結合實驗語音學和語料庫語言學，開創語料庫語音學研究方法，從感知角度出發標註口語語流韻律，兼顧語音產製與分析間的對應，透過量化語料的分析，從跨短語的基頻曲勢、音節時長、邊界資訊、邊界停頓和停延段，一一找到各韻律階層貢獻的證據及跨短語聲學語音模版組型，得出語流韻律語境的聲學語音具體的體現。這些模版組型的存在，對言語鏈裡語言產製和語音處理而言，是關鍵性的由上而下的證據；意謂著人在連續講話時，無論是說話人即時語音產製的前置規劃或聽話人即時語音處理的預設模版，都涉及句子以上的規劃單元及上限，而不僅是線性連接所必須顧及的平滑效應，因此也為句法以上的深層語言、語意結構及認知閾限提供共構語境的證據。再者，我們驗證了「階層式多短語語流韻律架構 HPG (Hierarchical Prosodic Phrase Group)」的說法，證明口語語流韻律如何包含了來自句法結構的句調以上的語篇上層訊息，造成短語間的系統性關連所造成的韻律語境。語流韻律最顯著的特徵，不在於各個單字的字調或單句的句調，而是多個短語如何因接受上層語篇語意的管轄，成組成段所表現的跨短句韻律語境。

我們因此論述，語流韻律也具有跨短語的深層結構。正因為跨短句的韻律語境與字調句調層層疊加，加總後語流韻律表面輸出所呈現的多變性，其實可以經過韻律階層的釐清，而回歸其系統性。我們認為，連同由跨短語基頻關聯性、音節時長組型、音強分布組型與邊界停延所構成的口語語流整體的樂律、節奏、響度分布以及不同程度的延長及停頓反映出相當程度的語音規劃與處理策略、單位與範圍，以及語流中各韻律層的到位程度。換言之，跨短句系統性的基頻走勢、時程配置與邊界停延兩項韻律特徵為語者在語言產製規劃過程以及聽者同步處理所使用，語者透過可資操弄的聲學媒介形成的語流韻律所傳達的，同時也是聽者用來處理與預測隨之而來語音訊號的由上而下的規劃與同步處理的依據。

我們更進一步論述，從語言學意義的角度，從口語語篇敘述得到的證據，一方面具有類似音位的對應；另一方面也具有語言共通性方面的意義。透過這些證據，我們不但釐清了字調、句調與語篇敘述韻律分別受詞義、句法和篇章訊息上系統性的階層對應、制約管轄及分層貢獻度，更進一步解釋了句調單位，只是篇章敘述語意裡下屬的句法單位。由句法結構制約的句調，在語篇敘述中必須接上層語意連貫銜接的管轄而做出調整，因而在語流中字調、句調都可能偏離其典型形式 (canonical form) 而呈現各種不規則的表象。但終歸其本，偏離只不過是調整後的體現，而調整有其系統性規則可循，因此表象的變異與不規則，實可回歸其所本的深層架構與組型。

總體來說，HPG 架構最顯著的特徵如下：

- (1) 具體說明三個 PG 相對位置：PG 首與 PG 尾次級個別 PPh，爲了要產生整體韻律，如何調整個別調型，以呈現在 PG 之下的相對位置以及詞義與句法的規格 (specification)。
- (2) 釐清並解釋「口語連續語流中句調變化為何不是隨機的」以及「什麼樣程度的 PPh 可能/不可能維持它的短語句調」的原因。
- (3) 透過所有相關聲學參數（基頻曲勢、音節時長與音強分布）的量化分析，證實了 PG 效應的存在。
- (4) 證明 PG 每一韻律層對輸出韻律皆有貢獻，層層疊加形成語流最終韻律輸出 (prosody output) 的結果。

在同樣的語料中，我們進一步透過聲學語音訊息中的音域變化與基頻重設位置，進行對應句法結構與語意銜接的分析，檢視 PG 如何組合成篇章敘述。我們的證據顯示：一個以上的 PG 連接時，或可直接串接；或可附加「篇章提語」和「韻律填充」，以致表面樂律性好似不規則，而其實亦可釐清，因此論述口語語流韻律形似變異而實無損其共通，對共通性提出更具體的證據。最後，順帶一提，新近研究結果與數學模型同時可應用在語音資訊工程科學實務層面的技術提升，特別是語音合成及非限定文本轉語音 (unlimited text-to-speech) 領域，足以證明基礎語音科學的研究可直接提供語音科技開發的參考。

總地來說，我們自 2003 年起逐步證明：語流韻律並非只是一句一句不相關的短語句調連接成串，短語句調也並非最大的韻律單位。反之，句調是語流韻律的次級單位，語流中各短語句調的變化呈系統性因此是可預測的。也就是說，短語或短句，除了是受句法管轄制約的單位，也同時是語篇的構成成分，因此也需接受語篇語意的制約及管轄。語流韻律所必須表達的韻律語境，是造成短語句調調整與變化的主因，也因此語流中的短語句調看似變化多端，非句法結構所能解釋。

正因為如此，我們也進一步論述：語流韻律反應了句法結構以外的上層語意規範框架及其疊加意義。語篇產製時，說話人規劃的單位不是個別的句法句，而是上層語篇語意的大範圍多短語語段；語流處理時，聽話人並非僅處理短語句調所表達句法訊息，也必須同時處理韻律語境所表達的上層的語篇訊息，才能正確的處理短語間的關係。因此在句法階層之上還有語意階層，線性語流遞增的其實是各韻律階層訊息不同範圍管轄制約互動後的輸出。

基於以上實驗結果，我們得到以下結論：

- (1) 語篇韻律研究，應包括由上而下的角度、以短語語段為單位，考慮韻律產製的規劃與韻律處理的範圍，推導跨短語句組的深層結構。
- (2) 語音學研究，如有研究方法的創新與配合，使用大量語料建構量化證據，強化語音變異的處理，則可從少量語音現象質性描述跨出，提升語音學研究的科學性，並向理論層面的解釋推進。
- (3) 實驗語音學的研究，不但可從小範圍語音單位擴及大範圍語音單位；也可從少量語料音質的個別描述，提升到大批語料以量化統計數據以提取語音特徵。
- (4) 語料的增加和量化處理更具實證科學的依據，可多次複現 (replicate) 的研究方法也使語音學的研究更具科學性，驗證理論架構的實驗所得到的證據，更大幅提高研究發現的理論意涵。

以上這些來自即時語言產製與語音處理的關係，與認知的關連較之與特定語言的關係為大，也就是說，儘管各語言的整體韻律中的樂律性及節奏會因語言特定性 (language-specific) 而不盡相同，但上層語篇韻律的產製與規劃來源則與認知範疇的關係較大，因此上層語篇所造成的語流韻律的深層結構，較有可能是放諸語言皆準的一項特性。我們會繼續從多語言、跨語言的角度，擴展語流韻律的語言共通性；並試圖從樂律的角度，探索口語中音樂成分的聲學語音具體體現。

引用文獻

- Chao, Yuan Ren. 1968. *A Grammar of Spoken Chinese*. Berkeley: University of California Press.
- Fujisaki, Hiroya, and Keikichi Hirose. 1984. Analysis of voice fundamental frequency contours for declarative sentences of Japanese. *The Journal of the Acoustical Society of Japan (E)* 5.4:233-241.
- Jurafsky, Daniel, and James H. Martin. 2000. *Speech and Language Processing: An Introduction to Natural Language Processing, Computational Linguistics, and Speech Recognition*. Upper Saddle River: Prentice Hall.
- Tseng, Chiu-yu. 1990. *An Acoustic Phonetic Study on Tones in Mandarin Chinese*. Taipei: Institute of History and Philology, Academia Sinica.
- Tseng, Chiu-yu. 2006a. Prosody analysis. *Advances in Chinese Spoken Language Processing*, ed. by Chin-Hui Lee, Haizhou Li, Lin-shan Lee, Ren-Hua Wang & Qiang Huo, 57-76. Singapore: World Scientific.

- Tseng, Chiu-yu. 2006b. *An Acoustic Phonetic Study on Tones in Mandarin Chinese*. Taipei: Institute of Linguistics, Academia Sinica. (2nd edition, CD-ROM)
- Tseng, Chiu-yu, and Fu-chiang Chou. 1999. Machine readable phonetic transcription system for Chinese dialects spoken in Taiwan. *The Journal of the Acoustical Society of Japan (E)* 20.3:215-223.
- Tseng, Chiu-yu, Shao-huang Pin, and Yeh-lin Lee. 2004. Speech prosody: issues, approaches and implications. *From Traditional Phonology to Modern Speech Processing* [語音學與言語處理前沿], ed. by Gunnar Fant, Hiroya Fujisaki, Jianfen Cao & Yi Xu, 417-437. Beijing: Foreign Language Teaching and Research Press (外語教學與研究出版社).
- 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.
- 王洪君. 1996. 〈漢語語音詞的韻律類型〉,《中國語文》1996.3:167-171。
- 吳為章, 田小琳. 2002.《漢語句群》。北京：商務印書館。
- 鄭秋豫, 李岳凌, 鄭雲卿. (排印中). 〈兩岸口語語流韻律初探——以音強及音節時程分布為例〉, 周薦、董琨主編《海峽兩岸語言與語言生活研究》。上海：商務印書館。

[Received 15 October 2007; revised 30 April 2008; accepted 6 June 2008]

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

Corpus Phonetic Investigations of Discourse Prosody and Higher Level Information

Chiu-yu Tseng

Academia Sinica

This paper argues Mandarin discourse prosody must take into prosodic context from higher-level information in addition to lexical tones and phrase intonations. We postulate a prosody framework of Hierarchical Prosodic Phrase Grouping (HPG) and demonstrate how cumulative contributions from each prosody layer add up to output prosody. Cross-phrase prosodic templates corresponding to each HPG layer can be derived to represent the relative cross-phrase prosodic associations by three HPG specified relative positions (-initial, -medial and -final). In addition, we also discuss (1) how higher-level information is crucial to fluent speech prosody; (2) why corpus phonetics provides quantitative evidence to qualitative descriptions; and (3) how relative phonetic information can be better understood.

Key words: discourse prosody, paragraph prosody, prosody production, prosody processing, HPG (Hierarchical Prosodic Phrase Grouping), PG (Prosodic Phrase Group), acoustic correlates, supra-segmental, higher level information, super-positional, corpus phonetics