

仿趙氏音高尺度之基週軌跡正規化方法及其應用

A Pitch-Contour Normalization Method Following Zhao's Pitch Scale and Its Application

古鴻炎# 張小芬* 吳俊欣#
Hung-Yan Gu# Hsiao-Fen Chang* Jiun-Hsin Wu#

#國立台灣科技大學資工系 *國立台灣海洋大學
#National Taiwan University of Science and Technology, *National Taiwan Ocean University
e-mail: { guhy@mail.ntust.edu.tw, joanne@ntou.edu.tw }
www: http://www.csie.ntust.edu.tw/

摘要

本文研究一種音節基週軌跡的正規化方法，將音高由絕對式的赫茲(Hz)尺度轉換至相對式的趙氏(趙元任)音高尺度，以便解決跨人之音高位準和音域差異的問題。在對數赫茲尺度上，這個方法先對各個語者求出個人之平均音高和音高標準差，再用以作個人內部的音高正規化處理，接著依據跨人之音高分佈所呈現之分佈參數，作音高尺度之對應而轉換至所模仿之趙氏音高尺度。為了驗證所提方法的效能，我們收集了聽障與耳聰學生的語詞發音，比較未作、有作音高正規化處理之基週軌跡曲線的差別，結果顯示所提之音高正規化方法，的確可大幅減低跨人之音高位準、音域差異的影響。此外，我們也以這個音高正規化方法，來分析、探討聽障學生的聲調發音，在人耳評分上的一些現象。

1. 前言

許多語音處理相關的應用裡，都需要量測出語音信號內攜帶的基週(pitch)資訊，例如國語語音辨識之研究，需要基週資料來作聲調辨識[1, 2]；文句翻語音(text-to-speech, TTS)之研究，需要訓練語句的基週資料來建立基週軌跡的產生模型[3, 4]；語言學家研究各種語言的特性，也需要擷取出語音信號裡的基週資料來作分析[5, 6]。量測語音信號內的基週數值的問題，過去已有許多人作過研究及提出不錯的量測方法[7, 8, 9]，一般來說，程式自動量測出的基週數值，已有很高的可信度(雖然偶而還是會有量測錯誤)。

然而不同的語音處理之應用裡，各自對基週資料在後續處理上的需求，是不太一樣的，例如聲調辨識之研究，需要考慮使用者個人的音高位準和音域的問題；TTS之基週軌跡模型，縱使不需考慮多個語者的音高差異問題，也仍需考慮同一語者在不同天所唸的訓練語句，音高位準存在變異的問題；趙元任先生提出的五度制調值標記法[5]，則是假設作調值標記的人，會對不同語者的音高、音域差異，作機動式的調適。

由前面提到的應用例子可知，對聲調語言的語音信號作處理，大多會牽涉到音高正規化的問題，不過，並不是一種應用裡發展出的音高正規化方法，就可直接使用於另一個語音處理的應用裡。以前我們研究 TTS 之基週軌跡產生模型，曾經發展了一種以語句為單位作整體調整的音高正規化方法[4]，雖然由合成語音之聽測實驗，顯示此正規化方法之效能很不錯，不過它並未解決跨人之音高、音域差異的問題。再者，本論文將

要考慮的語音處理應用是，聽障與耳聰學生的跨人聲調的分析，無可避免地需要解決不同學生之間音高位準與音域寬窄差異的問題。音高位準的差異是很明顯的，因為語者包含男、女學生，且同性別的學生中也包含了處於不同發育變音階段的人；至於音域指的是基週數值的變化範圍，這個也會因人而變，我們說某人說話像唱歌(音域較寬)，而某人說話很平淡(音域較窄)，音域就是這種感覺的一個重要的聲響(acoustic)線索。然而不管男、女生，講話抑揚明不明顯，我們都聽得懂他們講的話，這意味一般人都具有機動地調適音高位準、音域寬窄差異的能力。反觀電腦程式，要如何讓它也具有這樣的能力？

因此，我們開始思索如何把赫茲(Hz)尺度(scale)之絕對音高轉換成相對音高，並且音高尺度希望能夠仿倣趙元任先生的五度制(稱為趙氏音高尺度)作法，藉以解決跨性別、年齡之音高正規化的問題，就我們所知，文獻上還未看過電腦自動音高正規化以轉成趙氏音高的方法。後來我們發展了一種正規化的作法，這個作法精簡說來是，在量測出一個音節的一序列的基週資料後，由於序列的長度不一，因此要先作音長正規化的處理，然後將赫茲值之音高轉到對數(log)赫茲值之音高，接著依據各人的音高平均值與標準差值，作各人內部的音高正規化處理，再依跨人的音高值分佈的統計，取得尺度轉換之參數，而據以轉換至趙氏音高尺度。基週量測，音長正規化，及音高正規化的較詳細作法，在第二、三節中說明，之後第四、五則說明所提之正規化方法應用於聽障與耳聰學生之跨人聲調分析的結果。

2. 基週量測與音長正規化

在作音高正規化之前，有幾個相關的前置處理步驟要作，如圖 1 之整體處理流程所示，從一個詞彙的錄音檔案，取樣率 22,050Hz，16bits/sample，取出語音樣本，首先作

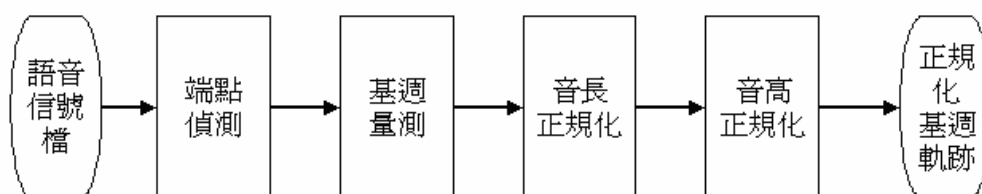


圖1 整體之處理流程

端點偵測，以決定各個音節的起點與終點；將起點與終點之間的語音樣本，切割成一序列的音框(frame)，以量測各音框內信號的基週數值，音框長度為 25ms 且相鄰音框重疊 12.5ms；由於各音節的有聲(voiced)音框數量有多有少，所以接著要作音長正規化之處理，以使用固定 16 維(dimensions)之頻率向量來表示一個音節的基週軌跡；之後，進行主要的音高正規化處理，以轉換成仿趙氏音高尺度所表示之基週軌跡。

關於圖 1 裡的處理方塊的細部說明，”音高正規化”將於下一節說明，本節就對”基週量測”和”音長正規化”來作進一步說明，較基本的”端點偵測”，則可參考語音處理之文獻[9, 10]，不過有一點需注意的是，詞彙發音時，兩音節之間可能會有共發聲(coarticulation)現象，而使端點偵測發生錯誤，這時就需要人工來作更正。

2.1 基週量測

依據端點偵測得到之起點與終點，將兩點之間的語音樣本，切割成一序列的音框，再對各音框來作基週量測。音框內的信號可能是無週期性的(如無聲子音部分)，我們能夠偵測出來，並直接設定基週值為一個特殊值；若為週期性的，則要量測出週期的長度值。

我們所用的偵測方法如下。先計算自相關(autocorrelation)函數 $R(k)$ 及平均振幅差距(average magnitude difference)函數 $M(k)$ ，計算公式分別是

$$R(k) = \frac{1}{N} \sum_{n=0}^{N-1} x(n)x(n+k) \quad , \quad k = k_{min}, \dots, k_{max} \quad (1)$$

$$M(k) = \frac{1}{N} \sum_{n=0}^{N-1} |x(n) - x(n+k)| \quad , \quad k = k_{min}, \dots, k_{max} \quad (2)$$

其中 $x(n)$ 表示語音信號樣本， N 表示一個音框內的樣本點數， k 表示可能的基週長度之點數。我們先設定最低及最高的音高值分別為 60 及 750 赫茲，再據以設定 k_{min} 及 k_{max} 的數值。在公式(1)中，當 $k = 0$ ， $R(0)$ 就是短時能量，我們也要計算此值，然後採用 Kim 等人提出的判斷規則[11]，用以判斷本音框內的信號，是無週期性或是有週期性的信號，兩個規則的內容是：

規則(1): if $\text{Max}_{k_{min} \leq k \leq k_{max}} (R(k)) < R(0)/4$ then not periodic

規則(2): if $\text{Max}_{k_{min} \leq k \leq k_{max}} (M(k)) / \text{Min}_{k_{min} \leq k \leq k_{max}} (M(k)) < E$ then not periodic

其中 E 為門檻值，我們依實驗的結果設定其值為 2.1。通過判斷規則後，再去計算基週數值，計算方式如下[11]：

$$f = \frac{22,050}{\text{Max}_{k_{min} \leq k \leq k_{max}} \left(\frac{R(k)}{M(k)+1} \right)} \quad (3)$$

其中 22,050 是取樣率。

2.2 音長正規化

每個人講話的速度，或同一人在不同時間的說話速度都會不相同，導致音節內有聲音框的數量也呈現出有的較多有的較少，但是為了方便作音高正規化之處理，所以在先作音長正規化，以使用固定 16 維(dimensions)之頻率向量來表示一個音節的基週軌跡。

我們的作法是，在時間軸上均勻放置 16 個音高之取樣點，然後以內差的方式來求出各點上的音高。設本次分析的音節裡有 NF 個有週期性之音框，則第 k 個音高取樣點放置的時間位置(以音框為單位)是：

$$t_k = \frac{NF-1}{16-1} \times k, \quad k = 0, \dots, 15 \quad (4)$$

以圖 2 為例，設時刻 t_k 時的音高值 g_k 是所要求取的，則我們可取 t_k 前後各兩個音框所量測出的音高值來作 Lagrange 內差，而求得 g_k 的值。實作上，令圖 2 裡的 i 表示第 i 個音框，它的值由 t_k 決定，即 $i = \lfloor t_k \rfloor - 1$ (取整數後減 1)，若 i 的值小於 0，則改設 i 的值為 0； f_i 表示第 i 個音框內量得的音高值；接著使用如下之 Lagrange 內差公式[12]

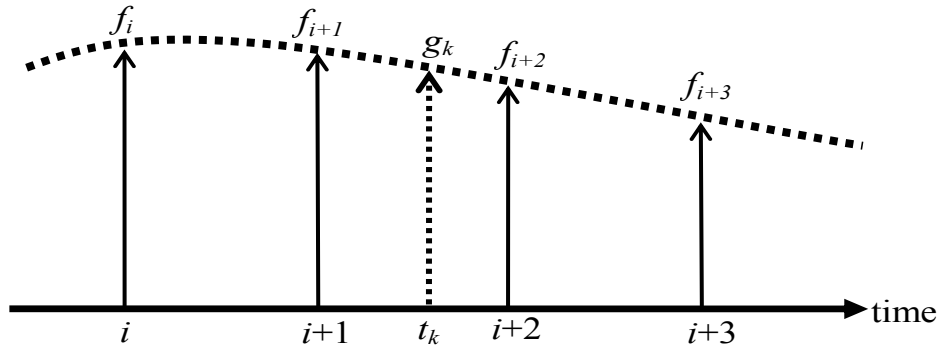


圖 2 音高內差示意圖

來求 g_k 之值：

$$g_k = \sum_{n=0}^3 L_n(t_k) \cdot f_{i+n} \quad , \quad L_n(t) = \prod_{\substack{j=0 \\ j \neq n}}^3 \frac{t - (j+i)}{(n+i) - (j+i)} \quad (5)$$

其中 i 就是圖 2 裡由 t_k 決定之 i 值。

3. 音高正規化

求出一個音節的 16 點時間正規化之基週軌跡 g_0, g_1, \dots, g_{15} 後，接著就將音高值轉換至對數尺度，即令 $h_k = \log_{10}(g_k), k=0, \dots, 15$ ，這樣做是要仿倣人耳對音高的知覺，就如同音樂裡以十二平均律來排列音階，或語音辨識裡常用的梅爾(mel)尺度。然後，我們按照圖 3 的流程來作音高正規化之處理，圖 3 的“個人平均音高”方塊，分別對各個

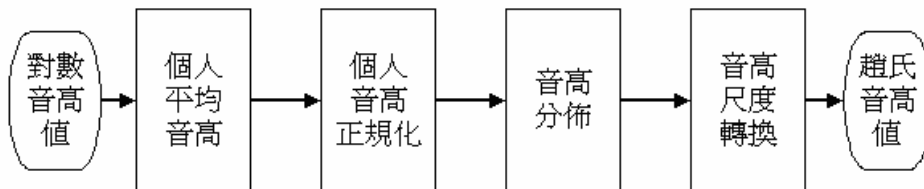


圖 3 音高正規化之處理流程

語者的音高資料去算出各個語者個人的平均音高 μ 和音高標準差 σ ，計算公式如下：

$$\mu = \frac{1}{NS \times 16} \sum_{n=0}^{NS-1} \sum_{k=0}^{15} h_k^n \quad (6)$$

$$\sigma = \left[\frac{1}{NS \times 16} \sum_{n=0}^{NS-1} \sum_{k=0}^{15} (h_k^n - \mu)^2 \right]^{1/2} \quad (7)$$

其中 NS 表示本次分析之語者所唸的音節個數， h_k^n 表示第 n 個音節的第 k 個基週軌跡點上的音高值。

求得一個語者的平均音高和音高標準差後，接著在圖 3 的“個人音高正規化”方塊，就依如下公式

$$\alpha_k^n = (h_k^n - \mu) / \sigma \quad (8)$$

來作個人內部的音高正規化處理， α_k^n 表示第 n 個音節的第 k 點上的內部音高正規化後的值。 h_k^n 扣掉 μ 就可把絕對式音高轉變成相對式音高，而解決了不同語者之間音高位置差異的問題；接著除以 σ ，則用以解決語者之間音域差異的問題。

做完每個語者的個人內部音高正規化處理，接著在圖 3 的“音高分佈”方塊，我們把所有語者唸的所有音節之音高資料 α_k^n 放在一起，作 Histogram 處理，由於大於 5(5 倍標準差)或小於 -5 的數值幾乎沒有，因此我們把 α_k^n 數值大於 5 的只當作一個區間，小於 -5 的也只當作一個區間，而在 5 與-5 之間，則切割成 100 個區間，如此我們得到如圖 4 所示的音高分佈圖，由此圖可看出，音高數值主要出現於橫軸第 20 至 80 區間之

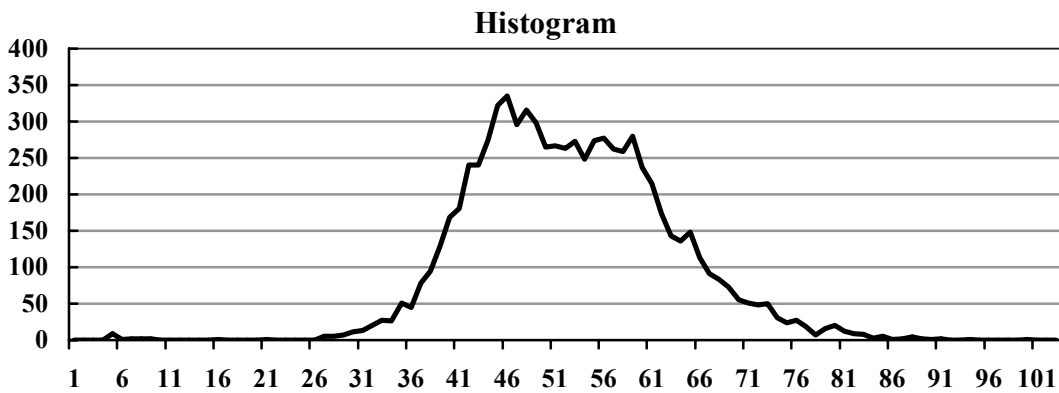


圖 4 跨人之正規化音高值分佈

間，也就是 α_k^n 的數值主要出現於-3 至+3 之間，並且分佈圖呈現出雙峰的現象，我們覺得這是可以解釋的，因為國語裡第一聲的基週軌跡音高很高且相當平坦，所以造就右邊的峰，而第三聲大多唸前半上且音調較低，且詞尾位置的第二聲，音調也較低，所以形成左邊的峰。

依據圖 4 的音高值分佈，接著在圖 3 的“音高尺度轉換”方塊，我們考慮把 α_k^n 的數值 0 對應至趙氏音高尺度的中間音高，即數值 3，這是加 3 的關係。至於數值範圍的對應，趙氏音高尺度的數值範圍是 1 至 5，而我們的 α_k^n 的數值範圍主要是在-3 至+3 之間，範圍差距不大，因次就不作數值範圍的減縮調整，如此我們決定使用如下之音高尺度轉換公式：

$$\beta = \begin{cases} -3, & \text{if } \alpha < -5 \\ \alpha + 3, & \text{if } -5 \leq \alpha \leq 5 \\ 9, & \text{if } \alpha > 5 \end{cases} \quad (9)$$

來把個人內部正規化後之音高值 α_k^n 轉換至趙氏尺度的音高值 β_k^n 。公式(9)中， α 值小於-5 或大於 5 的特殊情況，是用以處理聽障生所唸音節，有時會音調過高或過低。

4. 本方法之實驗驗證

對於第二、三節說明之仿趙氏尺度的音高正規化方法，在此節我們以實驗的結果來檢驗它的效用。這裡參與實驗的發音語者共計 14 名耳聰學生，均係說話清晰且國語發音標準者，其中 9 名為國小生 5 名為國中生，性別上則是男生 10 名女生 4 名。選用的發音詞彙為：‘冰棒’、‘熨斗’、‘帆船’、‘肥皂’、‘小熊’、‘草地’、‘崖谷’、‘愛心’、‘卡車’、‘燈塔’、‘西瓜’、‘圍巾’、‘照片’、‘拼圖’、‘氣球’等共 15 個，滿足雙字詞的所有聲調組合，且詞中第二字都含有子音聲母，以方便程式自動作音節切割。各詞彙以隨機出現的方式讓每人唸三遍，同時錄音存檔，之後選擇其中一遍聽起來聲調最準確且聲波振幅較佳者來作本節的實驗。

如果只作圖 1 中前三個方塊的處理，即不作音高正規化，且把各個聲調的基週軌跡分別集合起來繪圖，則我們得到如圖 5, 6, 7, 8 所示的四個聲調之基週軌跡帶狀圖，其中較粗且有菱形標記的實線為平均的基週軌跡曲線。接著如果進行音高正規化的處理，則圖 5, 6, 7, 8 之基週軌跡帶狀圖，會分別變成如圖 9, 10, 11, 12 所示的情況。比較圖 9 至圖 12、和圖 5 至圖 8 這兩組基週軌跡帶狀圖，整體而言，音高正規化之處理可以讓軌跡線條由鬆散變得較為集中；再者，可觀察到第一、二聲的軌跡，在作過音高正規化的圖 9、圖 10 之間，有較明顯的音調高度、與軌跡彎曲度的差異，而在圖 5、圖 6 之間，高度與彎曲度的差異則較不明顯；此外，觀察第三、四聲的軌跡，在作過音高正規化的圖 11、圖 12 之間，有較明顯的軌跡斜率的差異，而在圖 7、圖 8 之間，軌跡斜率的差異較不明顯。所以，本論文研究之音高正規化方法，的確可用以大幅減低個人音高位準、音域寬窄等因素所造成的影響。

對於作過音高正規化的圖 9, 10, 11, 12 裡，各聲調的軌跡線條，仍然表現出一些寬度的帶子形狀，而不是十分密集，這種情形是可以解釋的，我們所錄的 15 個雙字詞發音，各聲調都會均勻地出現在詞頭和詞尾，並且就同一個聲調而言，出現在詞頭的一般來說會比出現在詞尾的音調較高(即下傾現象)，所以，雖然是同一個聲調，音調高低仍然會存在著明顯的變異。

5. 聽障生之語詞聲調分析

這裡參與實驗的聽障生人數、年齡、性別都和耳聰生一樣，所以一共有 14 人，9 名國小生加 5 名國中生，而性別上則是男生 10 名女生 4 名。參與的聽障生，都是語言學習前失聰，包括中度 2 名、重度 6 名，極重度 6 名，且都有接受聽障班或聽障資源班之聽覺口語訓練。在此選用的發音詞彙和耳聰生的完全一樣，並且各詞彙也是以隨機出現的方式讓每人唸三遍，同時作錄音存檔，之後選擇其中一遍聽起來聲調最準確且聲波振幅較佳者來作本節的實驗。

聽障生所唸的雙字詞發音，基本上也是按照圖 1 的流程來作處理，不過在圖 3 的“個人音高正規化”方塊，我們認為聽障生的音高標準差值，有很大的可能性是不正確的，因為聽障生的一種發音障礙情形是，分不清四個國語聲調的相對音調高低，因此，我們在處理聽障生的個人音高正規化時，音高標準差值是直接設定為 14 個耳聰生的音高標準差的平均值。作完音高正規化的計算後，各個音節的基週軌跡已是在趙氏音高尺度上

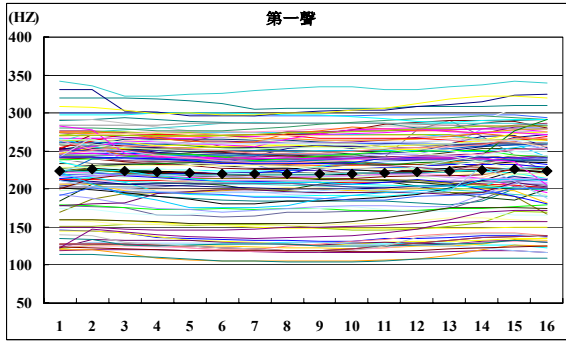


圖 5 第一聲基週軌跡(未音高正規化)

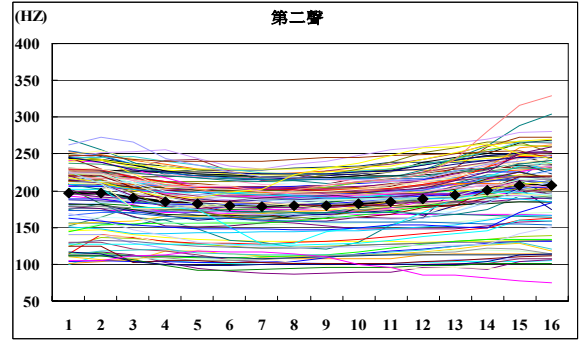


圖 6 第二聲基週軌跡(未音高正規化)

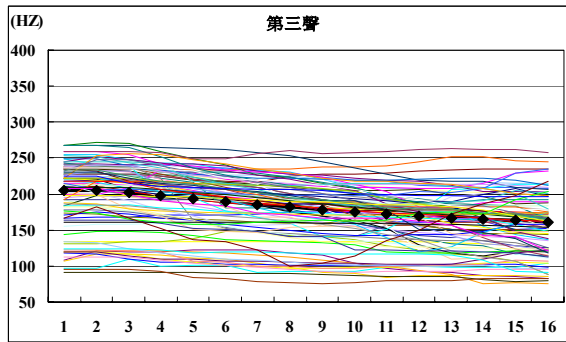


圖 7 第三聲基週軌跡(未音高正規化)

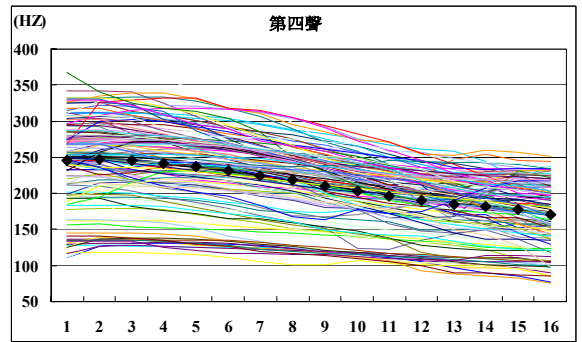


圖 8 第四聲基週軌跡(未音高正規化)

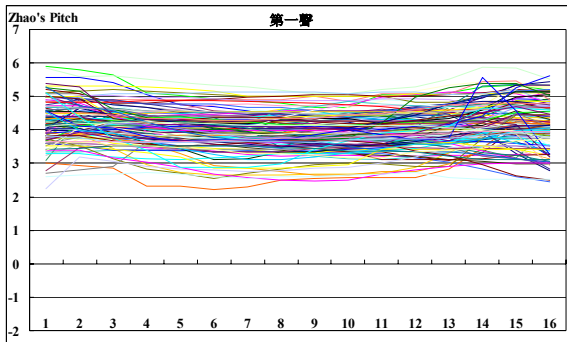


圖 9 第一聲基週軌跡(音高正規化後)

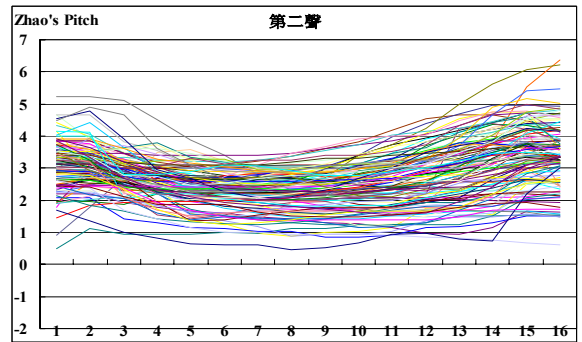


圖 10 第二聲基週軌跡(音高正規化後)

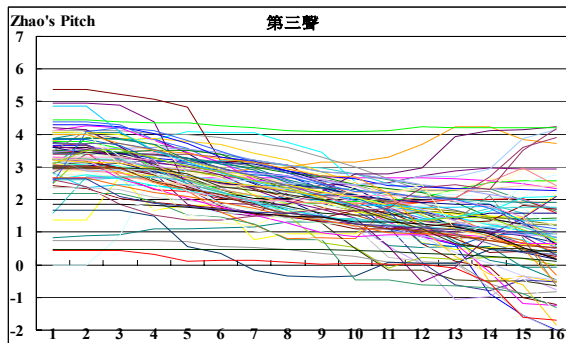


圖 11 第三聲基週軌跡(音高正規化後)

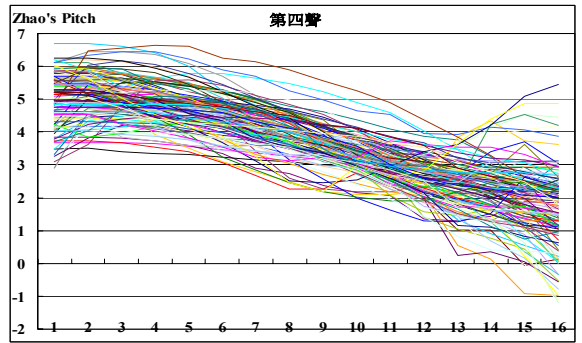


圖 12 第四聲基週軌跡(音高正規化後)

的數值，接著把各個聲調的基週軌跡分別集合起來，再依據詞頭與詞尾之發音位置分成兩組，各自去計算平均的基週軌跡，用以分析各個聲調的基週軌跡是否和詞頭、詞尾之位置有相關性。我們依照聲調及詞內發音位置，分別把聽障生和耳聰生的平均基週軌跡畫在同一圖裡，以方便作比較，結果得到如圖 13 ~ 圖 20 所示之平均基週軌跡圖形。

由圖 13 ~ 圖 20 可觀察得知，就耳聰生而言，第一、三、四聲的基週軌跡形狀，和詞內的發音位置並無明顯的相關性，但是第二聲的基週軌跡形狀則和發音位置有非常明顯的關係存在，即在詞頭位置所發的第二聲，後端具有大幅度的上揚趨勢，而在詞尾位置所發的第二聲，後端上揚的趨勢很和緩，反而前端下降的趨勢更為明顯。此外，對於第一、二、三、四聲都共同具有的現象是，同一個聲調的兩個發音位置，在詞頭位置的都會具有比較高的音高。就聽障生而言，各個聲調的基週軌跡形狀，都和詞內發音之位置沒有明顯的相關性。此外，聽障生對於兩個發音位置，在詞頭位置也會表現出比詞尾位置較高一些的音高。

另一方面，從圖 13 ~ 圖 20 也可觀察到耳聰生和聽障生的不少差異的地方。例如第二聲的基週軌跡(圖 15、16)，耳聰生的兩個位置的軌跡都會先下降再上揚，反觀聽障生的軌跡則都只有下降，而無上揚的趨勢，在軌跡的尾端，耳聰生和聽障生的軌跡甚至於呈現完全相反(一個上升一個下降)的走勢。再者，觀察第三聲的基週軌跡(圖 17、18)，聽障生和耳聰生之間也有明顯的差異，聽障生的軌跡較為平緩，下降速度明顯地比耳聰生的要緩慢很多，以致於聽障生本身的第二聲與第三聲的基週軌跡，無論在軌跡形狀和音高方面，幾乎是很難加以區分。這與潘奕陵[13]與鍾玉梅[14]之研究相呼應，即聽障兒童發音方面構音的轉換速度太慢，由此可見，聽障學生在第二聲和第三聲的發音最無法區辨，也就是第二聲、第三聲的發音最為困難。

從人耳對聽障生唸的聲調所作的評分，也發現聽障生四個聲調得分會因聲調不同而有差異存在。人耳評量係由修習教育學程的三位研究生擔任，在評量前均接受過講習，講習重點除說明評量標準與計分方式，也實際就語音樣本進行演練，確認給分標準與評鑑項目之後才分開進行評分工作，每位受試之各個語詞聲調得分，給分標準必須符合三人中至少有兩人給分相同，若任何一項得分不一致時，需三人同時再聽一次，重新確定真正得分。從評分結果我們得到，平均分數上第二聲和第三聲的分數最低，而第一聲和第四聲的分數較高，這樣的評量結果，與過去有關聽障學生之聲調研究的結果是一致的[15, 16]。所以，人耳評分的結果確實可由圖 13 ~ 圖 20 之音高正規化的基週軌跡曲線來作聲學分析上的佐證。第一聲和第四聲的得分會較高許多，這個也可從圖 13、14 和圖 19、20 的基週軌跡曲線來分析，因為發第一聲時，聽障生和耳聰生的軌跡都相當平坦且音高夠高，雖然在軌跡的左右兩端，聽障生和耳聰生的軌跡仍存在一點差異；至於唸第四聲時，聽障生和耳聰生的軌跡都呈現出劇烈的下降走勢，音高的高低落差至少是縱座標的兩個單位以上，所以可讓人耳感覺出是第四聲，不過聽障生的軌跡的下降幅度明顯地是比耳聰生的少許多。

6. 結論

本文研究提出一種跨性別、年齡的音高正規化的方法，將音高由絕對式的 Hz 尺度轉換至相對式的趙氏(趙元任)音高尺度，如此用以解決不同語者之間音高位準和音域差

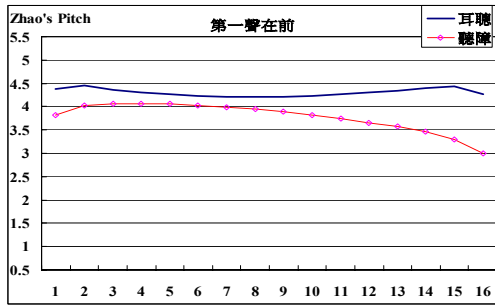


圖 13 第一聲在詞頭的平均基週軌跡

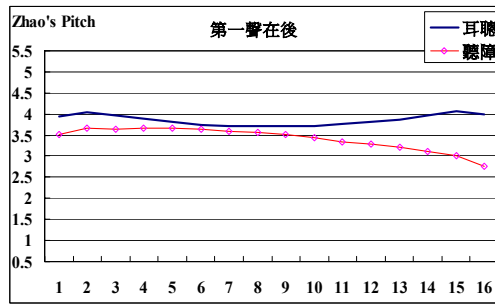


圖 14 第一聲在詞尾的平均基週軌跡

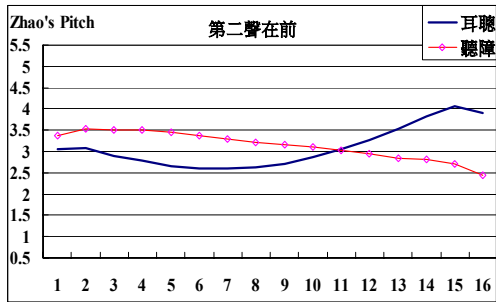


圖 15 第二聲在詞頭的平均基週軌跡

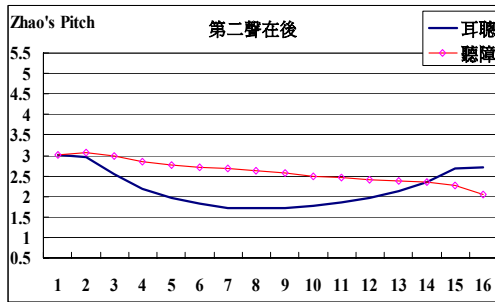


圖 16 第二聲在詞尾的平均基週軌跡

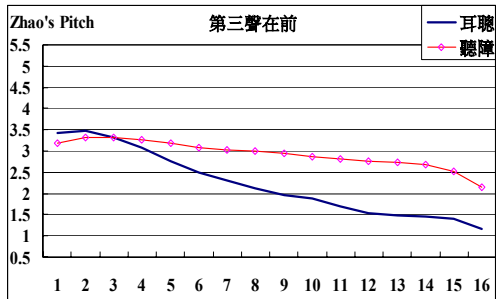


圖 17 第三聲在詞頭的平均基週軌跡

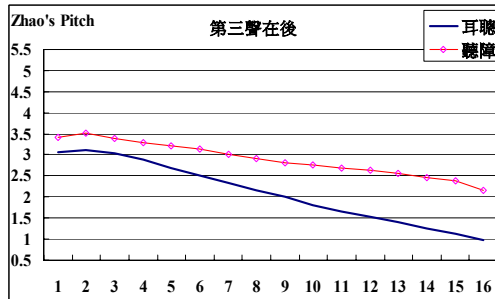


圖 18 第三聲在詞尾的平均基週軌跡

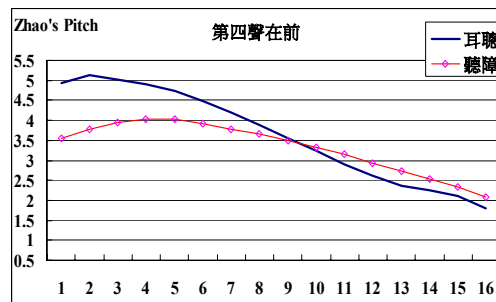


圖 19 第四聲在詞頭的平均基週軌跡

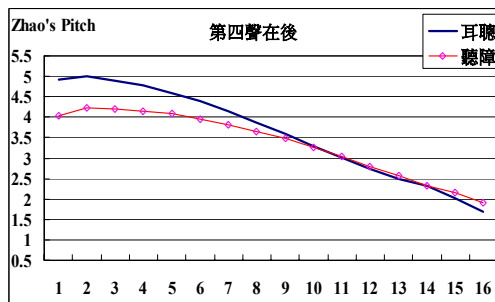


圖 20 第四聲在詞尾的平均基週軌跡

異的問題。趙氏音高尺度，是語言學家分析聲調語言最常用之音高尺度，但是過去只能靠人耳去調適、標記受測者講話的音高，難免有時會不甚精確，因此我們所提出的方法，也許仍有缺點，但至少是一種客觀、電腦自動作音高分析的方法。這個方法的處理流程也許看來很簡單，不過它的效能的確已經經過實驗的驗證。

對於所提出的仿趙氏音高正規化方法，我們以耳聰學生所錄的語詞發音，來比較未正規化前和正規化後的基週軌跡曲線，發現音高正規化處理可以讓軌跡線條由鬆散變得較為集中，並且第一、二聲之間的軌跡差異，在音調高度和軌跡彎曲度兩項上，會變得更為明顯；第三、四聲之間的軌跡斜率的差異，在作過音高正規化後，會變得更明顯而容易分辨。此外，我們也以所提之方法來對聽障生和耳聰生的基週軌跡作處理，之後再比較聽障生和耳聰生的軌跡差異，發現音高正規化後的軌跡，的確可用以解釋人耳對聽障生聲調評分的一些現象。

參考文獻

- [1] Yan, W. J., J. C. Lee, Y. C. Chang and H. C. Wang, "Hidden Markov Model for Mandarin Lexical Tone Recognition", IEEE trans. ASSP, Vol. 36(7), pp. 988-992, July 1988.
- [2] Gu, H. Y. and Lin-Shan Lee, A Study on a few Relevant Problems about Machine Dictation of Mandarin Speech, Ph.D. Dissertation, Department of Computer Science and Information Engineering, National Taiwan University, Taiwan, Jan. 1990.
- [3] Chen, S. H., S. H. Hwang and Y. R. Wang, "An RNN-Based Prosodic Information Synthesizer for Mandarin Text-to-Speech", IEEE trans. Speech and Audio Processing, Vol. 6(3), pp. 226-239, 1998.
- [4] Gu, H. Y. and C. C. Yang, "A Sentence Pitch Contour Generation Method Using VQ/HMM for Mandarin Text-to-speech", International Symposium on Chinese Spoken Language Processing (ISCSLP2000), Beijing, pp. 125-128, 2000.
- [5] 趙元任，中國話的文法，台北：敦煌書局，1981。
- [6] Tseng, C. Y., An Acoustic Phonetic Study on Tones in Mandarin Chinese, Special Publications No. 94, Institute of History & Philology, Academia Sinica, Taiwan, 1990.
- [7] Rabiner, L., M. Cheng, A. Rosenberg and C. McGonegal, "A Comparative Performance Study of Several Pitch Detection Algorithms", IEEE trans. ASSP, Vol. 24, pp. 399-418, Oct. 1976.
- [8] Medan, Y., E. Yair and D. Chazan, "Super Resolution Pitch Determination of Speech Signals", IEEE trans. Signal Processing, Vol. 39(1), pp. 40-48, Jan. 1991.
- [9] O'Shaughnessy D., Speech Communications: Human and Machine, 2nd ed., IEEE Press, 2000.
- [10] Rabiner, L. and B. H. Juang, Fundamentals of Speech Recognition, Prentice Hall, Englewood Cliffs, 1993.
- [11] Kim, H. Y., *et al.*, "Pitch Detection with Average Magnitude Difference Function Using Adaptive Threshold Algorithm for Estimating Shimmer and Jitter", Proc. of the 20th Annual International Conference of the IEEE, Engineering in Medicine and Biology Society, Vol. 6, pp. 3162-3164, 1998.
- [12] Stoer, J. and R. Bulirsch, Introduction to Numerical Analysis, 2nd ed., New York: Springer-Verlag, 1993.
- [13] 潘奕陵，聽覺障礙者語詞及句子層次的說話清晰度之知覺分析，碩士論文，特殊教育研究所，國立台灣高雄師範大學，1998。
- [14] 鍾玉梅，「聽障兒童的說話問題」，聽語會刊，第 10 期，第 72-79 頁，1994。
- [15] 張蓓莉，「聽覺障礙學生說話清晰度知覺分析研究」，特殊教育研究學刊，第 18 期，第 53-78 頁，2000。
- [16] 張淑品，國中重度聽障學生與耳聰學生國語單元音與聲調的聲學比較分析，碩士論文，特殊教育學系，國立台灣師範大學，1999。

致謝

感謝國科會計畫支援，計畫編號: NSC 91-2520-S-019-002.