

國語口語語篇韻律研究簡述

鄭秋豫

中央研究院語言學研究所

摘要

從 2003 年開始，我們提出口語語流韻律有多短語的預設基型，進行一系列口語韻律假說的驗證工作，逐步透過語料庫語音學的研究方法，以計量研究方式，分析大批語料的超音段聲學參數及停頓與邊界效應，採用多元迴歸統計方法，以計算結果直接驗證各韻律階層的分層貢獻度，得出疊加構成語流輸出的研究數據，得到階層式多短語韻律句群 HPG 一說的證據。在理論研究方面，我們證明語篇韻律是字調、句調和語篇調相加的代數和（Tseng 2006a）；對字調和句調而言，語篇訊息是造成動態變化的原因。我們因而論述，國語的口語韻律，絕非僅是字調和句調或語調的研究所能涵蓋或解釋，篇章訊息才是造成二者在語流裡動態變化的原因。在研究方法方面，我們以量化研究的方式，補強語音學一向偏重音質的描述但無力處理大量語料的缺陷，並經由量化的結果提出語篇韻律的深層結構，因而證明質與量的研究可以互補。以上研究也同時證明，口語的韻律預設結構具框架性意義，與口語產製的規劃範疇、自然語言處理或認知範疇都可能相關，對口語溝通的意義不容忽視。

關鍵詞：韻律結構，語篇訊息，階層式 HPG 韻律結構說，語料庫語音學，動態韻律，多元迴歸統計法，語篇韻律

1. 前言——中研院語言學研究所語音實驗室研究特點——從音段走向語流韻律

語音學的研究，向來以音段為主。漢語語音學的研究，則又加上聲調的研究。終其目標，不脫音段的特徵、對立及音韻系統的建立的這些方向。但自八〇年代開始的語音科技研究，於開發系統與技術時，卻發現音段、音韻系統的知識，無法解決語音合成輸出語音的語流韻律與自然度的瓶頸；語音識別的研究也發現，音段在自然口語中的多變，也無法以語音學和音韻系統

的知識來應付。自九〇年代起，語音科學（speech science）的研究，以口語溝通時的語音現象為主，更進一步的發現：語音的研究，不能再侷限於定量的語料與定性的描述，而必須解決進行式的口語溝通中，除音段、聲調外，同時存在於語音訊號中，但承載其他語言資訊的問題。換言之，研究語音的課題，不但必須包括語音中的超音段訊息，也不能再假設超音段訊息的次要性，因此從九〇年代開始，口語韻律成為語音研究的重點課題，至今已有大量的文獻。

我們認為，從語音學、聲學語音學的角度研究語流韻律時，一來必須在音段與超音段間求取進一步的知識；二來必須找出處理大量的語料樣本的研究方法。我們基於以上的原因，於 1997 年開始收集語音資料庫，開展實驗語音學與資料庫語言學研究結合的研究方法，發展出與傳統語音學研究極為不同的語料庫語音學研究取向。現今，本實驗室的研究特點如下：（一）在研究取向方面：從跨學科研究切入，兼顧語音學基礎研究與語音工程跨學科應用研究；（二）在研究方法方面：不以傳統語音學觀察並描述局部、少量語音現象為目的，而以解釋全面性口語語流韻律現象為目的，開創語料庫語音學的研究方法，同時兼顧語音產製與語音感知、量化處理大批口語語音現象為目標；（三）在研究課題方面：率先將漢語語音學研究，提升到口語韻律的上層語篇訊息如何在語音訊號中體現的全面性研究，不再侷限於單字調、短語單句調的片面現象研究；（四）在研究成果方面：以國語語料的研究證據，提出「階層式多短語韻律語流架構 HPG（Hierarchical Prosodic Phrase Grouping）」一說，並採用多元迴歸統計方法，用計算結果直接驗證各韻律階層的分層貢獻度，得出疊加構成語流輸出的研究數據，據此推論「HPG 多短語韻律預設基型」具有跨語言意義（Tseng et al 2004, 2005; Tseng 2006）。（五）在研究成果應用方面：以量化結果為基礎，提出語流韻律的聲學參數數學模型，此一模型共包括四個聲學模組：基頻框架、節奏樣版、能量分佈趨勢以及停延與停頓結構，可說是現今數學性的韻律模型中，最完整的聲學語音模型。量化結果直接以語音合成及語音識別作為研究成果測試工具；亦即以語音學的研究成果，直接與語音合成及語音識別技術接軌，作為測試及應用研究成果的研究方式，並且對當今語音科技裡「文句翻語音」（Text-to-Speech TTS）語流韻律的力有未逮之處，提出具體改進的建議。

綜合言之，我們的研究取向係以「語音學基礎研究為本、語音科技應用開發為用」，並以科學驗證取代內省式或直觀式的描述研究方法，不但將漢語語音學的研究視角從微觀走向宏觀，不再侷限於單字調、連讀變調、簡單句句調的研究，跨大了語音學（phonetics）與語音科學（speech science）的交集面，從聲學語音的參數，找到了語流韻律的各種超音段特徵，也因而開創了實驗語音學的研究取向，解決了語音學研究質與量間的落差，將量化的結果回歸重質的深層語言學意義。

2. 口語語流韻律研究與傳統語音分析之不同

口語語流韻律到底是什麼？其實就是語篇韻律（Discourse Prosody）。語篇韻律的特性，表達的是口語敘述裡各短語句間的韻律相關性，強調的是透過韻律相關性所表達的語篇語意關連與一致性。對國語而言，是字調和句調以上的韻律成分；對句調而言，是造成動態變化的原因。

我們的口語韻律研究假設是，口語的語流韻律在超音段的層面，除了字調與句調以外，在句子（sentence）之上，還透過句子間語意的連接（cohesion）與連貫（coherence）形成韻律語境以表達上層語篇訊息，因此，「多短語語段（multiple-phrase speech paragraph）」是形成口語篇章時不可或缺的韻律單位，同時，語篇訊息也是造成單句調系統性變化的原因。語段管轄並制約句法單句間的銜接關係，系統性的造成句調的變化進而達到傳遞語篇訊息的溝通目的。

口語語流韻律，若從語篇的上層訊息角度切入，語段是語篇的次級單位，字調與句調是語段的次級韻律單位，語流韻律實為語篇韻律。從韻律的階層性來說，字調由詞義規範、短語句調由句法結構規範，二者都是靜態韻律，規範及制約範圍都是孤立單位且均不及短語以上；語流韻律是由語篇訊息規範，是動態韻律，規範及制約範圍是短語句調間的關連性與系統性的變化，不能再以孤立單位的表現做為處理單位（Tseng 2006a）。

反觀傳統語音分析一向採「由下到上」的研究觀點，研究韻律的重點多偏重字調和短語句調。其實，從口語語流韻律的研究角度來說，字調的規範制約單位只包括單字，屬詞義韻律（lexical prosody）；句法結構只規範孤立句的句型結構，句調單位僅達簡單句調單位（intonation unit 簡稱 IU），屬句法韻律（syntactic prosody）。而連讀變調其實表達的是字調之間關連性，單位多半只達多字詞；連讀句調會引起什麼變化，則一向不是傳統語音

分析的研究重點。趙元任先生（1968）曾以「大波浪小波浪」比喻，描述了字調與語調之間的互動關係。然而，複雜句調型或短語群調型的描述則付之闕如，語篇中句調間如何的互動分析也一向未受重視，實為傳統語音研究之侷限。

然而，我們從事的口語語篇韻律研究不同於傳統語音分析，在語音單位、研究角度和研究重點上均提出更新的想法。我們證明，語流韻律的研究，在語音單位方面，不能只侷限於字調、詞而止於句型，必須有大於簡單句的語流韻律單位。研究角度方面，不能只採「由下而上」著重小單位的微觀研究角度，必須兼容「由上而下」進行較大單位分析的宏觀研究；並進一步釐清不同角度的研究取向如何互補，使我們更進一步的瞭解語篇韻律，強調研究重點不能只採取孤立語音或韻律單位的研究，必須放大語音訊號中的語境，檢視語音單位間的關連性，在複雜的表面變異中，找出大單位的基型——也就是語言學所謂的深層結構。在語料取樣方面，一旦收集了大筆的語料，聲學語音的分析，便無法僅止於觀察描述個別現象，而必須向量化的方式求助，用統計的方法提取韻律的聲學特徵。我們因而開創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，兼顧語音產製與分析間的對應，透過量化語料分析四個聲學語音參數，即跨短語的基頻曲勢、音節時長、邊界資訊、邊界停頓和停延段，一一找到 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。

再者，口語語流韻律最重要的特性，就是由於上層語篇訊息（語意）的管轄效應，還有多短語的韻律預設，會造成下層韻律單位（尤其是短句）之間產生大範圍、跨短語的韻律關連性，因此，除了各孤立音段、短語句調一個連一個串接起來時，語音產製層面必然涉及的協同發音外，還必須考慮線性串接以外的韻律訊息。語流中的「關連性」必須透過一個以上的單位形成韻律語境才能表達，因此觀察或分析語料時，就不能個別檢視任何單一的單位，也不能把下層單位視為各自不相關的孤立單位；而且各單位間的關係，不僅包括相鄰單位間因發音平滑所產生的協同發音而已，還必須包括上層訊息管轄所造成的跨單位關係。這是我們的研究與所有以 intonation unit IU 為單位，而無視 IU 間是否有韻律關連；或是從語篇中逕行取出 IU 或更小單位後，再當作是孤立單位來分析的研究方式最大的不同。我們提出的說法是：口語語流韻律中，直接受制語篇語意管轄的下層單位，是多短語語段

PG 而不是短語 IU；多短語語段的下層單位才是短語 IU。因此，如果在韻律階層圖裡由下往上看，短句的句調因受到語篇語意訊息管轄制約，進入多短語語段後，隨即成為語段的下層韻律單位，必須在韻律上與上層管轄進行相關的調整；而如果由上往下看，則可更清楚的看出管轄制約的來源是什麼。

而我們提出的三段式語篇韻律模板，其實就像句法界定的 *intonation patterns* 一樣，可以視為口語語篇韻律裡多短語句段或者是類似英語複雜句的預設基型。也就是說，（一）在句調方面，我們建立了跨短語語調的基型。在多短語語段一開始就會有一個新的基頻高點重設（*f0 reset*），暗示著 *new information* 即將由此開始。隨後並非一路下傾到底就結束，而是停在一個基頻低點，因為話才剛開始語意不能結束；然後停一下，再開一個短語的頭，但是這個頭的基頻不會重設到剛才開始的高點，而整個短語句調也不會一路下傾到底，如此可以依語速快慢相互調節，如此可多到將近十個短語，句調都偏低平但不下傾到最低點，短語之間有各級停頓和伴隨停頓的邊界延長。要結束前的最後一個短語前又會有較長的停頓，然後有基頻新高的重設，只是結尾的重設不會高過起始的重設，然後下傾到整段的最低點。短語的韻律表徵不能獨立於多短語語段而存在，離開了彼此，多短語間的相關性就不存在。一個 PG 的結束接著一個 PG 的開始，基頻高低與邊界停延停頓的效應則更加明顯。這就是我們說的跨短語語調的基型（如圖 1 所示）。

（二）在節奏方面，我們提出了跨短語的各層級韻律單位的音節長短配置模版，以及邊界停延段與邊界停頓的系統性證據及 PG 效應。（三）在響度分配方面，我們提出了以短語為單位才有統計意義的跨短語模版，也就是一個短語越長，起始所需的能量越大；對照比較三個韻律單位的響度，我們證明起始和結束的短語相比，能量分佈的統計差異最大（Tseng & Lee 2004）。

（四）證明語流中的停頓雖然是靜音段不含聲學訊息，但卻是構成語流韻律的一個重要元素，它們不但與邊界停延成系統性的對應關係（因與發聲的生理機制有關），而且是語流中韻律調節不可或缺的韻律單位（Tseng & Chang 2007）。

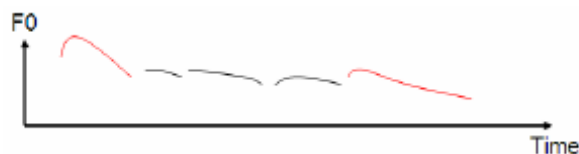


圖 1 跨短語語調基型示意圖

3. 口語語流韻律研究近年研究成果

從 1997 年開始的語料庫語音學研究，到 2003 年時，我們自行開發的語料庫語音學研究方法，已漸趨成熟，逐步透過量化研究方法，分析大批朗讀口語敘述語料的超音段聲學參數及邊界停頓效應，得到具體而有力的研究成果，獲得國際同行的青睞與肯定（Tseng 2003）。2004 年，參考 Keller 與 Zellner（1996 2001）有關法語時長的統計研究方法，以我們的語流韻律理論假說做為研究架構，修改多元迴歸模型，用來釐清上層訊息對於語流韻律的影響，從語流中跨短語音節時長分佈（語流的節奏）、邊界停延效應及響度的統計結果，證明語篇韻律是來自字調、韻律詞（詞義韻律）、韻律短語（句法韻律）和多短語語段（語篇語意）各韻律階層的規範與貢獻度的總和（Tseng et al. 2004），韻律短語的上層語篇訊息對總體韻律輸出確有貢獻度，我們提出的 HPG 假說因而得到支持，並進一步提出「階層式多短語韻律句群 Hierarchical Prosodic Phrase Grouping 簡稱 HPG」架構及跨短語節奏及響度模版。此說不但解釋了造成語流韻律的規範制約係來自語篇語意而非句法結構，並得以進一步推論：口語韻律絕非僅由字調及句法句調所能規範，也絕非語音合成觀點的字調與句調的線性串接及串接所涉及的平滑所能解釋。同年，我們亦得到音強響度與邊界停頓的證據（Tseng & Lee 2004），因而得以在 2005 年提出 HPG 跨短語基型及相對應的模組式聲學參數數學模型，詳見（Tseng, et al. 2005）。

2006 年，我們繼續從口語語流韻律獲得 HPG 的跨短語調型基型在基頻走勢的證據。採用 COSPRO 08（Tseng 2006a）承載語段（carrier paragraph）語料庫。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）。圖 2 中，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：爲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，圖 2 的目標字是一聲字「巴」），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。因此圖 2 中的三個目標字「巴」雖均爲第一聲，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。

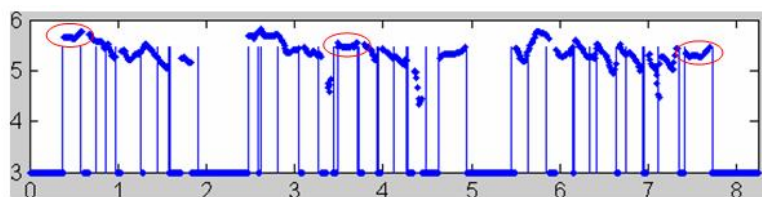


圖 2 承載語段基頻走勢圖（橢圓型小圈標出的是一聲目標字「巴」）

研究結果（表 1）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）爲 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 Ap 是變短的。PG 首與 PG 尾的係數反映出清楚的對比。總和來說，PPh 層僅構成韻律輸出近 40% 的貢獻度，而在 PG 層的上層訊息則能再多增加 25%。兩者加在一起，PPh 和 PG 層共組成韻律輸出的 65% 貢獻度。根據 HPG 階層架構的定義，其餘的 35% 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。此一研究結果證明：語流的樂律性也是來自字調、韻律詞、韻律短語和多短語韻律句群各韻律階層的規範與貢獻度的總和；再者，從基頻走勢的階層式貢獻度，得到韻律短語的上層語篇訊息對總體韻律輸出確有貢獻的證據（Tseng 2006a; Tseng 2006b; 鄭秋豫 審查中）。

表 1 有無 PG 效應 Ap 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

至此 HPG 一說已獲得每一個聲學語音層面的證據：即跨短語 F0 contour, syllable duration, intensity distribution, boundary lengthening 及 boundary breaks。我們因此得以進一步論述：1. 各韻律階層對韻律輸出的貢獻度可以得到量化研究證據，上層語意訊息對語流韻律的貢獻因而獲得證明，並系統性解釋語流中短語句調的變異來源；2. 各聲學語音參數（基頻軌跡、音節時程分佈、響度分佈、語流裡的韻律單位邊界停延與停頓）都存有跨短語預設基型（cross-phrase default bases form），如同語調，具有語言學及類似音位學的對比意義。

2007 年初，我們更進一步獲得以下二項新證據：（一）語流中韻律邊界效應的聲學特徵方面，語流停頓感知是透過邊界前後聲學參數對比以及停頓的系統性對比與互補所達成的，而非僅憑無聲停頓段的長短加以判別。在這些基礎上，我們將進一步檢視語流中各韻律單元的邊界（boundary）與接合（juncture）在聲學上的特徵與意義（Tseng & Chang 2007）。（二）透過分析不同韻律體裁（prosodic format）的語料，證明各韻律體裁既具有遵循韻律架構基型的一般規律，又具有區別於其他體裁的獨特之處；相對一致的韻律架構類型，它們不但證明了韻律基型的預設性，而且對語音表面的變異實來自共通的基底型式提出證據（Tseng & Su 2007）。因此，我們提出的口語語流韻律架構基型，不但具有類似音位的語言學上的意義；而且對語音科技開發、語言規劃提出由上而下思維的證據。

整體而言，口語語篇韻律研究近五年進行的語篇韻律研究重要成果，主要如下：

- （1）釐清連續口語語流韻律現象的規範制約來源。以量化研究資料證明：語流韻律的規範制約，其實來自語段或語篇的語意一致性（semantic cohesion）；最大的韻律特徵是韻律短語間的相對關連性（relative association）。口語溝通時，凡涉及敘述、討論、或

表達較複雜的概念時，單一孤立短語往往無法完整表達語意，而需以連續多個短語組成語意相關、口氣相連的語段。如果採傳統語音學的研究方式，只取孤立短語分析句調，其實並未包含任何語篇韻律訊息特徵；或者自語流中擷取短語，只做句調分析而不考慮所包含的上層語篇訊息，則無法解釋短語句調在語流中多變的表面形式因何而起。

- (2) 提出口語語流韻律即語篇韻律 (**Discourse Prosody**) 的說法，及六層的「階層式多短語韻律句群 **HPG** (**Hierarchical Prosodic Phrase Grouping**)」架構。「語篇」指的是語意完整的篇章敘述。透過量化分析大批語料的結果，證明語流韻律，除了包含詞義韻律 (**lexical prosody**, 即 **tone**)、句法韻律 (**syntactic prosody**, 即 **intonation**) 的訊息外，還包含了語篇語意 (**discourse semantics**) 的訊息。而來自上層語篇語意關連性訊息，形成由上而下的管轄效應，制約詞義韻律及句法韻律，達成語篇中大於句法句範圍的跨短語語意銜接作用，同時系統性地造成各級韻律單位的制約與互動。因此，我們認為語流韻律並非僅止於字調與句調的相加或抵消的單純效果，語篇範圍係由大於句調的韻律單位所構成，與邊界效應具有系統性關係存在；語流中的停頓是邊界效應停延後的必須成分，在語流中也承載韻律訊息。

在「階層式多短語韻律句群 **HPG**」架構中，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。**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是由下往上層層相疊，「自漣由波、推波助瀾、逐浪到潮」，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，其實是來自上層韻律單位系統性的規範制約。

- (3) 說明句調單位 IU 是口語語篇韻律的次級韻律單位，各短語是姊妹關係 (sister constituents)。即使同為敘述短句，成為 PG 的次級單位後，就必須依照 PG 指派的位置及基型進行修正調整，以產生大段落的韻律語流。在這個跨短語韻律語段的基型之上，還可以再附加焦點、強調、語氣等其他語音現象。換言之，「階層式多短語韻律句群 HPG」架構系統性的解釋了語流中短語句調變化多端的原因，並釐清短語句調在語流韻律中的變化皆成系統性關係，而非任意變異。
- (4) 提出對應 HPG 的跨短語調型、節拍、響度和停頓停延的預設基型模版 (default template)，及相對應的數學模型作為語音科技開發應用。簡單的說，HPG 將句法韻律 (短語句調)、詞義韻律、字調視為規範不同層級的下級韻律單位，只需在韻律短語上再加上一層 PG 韻律訊息，掌握三個韻律句群單位 PG-Initial, PG-Medial, PG-Final (亦即「PG 起始、中段與結束」) 的韻律特徵，便可系統性的解釋並預估多短語語段中短語句調、節奏、響度的調整，同時規範語流中的停頓與停延段 (也就是語流中的停頓邊界效應)，因此跨短語基型模版具有類似音位的語言學意義。
- (5) 強調口語語流韻律的共通性與變異性。以 HPG 的跨短語調型、節拍、響度和停頓停延的基型模版 (default template)，只需調節各韻律階層貢獻度的分佈，便可產生從朗讀白話文、氣象預報、古文朗誦、詩詞朗誦等不同的韻律體裁輸出。在語言認知與語言學的意義上，證明了此一基型的深層意義及語言發展上可能的意義；在語音科技開發方面，可直接應用到語音合成輸出，一個跨短語基型可產生多種語體的韻律輸出 (Tseng & Su 2007)。

4. 現階段研究重點

現階段的研究重點，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，檢視多短語句段 PG 如何組合成篇章敘述。初步的研究證據顯示：PG 連接時，或可直接串接；或可附加篇章提語 (discourse marker) 和韻律填充 (prosodic filler)，以致呈現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（Tseng et al. 2006）。我們因此提出語篇動態語流內含篇章提語及韻律填充的階層式結構（圖 3，Tseng 2006a），此一階層式結構伸縮自如，可因應語料的長短，將最上二層合併或展開。

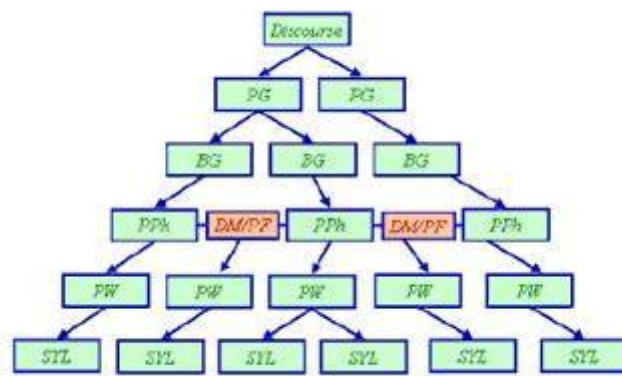


圖 3 階層式多短語韻律句群示意圖

我們也開始分析不同文體（stylistic）及韻律體裁（prosodic format）的語料，進一步提出變異無損共通的證據。我們分析朗讀不同文體（含唐詩、宋詞、樂府、古代散文）的韻律，發現每種文體在階層式的分層貢獻度上，呈現不同的分佈；韻律體裁越規律、短句以上管轄的規劃範圍越大、上層語篇語意訊息的貢獻度也越高。換言之，各種體裁，一方面既遵循韻律架構基型的一般規律；另一方面又具有區別於其他體裁、各自相對一致的韻律架構類型。因此，文體類型是建構於預設基型之上的，語流韻律因體裁的變異實為表象，其基底實乃共通，二者之間，毫不牴觸，不同文體的韻律模版，不過是以共通的預設基型為根據，在分層貢獻度上的做不同比例的系統性調整，因而造成不同韻律輸出（Tseng & Su 2007）。

5. 結論

本實驗室團隊從事語料庫語音學的研究十餘年來，以口語語流韻律研究為課題，相繼提出比短語語調更上層、更大範圍的語篇語意對語流韻律的貢獻證據。我們蒐集大批的口語語流語料庫，同時設計標註系統，從事量化研究，所有的假設，都用統計數據來支持，相較於傳統語音學研究不重視語料量的做法——只強調小範圍、細部現象的觀察，在研究方法上相當不同。在

語音單位、研究角度和研究重點上均提出更新的看法。以口語語篇中的多短語語段為主要研究單位，承襲了中國語言學之父趙元任先生以『大波浪小波浪』描述聲調與句調之間相互疊加或抵消關係的看法，強調各短語間因受來自上層訊息的管轄，而構築了相互間的韻律關係，因此語段裡的短語，好像兄弟姊妹一樣有了親屬關係，而不是將各短語之間視為不相關、各自擁有本身的句調的研究觀點。換言之，語段韻律如同句調，也有深層的多短語預設結構，表達跨短語的意義。從認知與生理角度出發，針對語流韻律中規劃模版、跨短句的抑揚頓挫、音樂性及節奏，層層找出證據，提出「階層式多短語韻律句群 HPG」架構，並以此一韻律架構提出跨句模組式的聲學模型，自由上到下（top-down）的研究取向，得到框架性的韻律組織在語音產製及處理二方面的意義，也將漢語語音的研究，走出字調與短語句調的研究範疇（鄭秋豫 2005）。

「階層式多短語韻律句群 HPG」架構，無論是解釋口語語流韻律現象，或是將研究結果應用到語音科技層面，在在說明了口語語流韻律最重要的特性：口語語流韻律必須包含句子以上的上層語篇語意訊息，語篇語意訊息管轄下層韻律單位（尤其是短句），並導致各短句間產生語氣上的銜接與關連性，而非傳統語言學研究將上層語篇語意訊息略去不看，僅把語篇段落視為各短句或短語的孤立句調特性個別串接起來所能比擬。

口語語流韻律研究，無論是與中文或西文研究語篇或韻律的研究相比（吳為章、田小琳 2000；馮勝利 1997, 2000, 2005；葉軍 2001；Hasan & Freis 1995），最大的特色在於：（一）以口語分析為主，文本分析為輔。先分析語料，提取韻律特徵，再進一步與文本分析結果對照。（二）研究方法係以 *corpus linguistics* 為主，證據完全來自量化分析的結果。（三）所選擇的口語語料，捨棄單獨唸出的語音小單位（如單獨唸出的單音節詞、多音節詞及短語句調），完全以語篇段落為主。（四）提出的研究結果解釋，以人說話時可能存在的認知閾限與語言產製時規劃的意義入手，而不僅從語言表面結構的字、詞或句法意義為主。

我們已證明，語流韻律的研究，（一）在語音單位方面，不能只侷限於字調、詞而止於句型，必須有大於簡單句的語流韻律單位。（二）研究角度方面，不能只採「由下而上」著重小單位的微觀研究角度，必須兼容「由上而下」進行較大單位分析的宏觀研究；並進一步釐清不同角度如何互補，使我們更進一步的瞭解語篇韻律。（三）研究重點不能只採取孤立語音或韻律

單位的研究，必須放大語音訊號中的語境，檢視語音單位間的關連性，在複雜的表面變異中，找出大單位的韻律基型。在未來的研究中，進一步的證據，則將推向跨語言的解釋性。（四）透過量化研究大量語料的方式，彌補傳統語音學只重語音細節音質描述、從音段的對立特徵建構音韻架構，卻不重觀察語料規模及口語語料中變異的研究限制。傳統聲學語音學研究，屬描述性的論述研究，取樣採定性音例，研究重點為個案描述，只針對特定類型語料進行分析；論述方式以哲學性推論為主，佐以少量音例，其研究的過程與結果無法複製（replicate），一旦碰到大量語料時，即顯出研究方法上的侷促，研究結果的個殊性無法解釋大量語料的普遍性。採語料庫語音學的研究方法時，蒐集大量語料後，不再以描述少數音例為目標，必須以採實證科學的方式，先有研究假設，再透過統計的方法進行預測，再以預測的結果驗證假設。如此的研究方法，整個過程都可以複製，而複製的信度則取決於量化的語料規模與過程。因此，採用語料庫語音學的方式，可以得出口語語流的特性與規則，同時具有跨語言上的意義。

整體來說，口語語流韻律研究的貢獻，除了說法上的原創和研究方法的創新，主要在於解答了語流中短語句調不明或多變的成因，還因為建置語音資料庫、處理大批語料，同時也彌補了語音學研究裡語料量不足的問題¹；又透過處理大批語料的量化程序，提取特徵與建立模版，同時也解決了語音現象時常因觀察語料量稍大便無法概括與找出操作基型的問題。口語韻律的研究，從此可以由上而下、從宏觀的角度整合微觀的觀察，強調各個語音單位在口語中的關連性，使語音學的研究不再是原子或分子式見樹不見林的研究，漢語的韻律研究，也不再只是字調（tone）與句調（intonation）的研究。

參考文獻

Chao, Y. R. (1968) A Grammar of Spoken Chinese. University of California Press, Berkeley and Los Angeles, California.

¹ 基於學術資源共享的理念，我們將 1995 年開始收集的口語語流韻律語料，共 9 組包含不同語流韻律特徵 10.58GB 的語料，連同我們開發的語音分析工作平台，統稱 Sinica Mandarin Continuous Speech Prosody Corpora（簡稱 COSPRO <http://www.myet.com/COSPRO>），於 2006 年初釋出。

- Hasan, R. and Fries, P. H. eds. (1995) *On Subject and Theme: a Discourse Functional Perspective*. John Benjamins Publishing Co., Amsterdam, The Netherlands.
- Keller, E., Zellner Keller, B. (1996) "A Timing model for Fast French", *York Papers in Linguistics*, 17, University of York. 53-75.
- Zellner Keller B, Keller E. (2001) "Representing Speech Rhythm" *Improvements in Speech Synthesis*.pp. 154-164. Chichester: John Wiley.
- Tseng, Chiu-yu. (2003) "Towards the organization of Mandarin speech prosody: Units, boundaries and their characteristics" *Proceedings of the 15th International Congress of Phonetic Science (ICPhS-2003)*, (Aug. 3-9, 2003), Barcelona, Spain, 599-602.
- Tseng, Chiu-yu, Pin, Shao-huang and Lee, Yeh-lin. (2004) "Speech prosody: Issues, approaches and implications" in *From Traditional Phonology to Modern Speech Processing (語音學與言語處理前沿)*, edited by Fant, G., Fujisaki, H., Cao, J. and Xu, Y., Foreign Language Teaching and Research Press (外語教學與研究出版社), 417-437, Beijing, China.
- Tseng, Chiu-yu and Lee, Yeh-lin. (2004) "Intensity in relation to prosody organization" *The Proceedings of the 4th International Symposium on Chinese Spoken Language Processing*, (Dec. 15-18, 2004), Hong Kong, 217-220.
- Tseng Chiu-yu, Pin, ShaoHuang, Lee, Yeh-lin, Wang, Hsin-min and Chen, Yong-cheng. (2005) "Fluent Speech Prosody: Framework and Modeling", *Speech Communication (Special Issue on Speech Prosody)*, Vol. 46:3-4, 284-309.
- Tseng, Chiu-yu. (2006a) "Prosody Analysis" in *Advances in Chinese Spoken Language Processing*, edited by Chin-Hui Lee, Haizhou Li, Lin-shan Lee, Ren-Hua Wang, Qiang Huo, World Scientific Publishing, 57-76, Singapore.
- Tseng, Chiu-yu. (2006b) "Higher Level Organization and Discourse Prosody" *Invited keynote paper, TAL 2006 (The Second International Symposium on Tonal Aspects of Languages)*, (April 27-29, 2006), La Rochelle, France. 23-34

- Tseng, Chiu-yu. (2006c) “Fluent Speech Prosody and Discourse Organization—Evidence of Top-down Governing and Implications to Speech Technology” Invited plenary keynote talk, *The 3rd International Conference on Speech Prosody 2006*, (May 2-5, 2006) , Dresden, Germany.
- Tseng, Chiu-yu, Su, Zhao-Yu, Chang, Chun-Hsiang and Tai, Chia-Hung. (2006) “Prosodic Fillers and Discourse Markers—Discourse Prosody and Text Prediction” TAL 2006 (The Second International Symposium on Tonal Aspects of Languages) , (April 27-29, 2006) , La Rochelle, France. 109-114.
- Tseng, Chiu-yu and Chang, Chun-Hsiang. (2007) “Boundaries and Boundary Properties in Mandarin Fluent Speech Prosody – the HPG Perspectives and Implications”, *Taiwan-Japan Joint Workshop on Speech Science and Technologies*, (March 29-30, 2007) , Taipei, Taiwan.
- Tseng, Chiu-yu and Su, Zhao-yu. (2007) “From One Base Form (HPG) to Multiple Prosody Outputs – Systematic but Dynamic Higher-Level Contributions and Distributions”, *Taiwan-Japan Joint Workshop on Speech Science and Technologies*, (March 29-30, 2007) , Taipei, Taiwan.
- 吳爲章、田小琳 (2000) , 《漢語句群》。北京：商務印書館。
- 葉 軍 (2001) , 《漢語語句韻律的語法功能》。上海：華東師範大學出版社。
- 馮勝利 (1997) , 《漢語的韻律、詞法與句法》。北京：北京大學出版社。
- 馮勝利 (2000) , 《漢語韻律句法學》。上海：上海教育出版社。
- 馮勝利 (2005) , 《漢語韻律語法研究》。北京：北京大學出版社。
- 鄭秋豫 (2005) , 〈「一氣呵成」是有意？還是有義？—淺談口語篇章的語流韻律〉。《中央研究院週報》Oct.6, 2005, 第 1040 期, 5-7, 台北。
- 鄭秋豫, (審查中) 〈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》。

鄭秋豫

cytling@sinica.edu.tw

Investigating Mandarin Spoken Discourse Prosody— Theoretical Significance and Methodological Advance

Chiu-Yu Tseng

Institute of Linguistics, Academia Sinica

Abstract

We analyzed speech corpora of read Mandarin Chinese discourses from a top-down perspective on perceived units and boundaries, and consistently identified speech paragraphs of multiple phrases that reflected discourse rather than sentence effects in fluent speech. A hierarchical multi-phrase framework is constructed to account for the top-down effect, with complimentary production and perceptual evidences. From derived cross-phrase F_0 and syllable duration templates we showed how the tune and rhythm characteristic to fluent speech prosody could be accounted for systematically, and argue for a prosody framework that specifies phrasal intonations as subjacent sister constituent subject to higher terms. Output fluent speech prosody is thus cumulative results of contributions from every prosodic layer. The corpus approach also narrowed if not bridged the gap between qualitative phonetic studies with quantitative evidences. To test our framework, we further constructed a modular prosody model of multiple-phrase grouping with 4 corresponding acoustic modules. The model could be applied directly to TTS and improve prosody simulation, and offer practical help to speech technology development.

Key words: prosody organization, discourse information, hierarchical PG framework, corpus phonetics, dynamic prosody