

不同言語速度下國語韻律的計時等時性

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：07009；初稿收件：2007年4月12日；第一次修正：2007年8月8日；正式接受：2007年8月13日
通訊作者：鄭靜宜 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)

本研究的主要目的在了解國語說話者在語音製造和知覺的韻律計時特性，考驗國語以音節計時的假設。實驗一使用的語音材料包括三十三個片語句組及兩篇短文，有三十位正常的說話者（男女各半）參與錄音實驗。由於言語中時長變項易受到語速影響，為控制語速在實驗一中先使用節拍器調節說話者的語速，再讓說者以同樣的速度說出語句。共有五種語速變化。以頻譜圖分析各說話者在各句中各音段的時長（子音、母音、停頓等）。使用成對變異指數（pairwise variability index, PVI）比較語句中兩種假設性等時單位（全音節單位、母音起始單位）的時長變異性。結果發現母音起始單位的PVI顯著小於以全音節為計時單位的PVI，顯示母音起始單位較具有等時性。以母音為計時單位起點正是所謂的語音知覺中心（perceptual center, P-center）的所在。實驗二使用較自然的語速控制方式，說者先聽語速示範帶後再說出句子。有二十一位正常的說話者（男女各半）參與錄音，並分析各語句音段時長後，同樣地以PVI考驗兩種計時單位的等時性，結果和實驗一相近似。實驗一和實驗二的結果顯示不管使用何種語速控制方式母音起始單位皆顯著地比全音節單位較具有等時性。在實驗三中有十八位受試者評估具有不同PVI句子的自然度，結果發現自然度分數和PVI值呈現高負相關，即PVI值愈大聽起來愈不自然，並發現使用母音起始單位計算的PVI值和自然度分數間的相關高於使用全音節單位PVI和自然度間的相關。無論是在言語的製造和知覺方面，本研究的結果支持母音起始單位為國語韻律計時的等時單位。

關鍵詞：韻律、言語計時、等時性、成對變異指數、知覺中心

緒論

就說話動作而言，說話者在製造語音動作時需要在時間向度上做精確地衡量與控制，各個相關結構間動作互相地協調，才能製造出自然、清晰可辨的語音。尤其是上呼吸道構音機制的動作需要與嗓音產生機制—喉部有協調性的動作合作，讓由喉頭出來的聲

源信號適時地受到構音機制的塑形而成為可被理解的語音信號。因此，就構音以及韻律調節的角度而言，說話者本身對於說話動作的精確計時能力是必須具備的。構音的動作的達成須要有各神經與肌肉協調性的控制，以達成各語音區辨特徵的對比性，如塞音送氣對比的表現需有精確的嗓音起始時間（VOT, voice onset time）控制，而且整體言語速度的控制也需有

致謝

感謝國科會專題計劃的補助，計畫編號 92-2413-H-024-019不同言語速度下國語各音段時長變化及其於連續語言中可預測性之探討，計畫編號 93-2413-H-017-014由語音的產生和知覺探討國語語音的計時特性。



精確的動作計時能力，以滿足言語韻律上速度調控的需要。因此，由說話動作的角度來看語言計時的特性似乎可推論說話者的運動神經機制中有一計時機制，在說話時做精確的動作計時調配，對每個動作的時間精確的掌控，以符合言語中音段（語音類別）和超音段（如說話速度、聲調、語調、強調等）的輸出需求，因此這個計時能力在言語行為應是不可或缺的一環。言語韻律中的等時性（isochrony）（Abercrombie, 1967; Pike, 1945）是假設在連續語音中有一些等時長間隔的基本單位存在，即等時單位，此種單位為言語節奏韻律感的來源。言語等時的單位可能隨著語言種類而不同，有些和重音有關，而有些和音節有關。令人好奇的是國語語流中韻律等時的單位為何？此計時單位的起點和終點又位於何處呢？

（一）言語的計時特性

語調韻律（prosody）屬於言語的超音段（suprasegmental）特性。就世界各語言而言，語音音段的長短在時間向度上的分配並非隨機化，而是如同音樂一樣，具有一些節奏規律性。在語言串流中似乎有一些固定相同的時間間隔存在著，即所謂的等時性。等時的單位可能隨著語言種類而有所不同，例如有些語言聽起來每個音節所分配到的時間相近，每個音節之間的時間間隔似乎等長，如法語、西班牙語和義大利語被認為是以音節計時（syllable timing）的語言。而有些語言則有輕、重音的區別。重音和輕音交替出現於各音節當中，重音節被分配到比較多的時間，而輕音節則較少。每個重音和重音間的間隔，稱為音步（foot），每個音步聽起來似乎長短較為一致，如英語、阿拉伯語被認為是以重音計時（stress timing）的語言。以音節計時的語言輕、重音的區別不明顯。大致而言，世界各語言韻律特性被區分為二類：一是以重音計時的語言，另一為以音節計時的語言。此外，另有日語被發現是使用一種比音節小的單位稱為mora來計時的（Beckman, 1982; Han, 1994; Port, Dalby, & O'Dell, 1986）。日語中mora間的時長相近似，被稱為以mora計時（mora-timing），和以音節計時的特性較近似。Dauer（1983）曾測量五種語言的重音間時距，發現重音間時距的長短並不能區分出以重音計時和音節計時的語言，而且說話者間的個別差異性很大。他認為語言韻律的計時特性只是一種知覺上相對的類別式分類，以重音計時和音節計時語言之間的差異主要在於母音的弱化（vowel reduction）的有無、重音節和輕音節間的差異性以及音節結構的不同。以重音計時的語言在非重音音節有母音弱化的現象，且音節結

構的種類較多樣化，而以音節計時的語言音節結構較整齊，且無母音弱化的現象，音節較為同質性，無明顯重音節和輕音節的區分。重音計時的語言語句中重音、輕音交替出現，通常重音節的時長較長，音量較強，音高較高。

有一些研究者（Classe, 1939; Dauer, 1983; Lehiste 1973; Lehiste 1977; Pointon, 1980; Roach, 1982; Wenk & Wioland, 1982; Van Santen & Shih, 2000）曾尋找聲學上時長測量的言語等時性（isochrony）證據，然而這些研究的結果卻沒有找到音節時長或是重音間時距等長的等時性證據，他們發現各單位間時長的差異性相當大，單位間隔也並不相等。例如在英語中發現只要重音和鄰重音之間的音節數目愈多，重音間時距就愈長，重音和重音間的時長並不相等（Dauer, 1983）。西班牙語一向被認為是以音節計時的語言，Pointon（1980）發現西班牙語音節的時長隨著音節的結構有相當大的變異性，而且音節或重音間等時性都不存在。Wenk與Wioland（1982）也提出否定法語是以音節計時的，他們發現法語音節的時長隨著音節的結構和強調（accented）與否有相當大的差異性，但它不像英語，沒有重音弱化的現象，提出法語是所謂修整計時（trailer-timed）的形式，它的節律架構主要是由片語詞群所組成。由於等時性在各語言無法得到聲學音段時長分析的客觀證據支持，於是有些研究者（例如：Couper-Kuhlen, 1990, 1993; Lehiste, 1977）推論提出語言韻律的等時性只是一種知覺主觀的現象，受限於人類聽知覺上的限制。節律上的等時性究竟是屬於語言製造時言語動作的特性或是語音訊號中存有的特性，亦或只是聽知覺上的主觀現象呢？Lehiste（1977）首先認為等時性來自人類聽知覺上的限制，認為人類聽知覺對時長敏感性並不大，音段間即使在聲學上時長有相當的差異，但人們聽起來還是一樣的，尤其是差異在30毫秒以下的刺激。因此他認為即使在英語重音間的間隔時長、音步的大小有相當大的差異，但聽起來仍然相似。Lehiste（1977）指出以重音計時的英語的等時性只是一種知覺上的現象，重音間時長聽起來似乎相等，但其實並無物理上的基礎。

國語有一字一音節的對應，意即口語中的單位——「音節」和書寫單位——「字」之間有一對一的對應關係。音節被認為是語音知覺和製造的自然單位或基本單位（曾進興，1997；Tseng, Huang, & Jeng, 1996; Tseng & Lee, 2004）。Tseng與Lee（2004）提出國語韻律的階層架構，包括有四種層級：音節、節律詞（prosodic words）、節律片語（prosodic phrases）和呼吸群（breath group），他們使用逐步回歸模式分析兩位說話者（一位快速語速、一位慢速語速）

的語料，發現音節層級對於時長的解釋力最高（約51%~60%），其他層級也有若干解釋力，其中以節律片語（prosodic phrases）對於時長的解釋力較高（約7%~14%），而語速的快慢對於各層級對時長的解釋力也有一些影響，他們推論國語的節律和音節單位有極密切的關係。音節若是國語的計時單位，那麼在聲學上音節是否具有時長等時性呢？

若是實際測量國語語句中各個音節的時長，在連續性的言語串流中各音節的時長是否相等呢？鄭靜宜（2004）曾對音節時長進行分析發現國語子音和母音的固有時長（intrinsic duration）是影響音節時長的重要因素，各類子音和母音各有不同的時長。其中具有摩擦音的音節最長，而帶有不送氣塞音的音節最短，而兩種音節時長的差距最大可達約200毫秒（在最慢速時）。即使是純粹比較在同一個句子中（具有相同的語速）的子音加上母音的純音節時長，這些純音節單位的時長變異性也是相當地大。事實上，音節時長受到許多因素的影響，包括語速、聲母種類、韻母種類、聲調、音節結構、句型和在語句的位置（如語尾）等因素。如在同一語速下，國語音節中帶有鼻韻的音節較長，不帶聲母的音節較短（鄭靜宜，2004），可見音節結構對音節時長的影響性。因此，很明顯地，說國語時語句中純粹的子音加上母音的音節（純音節單位）時長並不具有等時性。

由於語流中純粹的子音加上母音的音節時長變異性過大，不具等時性，分析時長等時性不應只計算分析純粹音節時長或是重音之間的時長，而不將音節（或重音）間的停頓或靜默時長列入考慮。說話者在計畫語音的輸出時應該會將所有音段的長度皆考慮進去以符合語速和節律的要求，也包括音節間的停頓或是靜默。鄭靜宜（2004）發現音節間的停頓或是靜默隨著語速變慢而增長，而且各聲母音節的平均音節時長加上音節前的停頓或靜默時長在各語速下各單位具有相似的時長，並發現在各語速下這個總和的時長很接近節拍器的調節週期。亦即在各語速下各聲母音節時長加上它們之前的停頓或靜默的總時長是相近似。在慢語速下說話者不僅拉長母音、子音時長，也拉長音節前的停頓或靜默時長來滿足慢節拍時的語速要求。在慢語速下不僅送氣塞音的VOT增長，VOT之前的靜默或稱閉鎖（closure）或空白（gap）音段也明顯增長許多。可知除了音節時長本身之外，音節和音節間的停頓或塞音、塞擦音之前的靜默音段也是計時單位的一部份，一樣具有調節語速的功能存在。況且就構音的動作而言，塞音之前的構音閉鎖（或靜默）階段也屬於塞音構音不可或缺的一個動作階段，此靜默階段應屬塞音或塞擦音動作的一部分，不應被忽略應

該也要納入計算。事實上，靜默和停頓時長除了受語速影響外，也受到許多因素的影響，如句型結構、強調、呼吸調節、詞語提取、話語遲疑等，而且說話者對於停頓時長似乎保有十分彈性的調整自由度，而有個別差異存在。

知覺的現象應該是建立於一定的物理或生理基礎之上的，時長的相等或相近應該是指在程度上的差異值較小，而此程度的特性或標準是可以經由實驗研究推估的。說話是屬於人類的動作行為之一，動作在執行時通常會有一定的變異性或是誤差性。因為即使是一樣的動作在每一次的動作時在時間和空間上也是會有些微的變異性（variability）存在，而這些變異性的大小範圍在何種程度下在知覺上會被認為是一樣或相近似的呢？比較相鄰單位之間音段時長的變異性即是由另一個方向來思考韻律等時性的問題，藉由單位時長變異性的比較可以讓我們更深入了解言語節律的等時性。

（二）成對變異指數

在統計上常使用標準差或變異數代表一個變項的變異性。在言語時長方面，成對變異指數（pairwise variability index, PVI）是用來計算時長變異性的指標，它主要在於計算語句中順序上相鄰兩單位間時長的差異（Grabe & Low, 2002; Low, Grabe, & Nolan, 2000）。以語句中各相鄰兩個單位之間時長差異的平均數來代表語句中各相鄰單位之間時長的變異性。以下為計算成對變異指數（PVI）的公式（Low, Grabe, & Nolan, 2000）。

$$PVI = \left[\sum_{k=1}^{m-1} |d_k - d_{k+1}| \right] / (m-1)$$

公式中的 m 是指音段單位數目， d_k 是指第 k 個音段單位的時長， d_{k+1} 是指第 $k+1$ 個音段單位的時長。PVI是計算一個語句中兩相鄰單位之間時長的差異絕對值的平均數，也就是使用相鄰單位間時長的差異來代表單位時長的變異性。假設要來比較幾種單位的優劣，若使用某種單位來計算可得到較低的PVI，即兩兩相鄰單位間時長差異值較小，單位間時長變異性較小，則此種單位即較具有等時性。

（三）語速的影響

語流中的時長變項最容易受到語速的影響。通常語速慢時音段時長變長，音段間時長的變異較大；

反之，語速快時各音段的時長變短，變異也較小。再者，語速對各音段時長的影響是非線性且非均質的，例如英語在快語速時，非重音音節的時長就會相對地被縮短較多。語速的差異不僅存在於個別的說話者間，相同說話者在不同時間時，甚至同一語句中也可能有語速不同的變化。因此在語音取樣時有必要對於語速加以規範，鄭靜宜（2004）使用節拍器調節說話者的語速，分析正常說話者在五種語速下說國語時的各種語音音段的時長，發現語速對於子音和母音的影響並不相同，對於母音的影響就比子音來得大。鄭靜宜（2004）並發現語速對於各種音段的影響也不盡相同。由於音段時長的變異隨著語速變慢而增加，且呈現非線性的趨勢，而國語的等時單位時長變異性或等時性是否隨著語速的不同而有不同呢？

（四）語音的知覺中心

語音中等時性單位的起點是否就在語音中每一音節（或重音節）的起始點呢？Allen（1972a, 1972b）要求受試者跟著錄音帶播放的連串的語音敲擊手指，發現聽者敲擊手指的位置大約是在重音節中母音起始的位置附近處。他因此提出一詞音中子音結束後母音起始點附近有知覺計時的記號。英語聽者似乎知覺到在重音節起始母音點附近的間隔相近似，而具有重音計時的韻律知覺。

Morton、Marcus及Frankish（1976）首先提出語音的知覺中心（perceptual centers, P-center）的概念。P-center為一假設性概念，代表在時間向度上聽者心理上一個語音所發生的位置。他們讓受試者調整連續音節播放的間隔使之具有韻律等時感覺，聽連續的「one、two、three、four....」等語音串時，聽者被要求調整兩音間的間隔使之成為規律的語音序列，他們發現聽者調整每段相等間隔所根據的位置大致是位於音節中母音開始後約十幾毫秒的位置。Morton、Marcus及Frankish（1976）發現在雙耳聆聽作業（dichotic listening task）中兩耳語音刺激序列若同時是以P-center的時間間隔播放，則兩耳接收的訊息互相干擾達最強。Morton、Marcus及Frankish（1976）指出知覺中心的位置絕不是在每個音的起點位置，而是在母音起始附近。事實上，音節中母音開始的剎那的位置通常是語音聲學信號變化最急遽的時刻，此點正位於子音和母音交接，也正是共振峰轉接（formant transition）之處，而共振峰轉接對於一個音節辨識也扮演著重要的角色（Assmann, 1995; Stevens & Blumstein, 1978）。此外，由於母音起始處通常也正是一個音節中音強最強的所在，因此在知覺計時向

度上可能代表著一個音所發生的時刻點。

Fowler（1979）認為 P-center和語音知覺以及語音製造有關，她發現Fowler（1979）要求受試者節律性地重複或輪替誦念一些單音節詞語，測量兩兩詞語起點的間隔時間，發現間隔時間和詞語的子音時長有高相關，推論P-center的位置就在子音結束後母音起點附近。她還發現若將連串的語音依照每一詞音的起點以相同的間隔距離來播放時，聽者卻不會產生語音等時規律的感覺。但若將它們以P-center的時間間隔播放時聽起來會具有等時規律性。認為P-center是聽者心理知覺中語音發生瞬間的位置和語音的韻律知覺有關。一個詞語的起始位置並非心理上此音發生的位置，而是位於稍微後方的母音起點。Buxton（1983）發現若是將個別語音依照P-center的間隔來播放，比起整齊劃一的間隔，聽者的辨識會較為正確。由於上述研究者所選用的實驗材料大多為英語單音節詞彙或由單音節詞組成的句子，對於多音節詞較多的英語，可能較有推論性的問題。而國語一字一音節，相較於英語，音節結構形式較為簡單，聽覺上音節的節律性也較強。是否國語的節律也和音節中的P-center有關呢？音節的母音起始點在國語韻律的等時性是否也扮演著重要的角色呢？

（五）研究目的

本研究欲藉由言語速度的變化分析在聲學層次上正常國語說話者語句中相鄰音節單位時長的差異性。探討國語連續語音中的等時計時特性，考驗兩種可能的計時單位的等時性。這兩種可能的等時單位一為全音節單位，另一為母音起始單位。全音節單位是除了純粹音節（子音加上母音）時長外，還包括音節之前的靜默或停頓時長。由於鄭靜宜（2004）發現各類聲母音節前的停頓或靜默時長型態卻正好和各類聲母音節的時長型態呈現互補的關係，即各語速下各類聲母音節的平均音節時長加上其前停頓平均時長的總合具有符合節拍器週期的一致性。因此語句中純粹音節時長加上其前的停頓或靜默時長（全音節單位），成為本研究中候選的等時單位之一。另一種可能的計時單位是母音起始單位，此單位的假設是來自P-center有關研究（Buxton, 1983; de Jong, 1994; Harsin, 1993; Harsin & Green, 1994; Fowler, 1979; Morton et al., 1976）。它的起迄點是以音節中的母音開始到下一個音節中母音起始前為止的間隔，包括母音（或韻母）時長和後接的靜默或停頓再加上後接音節的子音時長。母音起始單位實際上是跨音節單位，它跨越了語句中相鄰的兩個音節。

實驗一

實驗一的主要目的在比較五種語速下正常說話者在說片語和句子時兩種可能的言語計時單位時長的變異性。藉由節拍器的調整說話者語速，透過計算語句中各相鄰單位間的時長PVI分析兩種可能存在計時單位的等時性，並比較在各語速下兩種單位的等時性差異。PVI愈小代表該單位愈具等時性。考驗的兩種可能的等時單位一為全音節單位（包括音節之前的空白或停頓），另一為母音起始單位（包括母音時長和其後接的空白或停頓和子音時長）。

（一）方法

1. 參與者

三十位正常成年人（男15名，女15名）參與語音錄音的工作。這些參與者皆為台南大學大學部的學生，年齡範圍由19至24歲，皆以國語為母語，且其視力與聽力皆在正常範圍內，無已知的神經性疾患。

2. 語音材料

包括有三十三組的發語材料和兩篇短文，每組中各有一四音節片語與一句子（見附錄一和附錄二）。每個句子的長度為六或七音節。這些詞音的編選是依照聲母、聲調平衡為原則。將每個刺激詞句印製在3"×5"的紙卡上，護貝為提示卡。提示卡上除國字呈現外，並於每個字旁標有注音符號。兩篇短文內的詞語由三十三組發語詞句所組成，印製於A4紙頁上，並護貝成卡。

3. 錄音程序

所有的語音採樣錄音皆於聽力檢測用的隔音室中進行，採用數位錄音機 Marantz 670（取樣頻率設定為22.05kHz）與SONY EMC麥克風錄音。錄音時麥克風離說話者的嘴約7公分。主試者先示範五種不同的說話速度，包括中速、最快、稍快、稍慢和最慢，並使用音樂節拍器（Wave Metronome）提示參與者的說話速度。節拍器共有五種速度設定：中速（120次/分）、稍快（150次/分）、最快（180次/分）、稍慢（90次/分）、最慢（60次/分）。先讓受試者聽節拍器的節奏聲，然後練習跟著節拍器的速度練習將該組詞語唸一遍，要求說話者念時大約符合節拍器敲一聲說一個音的原則。練習唸完之後則停止節拍器，請受試者以類似的速度說出該組刺激詞語。三十三組詞語以隨機排序出現，保持每次發語時間間隔一秒。為避免混淆，速度的次序以中速、稍快、最快、稍慢和最慢的次序

進行。用以研究語音聲學分析的語音材料是採用在非使用節拍器時所產生的語音。

短文閱讀的錄音取樣程序和以上的四音節片語、句子閱讀時類似，遵循五種言語速度的設定。一開始使用節拍器調節語速練習，只練習文章前三個句子的閱讀，之後停止節拍器，依照該指定的速度由文章的第一個句子開始閱讀到最末尾處，並錄音。速度的次序以中速、稍快、最快、稍慢和最慢的次序進行。每個受試者完成整個實驗程序約在一小時左右。

4. 聲學資料分析

語音處理分析主要使用TF32（Milenkovic, 2004）語音頻譜分析軟體。使用TF32中提供的頻譜圖呈現功能，以寬頻頻譜圖（300Hz）呈現，以手動左右游標界線，逐一測量每個刺激語句中各語音音段的長度，包括子音（聲母）、韻母（母音或是母音加上鼻音）和音節間停頓或空白的時長。母音結束時標記標準則為同時在頻譜圖和波型圖上有明顯信號強度減弱趨於零的位置。

取每個片語組中的音段資料來計算兩種假設性的計時單位（全音節單位、母音起始單位）的PVI，即一組資料使用兩種單位的劃分得出PVI值。由於句子末尾音節因為可能出現拖長效果，末尾音節將不列入計算。取每個語句組中和短文句子中具有相同的音節音段單位做分析。由片語和語句中倒數第四個音節開始的三個單位。捨去最後一個音節的原因是因為句子末尾音節的母音時長可能有拖長的現象，而且以母音起始點的單位末尾音節母音並無法構成一個完整的單位。計算各個片語和語句中兩種計時單位的時長，各為全音節單位一、全音節單位二、全音節單位三、母音起始單位一、母音起始單位二、母音起始單位三（請見圖1）。如此使用兩種單位的計算，兩者間所包含韻母的部分其實是一樣的，兩者間的不同是在於單位計算的起始點和終點，一個始於音節之前的靜默或停頓，終至一音節母音的末尾；另一個在於音節中母音的起點，終至下一音節子音的終點。之後分別計算各個片語、句子和短文句子中那些順序上相鄰三單位之間時長差異的絕對值，後以各受試者的每個語句為單位計算累計時長差異的絕對值的平均數，得到片語、句子和短文句子的PVI指數。

（二）結果與討論

1. 兩種單位相鄰三單位的平均時長

比較兩種假設性單位的三連續單位的平均時長，在表1列出全部受試者在三十三組片語、句子和短文中

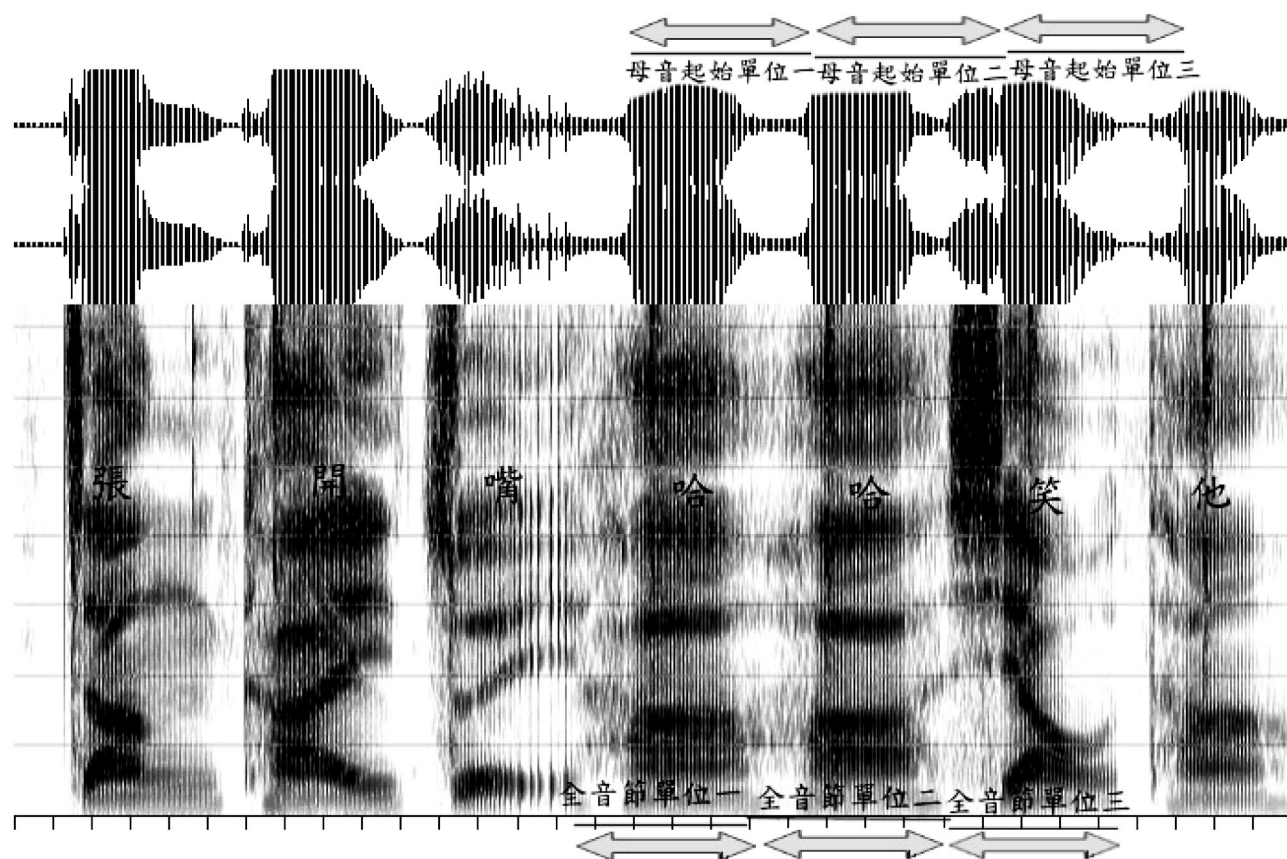


圖1：母音起始單位和全音節單位的界定舉例，在「張開嘴哈哈笑他」的句子中，各有三個連續測量的單位。母音起始單位以音節的母音為起點至下一音節的母音起始為止。全音節單位以音節的子音為起點至音節的末尾為終點

各以全音節和母音起始單位的相鄰三單位的平均時長和標準差。在最快語速下，兩種單位的時長相近，兩者在片語、句子和短文皆在300ms左右，隨著語速的變慢，平均單位時長漸增，到最慢速時平均單位時長在800、900ms左右，且兩種單位的差距變大，母音起始單位較全音節單位為長。若是三個連續單位比較，在慢語速下，全音節單位中的第三單位時長皆較短，和第一、第二單位的時長差距較大，反觀以母音起始為單位三個連續單位的時長則較接近。基本上，單由單位時長平均數的比較是無法比較兩種計時單位的變異性，而標準差的資料或可說明一部份的相鄰單位間時長的變異性。整體大致看來，無論在片語、句子或短文中，除了在最快速的語速下，以母音起始單位時長的標準差均小於以全音節為單位時長的標準差（見表1），這表示以母音為起點的單位時長變異性較以全音節為單位的為小，此趨勢在最慢速時尤為明顯。在最快速的語速時，兩種單位間的差異很小。在慢語速的短文以全音節為單位時單位時長變異性最大。

2. 兩種單位成對變異指數的比較

由PVI的比較就可以很容易的看出兩種單位的時長變異性。圖2呈現在片語、句子和短文中全音節和母音起始兩種單位的PVI。無論是在片語、句子和短文中，以母音起始單位的平均PVI皆小於以全音節為始單位的PVI，顯示以母音為起點的單位時長變異數較小，尤其是在慢速語速下更顯出此單位的優勢。其實語速的變化由最快、稍快、中速到稍慢的情況下，和全音節單位相較，以母音起始單位的PVI的增加其實不大。在最慢速的語速下，PVI則有非線性的增加趨勢，但是和全音節比較起來，以母音為起點單位的單位間時長變異量還是較小，顯示以母音為起點單位的穩定度較佳。圖2也顯示母音起始單位的PVI在跨片語、句子和短文的比較中較相近似，母音起始單位的等時性較不受發語單位的影響。

就片語而言，以計時的單位和言語速度為自變項，以PVI為依變項的受試者內二因子變異數分析結果顯示，單位和言語速度的效果皆達顯著，計時單位效果： $F(1, 29) = 86.27, p < .001, \eta^2 = .75$ ，言語速度

表1 以母音起始單位和全音節單位的連續相鄰三單位的平均單位時長(ms)和標準差

語速		最快		稍快		中速		稍慢		最慢	
單位		Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD
片語	母音起始單位一	285	49	347	55	425	67	575	90	857	159
	母音起始單位二	304	54	364	64	435	66	580	87	865	150
	母音起始單位三	316	44	370	51	436	59	574	82	859	153
平均	母音起始單位	302		360		432		576		860	
片語	全音節單位一	300	51	359	62	430	74	562	101	823	166
	全音節單位二	296	50	352	62	417	73	558	114	851	192
	全音節單位三	323	73	362	82	417	97	528	128	768	221
平均	全音節單位	306		358		421		549		814	
句子	母音起始單位1	280	54	341	65	416	68	569	93	852	157
	母音起始單位2	298	55	359	69	429	67	577	89	856	149
	母音起始單位3	304	46	361	61	429	61	577	83	857	149
平均	母音起始單位	294		354		425		574		855	
句子	全音節單位1	295	51	352	70	420	68	556	102	817	168
	全音節單位2	288	50	345	69	409	74	560	113	850	189
	全音節單位3	307	64	355	73	416	85	543	119	790	204
平均	全音節單位	297		351		415		553		819	
短文	母音起始單位一	289	62	356	64	462	72	585	103	898	177
	母音起始單位二	303	56	377	69	475	74	598	98	904	166
	母音起始單位三	304	51	373	59	469	70	590	92	906	172
平均	母音起始單位	299		368		469		591		903	
短文	全音節單位一	302	55	370	59	462	74	571	110	855	190
	全音節單位二	293	50	362	71	464	90	593	129	925	211
	全音節單位三	302	61	368	73	455	95	560	127	848	220
平均	全音節單位	299		367		460		574		876	

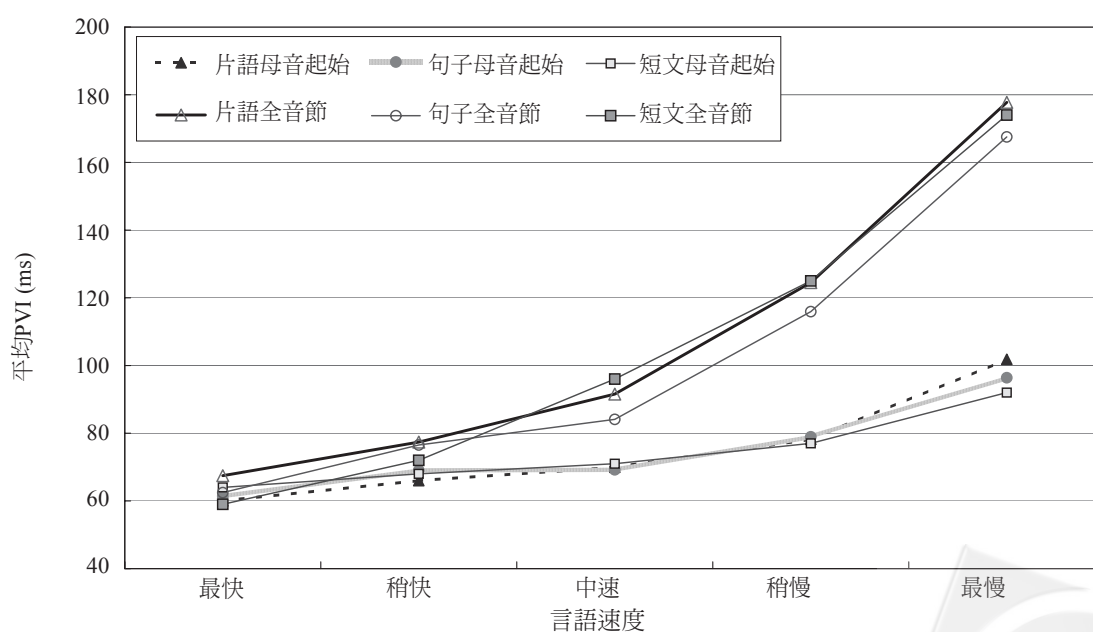


圖2：五種語速下片語、句子和短文的兩種計時單位（母音起始單位、全音節單位）時長的PVI

效果： $F(4, 116) = 180.37, p < .001, \eta^2 = .86$ ，言語速度與單位的交互作用亦達顯著， $F(16, 116) = 54.56, p < .001, \eta^2 = .65$ 。以母音起始單位的PVI顯著地小於以全音節為單位的PVI，且在慢語速下的PVI顯著大於在快語速下的PVI。交互作用的變異來源是以全音節單位和母音起始單位兩者間的差異在最快語速下很小，隨著語速的變慢兩者間PVI的差距逐漸變大，一直到最慢速時PVI差異達到最大。

在句子方面的分析結果和在片語中的相似。以計時的單位和言語速度為自變項，PVI為依變項的受試者內二因子變異數分析結果顯示，單位和言語速度的效果皆達顯著，計時單位效果單位效果： $F(1, 29) = 70.35, p < .001, \eta^2 = .71$ ，言語速度效果： $F(4, 116) = 136.63, p < .001, \eta^2 = .83$ ，言語速度與單位的交互作用亦達顯著， $F(16, 116) = 57.21, p < .001, \eta^2 = .66$ 。以母音起始單位的PVI顯著地小於以全音節為始單位的PVI，在慢語速下的PVI顯著大於在快語速下的PVI。兩變項間的交互作用來源是隨著語速的降低，兩種計時單位的PVI差距變大。在最慢語速時，全音節單位和母音起始單位間PVI差距最大。

在短文中的分析結果也和片語和句子中的結果相近似，短文中全音節單位時長的PVI隨著語速的變慢逐漸增加，尤其在慢語速時增加近最快語速時的三倍之多，而以母音為起點的計時單位時長PVI隨著語速也是略有增加，但是增加的速度較和緩的多，由最

快到稍慢其實增加不多，可見以母音為起點的計時單位的時長PVI是較為穩定的，較不會隨著語速變化而變動。

由以上在片語、句子和短文的分析可知言語速度與計時性單位顯著地影響單位時長PVI。比較兩種假設性單位，以母音起始的計時單位的時長的變異性顯著小於全音節的計時單位。在五種語速下以母音為起始單位的PVI顯著小於以全音節為始單位的PVI，也就是比起全音節，以母音起始單位之間時長的差異量較小，即較具有等時性。大體而言，隨著言語速度愈慢，PVI值漸增大，尤其是以全音節為單位的PVI受語速的影響比母音起始單位為大，尤其是在慢速下PVI增加最多。相對地，以母音起始為單位的PVI在慢速之下的增加較少，單位間時長變異性較小，穩定性較佳。以母音起始為單位的PVI不管是在各語速下的變動較小，亦即PVI受語速的影響較小。

是否言語計時有個別差異性呢？圖3呈現三十位受試說話者在說片語和句子的兩種假設性計時單位的五種語速的平均PVI。除了受試編號20、21和29以外，無論是在片語和句子中以母音起始單位的PVI皆小於以全音節為單位的PVI。顯示大多數的受試者是以母音起始為言語計時單位的趨勢較強，而其餘三位受試者的兩種計時單位的PVI很接近，較無法看出以母音起始單位的優勢性。圖3中還可看到以母音起始單位的PVI在說片語和句子時兩者的曲線十分近似，顯

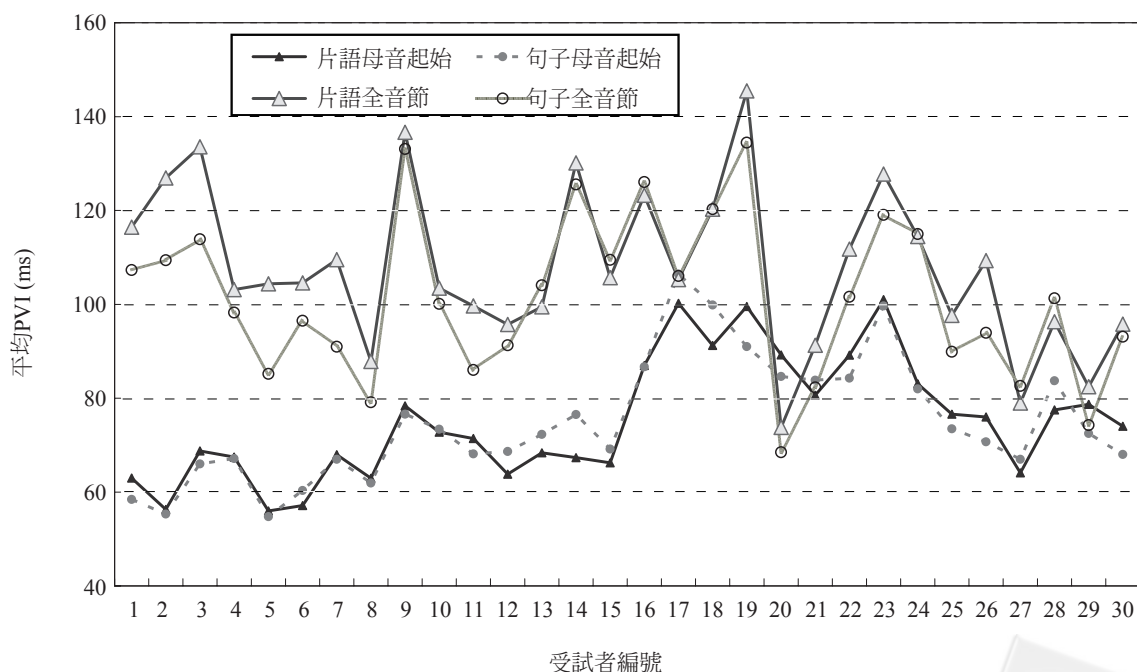


圖3：五種語速下三十位受試者在說片語和句子時的兩種計時單位（母音起始單位、全音節單位）時長的平均PVI比較

示具有相似的計時特性。以全音節為單位，在說片語和句子時兩條曲線較不重疊，如受試編號1、2、3、5、26等，顯示以全音節為單位的PVI受刺激詞語類型或發語長度的影響較大，而以母音起始為計時單位較不會受刺激詞語類型所影響。

實驗二

由於在實驗一中使用了節拍器調節說話的速度，雖然分析的語音是說話者在節拍器停止後產生的語音，但模仿節拍器的節奏難免可能會造成某些說話者的節律有些許不自然的情形。實驗二的主要目的是使用較自然的語速調節方式，說話者模仿示範錄音帶中的語速，再說出片語和句子。另外，在實驗的材料上也做了一些修改，加入五音節片語。由於實驗一中句子皆是六音節或是七音節，整體刺激材料中缺乏五音節的部分，因此在每一組中各加入一個五音節片語。

實驗二的目的和實驗一相同，在探究五種語速下正常說話者的語音輸出之等時單位的存在，分析兩種可能存在的等時單位的單位間的時長變異性指數（PVI）。這兩種可能的等時單位一為全音節單位（包括音節和音節之前的空白或停頓），另一為母音起始單位（包括母音時長和其後接的空白或停頓和子音時長）。

（一）方法

1. 參與者

共有二十一位正常成年人參與語音錄音的工作，其中男性10名，女性11名。這些參與者皆為台南大學大學部的學生，年齡範圍由19至24歲，皆以國語為母語，且其視力與聽力皆在正常範圍內，無已知的神經性疾患。

2. 語音材料

語音材料是由實驗一的材料延伸擴展而來，先由實驗一中使用的三十三個片語組中選取其中二十六組詞句組，這二十六組詞句組主要是開頭含有國語二十一個聲母的詞句組。每一詞句組中包括一個四片語組和一個句子。此外，每一詞語組中再加上一個五音節片語（詳見附錄三），五音節片語的構成主要是在四音節片語之前加上一個主詞、形容詞或副詞。如此每組詞語包括三種層次的發語單位，即各有四音節、五音節與七音節（或六音節）句子。實驗時使用的刺激詞卡的製作方式和實驗一相同。

3. 錄音程序

所有的語音採樣錄音皆於聽力測試隔音室中進行，採用數位錄音機（Marantz 670）、SONY EMC麥克風錄音。說話速度示範錄音帶由一說話清楚的大學女性事先錄製而成。錄音時遵循電子節拍器的速度。節拍器五種設定的速度和實驗一相同，各為中速（120次/分）、稍快（150次/分）、最快（180次/分）、稍慢（90次/分）、最慢（60次/分）。

正式實驗語音取樣時，主試者先解釋實驗進行步驟和語速的變化，使用Sony隨身錄放音機播放示範錄音帶呈現五種不同的說話速度，包括中速、最快、稍快、稍慢和最慢。請受試者以類似的五種速度說出該組刺激詞語，依照中速、最快、稍快、稍慢和最慢的順序，直到二十六組語句完成。各語句組以隨機排序出現，保持每次發語時間間隔一秒。正式錄音前練習三~五次。說話者音量需適中，話語需清晰、自然，即使在最快速時也須保持清晰，注意防止說話者音和音之間融混。在語句中若發現有唸錯或遲疑不連續的現象，則請受試者重唸一次。平均每個說話者錄音時間約為四十分鐘左右。

4. 聲學資料分析

和實驗一相同，使用TF32（Milenkovic, 2004）頻譜圖呈現分析各語句中各音段的時長，再取每個片語組中的音段資料做單位時長的變異分析。由於句子或片語末尾音節母音可能出現拖長效果，每句末尾母音將不列入計算。先計算每個句子中兩種可能單位的單位時長，再計算兩兩順序上相鄰單位（兩種）的時長的差異，並求其平均數，得到每個片語和句子的兩種單位的PVI。語音資料的聲學分析和處理大致和實驗一相同。唯一不同的是在計算PVI時，採用語句中較多的單位資料來計算。例如對於四音節片語採用每句中相鄰三個時長單位的平均差異；對於五音節片語採用每句中相鄰四個時長單位的平均差異；對於六、七音節的句子採用每句中相鄰五或六個時長單位的平均差異（依據各句句長而定）。使用兩種單位的方式計算，得兩種可能的計時單位（全音節單位、母音起始單位）的PVI。

（二）結果與討論

圖4呈現在五種語速下四音節、五音節片語和句子的PVI，無論是在何種語速下，以母音起始單位的PVI皆小於以全音節為單位。全音節單位在慢速時PVI增加較多，顯示受語速影響較大，以母音起始單位相對地受語速的影響較小。若跨四音節、五音節片語和

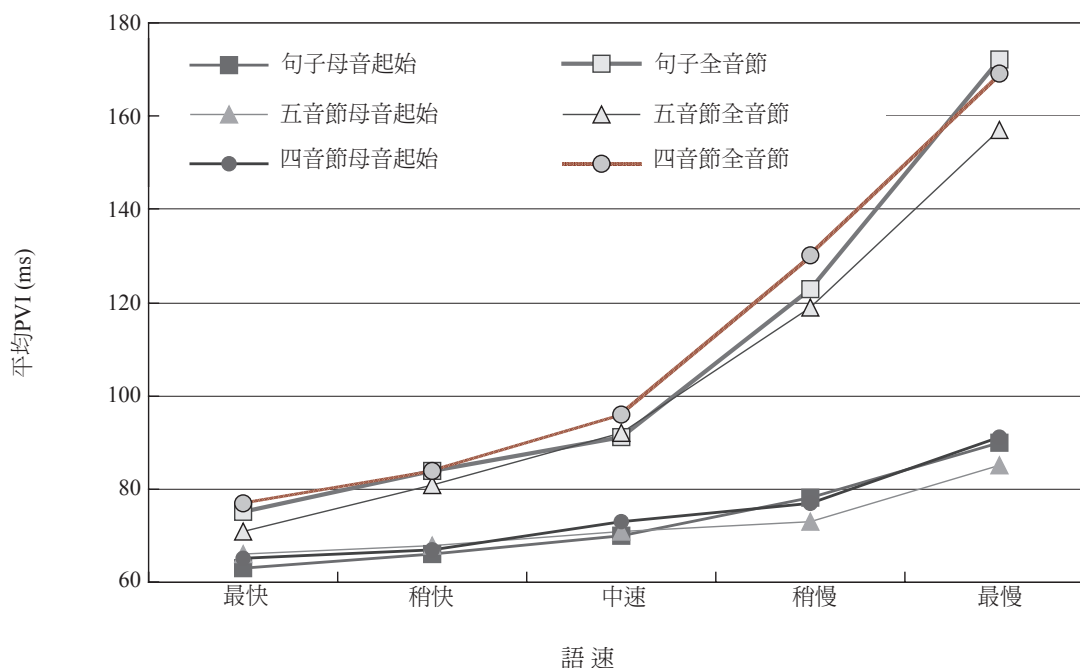


圖4：五種語速下兩種計時單位（母音起始單位、全音節單位）時長的平均PVI

句子比較，以母音起始單位PVI三曲線相靠近，在五語速下呈現相類似的形式，可見語音材料的單位大小對於以母音起始單位PVI的影響較小，相對地，以全音節為單位時受材料的影響較大。

在句子材料中，母音為起始單位和全音節兩種單位的PVI相比較，以受試者內二因子變異數分析變異數分析，結果顯示單位效果和言語速度的效果皆達顯著，單位效果： $F(1, 20) = 130.86, p < .001, \eta^2 = .87$ ，言語速度效果： $F(4, 80) = 67.18, p < .001, \eta^2 = .77$ 。言語速度與單位的交互作用亦達顯著， $F(4, 80) = 38.23, p < .001, \eta^2 = .66$ 。以母音為起始單位的PVI顯著地小於全音節單位的PVI。五種語速中在快速語速的PVI最小，隨著語速的降低PVI變大。言語速度與單位的交互作用顯著，表示在快語速下兩單位PVI的差異小，在慢語速時差異愈大。

在四音節和五音節片語部份，變異數分析結果和在六、七音節的短句中的十分類似。五音節部份：單位效果： $F(1, 20) = 55.88, p < .001, \eta^2 = .74$ ，言語速度效果： $F(4, 80) = 71.80, p < .001, \eta^2 = .78$ ，言語速度與單位的交互作用亦達顯著， $F(4, 80) = 44.56, p < .001, \eta^2 = .69$ 。四音節部份：單位效果： $F(1, 20) = 75.62, p < .001, \eta^2 = .79$ ，言語速度效果： $F(4, 80) = 52.66, p < .001, \eta^2 = .73$ 。言語速度與單位的交互作用亦達顯著， $F(4, 80) = 34.19, p < .001, \eta^2 = .63$ 。

在四音節、五音節片語和句子中皆發現以母音起始單位的PVI顯著小於以全音節為始單位的PVI，意即以母音起始單位的相鄰單位之間時長的差異較小，即以母音起始單位較具有等時性。也發現語速的影響，言語速度愈慢，PVI愈大。五種語速中，在慢速語速的PVI最大，亦即相鄰單位間時長變異最大，尤以最慢速時全音節PVI受語速的影響最大。比較兩種假設性單位，以母音起始單位的PVI在各語速下變化較小，數值較穩定。可見以母音起始為計時單位的時長變異性受語速的影響相對地較小。圖5呈現二十一位受試者說句子時兩種單位五種語速平均PVI。二十一位受試者以母音起始單位的PVI皆小於以全音節為單位的PVI。顯示全部受試者是以母音起始為言語計時單位的趨勢較強。雖然此次實驗中使用較不嚴格的語速控制方法，但PVI的趨勢卻和實驗一極為相似，可知不管使用何種語速控制方式皆較支持以母音為起點的計時單位較具有等時特性。

實驗三

實驗三的目的在聽知覺上比較具有不同時長變異性指數（PVI）句子的自然度。這些句子使用兩種計時單位規劃法計算各具有不同的PVI。這兩種等時單位一為全音節單位（包括音節和音節之前的空白或停

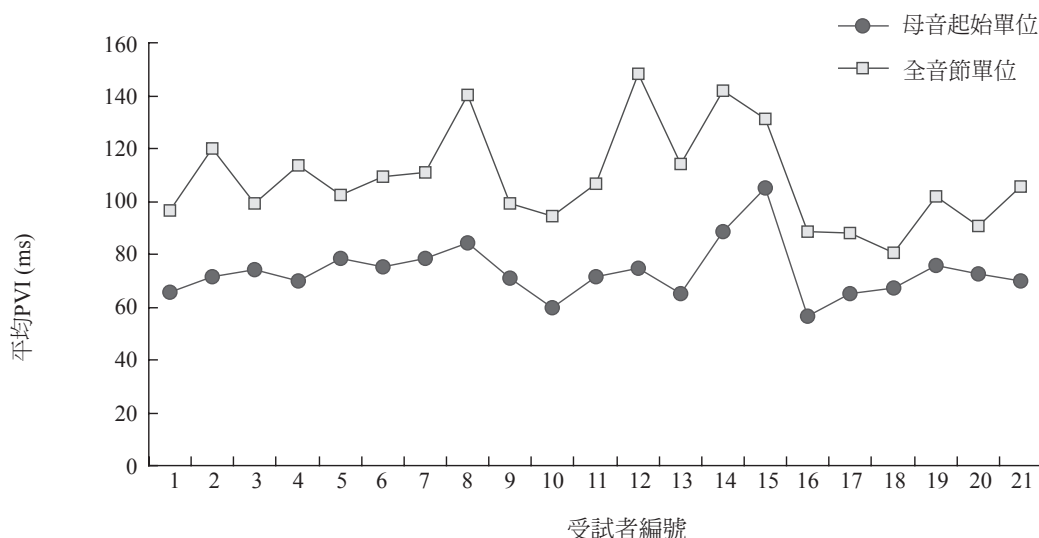


圖5：在五種語速下，二十一位受試者說句子時兩種計時單位（母音起始單位、全音節單位）時長的平均PVI比較

頓），另一為母音起始單位（包括母音時長和其後接的空白或停頓和子音時長）。究竟聽者的自然度判斷反應是和何種單位的PVI相關較大呢？可推測聽者判斷自然度時所根據的計時單位為何。

（一）方法

1. 參與者

共有十八位正常成年人（男、女各半）參與聽知覺語句自然度判斷作業。年齡範圍在二十至四十歲之間。這些參與者皆以國語為母語，他們的視力與聽力皆在正常範圍內，且無已知的神經性疾患。他們皆不曾參加過類似相關的實驗。

2. 語音材料

原語音刺激共有四句，由一位成年男性和一位成年女性錄製而成，語速約在中等速度。這四個句子分別為：「他低頭看地上螞蟻」、「那個小孩他聽不到」、「小企他踩破梳子了」、「他並不考慮要養那隻狗」。再使用Praat (Boersma & Weenick, 1999) 語音軟體中的語音操弄合成功能改變句中某些音段的時長。聽覺實驗使用的刺激材料有四組句子，每組中各有六個句子，共有二十四個句子。各組中句子的詞語皆是相同的，為同一個句子改變其中一些音段的時長而成，Praat提供的是增加或減少時長的倍數功能。因顧及子音音段改變可能會改變子音的類別，因此改變的主要是母音或是停頓時長，即增加或減少母音的

時長或是增添停頓時長。每組中各合成一些具有不同PVI值的刺激語句，每組中各有一個以全音節單位方式計算PVI接近零（或變小）的刺激，和以母音起始單位方式計算PVI接近零（或變小）的刺激，它們的單位時長保持約在500ms和600ms之間。因為Praat程式以視覺滑鼠拖曳的方式改變整個音段時長的倍數，音段時長的改變不可避免會有誤差存在。因此合成後再次測量這些語句的PVI，表2中列有各刺激句以兩種計時單位計算的平均PVI值。

3. 程序

所有的語音採樣錄音皆於聽力檢查隔音室中進行。首先向受試者解釋自然度的定義和量表的評估方式。所謂的「自然」就是句子的節奏聽起來很正常不會忽快忽慢，或有不正常的停頓。自然度量尺為六點量表，「一」是最不自然，「六」是最自然。共有四個自然度量尺列印於A4的答案卷以橫軸等距點量尺方式呈現，並將各點的自然度以文字描述標示。自然度「一」為「最不自然」；自然度「二」為「很不自然」；自然度「三」為「不自然」；自然度「四」為「不太自然」；自然度「五」為「稍不自然」；自然度「六」為「最自然」。每組中的各刺激需一一被分派一個的自然度分數，且組內各刺激所被分派到分數不得重複。此自然度判斷作業類似將各組中刺激以自然度作排序作業。語音刺激檔名列表呈現在電腦螢幕上，聽者在各組語音檔案夾中點擊檔案名稱撥放音檔，聽完後判斷該音檔的自然度，再於答案卷上該刺

激組的自然度量尺上一一標明該一音檔的語音自然度。組中每一音檔以隨機順序編號。語音音檔以Real Player程式撥放，刺激語音從耳機播放雙耳呈現。耳機型號為Sennheiser HD280。聽者可自由選擇組中的語音刺激檔播放，若一次的呈現無法判斷，聽者可自由重聽，直到完成為止。

(二) 結果與討論

和原始刺激語句相比，合成後低PVI值的刺激語句是否有增加自然度呢？表2呈現各刺激的母音起始單位PVI、全音節單位PVI以及自然度平均值和標準差。除了刺激組中的第三組外，其餘三組皆是以原刺激語句的自然度平均值為最高，可見經過合成改變後的語音刺激的自然度不如原始刺激。雖然聽起來不是很明顯，然而再合成的語音刺激在信號品質上的確總是比不上原始信號。雖然如此，可以看到四組刺激之中母音起始單位PVI較低的語音刺激的自然度皆是高於全音節單位PVI較低的語音刺激，以母音起始單位劃分較整齊的語音刺激聽起來較自然。

將受試者所評估對各刺激的自然度分數平均值和每個刺激的兩種計時單位的PVI值作相關分析，比較自然度分數和兩種計時母音起始單位PVI、全音節起始單位PVI值的相關係數，結果發現自然度分數和母音起始單位PVI值呈高相關，Pearson相關係數為 -0.80 ($p < .0001$)，而自然度分數和全音節單位PVI值呈現中高度相關，Pearson相關係數是 -0.70 ($p < .0001$)。顯示自然度分數和母音起始單位PVI值的相關較高，可見在聽知覺母音起始單位比全音節單位有較佳的心理真實性。圖6和圖7呈現各刺激語句的自然度分數和母音起始單位PVI散佈圖以及自然度分數和全音節單位PVI值的散佈圖，可以看到母音起始單位PVI對自然度分數分佈的線性趨勢較為明顯，而全音節單位PVI值分佈較為凌亂，尤其是在自然度分數較高的部分，較無線性趨勢，自然度分數無法用全音節單位劃分得到的PVI值來解釋。

綜合討論

本研究的前兩個實驗各使用不同的方式控制語速，讓國語說話者使用五種不同的速度說出片語、句子和短文，使用聲學的方式分析言語韻律單位的等時特性，並提出兩種假設性的言語計時單位，比較這兩種單位在語句中相鄰單位間時長的變異性。兩種計時單位分別是母音起始單位和全音節單位。實驗一的結果發現在母音起始單位的相鄰單位間時長的變異性顯

著小於全音節單位，母音起始單位較具有等時性。然而因顧及節拍器的提示使用，可能使有些說話者語音的產生較為呆板、不自然，因此在實驗二中使用較自然的語速調節方式，要求正常說話者依照示範帶的語速提示，產生五種語速變化的語句。實驗二的結果和實驗一的相近，均顯示以母音起始單位的PVI顯著小於以全音節為始單位的PVI，意即母音起始單位的相鄰單位間時長的差異性較小，以母音起始為度量單位的計時方式較以全音節計時方式為等時。另外，也發現語速的影響性，當言語速度愈慢，PVI值愈大。五種語速中，最慢速時的PVI達最大，即相鄰單位間時長變異最大，其中尤以最慢速時全音節單位PVI受語速的影響最大。相對地，以母音起始單位的PVI在各語速下變化較小，顯示母音起始單位受語速的影響相對較小。結果還顯示母音起始單位的等時性較不受發語單位的影響，即在四音節、五音節、句子和短文中以母音起始單位的PVI較一致，而全音節單位的PVI則受發語單位大小的影響較大。實驗三讓聽者使用聽知覺的自然度評估作業評估各具不同PVI的刺激語句，發現聽者的自然度評分和計時單位的PVI有中高度顯著相關，其中和以母音起始單位的PVI的相關較強。顯示在聽知覺上符合母音起始單位的整齊計時的刺激句比起全音節單位聽起來較為自然，較具有節律性。

本研究使用兩種語速控制的方式，讓說話者產生五種不同的語速變化。主要是因考慮到說話者在語句產生時語句中的個別計時單位時長會受整體語速的規範。說話者需顧及到整體語速和發語的單位而去調整每個計時單位的時長，以維持每個計時起點之間間隔的一致性。在一個語句中言語計時單位是由各音段的組合而來，時序上具有連續性，相鄰音段組合成一計時單位。可使用相同的一組音段時長資料來考驗不同的計時單位的組成的可能性。考驗說話者在各種不同的語速下是否有保持單位時長（相鄰音段總合）的一致性。比較實驗一和實驗二的結果，可以知道使用節拍器調節語速和錄音帶示範語速的結果非常相似，可見即使在無節拍器調節下語音也有計時的等時性，而且以母音起始單位的等時性不只存在於在節拍器調節下小心控制的言語行為當中，也出現於示範帶語速提示的片語和句子產生時。在未來相關研究中將可檢驗在一些自然情境下的自發性言語中是否也出現以母音起始單位的PVI的等時優勢。

母音起始單位PVI的優勢在最慢速時最為明顯，在最快速時最不明顯，似乎在最慢速時的情況有放大兩者間差距的趨勢。由此看來，語速的差別對於時長變項的影響只是在量上的不同而非質的差異。至於PVI是否有正規化的需要？由於時長的變異性有隨著

表2 各刺激語句的母音起始單位PVI、全音節單位PVI以及自然度平均值和標準差（單位：毫秒）

刺激組	NO.	全音節起始 單位PVI	母音起始 單位PVI	自然度分數 Mean	自然度分數 SD	備註
1	1	68	35	5.5	0.99	原始刺激
1	2	19	66	3.5	0.62	低全音節PVI
1	3	46	21	4.83	0.62	低韻母起始PVI
1	4	77	137	4.06	1.26	
1	5	227	169	1.94	0.54	
1	6	115	175	1.17	0.38	
2	1	20	15	5.89	0.47	原始刺激
2	2	39	10	4.67	0.84	低韻母起始PVI
2	3	12	35	4.22	0.55	低全音節PVI
2	4	132	152	2.83	0.71	
2	5	192	158	2.33	0.69	
2	6	250	191	1.06	0.24	
3	1	176	32	4.72	0.83	原始刺激
3	2	117	46	5.44	0.7	低韻母起始PVI
3	3	28	185	3.17	0.62	低全音節PVI
3	4	103	141	4.56	1.15	
3	5	325	260	1.67	0.59	
3	6	268	196	1.39	0.5	
4	1	99	91	6	0	原始刺激
4	2	81	46	4.78	0.55	低韻母起始PVI
4	3	16	70	4.11	0.32	低全音節PVI
4	4	108	111	3.11	0.47	
4	5	250	194	1.89	0.32	
4	6	220	116	1.11	0.32	

時長的增加，即語速變慢，而增加的趨勢，有些研究者（Low, Grabe, & Nolan, 2000）提出兩種有關時長PVI有正規化的方法。其中一種是將累積的時長的差異平均數除以句子中平均音節時長，再使之成為百分比分數。此算法較不適用於本研究，因為純粹子音加上母音的純音節並非本研究的假設等時單位。另有一種正規化的算法是將相鄰兩音間時長差異值除以兩兩成對單位時長的平均值。此算法對於小區段內的時長變異較敏感，易表現出相鄰時長的對比性（Low, Grabe, & Nolan, 2000），較適用於以重音的基礎語言如英語。若以本研究的結果來看比較五種語速下單位時長的PVI差異性，可以發現國語PVI在最快速、快速、中速和稍慢語速下，以母音起始單位時長的變異性其實並不大，在最慢速之下增加其實也不多，因此國語若以母音起始單位為計時單位，由於時長PVI不大，似乎較沒有正規化的需要，除非是在說話速度極

端緩慢的情形之下，然此情況在正常情況下發生的機率並不大。

在實驗一中研究者使用頻譜圖仔細觀察在節拍器調節時受試者同時說出的語句（雖這些語音並不列入聲學分析中），發現在大多數說話者的語句中節拍器的敲擊「答」聲所落的位置，雖然並非每次皆正好在CV音節母音的起始點，但多數落在一個CV音節母音的起始點附近，通常節拍器的敲擊「答」聲會比母音起始點稍微延宕二、三十毫秒，而在語句末尾音節時會延宕較多。而這種趨勢在當初錄音時並沒有刻意要受試者如此做，而是一種自然的趨勢。可推測在一音節中母音起始處具有構音上重要的指標意義，此處也正是所謂的「知覺中心」所在。而「知覺中心」是以聽知覺角度去看語音中音節重心安置的位置，若就言語製造的角度而言應稱為語音構音中心，即是構音時言語計時的中心點。說話者依據此點的位置來規劃

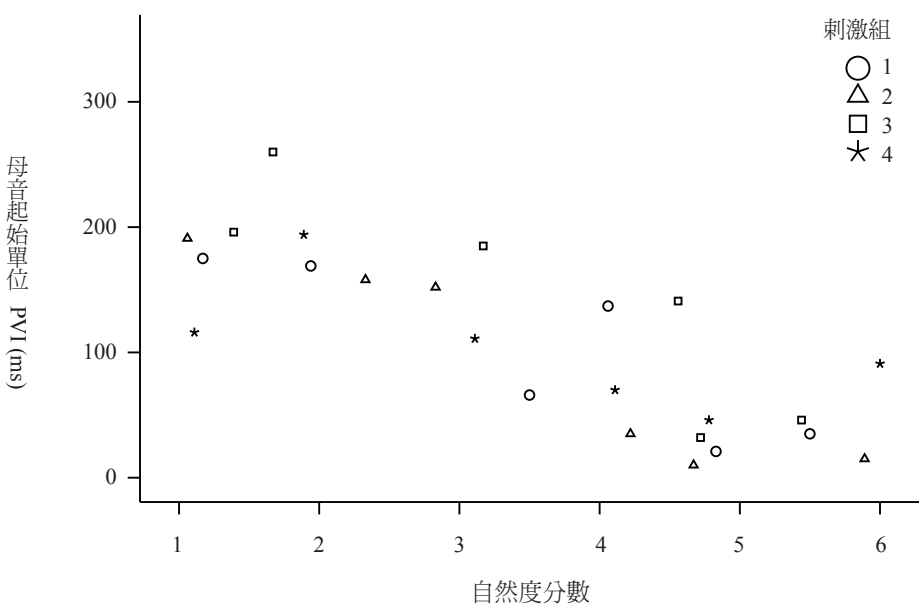


圖6：各刺激語句的母音起始單位PVI值對自然度分數的散佈圖

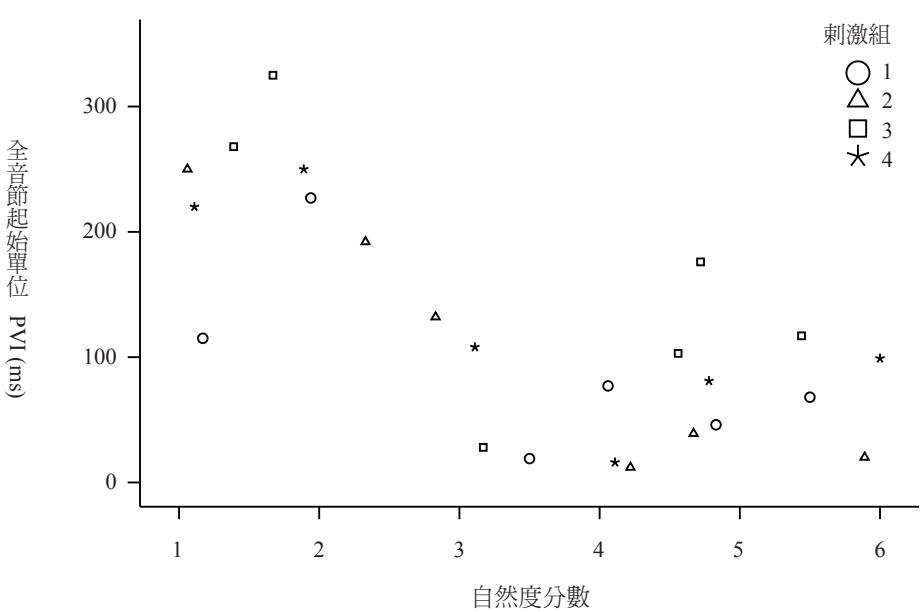


圖7：各刺激語句的全音節單位PVI值對自然度分數的散佈圖



或調節語句的速度或各音段時長的分配。由於母音皆為有聲音，母音的起始位置大都與一音節喉部聲帶的開始振動相吻合。對一個說話動作而言，喉部聲帶開始振動的時間處在每個語音計時單位中應該是具有需要標記的重要意義。在規劃每個言語製造的動作時，語句中每個聲帶開始振動的時間點應該是具有標記意義的重要地方，甚至整個言語動作計畫的時序架構是以此點作為規劃的中心，而加以延伸出來的。再者，由整個音節的音量分佈來看，母音起始點大多數也正是音節中音量最強的時刻點，也就是聲門下壓（subglottal pressure）最強之處；同時，在語音聽知覺上不僅是子音和母音之間共振峰轉折的所在，也是音強上響度最明顯的時刻點。

一個語音計時單位的組成包括語句中的子音、母音以及音節和音節之間的閉鎖（closure）或停頓期（pause）。閉鎖（closure）是發出塞音或是塞擦音之前的持阻動作期。對於一音節與一音節之間的閉鎖（closure）或停頓期（pause），到底這些音節間的閉鎖或停頓期應該歸屬於位於它之前的音節或是它接續下來的音節呢？就構音的觀點，塞音或塞擦音之前的閉鎖靜默期是構音子的持阻期，理論上應該也是該子音的一部份，屬於該音節的一部份。鄭靜宜（2004）發現就相同構音位置的子音而言，送氣塞音或塞擦音前的閉鎖靜默期小於不送氣塞音或塞擦音的閉鎖期，平均不送氣塞音大約平均多了四十毫秒左右。就構音的動作而言，其實送氣音需要大量釋放氣流，需要較多的口腔壓力，在閉鎖期時應該需要醞釀較多的壓力，而需要花費較長的時間，因此閉鎖期應較長，但事實上反而較為短。推測原因可能由於整體說話動作有維持單位等時性的需要，因為送氣音的VOT已經比不送氣音為長，因此送氣音閉鎖期就需要相對地較短。

一般而言，相較於子音，母音時長調整的彈性較大，常需要因應語境的不同而加以調整。若等時單位真的存在，在說話時為了保持單位時長的一致性，母音所負擔的調整責任應較大。若國語知覺的中心的位位置真是位於母音起始的位置，為了調整對齊母音起始的時刻，使得語句中各個母音起始單位的間隔具有等時性。音節的時長將會受到其前後相鄰音節的影響而需加以調整，此時母音的時長就應有相應的變化。母音後接子音若為較長的磨擦音，則該音節母音的時長應該相對的比後接其他子音的母音為短，或者也有可能是該磨擦音音節前面的音節母音會變得較短來抵銷後接磨擦音較長的音段。到底是磨擦音之前的變短或是磨擦音之後的變短？音節中母音時長會受到之前子音和後接音節子音的影響，音段時長依據等時規律的要求加以調整，何者需負擔較大的責任呢？為了維持

母音起始單位的等時性，在一個母音起始單位中，母音的時長是否會受到後接音節子音時長的影響，而有所增減呢？這些問題皆有待後續研究進一步的探討。

本研究藉由對分析發語詞中的各語音音段的時長，了解發語詞句主要的時序架構特性，考驗兩種可能的等時單位的相鄰變異性，得到支持母音起始單位時長的等時性質的證據。本研究的這些發現是否可以支持國語言語產生以音節計時的節律特性呢？這個答案是值得再深入探討的。因為母音起始單位應是屬於跨音節的單位，它跨越了傳統定義的音節間的界線。對於一個音節的定義傳統上通常是以一個子音開始加上母音或韻母為止。在實際上說話者語音串流中，計時單位的組成卻並非只是以子音為起點而是以母音為起點那樣單純。或許在語言製造過程中語言計畫層次和言語動作時序規劃層次是使用不同的組成單位或劃分，因為有不同層次因素的考量。語言計畫層次需考慮到的是詞義和音韻單位的匹配，而在言語時序動作規劃層次上需顧及音節和音節間動作上串接的和諧性、語速的均質性以及整體語調節律感等因素，會需要跨音節的單位來以維持說話者的語音串流中計時單位的等時性，而母音起始單位屬於在言語動作規劃層次上的計畫單位。由母音起始單位穿接起語句中一個接著一個的音節成為一個具有自然節律感的句子。

統整本研究三個實驗的結果可知無論是在語音製造方面或是語音知覺方面以母音為起始點的計時單位皆是較具有等時性優勢的。不管是在何種說話語速或發語單位下，以母音起始單位的計時皆較有等時的優勢性。可推測國語韻律的計時是以母音為中心，以母音為計時的起點，計時單位的劃分是跨音節性的，是由母音開始加上後續的停頓和後接音節的子音時長，而單位等時性主要是靠調整母音或是音節間停頓或是靜默的時長來維持。

參考文獻

- 曾進興（1997）。〈漢語以音節作為切割單位的新證據〉。《國家科學委員會研究彙刊：人文與社會科學》，7(4)，612-631。
- 鄭靜宜（2004）。〈不同發語單位、言語速度和發語詞位置對國語音段時長的影響〉。《南大學報》，39，161-185。
- Abercrombie, D. (1967). *Elements of General Phonetics*. Edinburgh: Edinburgh University Press.
- Allen, G. D. (1972a). The location of rhythmic stress beats in English: An experimental study I. *Language and Speech*, 15, 72-100.

- Allen, G. D. (1972b). The location of rhythmic stress beats in English: An experimental study II. *Language and Speech*, 15, 179-195.
- Assmann, P. F. (1995). The role of formant transitions in the perception of concurrent vowels. *Journal of the Acoustical Society of America*, 97, 575-584.
- Beckman, M. (1982). Segment duration and the 'mora' in Japanese. *Phonetica*, 39, 113-135.
- Boersma, P., & Weenick, D. (1999). Praat: Doing phonetics by computer. Retrieved from <http://www.fon.hum.uva.nl/praat>.
- Buxton, H. (1983). Temporal predictability in the perception of English speech. In A. Cutler & D. R. Ladd (Eds.), *Prosody: Models & Measurement* (pp. 111-121). Berlin: Springer-Verlag.
- Classe, A. (1939). *The Rhythm of English Prose*. Oxford: Basil Blackwell.
- Couper-Kuhlen, E. (1990). *Discovering rhythm in conversational English: Perceptual and acoustic approaches to the analysis of isochrony*, KontRI Working Paper No. 13, University of Konstanz, Fachgruppe Sprachwissenschaft.
- Couper-Kuhlen, E. (1993). *English Speech Rhythm: Form and Function in Everyday Verbal Interaction*. Amsterdam: Benjamins.
- Dauer, R. M. (1983). Stress-timing and syllable-timing reanalyzed. *Journal of Phonetics*, 11, 51-62.
- de Jong, K. J. (1994). The correlation of P-center adjustments with articulatory and acoustic events. *Perception & Psychophysics*, 56, 447-460.
- Fowler, C. A. (1979). "Perceptual centers" in speech production and perception. *Perception & Psychophysics*, 25, 375-388.
- Grabe, E., & Low, E. L. (2002). Durational variability in speech and the rhythm class hypothesis. In C. Gussenhoven & N. Warner (Eds.), *Laboratory Phonology VII* (pp.515-546). The Hague: Mouton de Gruyter.
- Han, M. S. (1994). Acoustic manifestations of mora timing in Japanese. *Journal of the Acoustical Society of America*, 96, 73-82.
- Harsin, C. A. (1993). Acoustics of perceptual centers in speech. *Journal of the Acoustical Society of America*, 94, 1864.
- Harsin, C. A., & Green, K. P. (1994). Perceptual centers as an index of speech rhythm. *Journal of the Acoustical Society of America*, 96, 3350.
- Lehiste, I. (1973). Rhythmic units and syntactic units in production and perception. *Journal of the Acoustic Society of America*, 54, 1228-1234.
- Lehiste, I. (1977). Isochrony reconsidered. *Journal of Phonetics*, 5, 253-263.
- Low, E. L., Grabe, E., & Nolan, F. (2000). Quantitative characterizations of speech rhythm: Syllable-timing in Singapore English. *Language and Speech*, 43(4), 377-401.
- Milenkovic, P. (2004). TF32 [Speech Acoustical Analysis Software]. University of Wisconsin-Madison.
- Morton, J., Marcus, S., & Frankish, C. (1976). Perceptual centers. *Psychological Review*, 83, 405-408.
- Pike, K. (1945). *The Intonation of American English*. (2nd ed.). Ann Arbor, MN: University of Michigan Press.
- Port, R. F., Dalby, J., & O'Dell, M. (1986). Evidence for mora timing in Japanese. *Journal of the Acoustical Society of America*, 81, 1574-1585.
- Pointon, G. E. (1980). Is Spanish really syllable-timed? *Journal of Phonetics*, 8, 293-304.
- Roach, P. (1982). On the distinction between 'stress-timed' and 'syllable-timed' languages. In D. Crystal (Ed.), *Linguistic Controversies* (pp. 73-79). London: Edward Arnold.
- Stevens, K. N., & Blumstein, S. E. (1978). Invariant cues for place of articulation in stop consonants. *Journal of the Acoustical Society of America*, 64, 1358-68.
- Tseng, C. -H., Huang, K.-Y., & Jeng, J.-Y. (1996). The role of syllable in perceiving spoken Chinese. *Proceeding of National Science Council (ROC) Part C: Humanity and Social Science*, 5(1), 181-204.
- Tseng, C.-Y., & Lee Y.-L. (2004). Speech rate and prosody units: Evidence of interaction from Mandarin Chinese. *Proceedings of the International Conference on Speech Prosody*, Nara, Japan, 251-254.
- Van Santen, J. P., & Shih, C. L. (2000). Suprasegmental and segmental timing models in Mandarin Chinese and American English. *Journal of the Acoustical Society of America*, 107(2), 1012-1026.
- Wen, B., & Wioland, F. (1982). Is French really syllable-timed? *Journal of Phonetics*, 10, 193-216.

附錄一：實驗一的語料列表

編號	始音素	四音節	句子/ 多音節片語
1	ㄣ	背上肩去	把弟弟背上肩去
2	ㄣ	潑到人家	潑水潑到人家
3	ㄇ	媽媽好心	媽媽好辛苦
4	ㄘ	飛的好高	小鳥飛的好高
5	ㄨ	低頭看地	低頭看地上螞蟻
6	ㄘ	特別愛看	特別愛看小紅花
7	ㄗ	泥土味道	充滿了泥土味道
8	ㄌ	綠油油的	綠油油的草地
9	ㄍ	給我錢買	給我錢買東西吃
10	ㄨ	褲子濕濕	弟弟的褲子濕濕
11	ㄆ	哈哈笑我	張開嘴哈哈笑我
12	ㄌ	雞腿肉香	雞腿肉香噴噴
13	ㄎ	汽水好喝	汽水好喝極了
14	ㄒ	吸一吸奶	給他吸一吸奶嘴
15	ㄘ	蜘蛛網多	角落裡蜘蛛網多
16	ㄨ	吃飯時間	中午吃飯時間到
17	ㄆ	時間到了	回家時間到了
18	ㄉ	乳酪好香	乳酪好香也好濃
19	ㄆ	紫色花兒	紫色花兒開滿地
20	ㄘ	刺到腳底	被圖釘刺到腳底
21	ㄌ	撕破簿子	不小心撕破簿子
22	ㄌ	魚肉多刺	魚肉多刺要小心
23	ㄨ	烏魚子飯	烏魚子飯也好棒
24	ㄨ	一起走到	一起走到動物園
25	ㄣ	安全門口	安全門口的暢通
26	ㄨ	哀嚎聲起	姐姐滑倒哀嚎聲起
27	ㄨ	凹凸的地	凹凸的地面不平
28	ㄨ	歐式花園	歐式花園好漂亮
29	ㄩ	阿里山上	阿里山上看日出
30	ㄌ	耳朵聾了	耳朵聾了聽不到
31	ㄨ	盆子外面	盆子外面很漂亮
32	ㄨ	骯髒的狗	骯髒的狗惹人嫌
33	ㄌ	蜂巢旁的	蜂巢旁的小蜜蜂

附錄二：實驗一的兩篇短文篇章

忙碌的星期日：一大早，弟弟的褲子濕濕，原來他尿床了。媽媽給他吸一吸奶嘴。今天全家決定大掃除。媽媽好辛苦。角落裡蜘蛛網多，爸爸把弟弟背上肩去清理，還要保持安全門口的暢通。一不小心，哥哥潑水潑到人家。妹妹被圖釘刺到腳底；姐姐滑倒哀嚎聲起，哥哥張開嘴哈哈笑她。小狗不小心撕破簿子了。今天家裡的每個人都很忙碌，真是一個忙碌的星期日啊！

動物園之旅：今天早上全家一起來到動物園，看到蜂巢旁的小蜜蜂，也有小鳥飛的好高。動物園裡的歐式花園好漂亮，雖然凹凸的地面不平，卻充滿了泥土味道。綠油油的草地上，紫色花兒開滿地，但是我特別愛看小紅花，哥哥正低頭看地上螞蟻。中午吃飯時間到，媽媽給我錢買東西吃，雞腿肉香噴噴，乳酪好香也好濃，烏魚子飯也好棒，但是魚肉多刺要小心。飯後汽水好喝極了！吃完飯後，回家時間到了，我們依依不捨地離開動物園！

附錄三：實驗二的二十六組語料列表

	四音節	五音節	句子		四音節	五音節	句子
1	背上肩去	他背上肩去	把弟弟 <u>背上肩去</u>	14	吸一吸奶	他吸一吸奶	給他 <u>吸一吸奶嘴</u>
2	潑到人家	他潑到人家	潑水潑到人家	15	蜘蛛網多	那蜘蛛網多	角落裡 <u>蜘蛛網多</u>
3	媽媽好心	他媽媽好心	他媽媽好辛苦	16	吃飯時間	到吃飯時間	中午 <u>吃飯時間到</u>
4	飛的好高	它飛的好高	小鳥飛得好高	17	時間到了	他時間到了	回家時間到了
5	低頭看地	他低頭看地	低頭看地上螞蟻	18	乳酪好香	那乳酪好香	乳酪好香也好濃
6	特別愛看	他特別愛看	特別愛看小紅花	19	紫色花兒	那紫色花兒	紫色花兒開滿地
7	泥土味道	那泥土味道	充滿了 <u>泥土味道</u>	20	刺到腳底	他刺到腳底	被圖釘 <u>刺到腳底</u>
8	綠油油的	那綠油油的	綠油油的草地	21	撕破簿子	他撕破簿子	不小心 <u>撕破簿子</u>
9	給我錢買	他給我錢買	給我錢買東西吃	22	魚肉多刺	那魚肉多刺	魚肉多刺要小心
10	褲子濕濕	他褲子濕濕	弟弟的褲子濕濕	23	烏魚子飯	那烏魚子飯	烏魚子飯也好棒
11	哈哈笑他	她哈哈笑他	張開嘴 <u>哈哈笑他</u>	24	一起走到	他一起走到	一起走到動物園
12	雞腿肉香	那雞腿肉香	雞腿肉香噴噴	25	安全門口	那安全門口	安全門口的暢通
13	汽水好喝	那汽水好喝	汽水好喝極了	26	蜂巢旁的	那蜂巢旁的	蜂巢旁的小蜜蜂

The Timing Isochrony of Spoken Mandarin at Different Speaking Rates

Jing-Yi Jeng

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

The behavior of speech timing plays an important role in the process of speech perception and production. The source of speech rhythm is based on the “isochrony”. Isochrony unit is a hypothetical unit which has an equal duration interval existing in the speech stream. The users of a language base on the isochrony unit to process the speech information during speech process. The purpose of this study is to investigate two candidates of speech timing units in connected Mandarin. The first candidate is “the whole syllables unit (pause/gap + CV)” and the other is “the rime onset unit (rime + pause/gap + C)”. This study consists of three experiments; the first two experiments are production tasks, and the third is a perceptual task. In the first experiment, 30 mandarin speakers read 33 sets of Chinese phrases and sentences and two passages at five speaking rates by using a metronome to modulate the speaking rates. Segmental durations were obtained via acoustic analysis, and then, the pairwise variation index (PVI) was computed for the two candidate units to compare the degree of isochrony. The results show that the rime onset unit had significantly less PVI value than the whole syllable unit, and the two kinds of units had the maximum difference at the slow speaking rates, but at the fast speaking rates there was no difference between them. The second experiment is the replicate of the first one, and the major difference is the method of speaking rate control. The second experiment had similar results as the first one, again, the rime onset unit had the less PVI. In the third experiment, 18 listeners estimated the degree of naturalness of 24 sentences with different PVI values. The results show that the degree of naturalness is highly correlated with PVI values; moreover, for the PVI calculated by the rime onset unit, the correlation is stronger than that by whole syllable unit. These findings support the rime onset unit is the isochrony unit in speech timing process of Mandarin. The results suggest the basic unit of Mandarin speech timing is related with rime onset, where is also known as the locus of “perceptual center”.

Keywords: *rhythm, speech timing, isochrony, pairwise variation index (PVI), perceptual center*

