

語音單調化對國語詞語和句子的辨識 及語音清晰度的影響

鄭靜宜

國立台南師範學院特殊教育學系

摘 要

國語語音的基本頻率攜帶著大部分的聲調和語調消息，如果基頻起伏變化的訊息消失，是否會對語音的辨識產生干擾，破壞語音的清晰度呢？本研究運用語音再合成將語音刺激中的基頻起伏變化變成一種平坦、無起伏的單調化語音，探討基頻起伏訊息的消失對國語雙音節詞語及句子的詞彙和聲調判斷以及語音清晰度的影響。

實驗一的語音材料為四十五個國語聲調平衡的雙音節詞，依照語音單調化(基頻去起伏)後詞語的性質分為三組，第一組為「去起伏後有一聲對應詞組」，第二組為「去起伏後無一聲對應詞的非音節組」，第三組為「去起伏後無一聲對應詞的非詞組」。受試者對這些語音刺激做聽覺詞彙判斷作業與詞語聽寫作業。結果顯示單調化語音辨識正確率與語音清晰度降低。三類型雙音節詞間在基頻單調化後語音清晰度減少與詞彙判斷作業反應時間的具顯著差異，其中對於「無一聲對應詞的非音節組」的影響相對最小。

實驗二是在於兩個語音單位層次下（雙音節詞與句子）探究無起伏基頻的影響，發現不管是雙音節詞或是句子，單調化語音均具有較低的語音清晰度；且基頻無起伏對雙音節詞和語意新奇的拼湊句的語音清晰度降低量十分相近。而兩種單調化句子間，非音節拼湊句的清晰度顯著高於非詞拼湊句。

由以上兩實驗的結果可知語音單調化對於三類雙音節詞的辨識具有非同質性的影響，而且語音單調化對雙音節詞和句子的語音清晰度損害程度相當。聽者極難辨識單調化的詞彙與句子，基頻起伏訊息的喪失對於國語詞彙與句子的聽覺辨識影響甚大。

關鍵詞：語音單調化、基頻去起伏、語音清晰度、聲調、聽覺詞彙判斷

緒 論

國語是屬於聲調語言 (tone language)，聲調與語調同屬於語音的超音段 (suprasegmental) 性質。國語聲調為語音的超音節屬性。國語的四種基本聲調類型各有其獨特的聲學特性，最主要的區別在於其各自獨特的基頻型態 (F0 contours)。語音的基本頻率 (fundamental frequency, F0)，簡稱「基頻」，為說話者聲帶振動的頻率，基頻為語者語音音波中最基本的頻率，任何有聲 (voiced) 的語音皆攜帶著基頻訊息。基頻乃是語音音波中最基本的頻率，乃是複雜波成分中強度最大的，語音頻率成分最低者。基頻被認為攜帶著大部分聲調的訊息，為聲調辨識的主要線索 (Fu, Zeng, Shannon & Soli, 1998; Tseng, 1990; Howie, 1976)。一般聽者對於基頻的覺知也最為容易，甚至對聽障者而言，基頻的接收也相對地容易，如 Liu, Hsu, 和 Horng (2000) 的研究顯示輕、中度的聽障者仍保留大部分的聲調偵測能力。除了基頻外，各種聲調也在音段時長和聲音強度上有所差異，Whalen 和 Xu (1992) 以及 Fu, Zeng, Shannon 和 Soli (1998) 曾使用語音合成方式使語音刺激的振幅外型 (amplitude contour) 均一化，發現使聲調辨識率略降低為 70%~80%，因此認為音段時長和聲音強度也攜帶著部份聲調的訊息。Tseng (1990) 指出儘管除了基本頻率，尚有時長、強度等音節參數具有分辨各字調的功能，但是基頻 (F0) 所攜帶的聲調訊息最多。因此，在語音聲學特徵上，知覺聲調不同主要在於語音的基頻型態或基頻走勢 (trajectory) 的不同，一旦基頻起伏的型態被破壞，聲調的辨識應會受到相當大的影響。

語音清晰度是指說話者要表達的意思能夠被聽話者瞭解的程度，清晰度可定義為聽話者所得到的正確訊息的比例 (Tikofsky & Tikofsky, 1964)。聽寫法 (transcribing method) 是測量語音清晰度常用的方法。聽話者，即清晰度評分者，需把所聽到的語音寫下來，或是由一組選擇項中選出認為是正確語音的選項，然後計算他們反應的正確率，以百分比表示，此即為語音清晰度分數，代表著語音刺激中能被聽者辨識的比例。所謂的語音辨識即是指個體的聽知覺對於語音刺激的語言性反應，語言性反應是指認出語音中的詞彙與所蘊含的語義。另一種常用來測量清晰度的方法為量尺式估計法 (scaling method)，為心理量尺的估計法，代表心理知覺上語音清晰程度的高低。一般認為使用聽寫法較量尺估計法所得到的語音清晰度較為客觀、實際。

De Bodt, Hernandez-Diaz 和 Van De Heyning (2002) 使用線性回歸模式分析啞吃者的語音清晰度，以四方面的變項來預測語音清晰度：構音、噪音音質、鼻音性與調律 (prosody)，發現語音清晰度和構音向度的相關最高，調律 (prosody) 因素其次，再來是噪音音質，而鼻音性的相關最低。可知對於英語語音，除了構音因素外，調律因素也對語音清晰度有顯著的影響。而對於聲調語言的國語而言，調律因素的影響應比英語更多才對，因為聲調的變化與語意關係更為直接。國語的語音中有許多音具有相同的聲母、韻母的音節，但就是聲調的不同造成意義的差異。構音的準確性無疑地是最主要影響語音清晰度的因素，因此屬於音段性質的聲母、韻母在語音清晰度的角色是無庸置疑的，但是屬於超音段性質的聲調與語調對於語音清晰度的貢獻卻一直被忽略，而屬於聲調語言的國語，聲調、語調訊息在語音清晰度評估的重要性也尚未被深入探究。

詞彙判斷作業 (lexical decision task) 是常用來檢視詞語辨識歷程所用的實驗作業程序。基本的程序是受試者針對刺激的詞語性質做最快的判斷和反應 (通常為按鍵)，受試者快速判斷刺激是否為有意義的詞語，藉由操弄刺激詞彙的語言性相關性質 (如語義、語音、語形等)，檢查反應正確率和反應時間的變化來推論個體的詞彙辨識歷程或語言運作機制的性質。詞彙判斷作

業常被用來研究文字閱讀歷程中的詞語辨識的詞頻效果 (word frequency effect)、促發效果 (priming effect) (楊牧貞和鄭昭明, 民 81; 楊牧貞和方佩君, 民 85; 吳瑞屯和陳欣進, 民 89) 與脈絡效果 (context effect) (如 Meyer, & Schvaneveldt, 1971) 等。在聽覺語音辨識的研究領域中, 詞彙判斷作業也常被使用了解口語語音的辨識歷程與心理詞彙庫 (mental lexicon) 性質 (Goldinger, 1996; Jakimik, Cole & Rudnicky, 1985), 也發現聽覺詞語辨識具有詞頻效果 (Marslen-Wilson, 1990; Slowiaczek & Pisoni, 1986)、鄰近密度效果 (Luce, Pisoni & Goldinger, 1990)、重複促進效果 (Slowiaczek & Pisoni, 1986)、語意促進效果 (Moss, Ostin, Tyler & Marslen-Wilson, 1995)、音韻促進效果 (Goldinger, Luce, Pisoni & Marcario, 1992) 與脈絡效果 (Seidenberg, Waters, Sanders & Langer, 1984) 等性質。在國語詞彙的聽覺辨識方面研究較少, 曾進興、陳振宇和黃國佑 (民 93) 使用聽覺詞彙判斷作業研究國語詞彙的促發效果。本研究除了使用聽寫式清晰度評估來了解語音單調化效果外, 也使用聽覺詞彙判斷作業來推論語音單調化對語音辨識的影響。

由於語音基頻攜帶著聲調和語調的主要消息, 如果基頻起伏、變化的訊息消失, 是否會妨礙我們正常語音的知覺與辨識, 進而妨礙語音的接收, 降低語音的清晰度呢? 基頻在語音辨識中所扮演的角色一直是讓人感到好奇的, Laures 和 Weismer (1998) 以及 Bunton, Kent, Kent 和 Duff (2001) 曾將英語句子的基頻單調化, 探究英語句子語調對語音清晰度的影響, 結果發現基頻去起伏化的句子語調使句子語音清晰度下降。可推知基頻對於並非聲調語言的英語的語音清晰度或語音辨識亦具有一些隱而不見的重要功能。對於屬於聲調語言的國語而言, 基頻起伏型態的破壞是否對語音辨識或語音清晰度的影響更大呢?

基頻在語音辨識或是語音清晰度的所載負的功能可能會隨著語音辨識的單位 (如單音節詞、雙音節詞、片語或句子) 或語音脈絡 (context) 不同而有所差異。這與語音訊號中訊息的贅餘性 (redundancy) 有很大的關係。在缺乏上下文或語音脈絡的情況下 (如單音節詞語音), 聲調無疑地將負有極重要的辨義角色, 因為在聲調語言中聲調的不同將導致詞義的改變。但是在有語音脈絡的情況下 (如多音節詞音或句子), 聲調所負有的辨義角色可能會大打折扣, 因為聽者可以依靠聲調以外的其他訊息得到整個詞或句子的意思。日常生活中所傳遞的語音訊息一般具有高度的贅餘性, 此外, 在多音節詞或句子中, 句子的語調與音節的聲調又有交互作用的關係, 如變調 (tone sandhi) 等因素。在句子語調的規範下, 基頻的變化將不如在單音節中明顯, 因為基頻會受到其前後鄰近音節協同發音 (coarticulation) 的影響 (鄭靜宜 張友馨, 民 92; Xu, 1997; Xu, 1994)。此時, 基本頻率在國語語音辨識中所擔任的角色, 是否會減弱? 是處於可有可無的狀態下? 亦或是仍舊扮演著重要的促進角色呢?

本研究的主要目的在於探討基本頻率在國語語音辨識中的角色, 探究在不同辨識單位 (雙音節詞、句子) 下, 單調化語音對清晰度的影響。本研究所使用的單調化語音, 是透過語音再合成的方式, 將各個語音刺激的聲調起伏去除, 以說話者的平均基頻取代之。本研究的目的是可簡單歸納為以下三點: (1) 調查單調化語音使語音清晰度下降的程度, (2) 比較語音單調化後對三種雙音節詞語辨識 (包括詞彙判斷與語音清晰度) 的影響, (3) 分析並比較兩種不同單位層次下 (雙音節詞、句子) 單調化對國語語音清晰度的影響。

實驗一

方 法

一、參與者

有四位正常成年人參與語音錄音的工作錄音者，包括二位男性、兩位女性，年齡範圍在二十至二十二歲。三十二位正常成年人(男女比例各半)參與聽覺詞彙判斷與詞語語音清晰度評量的工作，他們的年齡範圍由 19 至 40 歲。除了兩位為台南師院的教授外，其餘皆為台南師院大學部學生。這些參與者皆以為以國語為母語，聽力正常、無神經性疾病的正常成人。這些參與者皆為自願參與。

二、詞語材料

語音材料包括四十五個國語雙音節詞，為聲調平衡雙音節詞語。雙音節詞選取是依照詞中兩個音節字調組合原則，第一與第二個音節各有四種聲調（4×4），在去除「一聲接一聲」的組合後，共有十五種聲調組合（16-1=15）。這些雙音節詞語依照語音單調化(基頻去起伏)後詞語的性質分為三組，第一組為「去起伏後有一聲對應詞組」，第二組為「去起伏後無一聲對應詞的非音節組」，第三組為「去起伏後無一聲對應詞的非詞組」，每組各有十五個雙音節詞，第一組詞為「有一聲對應詞組」，第二種與第三組為無一聲對應詞組。第一組詞於單調化後產生語意上的變化，單調化後會被知覺成一聲，有其他不同的意義，例如，「供給」在語音單調化後可能被聽成「公雞」。第二組與第三組的詞於單調化後變成無意義詞語，第二組雙音節詞在於單調化後雙音節中至少有一個音節為非音節，即該音節不曾出現於國語的語音當中，為不合法音節，因此將單調化後的第二組詞組稱為「非音節組」詞語。第三組的詞語於聲調去起伏後變成無意義的兩個音節的串連，在此將單調化後的第三組稱為「非詞組」詞語。這些詞彙的選擇除了需滿足以上所述基頻去起伏後被知覺為一聲時的音節與詞彙狀態外，還顧及詞語的聲調及語音類別的平衡，另外雙字詞的選取也盡量考量詞頻平衡因素。三組雙音節詞語材料列於附錄一。

三、語音錄音程序

對四位說話者的語音錄音工作於隔音室內進行，語音收錄採用 Marantz PMD201 錄音機與 SHURE BG 2.1 型麥克風。提供給受試者的刺激詞提示卡上印有國字以及注音符號，而刺激字的順序以隨機方式出現，請受試者看提示卡，並依造主試者的指示唸出詞音。

四、聲調基頻的分析與基頻去起伏程序

每位說話者所說的四十五個國語雙音節詞皆經過語音再合成(resynthesis)程序處理過，之後每位說話者的語音形成兩套版本：一套為基頻正常的再合成語音，另一套為基頻去起伏的再合成語音，兩套語音皆是再合成語音。語音再合成處理主要使用 32 位元電腦語音卡(Creative sound blaster)與 Cspeech 電腦程式(Milenkovic,1996)，將語音訊號數位化（聲音取樣頻率為 22.05Hz）後，再以 Cspeech 程式中提供的語音分析與再合成功能，先分離出語音中的週期波成分(基頻與諧波成分)與其餘的非週期波的噪音成分，然後再將這些週期波與非週期波相加即為再合成。基頻的去起伏在於改變週期波成分中的基頻與諧波成分，再將語音中原來具有的高頻成分與基頻去起伏週期波成分再合成為單調化語音。而單調化語音的基頻參數則以該說話者的平均基頻取代之。各語者平均基頻的獲得則是以 Cspeech 提供的音頻追蹤 (pitch tracking) 功能進行基頻分析，計算出隨機抽取出每個語者八個語詞的基頻值，再加以平均，而得各語者的平均基頻值。

本研究使用的基頻去起伏與再合成的程序曾被 Laures and Weismer (1998) 及 Bunton 等人(2001)的研究所採用。此程序的特色是以線性預測入碼(linear predictive coding, LPC)為基礎，可同時改變語音基頻成分及其諧波(harmonics)成分，再合成語音除了在基頻及其諧波成份的改變外，此程序並不改變語音的音段性質，如共振峰頻率，且保留原語音的時長及相對音強等特性。基頻去起伏與再合成後所得到的語音聽起來是平板的單調化的語音。為了排除語音再合成程序可能降低語音清晰度的因素 (Massen et al., 1984,1985) 或使音量略減等因素，所有基頻正常的語音項目亦經過 Cspeech 程式的再合成程序處理過，再合成的程序中週期波成分仍維持原狀。

五、聽覺詞彙判斷作業程序

在實驗一中包含兩項作業：詞彙判斷作業與語音清晰度評量。實驗在防音室中以一對一施測方式進行，兩個喇叭的音量調至適中。在詞彙判斷作業中將個別的語音刺激，透過 Cspeech 語音播放計時程式與喇叭隨機呈現，請聽者逐一判斷所聽到的聲音是否為國語中有意義的詞語語音，並以按鍵表示。聽者被告知需要用最快的速度判斷。受試者將被隨機分派去聽四位說話者其中之一的語音，於是每個說話者各有八位聽者。在正式作業前有十個嘗試練習。正式作業中共有九十個嘗試，分為兩個部分，兩個部分中間有三分鐘的休息。聽者所聽到的刺激是由基頻單調化的語音與基頻正常的再合成的語音所隨機混合的順序組合而成。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，由程式自動紀錄。

六、語音清晰度的評量程序

在語音清晰度評量中，語音刺激材料與以上詞彙判斷作業相同，順序則另有隨機性改變，依照基頻的完整性分為兩大部分，第一部分是基頻去起伏的刺激，第二部分是基頻完整的再合成語音刺激，聽者在聽了每個語音刺激後需逐一用國字寫下所聽到的語音，若無法找到國字在用注音符號寫下來。在第一部分單調化刺激聽寫作業中，受試者在指導語中被告知語音經過基頻去起伏程序，受試者需將個別的已去起伏的語音，嘗試著回復為原來有意義的詞彙，並寫下來。

語音清晰度的計分是以音節為單位，計算每個詞的反應正確率，說話者的語音清晰度分數

則為所有詞語的正確反應百分比。所有的同音字皆算正確反應，或以正確注音符號¹標示音節者亦算正確反應。

結果與討論

一、反應正確率

表一列出聽覺詞彙判斷作業反應的正確率。所謂「反應的正確性」是需將基頻完整再合成語音和基頻去起伏語音的第一組刺激視為「有意義詞語」，而將基頻去起伏語音的第二組和第三組刺激視為「無意義詞語」。依此計算，三十二位受試者的詞彙判斷作業的正確率對於基頻完整的語音刺激的總正確率為91.80%，標準差 6.99，而對於基頻去起伏的刺激的總正確率為64.58%，標準差 8.28。基頻完整的語音刺激的總正確率顯著的高於基頻去起伏的刺激 [paired $t(31)=13.22, p<0.001$]，可知受試者在詞彙判斷作業中對於單調化刺激有顯著較多的錯誤，即語音的單調化對聽者的詞彙判斷產生了嚴重的干擾，使正確率下降 27.22%。

無論在基頻去起伏或是基頻完整的語音，第一組詞的正確率均顯著地較低(依次為 $\chi^2=171.84, p<0.05$; $\chi^2=10.20, p<0.05$)，而第二、三組詞的正確率相近。雖然第一組詞單調化後聽起來仍有意義，但聽者的反應中過半數為「無意義」(59%)。而另一方面，即使對於基頻完整的第一組詞，比起其他兩組詞聽者也有較多的「無意義」反應，可見受試者似乎有一種反應的偏向，傾向對於平調的刺激做較多的「無意義」反應，即將單調化刺激認為是「無意義」的語音，可能受試者反應的偏誤所致，對於單調的語音有預期為「無意義」反應。另外，也可能是因完全平板單調的語音聽起來較不自然，事實上正常語音中具有一聲聲調的語音也有小幅度的基頻變動。

表一 基頻完整與基頻去起伏的語音在詞彙判斷作業的正確率比較

刺激	基頻去起伏語音				基頻完整再合成語音			
	詞	非音節	非詞	小計	詞第一組	詞第二組	詞第三組	小計
正確反應數	198	357	375	930	425	449	448	1322
正確率	(41%)	(74%)	(78%)	(65%)	(89%)	(94%)	(93%)	(92%)
錯誤反應數	281	123	105	509	55	31	32	118
錯誤率	(59%)	(26%)	(22%)	(35%)	(11%)	(6%)	(7%)	(8%)

1 以注音符號標示者需聲母、韻母和聲調三者正確才算正確，但為符合現代台灣國語的現況，捲舌音與非捲舌音混淆及ㄅ、ㄆ混淆予以計分，例如ㄅ/ㄆ（如驟/咒）、ㄅ/ㄆ（輕/親）、ㄆ/ㄆ（殊/蘇）、ㄅ/ㄆ（腸/藏）不分，予以計分。

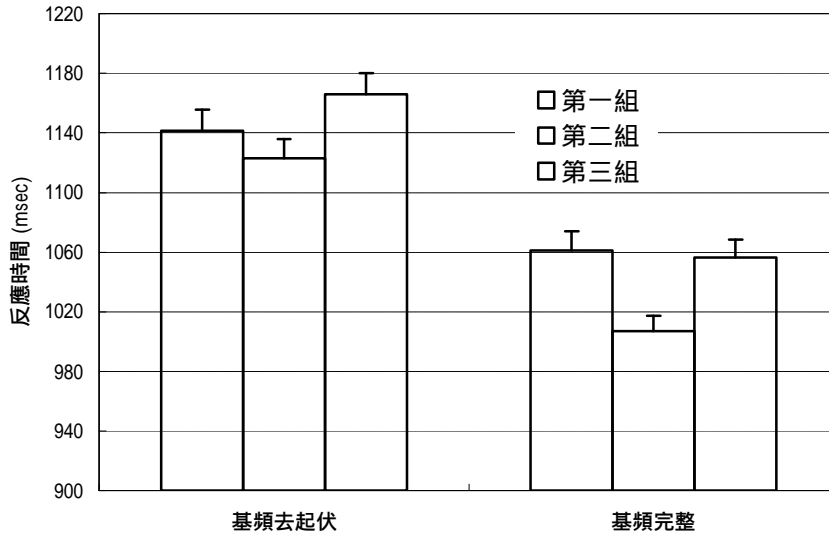
二、聽覺詞彙判斷反應時間

? 覺詞彙判斷反應時間資料的先行處理是將反應時間不足 500ms 與超過 2500ms 的資料²去除後將受試者所有的反應時間列入分析。圖一呈現對於三組語音刺激的基頻完整與基頻去起伏語音的聽者詞彙判斷的平均反應時間。在詞彙判斷作業中，比起基頻正常語音，聽者對於單調化刺激有較長的反應時間。對於基頻去起伏的刺激，聽者的平均反應時間為 1137 毫秒，標準差 296 毫秒，對於基頻完整的再合成語音聽者的平均反應時間為 1037 毫秒，標準差 257 毫秒，平均兩者差距約 100 毫秒之多。

受試者內二因子變異數分析的結果顯示基頻狀態效果（去起伏/完整）與音節詞彙效果（三組刺激間）達顯著 [依次為 $F(1, 413)=102.77, p<.0001, \eta^2=.20, F(2, 412)=8.53, p<.0001, \eta^2=.04$]，但交互作用卻未達顯著 [$F(2, 412)=1.43, p=.24$]。對於基頻去起伏的刺激，聽者的平均反應時間顯著大於基頻完整的再合成語音，也就是失去基頻起伏型態的語音聽者詞彙判斷的反應時間明顯地較長。事後考驗結果顯示第二組刺激（去起伏非音節）的反應時間顯著小於第一組（詞）與第三組刺激（去起伏為非詞）($p<.05$)，即第二組刺激無論在基頻完整或基頻去起伏情況下，受試者反應皆較快。而失去基頻起伏消息的對於三類詞語的影響為何？若比較三組語音刺激在基頻去起伏與基頻完整的狀態反應時間的差距可以發現第一組詞語的平均反應時間差距最小（80ms），而第二組與第三組刺激的反應時間差距較大（依次為 116ms, 110ms），也就是基頻去起伏程序對於第二組與第三組刺激反應時間的增加比起第一組的為多。

雖說第一組刺激於單調化後仍是「有意義的詞彙」，但平均反應時間仍比原刺激平均增加了 80ms，可能是「一聲接一聲」詞的基頻型態在自然的語音基頻仍有些自然的起伏變異性，與單調化語音並不全類似，而基頻的完全水平可能使得語音聽起來較怪異，且較不自然，因而增加了詞彙判斷的時間。事實上，對於第一組詞語而言，單調化的效果並不止於詞彙判斷的反應時間增加和錯誤率提高，更重要的是詞義的轉變。對聽話者而言，單調化的第一組詞語詞義已發生根本的改變，這可由聽寫式的語音清晰度評估得知。

2 檢視全部受試者的反應時間資料，若落在有效範圍外(反應時間不足 500ms 與超過 2500ms)即予以剔除，遭剔除資料佔全部原資料的 3%。反應時間是由刺激撥放的起始點開始計算，由於雙音節刺激較單音節較長，統計平均語音刺激時長 860ms，範圍由 560ms 至 1200ms，因此反應時間不足 500ms 與超過 2500ms 為不合理值，予以剔除。



圖一 三組語音刺激的基頻完整與基頻去起伏語音的平均反應時間

三、雙音節詞語清晰度

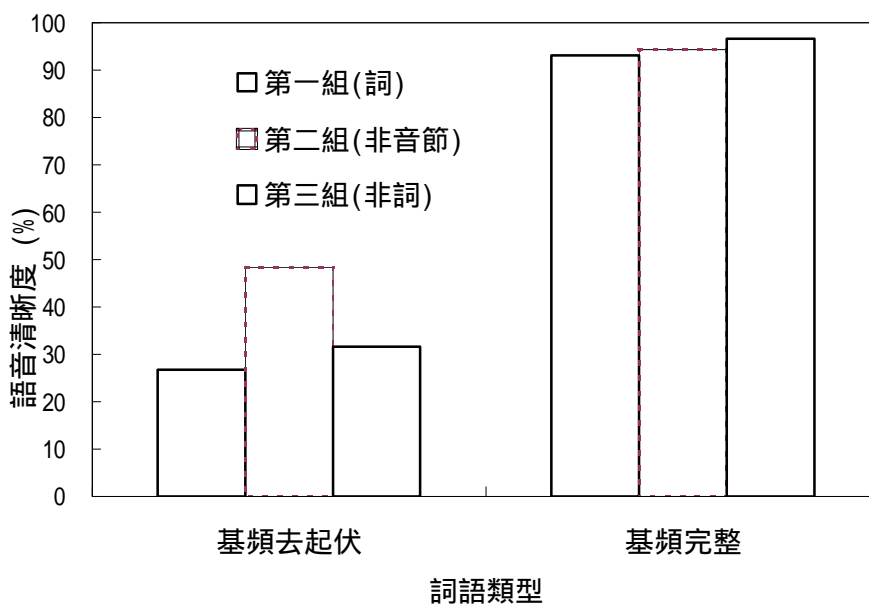
雙音節詞的清晰度評估是採用聽寫法進行，經由三十二位聽者評判後計算平均總正確率。結果對於基頻正常的再合成語音平均正確率為 94.72 %，標準差 4.37。而對於基頻去起伏的語音平均正確率為 35.56 %，標準差 13.26。可見對於相同的詞語，基頻去起伏導致語音清晰度下降了 60%之多。對於基頻去起伏的語音，三十二位聽者的分數範圍由最低 11% 到 58%，可見聽者間個別差異性頗大。圖二呈現三組語音刺激的基頻正常與基頻去起伏語音的平均語音清晰度。很明顯地，基頻正常的再合成語音語音清晰度高於基頻去起伏的語音。基頻正常再合成的語音雖然經過再合成，語音清晰度仍舊十分地高，可見語音的再合成程序並不損害語音清晰度，而單調化再合成語音的語音清晰度卻很低，可見語音清晰度的降低原因不在於語音再合成程序，而是在於語音基頻起伏形態的完整性。

詞語清晰度的受試者內二因子變異數分析的結果顯示基頻狀態效果（去起伏/完整）、音節詞彙效果（三組刺激）以及達交互作用達顯著 [依次為 $F(1, 31)=729.10, p<.0001, \eta^2=.96$; $F(2, 30)=20.47, p<.0001, \eta^2=.58$; $F(2, 30)=23.68, p<.0001, \eta^2=.61$]，基頻去起伏刺激的清晰度顯著大於基頻完整的再合成語音，交互作用達顯著代表著基平去起伏在三種詞彙中有不同效果。而失去基頻起伏消息的對於三類詞語的影響為何？

考驗三組基頻去起伏刺激的差異，受試者內單因子變異數分析的結果顯示基頻狀態效果詞語類別效果（三組刺激）達顯著 [$F(2, 30)=23.68, p<.0001, \eta^2=.612$]，事後考驗結果三組間配對比較皆達顯著，即第二組刺激（去起伏非音節）的清晰度顯著大於第一組（詞）與第三組刺激（去起伏為非詞）($p<.05$)，而其他組間無差異。

在三組基頻去起伏的語音刺激中，第一組基頻去起伏的語音清晰度明顯低於其他二組，因為第一組詞彙有一聲對應詞的混淆之故，即使受試者在實驗前指導語中被告知語音刺激的基頻

去起伏事實，但受試者還是依據一聲的聲調做反應，無法找到原本非一聲的詞語，而產生許多的錯誤。由此可知，單調化後的第一組詞語的知覺已發生根本的改變，無法再由單調化後的語音再取得詞義。相對地，第二組基頻去起伏的語音清晰度明顯高於其他二組，亦即第二組基頻去起伏的非音節語音的詞語辨識相對地較容易。在去起伏情況下，第二組刺激（去起伏非音節）的清晰度顯著大於其他組，然而第二組刺激卻在基頻完整情況下清晰度與其他組間並無差異 ($p>.05$)，可見單調化對第二組刺激語音清晰度的影響明顯較小，而對於第一、三組相對較大。單調化對第一組刺激的影響是詞意的改變，而對於第三組的影響則是語音清晰度的降低。



圖二 三組語音刺激的基頻正常與基頻去起伏語音的平均語音清晰度。

實驗二

方 法

一、參與者

三十二位正常成年人(男女比例各半)參與語音聲調判斷與語音清晰度評量的工作，他們的年齡範圍由 20 至 30 歲。這些參與者皆是以國語為母語，聽力正常、無神經性疾病的正常成人。

這三十二位受試者皆為台南師院大學部學生，且均未參與過實驗一。受試者皆為自願性參與。

二、語音材料

語音材料包括雙音節詞與句子，四十五個國語雙音節詞語與實驗一同，句子有八個。句子的詞語由以上雙音節詞所拼湊成，長度約在七到十個音節左右，語句的語意皆是新奇的、缺少上下文的關聯性，如「床上桃子成熟了」。八個句子中有四個句子是由第二組的雙音節詞所拼湊而成，基頻去起伏後句中至少四個音節成為非音節，是為「非音節拼湊句」，另四個句子由第三組的雙音節詞所拼湊而成，基頻去起伏後句中這些音將成為非詞，即為無意義的雙音節詞，是為「非詞拼湊句」。

實驗二的雙音節語音刺激是沿用實驗一中的語音刺激材料。八個國語句子同樣由實驗一的四位說話者所說。句子的語音再合成與基頻去起伏程序也與實驗一中所述相同，最後產生兩組刺激，一組為基頻正常的再合成句子，另一組為基頻去起伏的再合成句子。

三、聲調辨識作業程序

實驗均在隔音室中進行，兩個喇叭的音量適中。在實驗二中受試者需完成二項作業：聲調辨識作業與語音清晰度評量。聲調辨識作業是要求聽者辨識出語音刺激中的聲調類別。聲調辨識作業又分兩個部分，第一個部分為聲調回復作業，第二個部分為聲調偵測作業。兩項作業各使用不同的語音刺激，在聲調回復的辨識作業使用基頻去起伏的語音刺激，在聲調辨識作業使用基頻正常的再合成語音。在聲調回復作業中，受試者在指導語中被告知基頻去起伏程序，聽者所聽到的刺激是由基頻單調化的語音刺激所隨機混合的順序組合而成。聽者需憑藉著基頻以外的線索辨識出基頻尚未單調化時的詞語，並以最快的速度按鍵回答原來具聲調詞語的第一音節的聲調。受試者需將個別的基頻去起伏語音，回復為原來有意義詞彙的聲調，並按數字鍵表示詞語第一音節的聲調。聽者被告知需要用最快的速度反應。受試者將被隨機分派去聽四位說話者其中之一的語音，每個說話者各有八位聽者。在正式作業前有十個嘗試練習。正式作業中有四十五個嘗試。第二部分為聲調偵測作業，語音刺激為四十五個基頻完整的再合成雙音節詞，作業的情況與以上所述的聲調回復作業相似，聽者須以最快的速度決定雙音節詞語的第一音節的聲調。四十五個雙音節詞順序以隨機化出現，由程式自動紀錄受試者按鍵的反應時間。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，由程式自動紀錄，

四、語音清晰度評量程序

在雙音節語音清晰度評量中，語音刺激材料與以上聲調回復作業相同，順序則另有隨機式的改變，受試者所聽到的皆是單調化的刺激，使用聽寫方式，他們在聽了每個語音刺激後需逐一用國字寫下所聽到的語音。在句子語音清晰度評量中，每一位聽者先聽寫八個基頻去起伏的句子，再聽寫同一說話者基頻完整的再合成的八個句子，句子的順序皆以隨機方式出現。

在雙音節詞的語音清晰度計分部分，所有的同音字皆算正確反應，或以正確注音符號標示亦算正確反應。句子的語音清晰度評量亦以同法得之，句子的語音清晰度的評分原則與詞語的評

分原則相同。句子的清晰度分數以平均每句中寫對的字數的百分比計算，而個別聽者的語音清晰度分數則為其所聽到所有句子正確率的平均值。

結果與討論

一、聲調辨識作業的正確率

受試者在聲調辨識作業中的目標是辨識出語音刺激詞語的第一音節聲調。對於基頻「去起伏」的語音刺激的聲調回復辨識作業的三十二位受試者平均正確率為 35.0%，標準差 6.98，而對於基頻「正常」的再合成語音的聲調辨識作業平均正確率達 93.1%，標準差 6.60。顯示聽者對於基頻去起伏的語音刺激聲調回復的困難度極大，也就是大部份的情況下，聽者無法只藉由雙音節組合的音段線索來辨識詞語，而對於基頻再合成語音的聲調偵測卻顯得十分容易。

表二列出聽者在聲調回復作業（去起伏語音）與聲調偵測作業（基頻正常語音）的聲調反應情形。在聲調偵測作業（基頻正常語音）中聽者反應的正確率十分高，稍多一點的錯誤是「三聲聽成二聲」、「二聲聽成一聲」和「二聲聽成三聲」。三聲聽成二聲的錯誤可能與國語特有的「三聲接三聲」的變調規則有關，即在三聲連接三聲的詞語前面音節的三聲的基頻型態會轉變為二聲的型態，事實上本研究四位錄音者的語音在基頻軌跡的檢查中皆出現三聲變調的現象（請見鄭靜宜、張友馨（民 92）一文即是分析此實驗四位錄音者的基頻共構現象）。然而多數聽者對於三聲接三聲詞語的第一音節聲調反應依舊還是三聲（占 81.25%），聽成二聲的只佔 17.71%，可見三聲變調詞語的表面聲調還是三聲，可能為高層次音韻規則運作之故。

聲調回復作業中聽者反應的錯誤則大多是一聲反應，因為大多數情況下聽者無法將單調化的語音恢復為有聲調變化的詞語，因此聽者對刺激的反應還是為基頻較無起伏的一聲，而四種聲調中四聲類的反應最少，可能是因為原本四聲基頻的變化應是最劇烈且最為明顯的，而單調化刺激的聲調形式與四聲最不像，最不容易激起四聲的知覺，因此反應為四聲的數量最少。對於原本為四聲的刺激則有較多的一聲與二聲反應。

表二 對基頻去起伏與基頻正常的語音刺激聽者反應的聲調混淆矩陣

數量 刺激原聲調	對基頻去起伏語音反應				小計	基頻正常語音反應				小計
	一聲	二聲	三聲	四聲		一聲	二聲	三聲	四聲	
一聲	174	40	36	38	288	272	7	8	1	288
二聲	146	117	60	61	384	19	347	16	2	384
三聲	159	63	124	38	384	3	25	351	5	384
四聲	167	80	48	89	384	5	6	3	370	384
小計	646	300	268	226	1440	299	385	378	378	1440

表三列出聽者在聲調回復作業與聲調偵測作業的聲調反應正確率，第一組詞語為「具有一聲混淆詞」的詞語組，比較三種雙音節語音刺激，較多的錯誤出現於第一組。可能因為第一組

詞語本身在基頻去起伏後聽起來還是詞語，只是意義改變了，第一組刺激在基頻去起伏後由於還是具有詞的地位，聽者對於此類語音較無法將之恢復為其他有聲調的詞，因此正確率較低。對於第二組刺激（去起伏後為非音節），聽者將其聲調恢復的正確率較高（40%）。

表三 對三組基頻去起伏與基頻完整的語音刺激聽者聲調反應的正確率

數量	基頻去起伏語音				基頻完整語音			
	詞	非音節	非詞	小計	詞第一組	詞第二組	詞第三組	小計
錯誤	347 (72%)	286 (60%)	303 (63%)	936 (65%)	34 (7%)	25 (5%)	41(9%)	100 (7%)
正確	133 (28%)	194 (40%)	177 (37%)	504 (35%)	446 (93%)	455 (95%)	439 (91%)	1340 (93%)

表四列出對基頻去起伏與基頻完整的語音刺激，聽者的聲調反應數量。對於基頻完整的語音，聽者的聲調反應數量是平均散佈在各個聲調，而一聲刺激部份由於少了一聲加一聲的組合，因此總數略少。對於單調化的語音，聽者的聲調反應數目散佈在各個聲調極不平均，其中以一聲的反應數量最多，而三組刺激中，以「第一組」（詞）的一聲反應為最多，其次為「非詞」，而「第三組」（「非音節」）的一聲反應最少。

表四 對基頻去起伏與基頻完整的語音刺激聽者的聲調反應數

數量	基頻去起伏語音				基頻完整語音			
	詞	非音節	非詞	小計	詞第一組	詞第二組	詞第三組	小計
一聲	268	180	198	646	97	99	103	299
二聲	79	117	104	300	128	128	129	385
三聲	64	98	106	268	128	128	122	378
四聲	69	85	72	226	127	125	126	378
小計	480	480	480	1440	480	480	480	1440

二、聲調辨識的反應時間

聲調判斷的反應時間資料的先行處理是將反應時間不足 500ms 與超過 5000ms 的資料³去除後將受試者所有的反應時間列入分析。比較聲調回復作業與聲調偵測作業，聽者對於單調化語音在決定聲調時有較長的反應時間。對於基頻去起伏的刺激聽者的平均反應時間為 2030 毫秒，標準差 709 毫秒，對於基頻完整的再合成語音聽者的平均反應時間為 1613 毫秒，標準差 603 毫秒，平均兩者差距約 417 毫秒。受試者內單因子變異數分析的結果顯示基頻狀態效果（去起伏/完整）達顯著[F(1, 1379)=300.63, $p < .0001$, $\eta^2 = .179$]，對於基頻去起伏的刺激，聽者的平均反應時間顯著大於基頻完整的再合成語音；也就是對於失去基頻起伏型態的語音聲調判斷的反應時間明顯地長於基頻完整的再合成語音，對於失去基頻起伏型態的語音聽者難於判別其原來的聲

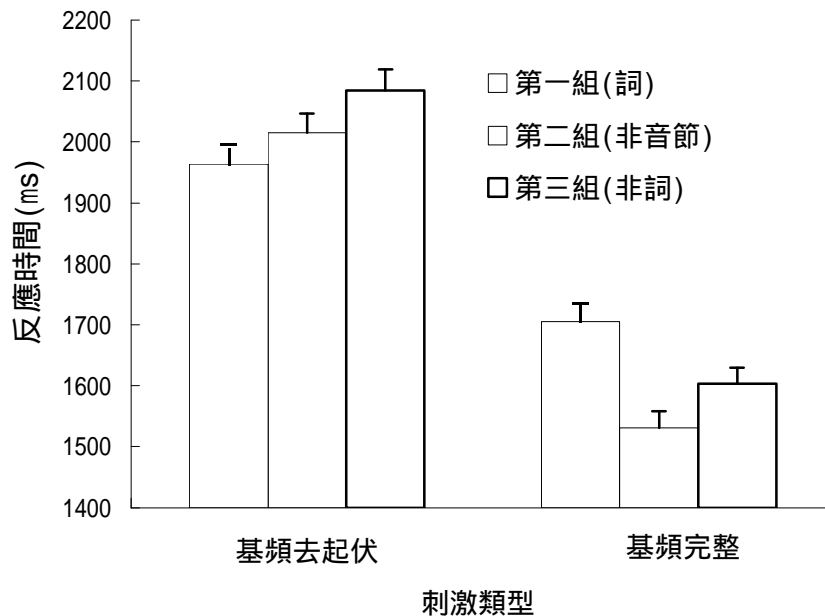
3 反應時間是由刺激播放起始點開始計算。而全部反應時間資料因落在範圍外(反應時間不足 500ms 與超過 5000ms 的資料去除)遭去除的比率為 2.2%。

調。圖三是聽者對於三組刺激在聲調回復作業與聲調偵測作業的反應時間比較。

在對於基頻去起伏刺激的聲調回復作業中，對於三組詞語反應時間比較，對於第三組（非詞）的反應時間較長（平均反應時間為 2071 毫秒），而對於第一組（詞）有較短的反應時間（平均反應時間為 1942 毫秒），對第二組的反應時間（平均反應時間為 1996 毫秒）則介於兩者之間。受試者內變異數分析的結果顯示詞型效果達顯著差異 $[F(2, 438)=5.09, p=.007, \eta^2=.023]$ 。事後考驗結果顯示第一組刺激的反應時間顯著小於第三組（非詞）（ $p<.05$ ），而其他組間則無差異。

在對於基頻完整刺激的聲調偵測作業中，對於三組基頻去起伏刺激的反應時間比較，聽者對於第一組（詞）有較長的反應時間（平均反應時間為 1707 毫秒），而對於第二組（詞）有較短的反應時間（平均反應時間為 1530 毫秒），對第三組（非詞）的反應時間則介於兩者之間（平均反應時間為 1598 毫秒）受試者內變異數分析的結果顯示詞型效果達顯著差異 $[F(2,464)=11.78, p<.0001, \eta^2=.048]$ 。事後考驗結果顯示第一組刺激的反應時間顯著大於第二組（非音節）與第三組（非詞）的（ $p<.05$ ），而二、三組間無差異。

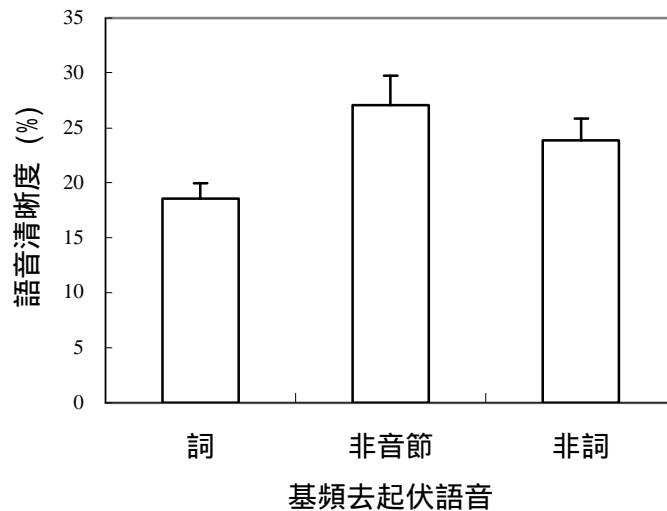
由以上在聲調回復作業與聲調偵測作業反應時間的分析，可知對於相同的語音刺激，基頻的完整性對於三組間反應時間的型態有戲劇化的轉變，尤其對於第一組刺激，原本在對於基頻完整刺激的聲調偵測作業中，第一組刺激的反應時間最慢，但在聲調回復作業中相對地卻是最快，原因可能是第一組刺激在去掉基頻起伏後仍可能被聽成有意義的詞語，而詞彙的意義性對於聲調判斷的反應具有促進效果。



圖三 對於三組刺激聲調回復作業與聲調偵測作業的反應時間。

三、語音清晰度評量

在詞語清晰度方面，因為所有的詞語皆為單調化語音，聽寫的平均正確率偏低，三十二位聽者的平均聽寫式的詞語清晰度分數是 23.16%，標準差 8.96，分數範圍由 8% 到 41%。比較三組詞的詞語清晰度，第一組詞有最低的詞語清晰度，而非音節詞組有最高的詞語清晰度（見圖四），此趨勢與實驗一所得的結果十分類似。受試者內變異數分析的結果顯示詞組效果達顯著 [$F(1, 30)=5.48, p=.009, \eta^2=.267$]，事後比較顯示第一組「詞」與第二組「非音節」間以及第一組「詞」與第三組「非詞」間達顯著 ($p<.05$)，而「非音節」與「非詞」間清晰度分數未達顯著 ($p>.05$)。此型態也與實驗一的結果類似。



圖四 三種基頻去起伏詞組的詞語清晰度比較

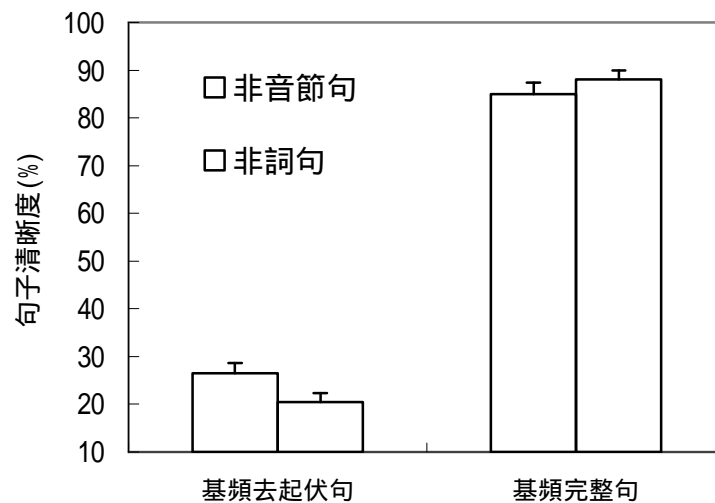
在句子清晰度方面，平均受試者的基頻去起伏句子清晰度為 23.41%，標準差為 23.39，可知基頻去起伏句子的清晰度與基頻去起伏詞的詞語清晰度是十分相近似的。平均受試者的基頻完整句子的清晰度為 86.60%，標準差為 23.39。可見即使語意新奇，語音再合成的拼湊句仍具有高清晰度。基頻完整句子的清晰度則明顯地高於基頻去起伏句子的清晰度，而語音單調化降低了約百分之六十三的句子清晰度。若比較兩種基頻去起伏句子（非音節拼湊句、非詞拼湊句）（請見圖五），非音節拼湊句的清晰度顯著高於非詞拼湊句。

在以基頻完整性和句子種類（非音節拼湊句、非詞拼湊句）為受試者內獨變項，以句子清晰度為依變項的變異數考驗中，基頻完整性效果達顯著 [$F(1, 127)=877.22, p<.0001, \eta^2=.874$]，句子種類效果未達顯著 [$F(1, 127)=.66, p=.42$]，基頻完整性和句子種類效果的交互作用達顯著 [$F(1, 127)=7.20, p=.008, \eta^2=.054$]，可見基頻起伏訊息的消失對於新奇句子的辨識影響很大，基頻去起伏後平均句子清晰度下降了 63% 之多，而句子種類效果未達顯著且兩變項交互作用達顯著可能是因為在基頻完整的正常句子中兩種句子並無差異而中和了句子種類效果。

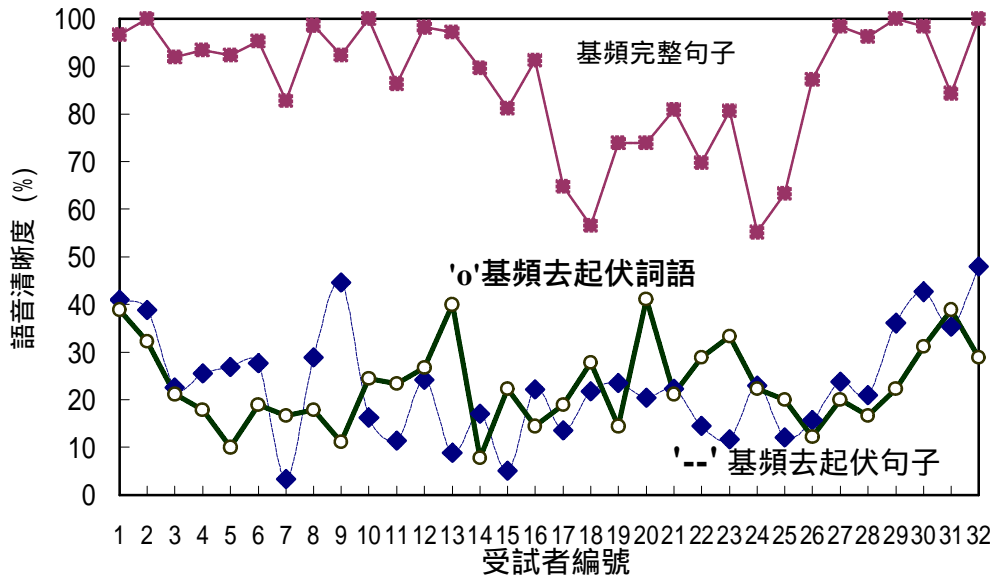
若單純比較單調化句子，對兩種句子（非音節拼湊句、非詞拼湊句）的清晰度的 t 考驗顯示，非音節拼湊句的清晰度顯著高於非詞拼湊句 [$t(127)=2.268, p=.025$]，然而在基頻完整句子方面，

兩種句子間的清晰度卻無顯著差異 [$t(127)=-1.412, p=.16$]。可見，對於基頻完整的句子而言，兩種句子間的清晰度本無差異，但在單調化後，兩種句子的差異性顯現出來。由於非音節拼湊句是由非音節拼湊而成失去基頻後詞語較易復原，因此辨識率較高。事實上，在單調化句子方面，非音節拼湊句的清晰度高於非詞拼湊句的清晰度的情況與詞語刺激中第二組詞（非音節）的語音清晰度高於第三組（非詞）的趨勢十分類似。

若觀察本實驗中三十二位聽者的三種清晰度（基頻去起伏詞語清晰度、句子的清晰度與基頻完整句子的清晰度）的一致性（請見圖六），可以發現大多數受試者的語音清晰度分數頗具一致性。基頻去起伏刺激，不管是句子或是詞語語音清晰度皆未有超過 50% 的，可見所有聽者的語音辨識皆受單調化的影響無法順利辨識出原來的刺激詞語或句子。基頻完整句子方面有四位聽者清晰度低於 70%，在基頻去起伏句子方面有三位低於 10%，可說是變異性較大的個案。若比較受試者的基頻去起伏詞語清晰度、句子的清晰度與基頻完整句子的清晰度三條曲線可發現每位受試者基頻去起伏的句子與基頻去起伏詞語的清晰度分數十分近似，可見對於語意新奇的拼湊句（本實驗所採用的句子）而言，基頻去起伏的降低清晰度效果與雙音節詞語的情況十分類似。就新奇拼湊句而言，由於語意脈絡效果十分有限，聽者無法藉由上下文關係來猜測語意，因此即使為句子，雖具有語法訊息輔助，其基頻去起伏版的清晰度卻與基頻去起伏的雙音節詞語相當。事實上，本實驗所得到的兩者之平均清晰度也十分相近，均在 23% 左右。



圖五 基頻去起伏與基頻完整的句子清晰度比較，非音節拼湊句和非詞拼湊句句子清晰度比較。



圖六 個別受試者的基頻去起伏詞語和句子的清晰度與基頻完整句子清晰度的比較。

綜合討論

本研究想藉由操弄基頻起伏的型態探究基頻和聲調在國語語音知覺中的角色，本研究的結果可簡單歸納為以下三點：

- (1) 不管是在哪一種刺激層面下（雙音節詞、句子），單調化語音導致詞彙判斷和聲調判斷的反應時間增加，錯誤率增加，並嚴重降低語音清晰度。
- (2) 在三種單調化雙音節刺激中，第二組詞語(非音節組)的辨識優於其他二組(非詞組、具一聲對應詞組)。
- (3) 語音單調化對兩種不同單位（雙音節詞、新奇句子）的語音清晰度下降的程度相似。

本研究的結果顯示儘管單調化的語音還保留著音段時長、音強等可能與聲調對比有關的訊息，然而由本研究中幾個作業：詞彙判斷、語音聽寫辨認(清晰度評估)、聲調判斷的作業結果均顯示聽者對於失去基頻起伏型態的語音在辨識上有極大的困難，且在大部分的情況下是無法將他們順利辨識出來的。而且，失去基頻起伏的消息造成詞彙判斷和聲調判斷反應時間變長，語音清晰度大幅下降。因此，基頻起伏的消息對於國語語音的辨識是十分必要的。失去基頻起伏的消息無論對於雙音節詞或是語意新奇的句子均有極大的清晰度損害，且對於雙音節詞與新奇句的清晰度損害程度十分接近。

對於不同辨識單位的比較，研究原先的假設是基頻單調化對雙音節詞的影響會較大，聲調在雙音節詞的辨識中占有舉足輕重的地位，但對於句子刺激的影響可能較小，因為句子有語法結構，且可能攜帶有重複性的脈絡訊息，因此聲調的影響力相對較弱。然而本研究在兩個辨識單位的比較下，兩者單調化的語音清晰度卻得到相近似的結果，兩者清晰度皆在百分之二十三左右，且語音單調化降低語音清晰度的程度在兩者也十分相近，實驗一中基頻去起伏將雙音節詞語的清晰度下降了百分之六十之多，而實驗二中基頻去起伏也降低了約百分之六十三的句子清晰度。由於本研究使用的句子為語意新奇的拼湊句，這些句子缺乏有上下文相關連的脈絡效果，即使有語法的訊息，語音清晰度仍和詞語相當。可見聽者聽取的單位改變，若無上下文相關的脈絡效果，句子的清晰度並沒有獲得額外的提升，而是和雙音節詞一樣。

在三種類型詞彙的比較方面，詞彙判斷的反應時間似乎較無法反映出三者間的差異，但在清晰度方面卻顯示單調化對三種類型詞彙的影響各有不同。單調化第二組刺激語音清晰度的影響明顯較小，而對於第一、三組相對較大。第二組詞語(非音節組)因為在於單調化後雙音節中至少有一個音節為非音節，即為不合法音節，其聲調為一聲的音段因為不曾出現於國語的語音當中，聽者較容易抑制掉「一聲」的反應，而尋得具其他聲調的詞彙，將他們恢復為有聲調的詞語，例如「ㄍㄨ 渦」恢復為「給我」，「ㄇㄨ ㄇㄨ ㄨㄨ ㄨㄨ」恢復為「面容」。第三組(非詞組)的詞語於單調化後變成無意義的兩個音節的串聯，此無意義詞語的兩個各別的音節，若知覺為一聲，雖然兩個音節各自的意義並無法組湊成一個國語中有意義的詞彙，在國語語音中各別的單音節還是有意義性的，聽者較難抑制住「一聲」的反應，而尋得具其他聲調的詞彙，因此詞語回復的機率與反應時間大幅下降。而語音的單調化對第一組刺激是產生詞意上的改變，單調化的雙音節詞若知覺為一聲則詞意改變為另一種意義。由於單調化語音既已存有一個詞彙的入口(entry)，即使意識上被告知語音經過基頻去起伏程序的處理，卻難以改變先入為主的詞語知覺，聽者無法抑制「一聲」的詞彙反應，因此語音清晰度大幅下降。

綜而言之，單調化對三組詞與各有不同的影響，對第一組刺激的影響是產生詞意上的改變，可說是「質」的改變，而對於第三組的影響則是語音清晰度的降低，可說是「量」的改變，單調化對於第二組(非音節詞)的影響相對較小。基頻所蘊含的訊息和語意的取得應有直接或間接的重要關係。單調化對於國語語音的影響並非同質性的，也就是聽者對於基頻訊息的依賴性會隨著詞語的性質而改變。後續的研究可加以探討單調化語音對常用的語句的影響，可使用含有不同脈絡程度的句子，觀察單調化對語音清晰度的影響。

比較實驗一與實驗二的反應時間，雖然所用的語音刺激相同，但在實驗二中反應時間卻較長，應與作業性質不同有關。實驗一為詞彙判斷作業，受試者需辨認出詞彙，做刺激的意義度判斷即可，而實驗二為聲調辨識作業，受試者除了需要辨認出詞彙外還需決定詞彙第一音節的聲調，因此聲調辨識作業應是屬於聲調覺知性質。按理聽者須辨認出詞語後再覺知其聲調，做聲調類別性的反應。很明顯地，實驗結果顯示聲調辨識作業的反應時間比詞彙判斷時間反應時間來的長，且平均差距約有 700ms 之多，可見詞彙判斷作業較聲調辨識作業簡單，而聲調的覺知是發生於詞彙辨識之後的聲調類別處理。本研究中兩種聲調辨識作業中，以對於單調語音的聲調回復作業最困難，反應時間也最長。聲調回復作業的刺激是單調化的雙音節詞語音，聽者只有憑藉著雙音節的音段語音線索認出其基頻尚未去起伏時的詞語，並覺知其原來的聲調。換言之，若聽者無法辨認出原來的詞語，就無法回答出其原來的聲調。實驗二的結果顯示聲調回復作業的正確率只有百分之三十五，即大多數的單調化詞語無法被回復，此比率與實驗一的語音清晰度相當(36%)，但高於實驗二的語音清晰度(23%)。

將本研所得語音清晰度的結果與相關的英語研究相比較，在 Laures and Weismer (1998)

研究中將英語句子的基頻單調化，使正常說話者的句子語音清晰度減少(經推算)約百分之二十五左右。在 Bunton 等人(2001)的研究中英語句子的基頻單調化，使啞吃說話者的句子語音清晰度減少(經估計)約在百分之六至百分之二十三左右。本研究基頻單調化使正常說話者的句子和詞語的語音清晰度均下降約百分之六十。可見，基頻單調化對中文語音辨識影響遠較與非聲調語言的英語為大，基頻起伏型態的訊息不管是在國語的詞語或是句子層次，對於語音辨識或是語音的清晰度皆有關鍵性的地位，破壞語音自然的基頻起伏型態將嚴重干擾語音辨識，並大幅降低語音清晰度。

參考書目

中文部分

- 曾進興、陳振宇、黃國佑(民93)：國語口語詞語辨識的新發現，中華民國聽力語言學會92年度學術研討會。
- 鄭靜宜、張友馨(民92)：國語雙字詞之基本頻率(F0)型態與共構特性，中華民國音響學會第十六屆學術研討會論文集，p75-80。
- 楊牧貞、鄭昭明(民81)：中文字彙知識的腦側化，中華心理學刊，34卷，1-7。
- 楊牧貞、方佩君(民85)：字彙知識在字彙觸接歷程中之作用，中華心理學刊，38卷，1-9。
- 吳瑞屯、陳欣進(民89)：字音與字義促發效果之比較，中華心理學刊，42卷，65-86。

英文部分

- Bunton, K, Kent, R., Kent, J., & Duff, J. (2001). The effects of a flattened fundamental frequency contours on sentence intelligibility in speakers with dysarthria. *Journal of Clinical Linguistics & Phonetics*, 15 (3), 181-193.
- De Bodt, M.S., Hernandez-Diaz Huici M.E., Van De Heyning, P. H. (2002) Intelligibility as a linear combination of dimension in dysarthric speech. *Journal of Communication Disorders*, 35, 283-292.
- Fu, Q-J., Zeng, F-G., Shannon, R. V. & Soli, S.D. (1998). Importance of tonal envelope cues in Chinese speech recognition. *Journal of Acoustical Society of America*, 104(1), 505-510.
- Goldinger, S. D. (1996). Auditory Lexical Decision. *Language & Cognitive Processes*, 11(6), 559-569.
- Goldinger, S.D., Luce, P. A., Pisoni, D.B. & Marcario, J.K.(1992) Form-based spoken word recognition: The roles of competition and bias. *Journal of Experimental Psychology: Learning, Memory and Cognition*, 18, 1210-1238.
- Howie, J.M. (1976). *Acoustical studies of Mandarin Vowels and tones*. Cambridge: Cambridge University Press.
- Jakimik, J., Cole, R. A. & Rudnick, A. I. (1985). Sound and spelling in spoken word recognition. *Journal of Memory and Language*, 24, 165-178.

- Laures, J., & Weismer (1999). The effects of a flattened fundamental frequency on intelligibility at the sentence level. *Journal of Speech and Hearing Research*, 42, 1148-1158.
- Liu, T-C, Hsu, C-J & Horng, M-J (2000). Tone detection in Mandarin-speaking hearing-impaired subjects. *Audiology*, 39, 106-109.
- Luce, P.A., Pisoni, D.B. & Goldinger, S.D. (1990). Similarity neighborhoods of spoken words. In G.T.M. Altmann (Ed.), *Cognitive Models of Speech Processing*, 122-147. Cambridge, MA: MIT Press.
- Marslen-Wilson, W.D.(1990). Activation, Competition, and frequency in lexical access. In G.T.M. Altmann (Ed.), *Cognitive Models of Speech Processing*, 148-172. Cambridge, MA: MIT Press.
- Massen, B. & Povel, D. J. (1984). The effect of correcting fundamental frequency on the intelligibility of deaf speech and its interaction with temporal aspects. *Journal of Acoustical Society of America*, 76, 1673-1681.
- Massen, B. & Povel, D. J. (1985). The effect of segmental and suprasegmental corrections on the intelligibility of deaf speech. *Journal of Acoustical Society of America*, 78, 877-886.
- Meyer, D.E. & Schvaneveldt, R.W. (1971). Facilitation in recognizing pairs of words. Evidence of a dependence between retrieval operations. *Journal of Experimental Psychology*, 90, 227-234.
- Milenkovic, P.(1996). *Cspeech Version 4* [computer Program]. Madison, WI: University of Wisconsin Madison, Department of Electrical Engineering.
- Moss, H. M., Ostin, R.K., Tyler, L. K., & Marslen-Wilson, W.D. (1995). Accessing different types of lexical semantic information: Evidence from priming. *Journal of Experimental Psychology: Learning, Memory and Cognition*, 21, 271-281.
- Tikofsky, R. S. & Tikofsky, R. P. (1964). Intelligibility measures of dysarthric speech. *Journal of Speech and Hearing Research*, 7, 325-333.
- Seidenberg, M.S., Waters, G., Sanders, M. & Langer, P. S. (1984). Pre- and postlexical loci of contextual effects on word recognition. *Memory and Cognition*, 12(4), 315-328.
- Slowiaczek, L.M. & Pisoni, D.B. (1986). Effects of phonological similarity on priming in auditory lexical decision. *Memory and Cognition*, 14, 230-237.
- Tseng, C-Y. (1990). *An Acoustic Phonetic Study on Tones in Mandarin Chinese*. Institute of History & Philology Academia Sinica, Special Publications No. 94. Taipei, Taiwan.
- Whalen, D.H. & Xu, Y. (1992). Information for Mandarin tones in the amplitude contour and in brief segments. *Phonetica*, 49, 25-47.
- Weismer, G. & Martin, R. E. (1992). Acoustic and perceptual approaches to the study of intelligibility. In R. D. Kent (Ed.), *Intelligibility in speech disorders: Theory, measurement and management* (pp. 67-118). Amsterdam/ Philadelphia: John Benjamins.
- Xu, Y. (1997). Contextual tonal variations in Mandarin. *Journal of Phonetics*, 25, 61-83.
- Xu, Y. (1994). Production and perception of coarticulated tones. *Journal of Acoustical Society of America*, 94-4, 2240-2252.

投稿日期：92年10月31日

修正日期：93年 3月 9日

接受日期：93年 3月16日

附錄一 實驗所用雙音節詞材料

第一組：有一聲對應詞組

	Tone2 & tone1 : 琴聲--親生	Tone3 & tone1 : 混音--婚姻	Tone4 & tone1 : 化身--花身
Tone1 & tone2 : 班級 - 班機	Tone2 & tone2 : 無疑--巫醫	Tone3 & tone2 : 等級 - 登機	Tone4 & tone2 : 變質--編織
Tone1 & tone3 : 供給 - 公雞	Tone2 & tone3 : 馳騁 - 吃撐	Tone3 & tone3 : 主講--珠江	Tone4 & tone3 : 辦法 - 頒發
Tone1 & tone4 : 書架--輸家	Tone2 & tone4 : 福氣 - 夫妻	Tone3 & tone4 : 苦幹--枯乾	Tone4 & tone4 : 會議--灰衣

第二組詞組：基頻去起伏後為「非音節組」

	Tone2 & tone1 : 來賓 - ㄉㄛ賓	Tone3 & tone1 : 牡丹 - ㄇㄨㄢ單	Tone4 & tone1 : 特殊 - ㄉㄨ書
Tone1 & tone2 : 安寧 - ㄉㄤㄣ	Tone2 & tone2 : 名人 - ㄇㄣ	Tone3 & tone2 : 旅途 - ㄌㄨ	Tone4 & tone2 : 面容 - ㄇㄣㄣ
Tone1 & tone3 : 溫暖 - ㄨㄢㄨㄢ	Tone2 & tone3 : 瓶口 - ㄅㄧㄣ	Tone3 & tone3 : 給我 - ㄍㄟ	Tone4 & tone3 : 拜訪 - ㄅㄞ
Tone1 & tone4 : 輕快 - ㄑㄩㄥ	Tone2 & tone4 : 疼痛 - ㄉㄤ	Tone3 & tone4 : 碗筷 - ㄨㄢ	Tone4 & tone4 : 壞掉 - ㄨㄞ

第三組詞組：基頻去起伏後為「非詞組」

	Tone2 & tone1 : 成婚 - ㄉㄥ	Tone3 & tone1 : 簡單 - ㄉㄢ	Tone4 & tone1 : 正當 - ㄉㄤ
Tone1 & tone2 : 新聞 - ㄉㄤ	Tone2 & tone2 : 及格 - ㄉㄤ	Tone3 & tone2 : 小腸 - ㄉㄤ	Tone4 & tone2 : 課文 - ㄉㄤ
Tone1 & tone3 : 優雅 - ㄉㄤ	Tone2 & tone3 : 王子 - ㄉㄤ	Tone3 & tone3 : 五角 - ㄉㄤ	Tone4 & tone3 : 放榜 - ㄉㄤ
Tone1 & tone4 : 交易 - ㄉㄤ	Tone2 & tone4 : 獨奏 - ㄉㄤ	Tone3 & tone4 : 感動 - ㄉㄤ	Tone4 & tone4 : 病患 - ㄉㄤ

The influence of Monotone on Mandarin Lexical Identification and Speech Intelligibility

Jing-Yi Jeng

Department of Special Education
National Tainan Teachers College, Taiwan

Abstract

In Mandarin Chinese, the information of lexical tone is mainly conveyed by fundamental frequency (*F0*). What would happen, if the *F0* contours were flattened to be a monotone speech? The purpose of this study is to investigate the effect of monotone for Mandarin lexical identification and intelligibility evaluation. All stimuli were derived from re-synthesized speech, and half of them were tone-flattened after *F0*-flattened resynthesized procedures. According to the meaningfulness after *F0* being flattened, the disyllabic stimuli can be divided as three groups: type I, meaningful words, type II, meaningless non-syllables, and type III, disyllabic nonwords. The first section of experiment I is an auditory lexical decision task, and the second section is a transcribing intelligibility task. The results show that the *F0*-flattened stimuli had significantly longer reaction time (RT), greater error rates, and lower intelligibility scores than tone-intact resynthetic stimuli. Among three types of stimuli, type I (word) had lower intelligibility and less RT difference between *F0*-flattened and tone-intact, while Type III (nonword) had longer RT.

In experiment II, the listeners were asked to restore the original tone for the *F0*-flattened stimulus, and detected the tone of the *F0* intact resynthetic stimuli. The second section of experiment II was to transcribe eight sentences, which also had two versions: *F0* flattened and *F0* intact resynthetic sentences. The results show that listeners had great difficulty to restore the original tone information for the *F0*-flattened stimuli. Both the correct rates of tone restoration and intelligibility scores were pretty low. In addition, the RT and correct rates of tone restoration were found different for the three kinds of disyllable stimulus. The intelligibility scores for *F0*-flattened sentences were significantly lower than those for *F0*-intact sentences. The amount of intelligibility declination was also observed differently in two types of *F0* flattened sentences. The intelligibility scores of 'nonsyllable sentences' were higher than those of 'nonword sentences'.

Key words: monotone speech, *F0* flattened, speech intelligibility, fundamental frequency, auditory lexical decision