

華語捲舌音對比的聽覺辨識與頻譜動差分析

鄭靜宜

國立高雄師範大學聽力學與語言治療研究所

論文編號：07039；初稿收件：2007年11月8日；第一次修正：2009年5月7日；完成修正：2009年5月11日；正式接受：2009年5月12日

通訊作者：鄭靜宜 802 高雄市苓雅區和平一路116號 國立高雄師範大學聽力學與語言治療研究所 (E-mail: jjeng@nknuc.nknu.edu.tw)

華語中有三對摩擦音和塞擦音的捲舌/非捲舌音對比，華語捲舌音具有特殊的聲學和聽知覺特徵。本研究的目的主要在探討正常成年人對於單音節和雙音節詞中捲舌音對比的聽覺區辨以及聲學頻譜動差參數和捲舌音與非捲舌音聽辨之間的關係。共有三十個正常成人參與捲舌音判斷的實驗，實驗包含兩個作業，第一個作業使用具有三種捲舌程度的單音節詞為聽覺刺激，包括不捲舌、稍捲舌和較捲舌三類，要求聽者做是否為捲舌音的判斷。第二個作業刺激材料為雙音節詞，聽者根據第一音節的聽知覺捲舌性做判斷。單音節方面的結果發現，若稍捲舌音不計，聽者的總正確率為78%。對三種捲舌程度刺激的反應時間比較，對最捲舌刺激的反應時間顯著最短；而稍捲舌和不捲舌之間則無顯著差異。根據聽者的反應定義刺激的捲舌對比類別，結果顯示捲舌音類比起非捲舌音類刺激有顯著較低的M1，較高的M2，較高的M3，和較低的M4，這些趨勢和之前的捲舌音製造聲學分析研究（鄭靜宜，2006）的資料相吻合。四個頻譜動差參數中捲舌音反應百分比和M1的相關值最高，達-.61 ($p < .001$)。在作業二中使用雙音節刺激來檢視詞彙脈絡效果，結果發現聽者的捲舌音判斷既非完全是根據詞彙線索，也非單就聲學線索，而是兩者因素的綜合，而聲學線索中第一個頻譜動差參數（M1）在區分華語雙音節詞捲舌音對比仍為有效的參數。

關鍵詞：M1、動差分析、捲舌音、頻譜、聽覺辨識

前言

華語的捲舌音（retroflexes），又稱為翹舌音，具有獨特的聽覺和構音的特性。世界語言中也有一些語言中有捲舌音類別，如印度語、波蘭語和俄語，但這些統稱為「捲舌音」類的語音卻有著不盡相同的構音和聽覺特性。華語中有三對摩擦音和塞擦音的捲

舌子音對比：ㄗ與ㄛ的對比、ㄗ與ㄘ的對比以及ㄗ與ㄌ的對比。這些對比音之間在聽知覺方面很容易互相混淆。構音方面也較為困難。華語捲舌音屬於舌尖後音，構音時需將舌尖往上翹起或捲起靠近齒槽後部或上硬顎前部的位置，使舌尖或舌尖背與上硬顎形成壓縮的通道，肺部氣流通過時受擠壓造成高頻摩擦的噪音，具有特殊的聲學和聽覺特質。捲舌

致謝

感謝國科會專案計畫的支持，計畫編號NSC93-2413-H-024-020。

音對比的差異主要在構音位置，捲舌音的構音位置較為後方。目前IPA（international phonetic alphabet, International Phonetic Association, 2005）界定捲舌音構音位置為舌尖位置介於齒槽緣（alveolar ridge）後方和硬顎前方，亦即介於在發齒槽音和硬顎音之間的前硬顎範圍。在此範圍中舌尖位置愈往後翹起接近硬顎，舌尖背接近上顎的靠近面積愈大，則捲舌的程度愈大，發出捲舌音的特徵也就愈強。然而，華語捲舌音在構音時舌尖位置相對於上硬顎的位置有著相當大的變異性，只要是在這個構音位置範圍發出來的音，不管實際上舌尖是否真有捲曲動作皆被視為捲舌音。Ladefoged與Wu（1984）就曾使用X光造影調查三位說話者的說北京話的摩擦和塞擦捲舌音時的構音動作，發現這三位北京話說話者發出捲舌音時，舌尖並沒有捲起向後以舌背面接近上顎的情況，而是依舊是以舌尖略接近上顎的齒槽緣（alveolar ridge）的位置，而舌根位置則比發 /s/ 時較為後方。事實上，華語捲舌音在構音上存在著個別差異性與地區差異性，例如在台灣所說的華語捲舌音捲舌程度就不及在北京所說普通話的捲舌音的捲舌程度強。華語捲舌音的構音存在有不同程度的「捲舌」，構音時舌頭的位置和捲舌的程度也存在著不小的個別內和個別外的差異性（鄭靜宜，2006；謝國平，1998）。所謂個別內的差異性是有些說話者也會隨著場合、溝通對象或溝通需要而調整說話時的捲舌程度。鄭靜宜（2006）即發現說話者對於同一捲舌音在單音節詞朗讀和機智問答時有不同程度的捲舌產出，在另一個實驗中說話者可以製造出三種程度的捲舌音來。

在聲學方面，華語摩擦和塞擦捲舌音的信號屬於高頻噪音，在頻譜圖呈現為一段高頻的非週期噪音。頻譜尖頂頻率或是摩噪音下限頻率常是用來分析此類信號頻率特性的指標。例如翁秀民、沈牧璋及郭德惠（2006）分析華語捲舌音對比的摩擦音段頻譜尖頂頻率，發現捲舌音的尖頂頻率低於非捲舌音的頻率。謝國平（1998）調查臺灣年輕人在閱讀短文時的捲舌音對比的聲學特徵，測量擦噪音（noise frication）強度集中區的下限頻率，發現女性說話者擦噪音下限頻率在捲舌音與非捲舌音間有顯著差異，捲舌音下限頻率較齒槽音的為低，但在男性方面卻無差異。摩擦音與送氣塞擦音的捲舌音對比在擦噪音下限頻率有顯著差異，但不送氣塞擦音的捲舌音對比則無差異。他的結論是就擦噪音下限頻率此變項而言，臺灣年輕人說話時除了在ㄗ與ㄘ對比無區辨性外，而其餘對比（ㄋ與ㄌ對比、ㄖ與ㄣ對比）之間則仍維持著聲學上的差異。由於摩擦噪音在時間向度比起其他子音通常延續有一段較長的時間，而噪音的能量分佈在時間和頻率

向度上常呈現動態性的變化。尤其是塞擦音在噪音起始位置和終點的噪音能量分佈型態常呈現極大差異性。以上這些測量若沒有根據整段噪音的平均頻譜特性來分析，會出現測量定點取樣的偏誤性，因為噪音音段時長中時間上的任一點取樣皆無法代表此段噪音的聲學特質。因此，若能有聲學參數可整合一整段噪音音段的特性即能較客觀地代表此語音的聲學特性，而對於噪音音段進行頻譜動差分析即能對一整段噪音音段的特性做時間和頻率向度的整合。

頻譜動差分析

摩擦與塞擦捲舌音的聲學性質屬於高頻噪音，在聲音能量在頻率上有著特殊的分佈型態。頻譜動差分析（moments analysis）是將一段噪音音段頻譜特徵量化的方式之一，可被用來分析語音中非週期波的頻譜特性。動差分析是將摩擦噪音的FFT（fast Fourier transformation）量能頻譜（power spectrum）曲線做統計的動差分析，亦即將噪音能量的頻率分佈視為一種統計機率分配（Kent & Read, 2002），經由演算得到四個動差參數：一級動差為平均數（mean）；二級動差為變異數（variance）；三級動差代表偏態（skewness）；四級動差代表峰度（kurtosis）。動差分析是將語音音段的時間斷面頻譜視為一種統計的分布圖，動差參數描述頻譜能量分配的情形。Kent與Read（2002）提出頻譜的動差參數中M1可代表頻譜的重心（central gravity of spectrum），M2則代表頻譜能量的分散性（diversity），M3代表頻譜的傾斜（spectral tilt），而M4代表著頻譜的頂峰尖聳性。四個頻譜動差參數再加上音段時長即可對語音的噪音音段有較完整的描述。

目前動差分析已廣被使用來分析具有非週期波性質的語音（子音），例如Forrest、Weismer、Milenkovic及Dougall（1988）使用頻譜動差分析在英語詞首無聲塞音與擦音，結果發現使用動差參數可區分出三個不同構音部位的無聲塞音，正確率達80%以上。Jongman、Wayland及Wong（2000）使用頻譜動差分析不同構音部位的摩擦音，區辨分析結果顯示使用頻譜動差參數得到相當高的區分正確百分比。Shriberg等人（2003）測量患中耳炎言語遲緩兒童，發現他們的齒槽塞音和嘶擦噪音的M1較正常兒童為低，認為這和他們的構音後置化（顎音化）有關，並導致語音清晰度降低，認為動差分析參數可作為言語遲緩兒後置音化的診斷聲學指標，具有敏感偵測性。對於區分正常構音或錯誤構音具有敏感性，可成為有效的構音能力指標，具有診斷性（Forrest,

Weismer, Elbert, & Dinnsen, 1994; Forrest, Weismer, Hodge, Dinnsen, & Elbert, 1990; Shriberg et al., 2003) 價值。在華語方面，鄭靜宜（2006）曾使用動差分析比較華語捲舌和非捲舌摩擦音與塞擦音對比，發現捲舌音比非捲舌音具有較低的M1和M4以及較高的M2和M3，亦即捲舌音的頻譜重心頻率較低，能量分佈的頻率範圍較散，能量分佈集中偏於低頻區，能量分佈的峰態較為平坦。反之，非捲舌音的噪音能量較集中於高頻區，能量分佈較集中，能量分佈偏於高頻區，能量集中度較強。圖1是使用TF32（Milenkovic, 2004）分析華語語音「刺」中的非捲舌塞擦音和「翅」中的捲舌塞擦音的擦噪音音段的頻譜圖和頻譜分析舉例，擦噪音音段頻譜分析使用FFT量能頻譜以及動差參數曲線，動差分析所得的動差模式曲線大致吻合地罩在FFT曲線上，可見動差分析對於摩擦噪音的FFT有良好的預測功能。鄭靜宜發現四個動差中，M1是最有效區分捲舌音和非捲舌音的聲學指標，M1值和說話者捲舌程度有較強的負相關，說話時捲舌程度愈大，則M1值愈低。此外，該研究中也發現語音M1值的個別差異性頗大，有些說話者捲舌音與非捲舌音間M1的差距大，對比性極強，而有些說話者則否，目前在台灣的華語捲舌音的構音的確存在著不小的個別差異性。這些研究皆較偏於語音的製造方面，在聽知覺方面，到底四個動差參數是否捲舌音特徵的辨識有相關存在呢？

摩擦音和塞擦音一向是兒童音韻／構音能力發展過程中最後獲得的語音，而兒童的語音錯誤有極大部分屬於此類語音，此類語音中捲舌音一向被認為是構音最為困難的一類語音，也是一般兒童音素發展進程中最慢出現的語音（王南梅、費珮妮、黃恂、陳靜文，1987）。吳月瑛與邱瑜萱（2004）的行動研究指出國小一年級學生注音符號學習的困難主要並非是符號認讀的問題，而是在捲舌音「ㄘ、ㄙ、ㄞ」跟非捲舌的「ㄐ、ㄑ、ㄒ」聽覺分辨不清，捲舌音對比在分辨上錯誤率偏高所致，捲舌音區辨是小學生學習注音符號的主要困擾之一。事實上，有關華語捲舌音對比區辨的聽知覺研究目前十分稀少。由於捲舌音對比之區分在聽知覺上是屬於較為細微、不明顯的特徵，目前在台灣的所說的華語也漸出現捲舌音對比合流的現象（曹逢甫，2000），一般人在言語製造方面較不重視捲舌的動作；然而，在聽知覺方面是否也失去了對捲舌音特徵區辨的敏感性呢？到底在聽知覺層面上，一般人是否對於這些容易混淆的捲舌音對比有聽辨的困難呢？

本研究的目的主要在探討正常成年人對於捲舌音聽覺辨識的能力，以及根據聽者的反應分析聲學頻譜的動差參數和捲舌音特徵辨識的相關，尋找在聽覺上區分華語捲舌音對比特徵的有效頻譜動差參數，並比較捲舌音在單音節詞和雙音節詞聽知覺區辨的差異，探討高層次詞彙音境脈絡對捲舌音對比區辨的影響。

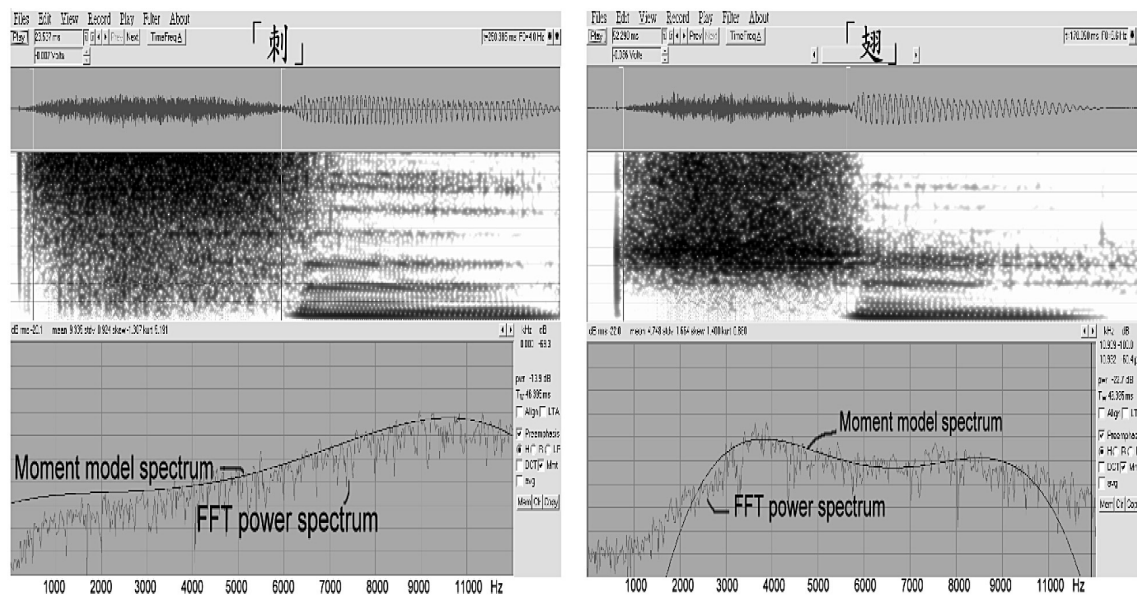


圖1：華語「刺」中的非捲舌塞擦音和「翅」中摩擦噪音音段的FFT量能頻譜與動差參數曲線比較，捲舌音的「翅」頻譜重心較低

方法

(一) 參與者

三十位正常成年人（男、女性各十五名）參與語音聽知覺實驗。這些參與者皆出生與受教育於台灣，能說流暢的華語。這些參與者皆為台南大學大學部的學生，年齡範圍為19至25歲，他們的視力與聽力皆在正常範圍之內，且無已知的神經性疾患。由受試者自我報告以及主試者於互動時觀察，這些參與者皆為視力、聽力與運動功能正常的大學生。

(二) 語音材料

知覺判斷實驗分兩部分，第一部分使用的語音刺激為單音節詞，第二部分為雙音節詞刺激。實驗一所使用的刺激材料單音節刺激項共計有274個，其中24個為練習刺激，正式實驗中有169項為具有不同捲舌音程度的單音節刺激，另有36個單音節刺激為第二部分雙音節詞擷取第一音節而來，用以檢驗詞彙語境效應。另有隨機挑選的45個刺激重複施測作為信度考驗的刺激項。所使用的語音刺激材料皆是來自鄭靜宜（2006）研究中錄音取樣中含有華語三對無聲摩擦音和塞擦音捲舌/非捲舌音對比的單音節語音和雙音節詞語音，語音詞彙材料的設計符合最小音素對比和韻母語音平衡的原則，所用的對比詞彙音節請見附錄1。鄭靜宜研究的實驗二要求說話者於這些詞彙捲舌音和非捲舌音對比之間變化捲舌程度說出四種不同捲

舌程度的語音，包括不捲舌、稍捲舌、中度捲舌和高度捲舌，請詳見該研究之方法部分。本研究使用的刺激選取不捲舌、稍捲舌、中度捲舌類的部分語音，定義為本研究中的三種不同的捲舌程度，包括不捲舌、稍捲舌及較捲舌。169項不同捲舌音程度的單音節刺激中稍捲舌類有102個，不捲舌類的有40個，較捲舌類的有27個。這些刺激的挑選除了根據說者當時捲舌程度和研究者的聽知覺標準外，還參考了刺激的第一動差值（M1），為使得M1介於最高和最低值之間能具有連續變化的數值系列，因此有較多「稍捲舌類」刺激被選用，目的在於尋找聽者捲舌音和非捲舌音知覺上和聲學相對應的分類界線所在。所選用的語音刺激是來自於二十二位台灣的女性說話者，這些語音係經過數位化（取樣頻率22.05kHz），並經過語音聲學的分析，聲學分析使用TF32（Milenkovic, 1999, 2003, 2004）進行語音噪音音段的頻譜分析得到噪音音段平均的四個頻譜動差參數（20ms Hamming window with 10ms step）並測量整體音節的時長。表1列出三類刺激項噪音音段的四個頻譜動差值以及音節時長的平均數和標準差。

第二部分使用華語雙音節詞語刺激，這些刺激詞彙的第一個音節都含有捲舌音對比，如社會、色彩等，所使用的雙音節刺激詞語列於附錄2。這些刺激是選自於鄭靜宜（2006）研究中的實驗一詞語朗讀錄得的捲舌對比詞語語音，這些刺激是來自於十三位台灣的女性說話者。選取的標準主要是根據第一動差參數（M1）的值。對雙音節詞彙刺激的選擇，為了區分詞彙與聲學因素對聽者雙音節詞捲舌音判斷的影

表1 三類刺激項噪音音段的四個頻譜動差和時長的平均值和標準差

語音類別		M1	M2	M3	M4	時長（ms）
不捲舌	Mean	7.956	1.56	-0.54	1.88	456
	SD	1.520	0.56	0.99	2.81	85.14
稍捲舌	Mean	6.398	2.03	0.32	-0.63	467
	SD	0.763	0.34	0.65	0.87	82.64
較捲舌	Mean	5.666	2.15	0.54	-0.02	448
	SD	1.020	0.34	0.99	1.32	76.66

響，原先在選擇上盡量保持兩群刺激（捲舌／非捲舌）的M1值相當，變異數分析的結果顯示所選取刺激中捲舌音和非捲舌音詞彙的第一音節噪音段的第一動差參數並無顯著差異， $F(1, 61) = .21, p = .65$ 。非捲舌詞語M1的平均值為6.386，而捲舌詞語的M1平均值為6.491，兩者M1相近。所使用的雙音節刺激共有82個，其中12個為練習刺激，正式測試使用的語音刺激有70個，其中包括有七組成對的詞彙音境合成刺激。

七組成對的詞彙音境合成刺激是將原本雙音節詞彙的第二音節剪接改變，使其在詞彙捲舌性方面發生改變，例如原本為「樹木」的雙音節將第二音節的「木」去掉，再接上原發語者說「速度」時的「度」音，聽起來變成「速度」，但第一音節仍舊是「樹」音。即兩者語音中第一音節的物理刺激是相同的，只是在第二音節接有不同的語音，檢驗詞彙捲舌性的改變是否會影響第一音節捲舌音判斷的反應。語音剪接使用Praat（Boersma & Weenink, 2003）語音聲學處理程式配合頻譜圖顯示，務使聽覺上自然且不突兀。使用的詞彙有初二、樹木、豬腳、粗細、速度及租金。

（三）捲舌音辨識作業

實驗在聽力檢查隔音室中進行，耳機音量調整為適中舒服的程度。耳機型號為Sennheiser HD280。聽知覺辨識實驗分兩部分進行，第一部分為對單音節詞捲舌音的判斷，第二部分為對雙音節詞的捲舌音判斷。使用筆記型電腦透過程式將個別語音刺激隨機依序呈現至聽者的耳機中。聽者所聽到的刺激是由隨機混合的順序組合而成。請受試者逐一判斷聽見的語音是否為捲舌音，並按鍵盤上標示有「捲舌」與「不捲舌」鍵做反應。在指導語中說明此實驗的大致進行過程以及捲舌音和非捲舌音的定義，並告知聽者一旦聽到刺激音時，要以最快的速度判斷並按鍵作答。為避免疲勞因素，作業中各有多次小休息，兩個作業間並有三至五分鐘的休息。在每部分的正式作業前各另有十二個練習嘗試刺激。反應時間定義為由刺激呈現的開始至受試者按鍵之間的時間間隔，由DMDX（Forster & Forster, 2004）程式自動紀錄按鍵與反應時間。反應時間小於100ms以及大於5000ms以上的資料不予分析。

第二部分的雙音節詞捲舌音判斷，聽者需判斷雙音節中的第一個音節是否為捲舌音。在指導語中告知受試者判斷雙音節詞音的第一個音節的捲舌性，並告知受試者需以實際說者製造語音的聽覺捲舌特性來判斷，即是單就聽到的音來判斷是否為捲舌音，而非該詞語是否應該為捲舌音。施測作答方式以及程序皆和

單音節的相同，同樣是在鍵盤上做作是否為捲舌音的判斷。整個實驗的施測過程約需40分鐘左右。

結果

（一）聽者反應的信度考驗

單音節部分的實驗刺激項中混入有45個重複刺激，檢驗三十位受試者對於這45個重複刺激的反應，分析前後兩次反應的一致性作為聽者反應的信度考驗。全體聽者前後兩次判斷為相同的比率介在百分之五十三到百分之九十三的範圍，平均為75.2%。整體上，這些聽者反應的信度在可接受的範圍中，但個別差異性頗大。其中有十一位受試者在百分之八十以上，有九位在百分之七十以下。

（二）對單音節詞的捲舌音判斷

第一個捲舌音判斷作業中的單音節刺激分有三類不同的捲舌程度：不捲舌、稍捲舌及較捲舌等三類，而聽者的判斷反應有兩類：捲舌音反應和非捲舌音反應。表2列出所有聽者對三種捲舌程度的單音節刺激的知覺判斷反應個數和在不同捲舌程度中所佔的百分比。卡方檢驗的結果達顯著水準， $\chi^2(2) = 677.00, p < .001$ ，聽者對三種捲舌程度的單音節刺激的反應有顯著的差異。「稍捲舌」類的語音刺激中61%被判斷為「捲舌」類，可見聽者對於「稍捲舌類」的語音刺激有稍多傾向於判斷為「捲舌類」。可見，雖然「稍捲舌」刺激是構音介於捲舌和非捲舌之間的模糊類（ambiguous）語音，聽者可能做「捲舌音反應」，也可能做「非捲舌音反應」，但其實聽者對於稍捲舌類的語音其實是具敏感性的，說話者的些微捲舌動作在聽覺上較多被聽者偵測認為是捲舌音。

在計算正確率時，由於「稍捲舌類刺激」是介於捲舌音和非捲舌音類，正確性較為模稜兩可的，因此不將「稍捲舌類刺激」列入正確率的計算中。所謂「正確反應」即是對「較捲舌音」刺激需有「捲舌音」反應，而對「不捲舌音」刺激需有「非捲舌音」反應。單音節刺激中，知覺判斷在「較捲舌」和「不捲舌」類別中的正確率分別為83%與73%。稍捲舌類刺激不計，全部聽者對「非捲舌」刺激和「較捲舌」刺激的捲舌判斷正確率為78%，標準差為16.21。比較「較捲舌音」類的錯誤和「不捲舌音」類的錯誤發現有不對稱的情形，「不捲舌音」類刺激有較多被聽成「捲舌」類音的錯誤。另外，若由三十位聽者個別來看，捲舌音對比區辨個別差異頗大，判斷正確率範圍

由最低的48%到100%，有十五位正確率在80%以上，但也有三位正確率在48%與50%之間。

(三) 單音節詞判斷的反應時間

在反應時間方面，整體反應時間平均數為1336ms，標準差為608。圖2顯示聽者對三類不同捲舌程度的單音節語音做捲舌音和非捲舌音判斷的反應時間。整體而言，對於較捲舌音的反應時間最快，對於不捲舌音的反應最慢，對於「稍捲舌」刺激的反應時間是介於其他兩類刺激之間。「正確反應」的反應時間是較快於「錯誤反應」的反應時間，然而對於捲舌性質較為模糊的「稍捲舌」刺激則無所謂的「正確反應」或「錯誤反應」。若以正確反應的反應時間（對稍捲舌音則不管是捲舌音反應或非捲舌反應皆納入計算）為依變項，語音捲舌性類別為自變項，單因子重複量數變異數分析的結果顯示聽者對三類刺激的反應時間達顯著差異， $F(2, 28) = 4.34, p = .018$ 。事後考驗的結果顯示對較捲舌音的反應時間顯著快於不捲舌音和稍捲舌音的反應時間（各為 $p = .025, p = .018$ ），即對於「較捲舌音」的判斷反應時間最快，對於不捲舌音的反應較慢，但對不捲舌音和稍捲舌音的反應時間兩者則未達顯著差異（ $p = .41$ ）。推測可能因為「較捲舌音」刺激的捲舌特徵最為明顯，聽者一旦發現音段中存在此特徵就可做出反應，因此反應速度較快。對於不捲舌音刺激則需要搜尋整段語音，待確定整個音段完全沒有捲舌音特徵才作反應，因此反應速度會較慢。由於稍捲舌刺激位於捲舌音和非捲舌音的分界帶中較為難判斷，判斷時還是需要搜尋整段的語音，因此對於「稍捲舌」刺激的反應較慢

於「較捲舌類」刺激，但和不捲舌音類之間的反應時間則無差異。事實上，對於稍捲舌刺激中聽者判斷為捲舌音的反應時間是顯著快於做出非捲舌音的判斷反應， $\text{paired } t(1, 29) = 5.19, p < .001$ 。推論對於稍捲舌類聽者的判斷仍然是以捲舌特徵偵測的方式反應，而非如知覺的板模（template）（Klatt, 1979）或典型（prototype）理論（Oden & Marssaro, 1978）認為聽者在媒合一些介於類別之間的模糊刺激時，會因缺乏適當的板模或型態可供比對而導致處理困難，造成處理時間的增加。

(四) 單音節捲舌音判斷與動差參數

在捲舌音製造的頻譜動差研究方面，由於之前鄭靜宜（2006）的研究顯示無論是閱讀詞語和問答說話時，捲舌音相較於非捲舌音均有較低的頻譜重心，即捲舌音的M1較非捲舌音的為小。然而，在語音知覺方面，捲舌音聽知覺判斷和刺激音的動差參數值的關係又是如何呢？依據三十位聽者對雙音節刺激所做捲舌音和非捲舌音的判斷，分析這些聽者判斷為捲舌音和非捲舌音類刺激音的動差參數值。重複量數 t 考驗的結果顯示被聽者判斷為非捲舌音的語音刺激的M1顯著大於判斷為捲舌音的M1， $\text{paired } t(29) = 8.63, p < .001$ 。在M2方面，判斷為非捲舌音的語音刺激的M2顯著小於判斷為捲舌音的， $\text{paired } t(29) = -9.21, p < .001$ 。在M3方面，判斷為非捲舌音的語音刺激的M3顯著小於判斷為捲舌音的M3， $\text{paired } t(29) = -7.22, p < .001$ 。在M4方面，判斷為非捲舌音的語音刺激的M4顯著大於判斷為捲舌音的M4， $\text{paired } t(29) = 6.52, p < .001$ 。

表2 聽者對三種捲舌程度單音節詞的捲舌音判斷個數和百分比

		說者捲舌的程度			
		不捲舌	稍捲舌	較捲舌	總計
聽知覺判斷	非捲舌	872	1191	136	2199
		(72.7%)	(38.9%)	(16.8%)	(43.4%)
	捲舌	328	1869	674	2871
		(27.3%)	(61.1%)	(83.2%)	(56.6%)
總計	個數	1200	3060	810	5070
		(100%)	(100%)	(100%)	(100%)

註： $\chi^2(2) = 677.00, p < .001$ ；括號內為百分比。

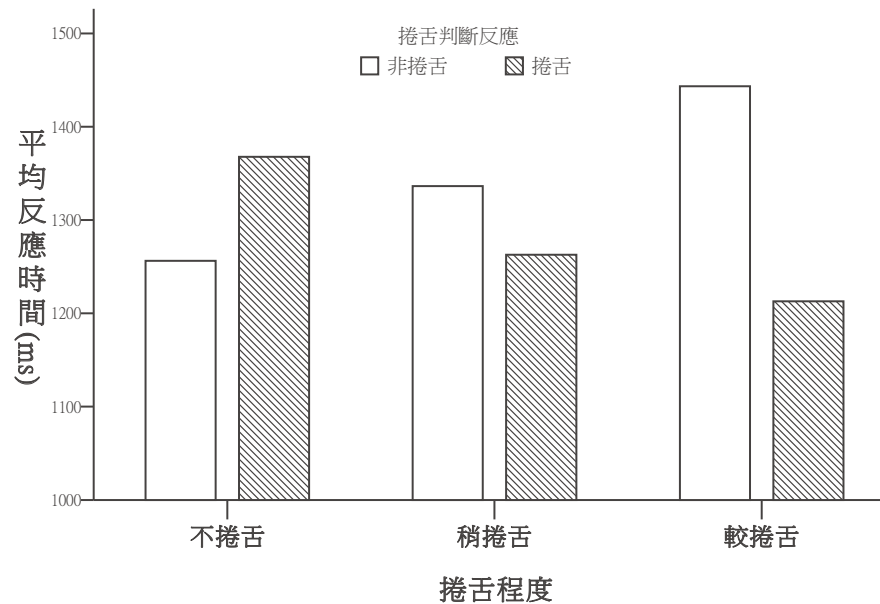


圖2：對於三種不同程度的捲舌音刺激的捲舌音判斷和非捲舌音判斷的平均反應時間(ms)

根據全體三十位聽者對這些單音節語音的知覺判斷，計算各刺激項的捲舌音和非捲舌音的反應百分比值。對於一個刺激項若三十位受試者中若有超過半數回答為非捲舌，則此刺激歸類為「非捲舌音類」。對於一個刺激項若三十位受試者中若有超過半數回答為捲舌，則此刺激歸類為「捲舌音類」。若一個刺激項的捲舌音反應百分比正好是50%則在此暫不列入計算。本實驗中共有三個刺激有50%捲舌音反應百分比，它們的M1值介於6,076Hz與6,665Hz之間。根據受試者的捲舌音知覺判斷來分析這些刺激的四個頻譜動差參數，將結果列於表3。聽者歸類於非捲舌音反應的刺激具有較高的M1、較低的M2、較低的M3以及較高的M4。和鄭靜宜（2006）研究的結果相比較，捲舌和非捲舌音的區別在M1、M2、M3及M4部分皆是相吻合的，即是和非捲舌音相較，捲舌音具有較低的M1和較高的M2、較高的M3和較低的M4，因此語音產生和知覺判斷這兩個實驗所得的結果是一致的。且和閱讀或問答時說話產出的資料（鄭靜宜，2006）相比較，發現在知覺實驗中捲舌音和非捲舌音之間的差距對比性更大，尤其是在M1、M2及M3方面。四個動差參數中和捲舌音判斷的比值間何者相關最強呢？M1和捲舌音反應百分比間相關係數達-0.61 ($p < .001$)，屬於中度負相關，若刺激的M1值愈低，則被歸為捲舌類的反應比率愈高。和其他動差參數的相關則皆為較低，和M2、M3及M4的相關係數各分別為0.33 ($p < .001$)、0.36 ($p < .001$)及-0.28 ($p < .001$)。

圖3呈現聽者對單音節刺激的捲舌音反應百分比和刺激第一動差值 (kHz) 關係，M1愈高回答為捲舌音的比例愈低，M1值和捲舌音的比例成反比趨勢。以捲舌音反應的百分比值為依變項，四個動差值為獨變項的逐步多元回歸分析的結果顯示M1和M3兩變項合起來預測捲舌音反應百分比的多元相關係數為0.604 ($p < .001$)，校正後的決定係數為0.358 ($p < .001$)。而單獨M1變項預測捲舌音反應百分比的多元相關係數為0.542 ($p < .001$)，校正後的決定係數為0.293 ($p < .001$)。使用四個動差值作為捲舌音對比分類的區分性分析結果顯示正確預測率為79.2% (Wilks' Lambda = .616, $p < .001$)。若只以M1為類別區分特徵，捲舌音類別與非捲舌音類別兩類別在第一動差的聽知覺切截點 (50%) 約在7.20kHz左右，M1值在7.20kHz以下較可能會被歸類為捲舌音，而在7.20kHz以上則較會被歸類為非捲舌音。M1為頻譜重心所在，由於非捲舌音的構音位置較捲舌音為前，因此噪音共鳴頻率會較高，因而提升了整個摩擦噪音音段的頻譜重心。

圖4顯示三十位聽者判斷為「非捲舌音」類和「非捲舌音」類的這兩類單音節刺激的平均M1值，可見除了五位聽者（編號12，18，24，25，30）「非捲舌音」類和「非捲舌音」類M1的區分性較小外，大多數的聽者「非捲舌音」類和「非捲舌音」類在M1上皆有不錯的區分對比。事實上，這五位聽者也正是捲舌音對比區分正確率最低的受試者，他們的捲舌音對

表3 根據聽者對捲舌與否兩類別反應的單音節之四個平均動差參數值

判斷反應		M1	M2	M3	M4
非捲舌	Mean	7.469	1.770	-0.238	0.679
	SD	1.333	0.510	0.917	2.513
捲舌	Mean	6.079	2.051	0.338	-0.403
	SD	0.834	0.357	0.779	1.101

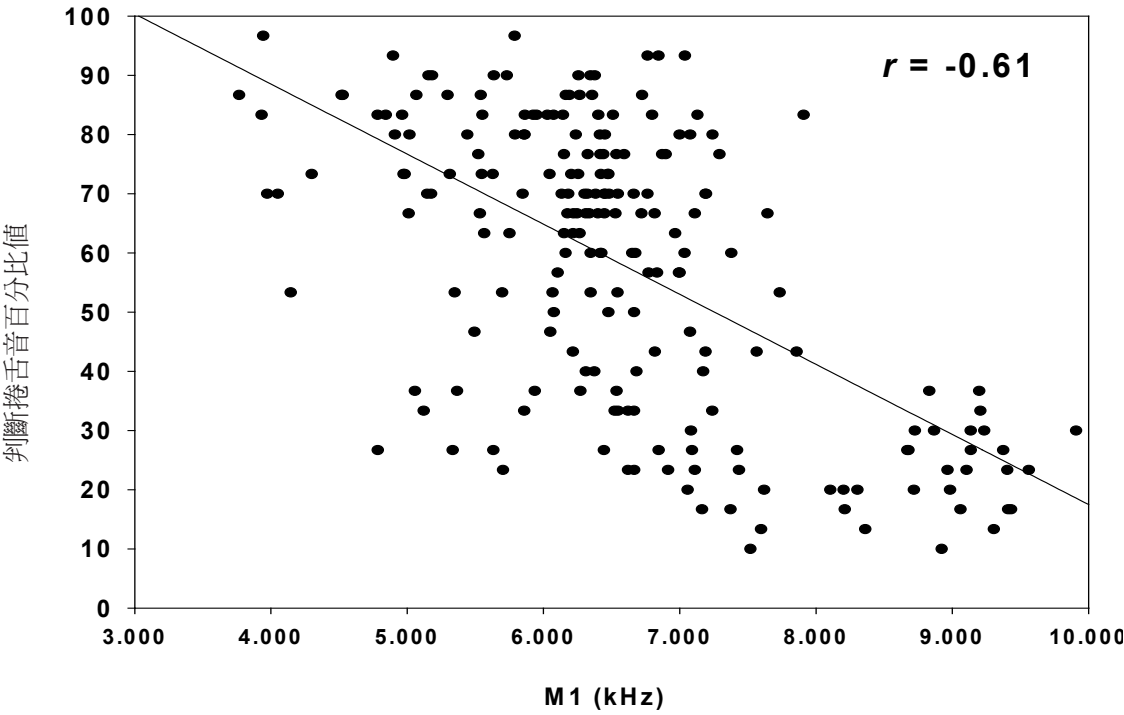


圖3：聽者對個別單音節刺激的捲舌音反應百分比對該刺激噪音段的M1（kHz）值分佈

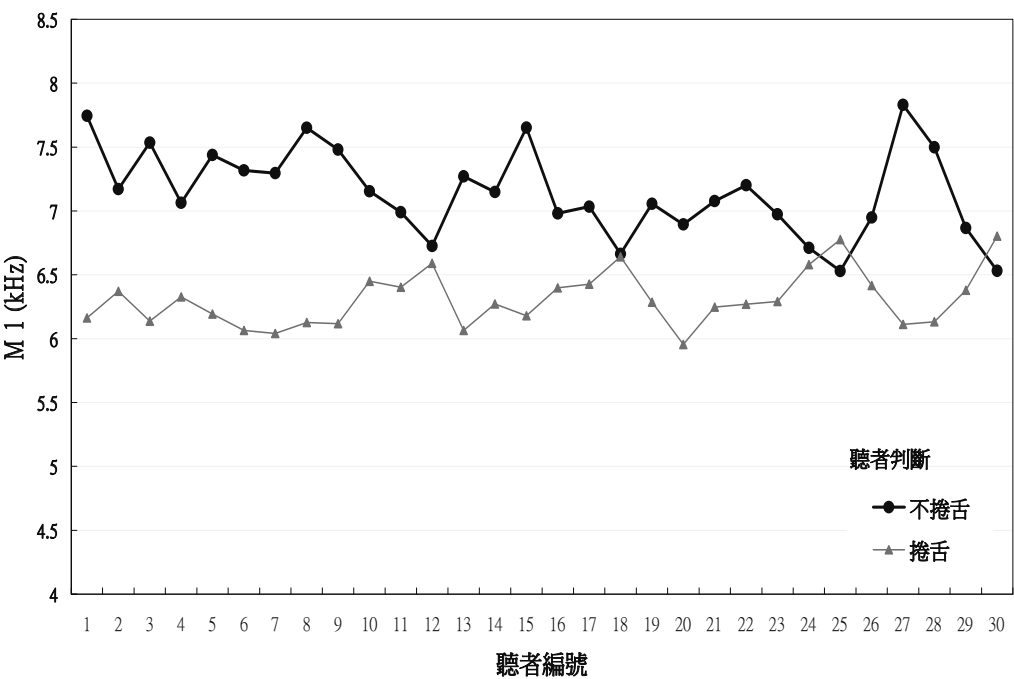


圖4：三十位聽者對單音節刺激的捲舌音對比判斷的刺激M1值（kHz）

表4 聽者對雙音節詞第一音節的捲舌性判斷反應個數以及對全部反應個數的百分比

		詞彙的捲舌性		
		不捲舌	捲舌	總計
聽知覺反應	非捲舌	564	243	807
		(29.8%)	(12.9%)	(42.7%)
	捲舌	306	777	1083
		(16.2%)	(41.1%)	(57.3%)
總計	個數	870	1020	1890
整體	百分比	(46.0%)	(54.0%)	(100%)

註：個數為受試者的反應個數， $\chi^2(1) = 322.65$ ， $p < .001$ ；括號內為百分比。

比辨識正確率平均只有在百分之五十左右。在此推論不使用M1為線索的聽者對捲舌音對比有較差的辨識率，也顯示M1在捲舌音對比的區辨上有著重要的角色。

（五）對雙音節詞的捲舌音判斷

使用雙音節刺激的主要目的在於測試詞彙脈絡

效果。在雙音節詞語方面，除去7個改造的合成刺激後，分析的方法和單音節詞的類似。表4列出依據雙音節第一音節的詞彙捲舌性對聽者聽知覺反應的判斷個數和百分比（在全部反應個數）。卡方檢驗的結果達顯著水準， $\chi^2(1) = 322.65$ ， $p < .001$ ，聽者對兩種第一音節的詞彙捲舌性雙音節刺激的反應有顯著差異。若不計聲學因素，只就雙音節詞彙的捲舌性來定義正確率，全體受試者對於63個雙音節詞語捲舌判斷平均

正確率為71%，標準差為12.39。各聽者間的判斷正確率個別差異頗大，判斷正確率範圍由最低的48%到90%，有十八位在70%以上，有一位低於50%。在此「捲舌判斷正確率」是純粹以詞彙捲舌性來定義的，然而事實上第一音節捲舌性和不捲舌性兩組詞彙在聲學上的M1是不具統計顯著差異（請見刺激材料的描述）。因此，事實上，聽者百分之七十一的判斷正確率乃是聽者受詞彙因素影響的偏移（bias）。然而，因為這比率並非百分之百，可知聽者並非完全根據詞彙捲舌性來判斷音節的捲舌性。

（六）雙音節捲舌音判斷與動差參數

若不管詞彙的捲舌性分類，純就依據聽者對雙音節刺激所做捲舌音和非捲舌音的判斷分類，計算聽者所判斷為捲舌音和非捲舌音類刺激的動差參數值。表5列出聽者對雙音節刺激判斷成捲舌與否兩類別語音的四個動差平均值。受試者內相依樣本 t 考驗的結果顯示判斷為捲舌音的語音刺激的M1顯著小於判斷為非捲舌音的M1，paired t (29) = 3.82， p = .001。在M2方面，判斷為捲舌音的語音刺激的M2顯著小於判斷為非捲舌音的，paired t (29) = 5.19， p < .001。在M3部分，判斷為捲舌音的語音刺激的M3和判斷為非捲舌音的M3差異則未達顯著，paired t (29) = 1.98， p = .06。在M4部分，判斷為捲舌音的語音刺激的M4顯著大於判斷為非捲舌音的M4值，paired t (29) = -3.77， p = .001。這些結果顯示若根據聽者對刺激音捲舌性的分類在動差參數上是有明顯差異的，而原本這些刺激音依據詞彙音韻所做的捲舌對比分類在聲學參數上是不顯著的（見刺激材料的部分）。可推論雖然原刺激的詞語有捲舌音和非捲舌的詞彙音韻性質，聽者的捲舌音判斷就如同指導語所要求的，還是有部分是根據刺激音的聲學捲舌音特徵，而聲學捲舌音顯著特徵即是頻譜動差參數。比較捲舌音和非捲舌音的動差參數，捲舌音反應的詞語具有較低的M1與M2，較高的M4。

比較雙音節和單音節的捲舌音判斷和四個頻譜動差參數的趨勢，則只有在M1是相符合的，在M2、M3及M4部分則有不吻合之處。在單音節詞時，捲舌音比非捲舌音具有較低的M1、較高的M2、較高的M3及較低的M4。但在雙音節詞時捲舌音的M2反而較小，M4則稍較高，在M3捲舌音和非捲舌音之間則無差異。究其原因，主要是在雙音節詞刺激本身聲學參數的對比性原本就較小，影響雙音節詞的捲舌音判斷除了音節本身的聲學線索外尚包含高層次的詞彙效果。整合雙音節和單音節的捲舌音判斷的結果，可見在聽知覺上，捲舌音和和非捲舌音之間的差異並不在M2、M3及M4部分上，而雙音節和單音節比較的結果皆支持M1為捲舌音和非捲舌音差異的主要關鍵參數。

分析對雙音節詞捲舌音反應的百分比值和刺激噪音段的M1的關係，得到對雙音節刺激的捲舌音反應百分比和刺激第一動差值兩者相關係數為-0.25（ p < .05），屬於低顯著相關，和M2的相關也屬於低顯著相關-0.31（ p < .05），其餘和M3以及M4的相關則不顯著，相關係數為各為-0.08（ p > .05）與0.169（ p > .05）。可見比起單音節詞，在雙音節詞知覺實驗中捲舌音和非捲舌音之間的對比性變小了，不過造成此結果可歸因於雙音節詞刺激選取的偏誤，另一原因可能是在雙音節詞語音製造時捲舌音和非捲舌音之間的對比性原本就較小，在聲學動差參數上雙音節詞的捲舌音對比性原本就較單音節詞的為弱，請見鄭靜宜（2006）對雙音節詞捲舌音的分析。

（七）雙音節詞判斷的反應時間

在反應時間方面，全部聽者整體反應時間的平均數為1561ms，標準差為625。圖5顯示聽者對兩類不同捲舌性雙音節詞語所做的捲舌音與否判斷反應的反應時間，可見對捲舌音詞語做捲舌音反應的反應時間是最快的，具有捲舌音特徵的刺激能較快引發捲舌音判斷。聽者做出捲舌音判斷的反應時間（ M = 1365ms，

表5 聽者對雙音節詞刺激的判斷成捲舌與否兩類別語音的四個動差平均值和標準差

判斷反應		M1	M2	M3	M4
非捲舌	Mean	6.5715	2.0879	-.0207	-.3877
	SD	.94358	.37850	.77761	1.25001
捲舌	Mean	6.3473	1.9717	-.0900	-.1872
	SD	.84837	.35469	.81793	1.13226

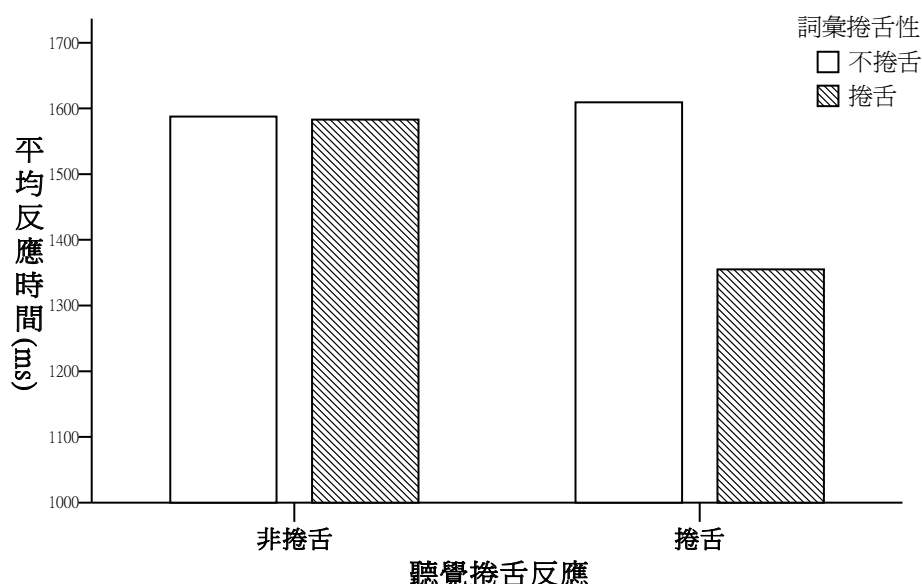


圖5：對於兩類詞彙捲舌性的雙音節刺激的捲舌音判斷和非捲舌音判斷的平均反應時間 (ms)

$SD = 239$) 顯著快於做非捲舌音判斷的反應時間 ($M = 1596\text{ms}$, $SD = 252$), $\text{paired } t(1, 29) = 7.07$, $p < .001$, 這和單音節詞的情形相似。整體上和單音節詞比較, 對於雙音節詞的判斷反應時間大幅地慢於單音節詞的判斷反應時間。推測可能聽者對於雙音節刺激做捲舌音判斷時, 雖然只需對第一音節做判斷, 但仍是傾向於聽完整個詞後再做判斷, 因此對雙音節刺激的反應速度會較慢。

(八) 相同音節在單、雙音節詞中的改變

為了觀察是否同一刺激出現在兩種作業(單音節詞 vs. 雙音節詞)導致捲舌判斷的變化。在第二部分作業的70個雙音節詞語其中有36個詞語的第一音節也納入於第一部分作業單音節刺激中, 36個詞語中依據詞彙的捲舌性, 非捲舌和捲舌類各佔一半。在分析每個刺激組反應的改變, 計算各組捲舌反應百分比的差異絕對值的結果後發現對於這些兩兩配對刺激的捲舌音回答率每對平均有15.57%的改變量, 這比率乃是加入詞彙線索(出現在雙音節)的平均改變量。

由於這些單音節刺激是由雙音節中擷取出來, 若純粹以在雙音節詞彙捲舌性來定義正確率, 表6為對於雙音節詞及單音節詞中的同一單音節刺激的捲舌判斷情形。對刺激音節的捲舌與否判斷反應是以三十位受試者對該刺激的類別反應百分比來分類, 若對

一刺激判斷為捲舌反應若超過50%則歸為捲舌類, 若低於50%則歸為非捲舌類。聽者對單音節詞的捲舌判斷位於顯著邊緣, $\chi^2(1) = 4.5$, $p = .034$, 而對於雙音節詞(第一音節)的捲舌判斷則達顯著差異, $\chi^2(1) = 22.05$, $p < .001$, 可見雙音節的詞彙音境對聽者的捲舌判斷產生了影響。若單純以詞彙的捲舌性為判斷正確與否的標準, 比較兩個作業中對這些具有相同第一音節的反應正確率, 比較兩組資料的結果顯示雙音節刺激較單音節刺激的正确率為高, 單音節詞知覺判斷正確率為66.7%, 而雙音節詞知覺判斷正確率為88.9%。同樣的語音音段在雙音節語境中正確率增加了22.2%, 可說是雙音節詞語境的促進效果。由表6中單、雙音節的比較中也可看到對於原本非捲舌性的音節, 在單音節時被誤判為捲舌的比例偏高, 在雙音節時則此種錯誤較少。因為單音節刺激為雙音節詞刺激中擷取下來的片段為同樣的物理刺激, 而可能由於雙音節詞有語境脈絡的促進而對於捲舌音判斷的正確性有提升作用。

對這些語音的捲舌音判斷反應時間, 對於這些單音節刺激的捲舌音判斷平均反應時間為1251ms ($SD = 98.91$), 對於具同樣第一音節的雙音節刺激的反應時間則較稍慢, 平均為1516ms ($SD = 191.77$), 單、雙音節間反應時間差異達顯著, $\text{paired } t(35) = 9.27$, $p < .001$ 。可能是聽者在聽雙音節時較傾向於聽完整個雙音節刺激後再做反應, 而非只聽到雙音節詞的第一音

表6 對於雙音節詞及單音節詞中的同一單音節刺激的捲舌音判斷個數和對全部反應個數的百分比

	對雙音節詞知覺判斷		對單音節詞知覺判斷		總計
	非捲舌	捲舌	非捲舌	捲舌	
	非捲舌	捲舌	非捲舌	捲舌	
詞彙的捲舌性	15	3	9	9	18
	(42%)	(8%)	(25%)	(25%)	
	捲舌	17	3	15	18
	(3%)	(47%)	(8%)	(42%)	
整體					100%

註：雙音節部分： $\chi^2(1) = 22.05$ ， $p < .001$ ；單音節詞部分： $\chi^2(1) = 4.5$ ， $p = .034$ ；括號內為百分比。

節的聲學刺激就作反應。由於雙音節詞的時長較長，因此雙音節詞反應時間較單音節時的為長。

（九）詞彙音境改變合成刺激

音節詞中有七個語音刺激是合成剪接的詞彙混淆性刺激，在於改變詞語第二音節的語音，第一個音節則保持不變，使其聽起來像是成為另一個詞語。有些是原本應為捲舌的詞語，但被剪接後成非捲舌的詞語；也有些是原本應為不捲舌的詞語，但被剪接後成捲舌的詞語。比較聽者對這些合成刺激的反應可和原來的語音刺激的反應，考驗聽者的捲舌音判斷是否會受到詞語脈絡線索的影響而有偏誤。計算各組捲舌音反應百分比的差異絕對值後，結果顯示對於這些合成刺激的捲舌音回答率每對平均有29.43%的改變量。此比率為受到第二音節詞改變的影響造成的捲舌音的判斷比例的偏誤，此效果比起同一音節在單、雙音節詞中比較得到的雙音節詞促進效果還要高。然而，由於這比例並未超過50%，即就平均而言並沒有大到足以發生捲舌性知覺類別的改變。就反應時間而言，對於這些合成刺激的反應時間平均增加87ms。雖然聽者在作業之前的指導語中被告知只需對第一音節的聲學捲舌性加以判斷，但判斷還是不免會受到第二音節語音影響。

討論

正常聽者對於無上下文脈絡線索的華語單音節刺激的聽覺辨識，由於只有短暫的聲學刺激，辨識率往往不能達到百分之百，尤其對於華語語音中一些聽起來較為相近的對比音，如在捲舌音對比，常有較高的

錯誤率。本研究的結果發現正常成人對於無上下文脈絡的華語單音節刺激的捲舌判斷平均正確率為78%。看似有些低落，其實就華語單音節刺激而言，Jeng（2000）語音清晰度研究中發現對正常成人說話者的華語單音節辨認正確率平均約在90%左右，但其中捲舌音對比部分的錯誤率較高，約25%，而此比率和本研究的捲舌判斷錯誤率頗為相近似。Jeng也發現相較於其他語音對比，捲舌音對比的辨識有著較高的錯誤率，而捲舌音和非捲舌音類的錯誤主要是來自於捲舌音對比之間的相互混淆。可知，正常成人對捲舌音對比的辨識能力是稍遜於其他語音對比。然而，就另一方向來看，正常成人對捲舌音對比的辨識雖較差，但也仍保有對捲舌音對比的基本區辨或辨識的能力。具捲舌音最小音素對比的這些語音由於具有相同的構音特徵，如構音方式皆為摩擦音或是塞擦音，構音部位又極為相近，皆在齒槽緣（alveolar ridge）附近，唯一的差異只在構音位置上，捲舌音只比非捲舌音稍微後方一點，兩類音的對比性和其他類音比較起來較為弱，因此較容易相混淆。在日常生活中，說話者在對比性的維持上往往也呈現著極大的個別差異性，受到如性別、地區性、職業、學習等因素的影響，這些皆會造成捲舌音對比辨識的困難，亦相互混淆。

本研究發現三十位個別受試者中，對捲舌音特徵的區辨存在著一些個別差異。其中有三位聽者對單音節詞的反應正確率不盡理想，正確率在50%左右，而且他們的反應型態和M1的相關極小。或許可以推論這些不採用M1為主要知覺線索的聽者捲舌音對比辨識率較差。然而，由於M1的解釋力只佔有約36%，影響單音節捲舌音和非捲舌音判斷的因素並非只有M1，而還有其他存在於摩擦噪音之內和於摩擦噪音之外的因素，如其他動差參數、噪音頻譜的特徵

性或型態性、頻譜頂峰頻率、噪音下限頻率值等，或是存在於噪音之外的因素，如子音和母音共振峰的過渡帶（transition）。Nowak（2006）進行一位說話者語音聲學分析，發現波蘭捲舌音的頻譜重心頻率（約5600Hz）介於齒音和齒槽顎音之間，而波蘭捲舌音有兩個頻譜頂峰頻率，一個位於頻譜重心附近，另一個則位於較低區域（約3300Hz）。Nowak調查了波蘭捲舌音的辨識影響因素發現聽者對於去掉母音音段的只留下摩擦噪音音段的刺激使得齒槽顎音和捲舌音間混淆性增加，然而造成辨識率的下降主要是齒槽顎音（不捲舌音）的錯誤而非在捲舌音上。同樣地，本研究的結果亦顯示捲舌音對比混淆的錯誤中混淆錯誤的方向也有不對稱的情形，捲舌音被誤聽為非捲舌音的情況較少於非捲舌音被誤聽為捲舌音的情況。就聽覺特徵而言，捲舌音和非捲舌音最小音素對比之間非常相似，唯一的不同只在捲舌音特殊的頻譜特徵，結果顯示對於捲舌音對比的區辨產生較多假警報(false alarm)的情形，即是聽者較容易把「非捲舌音」聽成爲「捲舌音」，而錯過（miss）的情況反而較少，較不會把「捲舌音」聽成爲「非捲舌音」。

到底影響聽者捲舌音對比辨識率的因素爲何？類別化的感知（categorical perception）爲語音知覺的重要現象（Pickett, 1999）。如同其它的語音對比類別一樣，捲舌音和非捲舌音對比間存在著語音類別的區分界線，到底捲舌音對比的區分界線線索爲何？捲舌音和非捲舌音之間的知覺界線是否存有一種所謂的量子（quantal）躍動性質。Stevens（1972，1998）對於「語音產生—語音聲學—語音知覺」三者間的關係提出量子理論。所謂量子的特性是指非線性（nonlinear）和非單調性（immonotonic）的關係，兩變項的對應關係中有時有一種急劇跳躍方式的相應改變關係，但在其他另一些對應範圍內的變化就相對較穩定。Stevens認爲介於口道形狀（vocal tract configurations）的參數與聲學參數，以及介於聲學參數與聽知覺的反應之間有一種量子的特性。即有時一些構音動作上的些微改變卻造成聲學參數的巨大改變，但有時在一段範圍中的一些構音動作的大改變，在信號聲學特性上改變卻很小。同樣地此量子關係也存在於聲學和語音聽知覺之間，在一段聲學參數範圍之內，聽知覺對於聲學參數的改變較爲敏感；然而在此段範圍以外，聽知覺對聲學參數的敏感度就較小。語音知覺具類別感知性質，類別對比之間似存有一界線或閾限（threshold）範圍存在，與類別對比相關的聲學參數在此界線範圍中，若有小量的改變即會產生顯著質的轉變，造成語音知覺上類別的改變，但在此界線範圍外的改變卻無法造成語音知覺類別上的轉

變。就捲舌音對比而言，聽者區分捲舌音和非捲舌音時所憑藉依賴的重要的線索，因本研究結果得到的噪音段頻譜動差參數中的M1和捲舌音反應率間的相關最強，相關係數達-0.61，或許在就M1的知覺軸線上有此一量子躍動的特性，在捲舌音類和非捲舌音類之間的知覺類別界線上存在一個量子急遽變動區域。本研究採用許多所謂的「稍捲舌音」類的刺激，即是欲探求此界線的特性。由於噪音段的四個頻譜動差參數爲統計上歸納噪音音段的能量頻率上分配統整的數值，聲學方面實難以合成操弄。對於對比分類界線的測試只能使用介於捲舌音和非捲舌音的所謂「稍捲舌音」刺激來測試聽者的反應。「稍捲舌類」刺激是在說話者構音時將舌尖置於「捲舌音」和「不捲舌音」類之間的位置。然而「稍捲舌類」產生過程中無疑地，也受到說者的聽覺類別知覺反應的影響，說話者的語音類別中是「全有或全無」，並無「介於中間」類語音存在，「稍捲舌類」會有稍偏於「捲舌類」的情形，而非真的「介於中間」。

對單音節和雙音節刺激的比較可突顯出上下文語境脈絡的影響性。上下文線索可以幫助解除捲舌音對比的模糊性，但也可能造成知覺預期上的偏誤。在判斷單音節詞時，受試者只需要憑著語音聲學線索，判斷的依據較爲單純。在判斷雙音節詞時，由於雙音節詞有加了第二音節的詞，受試者有了詞彙線索，來輔助第一音節捲舌音的判斷。當有多個知覺線索時涉及對各個線索的比較與評估時，對於雙音節詞語，受試者的捲舌音判斷是受音段聲學因素的影響較大呢？還是受詞彙脈絡線索的影響較大？若是詞彙線索和聲學線索相衝突時，受試者的反應會是根據那個線索呢？由本研究聽者對於雙音節詞的反應判斷可知聽者的反應既非完全根據「由上而下」的詞彙線索也非完全是「由下而上的」聲學線索，而是此兩方面線索的綜合。然而哪一方面的線索較有優勢呢？由本研究中一些結果分析可知「由上而下」的詞彙線索會造成捲舌音反應百分比的改變，但改變的量卻不大（每對音組平均約15.57%的改變量），這比率乃是加入詞彙線索（出現在雙音節時）的平均改變量。因此，整體上看來聽者還是較信任由「由下而上的」聽覺聲學線索。或許在不同的捲舌特徵強度的刺激中詞彙脈絡線索可能扮演不同的角色，例如在刺激中若存在充足的聲學線索，則詞彙脈絡線索的影響可能相對會較小。若是如此，那麼使用非詞刺激是否即可去除詞彙脈絡的影響呢？這些問題可留待日後再深入探討。

就反應時間來看，本研究所使用的單音節刺激平均時長爲400多毫秒，而判斷的反應時間（約1300ms）卻長過刺激時長許多，而三類刺激的平均時

長差異（見表1）其實也不大，約在十幾、二十幾毫秒之間。研究結果發現對「捲舌音」的判斷反應時間最快，對於「不捲舌音」的反應較慢，兩者反應時間的差異（約五十毫秒）又較大於刺激時長間的差異。可見這些刺激的處理和捲舌與否的判斷需要花費較多的時間。事實上，和後續相關研究的反應時間比較，也有發現國小兒童的捲舌音判斷反應時間較成人為慢，約1700ms，並且也發現兒童聽者對三類刺激間反應時間的差距，這些差距值又比成人的又大許多（約一百毫秒以上），可見捲舌音判斷反應時間資料可以顯示聽者對於刺激作業的熟練度或是難度。若就對單、雙音節的比較，雖然對於雙音節詞語整體的反應時間較單音節詞為慢，但受試者在實驗後普遍表示對於雙音節詞語的判斷較為簡單。推論受試聽者為何有此感覺的原因可能是因雙音節詞刺激線索較多，不需要對刺激聲學特徵作更精細的處理。但是由於反應時間的計算是由刺激開始播放的瞬間即開始計算，可能因受試者傾向將詞語聽完後再回答，而雙音節詞較單音節詞音長較長，因此對雙音節詞的反應時間較單音節的為長。此時若單純以反應時間的比較來推論作業的難易度則較沒有意義。

本研究發現若只使用M1為捲舌音類別與非捲舌音類別的區分特徵，兩種類別在第一動差的聽知覺切截點（50%）約在7.20kHz左右，M1值在7.20kHz以下較可能會被歸類為捲舌音，而在7.20kHz以上則較會被歸類為非捲舌音，此數值可作為使用頻譜動差分析的自動語音辨識系統設定的參考，但需注意的是捲舌音和非捲舌音分界的聽知覺切截點（50%）在7.20kHz左右的結論只適用對女性語音的知覺，對於男性語音的知覺切截點應該會稍低一點。由鄭靜宜（2006）的研究可知說話者的語音噪音音段的頻譜參數值會受性別因素的影響。男、女性因為口道的形狀、長度不同，共鳴的頻率會有差異。一般而言，女性因口道稍短，語音共鳴頻率會較男性的為高。由於本研究所用的語音刺激取自鄭靜宜的研究，該研究中的實驗二中參與者需製造不同程度的捲舌音，因男性說話者在此方面技巧較難掌握，參與者皆為女性說話者。為求數據單純化，避免性別因素的干擾，本研究的單、雙音節語音材料皆使用女性說話者的語音。此為本研究的限制所在，將來的研究應擴展使用男性說話者的語音為材料，以得到更為完整的捲舌音知覺面向。或是使用接節的方式將噪音後的母音換成男性的語音觀察是否捲舌音判斷的反應率是否會受到語音中母音訊息（如基頻、共振峰）的影響，可對語音的正規化（normalization）議題再深入探討。

由於捲舌音特徵為華語中既存有的真實語音特

徵，而此特徵和其他語音特徵相較在聽知覺區辨上較有難度，使用此語音特徵來評估聽者的語音聽知覺或區辨能力，將較有鑑別力。本研究的結果可作為一種小型的聽知覺常模，以做為一些語音聽知覺異常者的比較基礎，如聽覺障礙、老年失聰、中樞聽覺處理異常（CAPD）或音韻障礙者等，有臨床診斷的意義。事實上，比起其他較為不自然的合成語音的偵測，例如空白偵測（gap detection）或聲音數量計數（tone number counting），捲舌音特徵的偵測是較有功能意義的。另外，本研究發現動差參數M1和捲舌音的聽覺判斷相關最強，自動化語音辨識方面或可使用M1作為捲舌音對比判斷的線索，以提高語音辨識的成功率。其實M1的計算在信號的處理上手續並不繁瑣，是可行性相當高的一個方法。

參考文獻

- 王南梅、費珮妮、黃恂、陳靜文（1987）。〈三歲至六歲學齡前兒童華語語音發展結構〉。《聽語會刊》，1，12-15。
- 吳月瑛、邱瑜萱（2004）。《注音符號補救教學之行動研究》。摘自彰化縣花壇國民小學，<http://ws2.htes.chc.edu.tw/資訊種子/WEB/>
- 翁秀民、沈牧璋、郭德惠（2006）。〈華語捲舌子音及其相對非捲舌子音的比較研究〉。《高雄應用科技大學學報》，35，325-333。
- 曹逢甫（2000）。〈台式日語與台灣國語一百年來在台灣發生的兩個語言接觸的實例〉。《漢學研究》，18，273-297。
- 鄭靜宜（2006）。〈國語捲舌音和非捲舌音的聲學特性〉。《南大人文研究學報》，40，27-48。
- 謝國平（1998）。〈臺灣地區年輕人ㄘㄣ'ㄣ'與ㄗㄣ'ㄣ'真的不分嗎？〉。《華文世界》，90，1-7。
- Boersma, P., & Weenink, D. (2003). *Praat*. From <http://www.fon.hum.uva.nl/praat/>
- Forrest, K., Weismer, G., Elbert, M., & Dinnsen, D. A. (1994). Spectral analysis of target-appropriate /t/ and /k/ produced by phonologically disordered and normally articulating children. *Clinical Linguistics and Phonetics*, 8, 267-281.
- Forrest, K., Weismer, G., Hodge, M., Dinnsen, D. A., & Elbert, M. (1990). Statistical analysis of word-initial /k/ and /t/ produced by normal and phonologically disordered children. *Clinical Linguistics & Phonetics*, 4, 327-340.
- Forrest, K., Weismer, G., Milenkovic, P., & Dougall, R.

- (1988). Statistical analysis of word-initial voiceless obstruents: Preliminary data. *Journal of Acoustical Society of America*, 84, 115-123.
- Forster, K. I., & Forster, J. C. (2004). DMDX, ver 3. Monash University and University of Arizona.
- International Phonetic Association (2005). *International Phonetic Alphabet*. From <http://www.langsci.ucl.ac.uk/ipa/>
- Jeng, J. Y. (2000). *The speech intelligibility and acoustic characteristics of Mandarin speakers with cerebral palsy*. Unpublished Ph.D. dissertation, University of Wisconsin-Madison.
- Jongman, A., Wayland, R., & Wong, S. (2000). Acoustic characteristics of English fricatives. *Journal of the Acoustical Society of America*, 108, 1252-1263.
- Kent, R. D., & Read, C. (2002). *The acoustic analysis of speech*. San Diego: Singular Publishing.
- Klatt, D. H. (1979). Speech perception: A model of acoustic-phonetic analysis and lexical access. *Journal of Phonetics*, 7, 279-312.
- Ladefoged, P., & Wu, Z. J. (1984). Places of articulation: An investigation of Pekingese fricatives and affricates. *Journal of Phonetics*, 2, 267-278.
- Milenkovic, P. (1999). Analysis and synthesis of unvoiced speech using spectrum moments. *Workshop on spectrum moments measures for speech language and hearing research*. ASHA San Francisco.
- Milenkovic, P. (2003). *Moments: Batch speech spectrum moments analysis*. Retrieved from <http://www.medsch.wic.edu/~milenkvc/tools.html>
- Milenkovic, P. (2004). TF32 [Computer Program]. Madison, WI: University of Wisconsin-Madison, Department of Electrical Engineering.
- Nowak, P. M. (2006). The role of vowel transitions and friction noise in the perception of Polish sibilants. *Journal of Phonetics*, 34, 139-152.
- Oden, G. C. & Marssaro, D. W. (1978). Integration of featural information in speech perception. *Psychological Review*, 85, 172-191.
- Pickett, J. M. (1999). *The acoustics of speech communication: Fundamentals, speech perception theory, and technology*. Boston: Allyn and Bacon.
- Shriberg, L. D., Kent, R. D., Karlsson, H. B., McSweeney, J. L., Nadler, C. J., & Brown, R. L. (2003). A diagnostic marker for speech delay associated with otitis media with effusion: Backing of obstruents. *Clinical Linguistics & Phonetics*, 17, 529-548.
- Stevens, K. N. (1972). The quantal nature of speech: Evidence from articulatory-acoustic data. In P. B. Denes & E. E. David Jr. (Eds.), *Human communication: A unified view* (pp. 51-66). New York: McGraw-Hill.
- Stevens, K. N. (1998). *Acoustic phonetics*. Cambridge: MIT Press.

附錄1 單音節詞彙材料

對比編號	捲舌	非捲舌
1	樹	速
2	山	三
3	社	色
4	沙	撒
5	書	蘇
6	收	蒐
7	使	死
8	煞	薩
9	是	四
10	吃	疵
11	插	擦
12	翅	次
13	初	粗
14	徹	測
15	超	操
16	照	造
17	站	贊
18	智	自
19	知	滋
20	紙	紫
21	豬	租
22	皺	奏
23	鏟	慘
24	渣	紮
25	臭	湊
26	哲	責
27	主	組

附錄2 雙音節詞材料

捲舌	非捲舌
初二	粗細
社會	色彩
炸彈	擦洗
山頂	責任
站立	贊成
沙子	操場
獅子	思念
皺紋	湊和
豬腳	租金
樹木	速度
鏟子	慘劇
知識	
吃麵	

註：依照第一音節捲舌性分捲舌和非捲舌兩類

The Auditory Discrimination of Mandarin Retroflex Contrasts and Spectral Moment Analysis

Jing-Yi Jeng

Graduate Institute of Audiology and Speech Pathology, National Kaohsiung Normal University

There are three pairs of retroflex contrasts in Mandarin fricative and affricates, and perceptually, they are very similar and not easy to be distinguished, especially for non-skilled listeners. The purpose of this study is to investigate the auditory perception of retroflexation of Mandarin in adults and investigate the relationship between the spectral moments of frication noise and the auditory discrimination of the fricatives and affricates with retroflexed contrasts. There were thirty adults of native Mandarin speakers participating in the retroflex recognition tasks. There were two tasks of retroflexation judgment in the experiment. In the first one task, the monosyllabic words were used as stimuli and the listeners were asked to judge the retroflexation of Mandarin monosyllabic words. In the second task, the disyllabic words were used, and the retroflexation of the first syllables of words was required to be judged.

In the first task, the monosyllable stimuli had three degrees of retroflexation variations which was consciously made by the speakers, including nonretroflexation, slightly retroflexation, and more retroflexation. The results show that if the slightly retroflexed stimuli were excluded, the correctness percentage of responds was 78% for all listeners. The highly retroflexed consonants had the significantly shortest reaction time (RT), while the RTs between the nonretroflexed and the slightly retroflexed had no significant difference. According to the retroflexness categorical responses of the listeners, the spectral moment analysis of the stimuli showed that the retroflexed consonants had significantly lower M1, higher M2, higher M3, and lower M4, compared to the nonretroflexed consonants. The results are compatible with the data of the previous study of retroflex production (Jeng, 2005). Among the four spectral moments, M1 was the most robust parameter for discriminating the retroflex contrasts. The correlation coefficient between the retroflex response percentages and M1 of the stimuli was -0.61 ($p < .0001$). In the second task, the listeners were asked to judge the retroflexation of the first syllables of the disyllabic words based on their acoustical retroflexation, instead of their lexical retroflexation. However the results show that the retroflexation judgment of the listeners was inevitably influenced by the lexical information of the second syllables of the words. Among the four spectral moments, M1 still the most robust parameter for discriminating the retroflex contrasts. The interaction of the lexical information and acoustical information during the retroflexation judgment of the disyllabic words was discussed.

Keywords: auditory recognition, Mandarin fricatives, moments, spectral retroflex

