

人工電子耳兒童之國語聲調指認與表達： 以六歲以上兒童接受電腦化測驗評量的結果為例

江源泉

台北護理健康大學語言治療與聽力學系

摘 要

現有電子耳國語（Mandarin）的聲調研究結果常包含三歲以下或電子耳經驗僅數月兒童的聲調表現。本研究以聽力室內的電腦化測驗評量 12 位 6-15 歲、電子耳經驗兩年以上、習語前失聰兒童的聲調指認和表達。實驗一聲調指認，由兒童在聲調刺激與不同聲調圖片搭配組成的 12 情境下，以二選一方式指認單音節語詞的聲調，但需兩次指認嘗試皆正確才計分。所有語詞皆以對話形式錄製。實驗二聲調表達，將錄自實驗一電子耳兒童的語詞，由評審以四選一指認方式評量其聲調。兒童聲調指認的團體得分為 58%，低於文獻記載的 67-73%，且 2、4 聲的指認受測驗情境影響，並以 4 聲搭配 1、3 聲圖片之得分別為最高（72%）和最低（30%）。兒童聲調表達的團體得分為 67%，高於文獻記載的 47-53%，且 1、4 聲得分（79%、71%）高於 2、3 聲（59%、59%），且聲調間的混淆模式各有不同。個人表現方面，指認和表達得分呈高度相關，且近半數兒童，雖有 3-9 年的電子耳經驗，仍有對聲調教學介入的特殊需求。

關鍵詞：人工電子耳、國語、聲調辨識、聲調表達、習語前失聰

* 本研究受國科會專題計劃 NSC 97-2410-H-134-027、NSC 99-2410-H134-009 補助。承新竹縣竹東國小范珮筠、陳亮希、山崎國小蘇淑菁、史玉詩四位聽障巡迴教師、12 位參加學生及其家長的熱心參與協助，研究助理黃姿慎、李承龍的資料蒐集分析及電腦程式及介面設計，特此一併致謝。

壹、緒論

隨著助聽科技的進步，目前習語前失聰的兒童多借助人工電子耳創健聽能和口語。根據韓福榮、曹光文與江以文（2015）的報告，國內自 1985 年引進電子耳植入技術，三十餘年來接受植入者約兩千人，其中多數為先天性失聰兒童。雖然相對於使用助聽器，電子耳的植入讓多數兒童在口語學習（如 Tobey, Geers, & Brenner, 1994; Tye-Murray, Spencer, & Woodworth, 1995; Tye-Murray & Kirk, 1993）及清晰度的提升（Tye-Murray et al., 1995; Osberger, Robbins, Todd, & Riley, 1993）有更佳表現，但對國語、廣東話等超音段（suprasegmental）的聲調表現似乎助益有限（如 Ciocca, Francis, Aisha, & Wong, 2002; Peng et al., 2004; Han et al., 2007）。這項缺失值得重視，因聲調特徵擔負著區辨語意的功能，如國語的“ㄊㄨ-ㄊㄨ-ㄊㄨ'”（“湯-糖”）。雖然在安靜中，聲調的角色可因語句主題脈絡的線索而顯得多餘（Wang, Shu, Zhang, Liu, & Zhang, 2013），但在噪音中，即使是正常聽者仍靠對聲調的穩定感知以理解語意（Chen, Wong, & Hu, 2014）。因此，聲調的感知（或稱辨識）（perception）和表達（或稱產出）（production）直接影響溝通的效果。本研究以電腦化測驗評量 6-15 歲、使用電子耳兩年以上兒童的這兩項表現，目的在以適當的實驗控制和評量標準，記錄此學齡階段電子耳兒童的聲調表現水準，並了解他們聽語方面的特殊需求。

一、人工電子耳與聲調辨識

習語前失聰的電子耳兒童藉由電子耳對聽神經的刺激，開通了聽覺路徑，進而發展出聲調辨識和表達的能力。聲調辨識主要靠感知說話者的基頻（fundamental frequency，簡稱 *F0*）或解析出其諧音（harmonics）（Fu

& Zeng, 2000; Massaro, Cohen, & Tseng, 1985）。*F0* 是發聲機轉在成聲（phonation）時，聲帶每秒的開合次數。男、女性嗓音的 *F0* 分別約 100 和 200 Hz，屬於人類正常聽覺場域（auditory field）的低頻音，而國語聲調變化範圍也在此基頻附近。根據多項研究報導，國語電子耳兒童的團體聲調指認得分約為 70%（見下文），顯示他們尚未能充分利用說話者的聲調線索，而原因可能與電子耳電極束的長度有關。由於目前標準款電極束經植入後，深度只達耳蝸由外而內捲曲 2.75 圈中外圍的 1.25 圈（Boyd, 2011），而所對應的刺激頻率都在 500 Hz 以上（Greenwood, 1990）。如此一來，*F0* 訊息即落在電極無法提供刺激的頻率範圍，自然聲調訊息就無法抵達分布於耳蝸內基底膜下方的聽神經。在此限制下，使用者或許可經由解析 *F0* 在 500 Hz 以上的諧音（harmonics）而取得 *F0* 線索，但如此所得之頻率線索未必精準（Wei, Cao, & Zeng, 2004）。近年推出的加長型電極雖可把刺激延伸至全耳蝸，但對改善低頻信號接收的效益不明（Boyd, 2011; Landsberger, Mertens, Punte, & Heyning, 2014）。另一方面，文獻中卻不乏以標準款電極束即達到近乎完美的聲調辨識的案例（如 Peng et al., 2004）。由此可知，聲調辨識表現並非僅由電子耳科技決定。

二、國語聲調辨識與表達的正常發展

聲調辨識和表達兩項能力都須經發展而成熟，但所需時間只有出生後短暫的幾年。此兩項能力的正常發展可參考下列兩研究。Wong、Schwartz 與 Jenkins（2005）以二選一的圖片指認，評量三歲兒童對國語四聲調六對比（1-2、1-3、1-4、2-3、2-4、3-4 聲）的辨識，所得指認結果除 3 聲低於 70% 外，其他聲調都接近 90%。Zhou、Huang、Chen 與 Xu（2013）以百位以上 3-9 歲的兒童為研究

樣本，所得平均指認得分是 98.7%，顯示大多數兒童在 3 歲時的聲調辨識已達成熟。在表達方面，Wong 等所得也以 3 聲最差（44.04%），其他聲調都達 70% 以上，超越了 66.7% 的發展「穩定」水準（Clumeck, 1977; Zhu, 2002）。Zhou 等所得的平均得分為 94.8%（標準差 = 9%），但未提供個別聲調的資料。上述兩研究可得的結論是：（1）即使是正常兒童，此兩種能力的穩定發展最少也需三年；（2）三歲左右兒童聲調的聽和說表現，皆以 3 聲為最困難。

三、電子耳兒童的國語聲調辨識

國語電子耳兒童的聲調辨識，過去曾針對不同特徵的受試或測驗任務進行評量。若受試包含幼童，則以六聲調對比的圖片為刺激，而以二選一為指認任務。所得差異可見於團體得分和各聲調對比得分的高低排序。Peng 等（2004）是少數以現場口語施測的研究，也是探討電子耳兒童聲調聽與說的雙主題研究。參加兒童的年齡為 6-13 歲、電子耳平均使用 3.7 年（範圍 1.6-6.4 年）。所得平均指認得分是 73%，且整體趨勢是含 4 聲的對比得分高於其他對比。此後研究多以電腦化二選一測驗評量，且受試條件大幅放寬：年齡方面，下至 2.1-3.5 歲，上至 16.2-21.5 歲（Han et al., 2009; Xu et al., 2011; Zhou et al., 2013）；電子耳經驗方面，可縮短至 0.1 年（Xu et al., 2011; Zhou et al., 2013）。其中 Zhou 等（2013）是雙主題的大型研究，採用受試 107 位，但因多數（53%）年齡在 2-4 歲，所以仍是以幼童為主的研究。研究結果方面，除 Zhou 等外，以上各研究所得的團體表現都和 Peng 等人所得一致，也就是得分在 71%-82%，而六對比的得分排序也以含 4 聲的對比得分最高。其中 Han 等（2009）的得分排序雖以 1-3 對比最高，但其他五對比的排序也和 Peng 等相同。至於例外的 Zhou 等所得

平均表現：團體平均得分降為 67.3%，且聲調對比間並無顯著的得分差異。雖然以上各研究結果不盡一致，但都發現受試的指認表現與電子耳植入年齡有關。Zhou 等則另外發現，也與植入經驗長短有關。

上述二選一聲調指認研究提供兒童的嘗試次數不同，所以得分的解讀需經二項式（Binomial test）檢驗（Thornton & Raffin, 1978）。其中 Peng 等（2004）和 Zhou 等（2013）兩研究的機率水準皆為 60%，Han 等（2009）和 Xu 等（2011）為 63%。對照後以 Zhou 等大型研究的平均得分（67.3%）最接近機率。另外，根據該研究的中位數（median）64.6%，可推知 107 位兒童中有多於半數 54 位的得分低於平均（67.3%），且有接近 54 位的得分並未超出機率（60%）。因該研究是現有唯一的大型研究，所以結論不容忽視。

聲調評量方式除二選一外，另有四選一。此類評量適合年紀較長的受試。Chang、Chang、Lin 與 Luo（2016）評量 15 位 10-20 歲語前失聰、具一年以上電子耳經驗聽者的韻母聲調辨識。所得團體平均為 67.5%，與 Zhou 等（2013）大型研究的結果相似。個別聲調結果，包括以 4 聲得分最高（約 84%）、3 聲最低（約 42%）及 2、3 聲間混淆最頻繁等，都與其他二選一研究結果一致。

四、國語電子耳兒童的聲調表達

聲調表達曾以客觀的（一）聲學分析和主觀的（二）聽知覺評分結果評量。

（一）聲學分析

聲學分析是根據 F_0 曲線的起點和終點頻率在座標平面的分佈狀況，檢視語者變化聲調的能力（Barry & Blamey, 2004）。若各聲調的頻率分佈面積愈小且聲調間重疊部分愈少，則表示說話者愈有能力製造出可供區辨的聲調。Zhou 與 Xu（2008）和 Zhou 等（2013）兩項聲學分析研究以平均年齡約 5 歲、電子

耳平均經驗 1.3-1.7 年的兒童為受試。所得頻率數據點的分佈不僅重疊度高，分佈面積也比正常兒童的分佈面積大，且未隨年齡或電子耳經驗的增加而縮小等特徵。兩分析研究的共同結論是，電子耳兒童無法有效表達聲調間的差異。值得注意的是，兩研究的受試平均年齡和電子耳經驗都有限，甚至包含 3 歲以下、電子耳經驗極短的兒童 (0.3 年)。另一聲調表達評量方式是人工神經網路 (artificial neural network) (Zhou, Zhang, Lee, & Xu, 2008)。因輸入數據是來自 F_0 曲線，所以在此廣泛歸為聲學分析。Xu 等人 (2011) 以此方式為年齡在 2.1–21.5 歲、電子耳經驗在 0.1–8.1 年的使用者進行聲調評分，所得團體平均為 52%。這樣的水準雖不理想，但相較於上述 Zhou 與 Xu 及 Zhou 等兩聲學分析所得，聲調的可區辨性已明顯提升，而最可能提升的原因應是增加了 13-21 歲的受試和較長的電子耳平均經驗。上述三項客觀評量結果顯示聲調表達的評量需參照受試的年齡和使用經驗。

(二) 聽知覺研究

聲調表達的聽知覺研究都採錄音回播 (playback) 方式，即先以圖片或口頭示範誘發兒童單音節語詞並進行錄音，再將錄音內容製作為測驗，播放給常人進行主觀評量。國語電子耳聲調的聽知覺研究曾以 (1) 主觀等級評分或 (2) 電腦化四選一聲調指認方式評分。Peng 等人 (2004) 的等級評分以三位專業評審，在已知表達語詞的情況下，對兒童的聲調給予 1–5 分評量，再轉換成百分等級 (即 1=20%，5=100%)，所以評量的最低得分是 20%。電腦化指認評分則以 Han 等 (2007)、Zhou 與 Xu (2008) 及 Zhou 等 (2013) 三項研究為代表 (下稱三電腦化評量)。兩類評分方式所得平均表現雖然近似 (等級評分是 53%、三電腦化評量得分是 46.8%-48.7%)，

但各聲調的得分排序卻有明顯差異。等級評分結果分高、低兩組，高組是 1、4 聲 (得分 62.13%、62.22%)，低組是 2、3 聲 (得分 42.13%、45.89%)，兩組間得分差異顯著，但組內則無顯著差異。三電腦化評量所顯示的四聲調得分差異明顯，皆以 1 聲得分最高 (68.4%–71.8%)、4 聲次之 (50%–55.3%)、3 聲再降 (41.1%–48.1%)、而以 2 聲最低 (19.4%–26.5%) (機率水準為 25%)。此 2 聲顯著低於 3 聲的結果不僅少見，也與正常聲調發展資料不合 (Wong et., 2005)。整體而言，兩類評分結果除團體得分外，並不一致。

貳、研究目的

以上回顧的文獻除 Peng 等人 (2004) 的雙主體聲調研究外，都含不適合擔任測驗任務的兒童受試。如指認方面，雖然 Ching (1984) 已指出，即使聽力正常的兒童在四歲時仍無法可靠的重複指認單獨呈現的聲調，但 Han 等 (2009)、Xu 等 (2011)、Zhou 等 (2013) 的指認研究仍有年僅 2.1-3.5 歲的兒童。同樣的，在表達方面，即使已知正常兒童的聲調穩定表達出現在三歲以上，而 3 聲則要更久 (Clumeck, 1977; Zhu, 2002)，但三電腦化評量 (兼含兩項聲學分析) 依然採用了 2.4-2.9 歲的受試。另外，在電子耳經驗方面，三電腦化評量還用了植入僅 0.1-0.3 年的兒童。因此目前有待回答的問題是，如果三電腦化評量也以成熟度如 Peng 等的兒童為受試，是否所得將與 Peng 等的等級評量所得一致？

除受試條件差異引發的疑議外，上述文獻仍有多項研究方式的限制。例如：(1) 二選一的指認研究機率高達 50%，很可能是把平均得分推高到 70% 以上的主因，且除 Zhou 等 (2013) 外，各研究都僅提供每聲調刺激一次的指認機會，所以無法確認測驗信度。(2)

各二選一研究皆未探討搭配圖片對指認的可能影響。如各研究結果都無法回答「1 聲標的搭配 2 聲圖片」和「2 聲標的搭配 1 聲圖片」、或「1 聲標的搭配 2、3、或 4 聲圖片」的指認結果是否相同，而這些問題的答案將有助於理解「二選一」和「四選一」指認反應的差異處。(3) 有關 3 聲刺激的錄音形式，上述電腦化的指認研究皆未特別描述，因此推測都採用朗誦式的「全上」，但在真實對話情境，除非特別強調，否則連單字詞都是以 *F0* 曲線為低降式的「前上」的方式表達。因此，以「全上」為刺激的研究所得並不能廣泛描述電子耳兒童對 3 聲的辨識表現。另外，由於二選一的聲調指認都是在對比情境下評量，所以 3 聲的刺激型態可能也間接影響了其他聲調的指認表現。(4) 針對聽與說的相關性，Peng 等 (2004)、Xu 等 (2011) 及 Zhou 等三項雙主題研究所得差異頗大，*r* 值分別為 .44、.81 和 .56，且 Peng 等的資料若移除異數 (outliers) 後，甚至未達相關。雖然國語電子耳研究的聽、說相關例證不足，但英語句型的語調研究 (Peng, Tomblin, & Turner, 2008) 和日語語句的情緒研究 (Nakata, Trehub, & Kanda, 2012) 都支持聽與說的相關性。有鑑於電子耳兒童的聲調表達源於聽覺路徑的創健，以致於指認和表達結果的相關性是理論上不可迴避的議題，因此有必要再探討。

本研究以電腦化測驗再度評量六歲以上且電子耳經驗兩年以上兒童的聲調表現。具體目標在於回答下列問題：(1) 現有文獻約 70% 的平均聲調指認得分，是否隨本研究評分標準提高而降低？(2) 以「前上」式取代傳統「全上」式的 3 聲刺激，對指認得分有何影響？(3) 聲調表達結果與受試條件的關係如何？(4) 隨著聲調指認評分標準的改變，是否牽動聲調指認和表達結果的相關性？

參、研究方法

實驗一、聲調指認

(一) 受試

本研究受試共 12 位，含七男五女，年齡在 6.8-14.5 歲 (平均=10.8，標準差= 2.61)。其中七位為國小生，五位為國中生，皆就讀於普通班，但接受融合教育或聽覺障礙巡迴服務。所有兒童皆於兩歲前失聰且單耳接受電子耳植入，植入年齡在 2.6-9.3 歲 (平均= 4.60，標準差= 2.32)，至評量時所累積的使用經驗是 2.5 - 11.1 年 (平均= 6.0，標準差= 2.46)。兒童平日於另一耳皆配戴耳掛型助聽器，但在參與本實驗的過程中，都未配戴。兒童中有八位於植入後，接受過 1-6 年的語言訓練。兒童的父母都是聽人；平日溝通全部仰賴口語。除可矯正的近視外，兒童皆無智力、情緒、注意力/過動等障礙。本實驗另加測三位 6-7 歲聽力正常兒童，以估計六歲以上正常兒童之基本聲調指認水準。實驗進行前，皆向參與兒童及其家長解說研究目的和過程，並取得家長知情同意書。與研究相關的電子耳兒童的背景資料，列於表 1。

表 1

人工電子耳受試背景資料

受試	受測年齡 (年)	植入年齡 (年)	電子耳經驗 (年)	電子耳廠牌	復健(年/ 小時(週))
S1	9.3	2.6	6.7	Cochlear Nucleus	6/2
S2	13.5	4.4	9.1	Cochlear Nucleus	3/1
S3	9.5	4.1	5.4	Cochlear Nucleus	無
S4	10.6	4.3	6.3	Cochlear Nucleus	無
S5	7.8	2.5	5.3	Cochlear Nucleus	2/2
S6	8.8	3.2	5.6	Cochlear Nucleus	2/1
S7	13.5	5.6	7.9	Cochlear Nucleus	無
S8	13.5	9.3	4.2	Medel	無
S9	6.8	4.3	2.5	Cochlear Nucleus	2/2
S10	9.1	2.4	6.7	Medel	6/1
S11	12.3	9.0	3.3	Cochlear Nucleus	1/1
S12	14.5	3.4	11.1	Cochlear Nucleus	5/2

(二) 聲調指認刺激

本實驗採用二選一圖片指認的測驗任務。雖然最直接的任務是四選一的國字指認，但考量受試中國小一年級新生(S9)的識字量不足，若僅播放聽覺刺激，恐不易維持測驗時的專注力；若逕以國字搭配，又逢國語中「媽、麻、馬、罵」這類語意單純具體的四聲調組合，數量有限的問題，因此沿用圖片指認。指認的刺激(stimuli)皆為單音節語詞，挑選方式是根據國語四聲調的六種對比，每對比挑選五組語詞為代表(token)，共60個語詞(6對比 x 2聲調 x 5單音語詞)(見表2)。若以聲調區分，每聲調有15個語詞。實驗時，刺激是以聽覺刺激配合視覺圖片對比刺激方式呈現。聽覺刺激即60個語詞，每次嘗試(trial)播放一個，同時搭配圖片一組呈現於電腦螢幕，為視覺刺激。視覺刺激如同二選一的試題紙本，每題兩張圖片並列，所代表的兩個語詞皆為最小差異詞組(minimal pair)，即音節結構(聲母、韻母)完全相同，

但聲調不同，如圖1所示的“3(ㄇㄣˊ) - 傘(ㄇㄣˊˇ)”。

兩圖片之一對應的是所播放的聽覺刺激，也就是「標的」(target)選項，剩餘的一張則用以混淆。每「標的」的指認都在搭配其他三聲調圖片的情境下進行，例如，1聲標的指認包括1-2聲(如「湯-糖」)、1-3聲(如「3-傘」)和1-4聲(如「山-扇」)等三種圖片組合的測驗情境，以致於指認實驗共含12個測驗情境(4標的x3聲調圖片)。若就圖片而言，每種組合可供兩個聽覺刺激搭配，如圖一所示的「3-傘」出現在電腦螢幕時，喇叭播放的聽覺刺激可以是「ㄇㄣˊ」或「ㄇㄣˊˇ」。視覺刺激的圖片由分別對應60個聽覺刺激的60張彩色實物照片所組成。照片的挑選皆以兒童熟悉且圖象單純為原則，如“ㄊㄧˊㄙㄜˊ”是以“皮鞋”，而不以“拖鞋”或“高跟鞋”為代表。

表 2

聲調指認和表達的標的語詞（以最小差異詞組方式陳列）

組別 最小差異	一	二	三	四	五
1-2 聲	鴨-牙	豬-竹	窗-床	七-旗	湯-糖
1-3 聲	煙-眼	鉤-狗	針-枕	衣-椅	三-傘
1-4 聲	山-扇	蛙-襪	香-象	書-樹	腰-藥
2-3 聲	鞋-血	壺-虎	鼻-筆	鈴-領	魚-雨
2-4 聲	爐-鹿	橘-鋸	河-鶴	驢-綠	毛-帽
3-4 聲	剪-箭	五-霧	米-蜜	土-兔	腳-叫

（三）指認刺激的製作

60 個單音節語詞的聽覺刺激分男、女兩版本在隔音室內以數位錄音設備（Sony Hi-MD Walkman, MZ-NH1）分開錄製存檔。男（21 歲）、女（25 歲）兩位語者的國語構音正常，無特殊腔調。為維持刺激的真實性，各語詞皆以自然對話方式錄製，所以 3 聲只保留「前上」的頻率低降部分。各聲調的長度接近，約 300-400 毫秒。錄音時，每語音取樣三到五次，之後由實驗人員以語音分析軟體 CoolEdit 2000，根據波形、配合聽覺檢視，挑出效果最佳的一個，儲存成獨立電腦音檔。

（四）實驗步驟

測驗在含控制室和測驗室的隔音聽力檢查室（ETS-Lindgren Acoustic Systems）進行。聽覺刺激由位於控制室的筆記型電腦（Sony Viao）播放，輸入聽力檢測儀（GSI61）後，以雙頻道各 50 dB HL 的音量分送至測驗室內的兩個喇叭播放。喇叭位於受試者座位的前方左、右各 45 度、距離約 1 公尺處。由於測驗時，聽覺刺激同時自兩喇叭播放，故到達受試耳廓後電子耳麥克風的實際音量應接近 56 dB HL，符合一般距離一公尺的正常對話音量（Pearsons, Bennett, & Fidell, 1977）。測驗前，每位受試都在無反應回饋情況下嘗試 12 題（6 對比 × 2 標的聲調）與正式測驗

不同的反應練習，以熟悉測驗步驟及回答方式。練習前，實驗人員皆確認音量大小適當。

實驗以個別施測方式進行。測驗以二選一強迫選擇方式（two-alternative-forced-choice test paradigm, 2AFC）進行，全程由電腦控制。因考慮兒童對電腦介面操作能力不足，因此由實驗人員操作相關流程並鍵入選項。測驗時，受試者坐在聽力檢查室中央，面對座前矮桌上離面 30 公分的 19 吋電腦螢幕。測驗前，實驗人員朗讀指導語如下：「在下面的測驗，你會聽到很多詞。每聽到一個詞，請用手在電腦螢幕上出現的兩張圖片中，指出你聽到的那一張。如果不確定，也要好好的猜是哪一張。」每次嘗試從電腦播放聽覺刺激開始，電腦螢幕隨即呈現左右並排的視覺刺激圖片一組。再以圖 1 為例：當播放的聽覺刺激是「ㄇㄣˊ」或「ㄇㄣˊˇ」時，電腦螢幕同步出現「3」和「傘」兩圖片；若聽者判斷所聽到的是「ㄇㄣˊ」，則應選左圖「3」，否則應選右圖「傘」。圖組中央上方的對話視窗內標有箭頭，分別指向左邊「1」號和右邊「2」號選項。受試者根據所聽到的音節，直接在螢幕上以手點選對應的左或右方圖片，而在旁的實驗人員隨即以滑鼠在對話方塊內點選對應左圖的「1」或右圖的「2」。答案輸入後，下一次嘗試隨即啟動。測驗共含兩輪，以男、

女版刺激各為一輪，每輪內播放 60 個聽覺刺激各兩次，兩次播放時，對應的兩張視覺刺激圖片位置交換。每位受試共進行 240 次嘗試（60 聽覺刺激 × 2 次 × 2 輪）。兩輪的順

序和每輪內聽覺刺激的播放順序，都採隨機安排。完成一輪測驗所需時間約 10 分鐘，兩輪中間為休息時間，受試的反應結果由電腦並計分，測驗結束後，每位都獲得禮品一份。



圖 1 聲調指認實驗的視覺刺激範例。圖組中央上方的對話視窗內標有箭頭分別指向左邊「1」號選項“3”和右邊「2」號選項“傘”；同步播放的聽覺刺激則是「ㄇㄣˇ」或「ㄇㄣˇ」

實驗二、聲調表達

（一）聲調錄音

1. 受試

受試是實驗一聲調指認測驗的 12 位電子耳兒童，因本實驗以他們被誘發產出的語音為回播評量的語料，所以又稱他們為語者。

2. 聲調誘發

聲調的誘發以實驗一的 60 張圖片為刺激，圖片內容對應每聲調 15 個單音節語詞（見表 2）。每張圖片下方皆印有代表圖片內容的國字和注音。雖然圖片已用於實驗一，但在錄音前，實驗人員仍要求兒童對所有圖片以單字詞命名，以確認命名正確。

聲調誘發在聲調指認的兩輪測驗中的休息時間進行，並同時進行錄音。錄音前、後語者皆獲得充分休息時間。錄音場地在具單牆隔音的聽力檢查控制室內。圖片刺激以每頁六張橫向排列方式，複製在十頁 A4 尺寸的卡紙上，上下重疊平置於桌面。錄音時，語者與實驗人員並排坐於桌前，麥克風則置於語

者口部前方約 10 公分的桌面。錄音開始，實驗人員以手指點指任一頁卡紙上的任一張圖片，每點指一次，受試即唸一次。唸完三次後，實驗人員換指同頁的另一張圖片，並重複上述程序；六張圖片都錄完後，換另一頁，直到十頁 60 張圖片全部錄完，過程約費時 15 分鐘。所得每圖片的三次語詞錄音，於事後分割成三個獨立數位音檔儲存。

（二）聲調評分

1. 評審及篩選標準

評審由 20 位成年男、女大學生擔任。聽力狀況皆自訴正常，且對聽覺障礙者的口語毫無經驗。實驗前，全部以 100% 的正確率通過 20 題四選一的國語聲調指認測驗，其聲調刺激取自實驗一男聲版的語詞。

2. 評分語料

評分語料是錄自 12 位電子耳語者的單音節語詞。實驗人員以軟體 CoolEdit 2000 將錄存的每音節三個音檔，根據音節的波形及聽覺效果，選出最佳的一個代表，並經音量

正規化處理後，依語者分製成 12 個各含有 60 個單音節語詞的分測驗。

3. 電腦化聲調評分及實驗步驟

兒童聲調評分是根據評審對 12 個分測驗進行聲調指認的結果。測驗地點、設備擺設、及評審的座椅位置和音量控制都與實驗一相同。20 位評審對測驗內的語詞，以四選一強迫選擇方式（4AFC）指認其聲調。每語詞播放時，電腦螢幕上同時出現「1」、「2」、「3」、「4」，代表聲調的四個選項。評審所得的指令是：「請根據您所聽到的語音，以滑鼠在螢幕上點選它的聲調。若不能確定，也請

點選最近似的聲調。」每次點選完成後，即自動進入下一題嘗試，每題播放一次，60 題指認結束後，即進入下一分測驗。分測驗和內部語詞刺激的播放順序都採隨機安排。每位評審共嘗試 720 次（60 語詞 × 12 分測驗）。每分測驗所需時間約 3 分鐘，全部 12 個分測驗完成約需 30 分鐘，評審可在一天內參加的任何分測驗之間休息，選擇結果皆由電腦蒐集。評審於課餘時間參加，全部過程跨兩週完成。

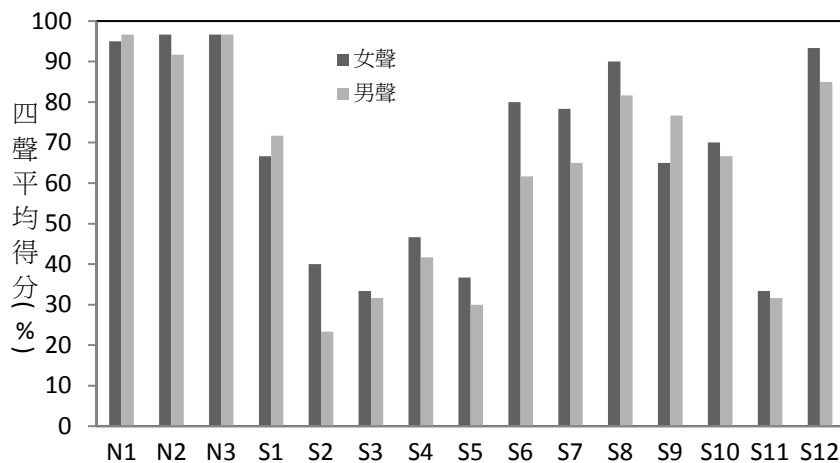


圖 2 三位正常兒童及 12 位電子耳受試對男、女聲調指認之個別得分 (%)

肆、結果

一、聲調指認

本實驗以二選一圖片為指認任務，但以每語詞的兩次嘗試都指認正確才計分。根據二選一任務的正確機率為 50% 計算，240 次嘗試的機率水準為 50.8%（二項式檢驗； $N = 240, p = 1/2, \alpha = 0.05$ ）。所得男聲版的平均得分是 55.55%（範圍在 23.33% - 85.00%），女聲版平均得分是 61.11%（範圍在 33.33% - 93.33%），兩版得分呈高度正相關（ $r = .93, p < 0.05$ ）。圖 2 展示的是個別受試對男（淺

色）、女聲調（深色）的指認結果。圖左（N1-N3）是三位 6-7 歲聽力正常兒童的得分，其餘（S1-S12）是電子耳兒童的得分。以下各項統計數據是先將兩聲版獨立計分，再平均算出個人得分。

（一）兒童的個別表現

根據圖 2，三位正常兒童的表現近乎完美，但電子耳兒童則差異明顯，得分範圍是 31.67% - 89.17%。其中未達機率（即 $< 50.8\%$ ）者有五位（S2、S3、S5、S11、S4），高於機率者有七位，其中得分在 60-79% 者有五位（S1、S6、S7、S9、S10），在 80-89% 有二位

(S8、S12)。指認表現與電子耳受試之年齡 ($r = .19, p = .56$)、植入年齡 ($r = -.00, p = .54$) 或使用經驗長短 ($r = .20, p = .99$)，都無顯著相關。

(二) 團體表現

1 至 4 聲的團體得分分別是 62.78%、54.44%、62.5%、53.61%，整體平均得分為 58.33% (標準差 = 21.92)，皆高於機率。福里曼單因子關聯樣本變異數分析 (One way

Friedman Test for correlated samples) 顯示聲調類別的主要效果 (main effect) 顯著 ($\chi^2 (3, N = 12) = 8.78, p = .03$)。事後魏克森符號等級檢定 (Wilcoxon signed ranked test) 顯示 4 聲的平均指認錯誤顯著多於 1 聲 ($p = .005, d = 3.15$)。根據 Bonferroni 法修正之 p 值，其他任何兩聲調間的指認表現，均無顯著差別 ($ps = .04 - .92$)。

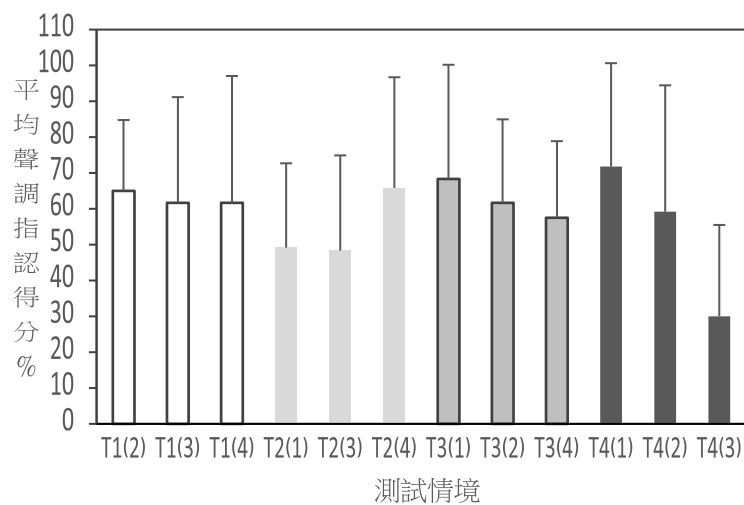


圖 3 1 至 4 聲在分別搭配三種聲調圖片情境下的團體指認平均得分 (%)。誤差線表示標準差。圖正下方各測試情境括弧內的數字為搭配圖片所代表的聲調

(三) 12 種測試情境得分

圖 3 陳列的是 1 至 4 聲在 12 種測試情境 (4 標聲調 × 3 聲調圖片) 的團體得分和標準差。白色、淺灰、深灰、黑色的長條分別代表各聲調標的三種情境，而各長條下方括弧內的數字則是搭配圖片所代表的聲調，如白色 T1 (2) 表示當指認標的是 1 聲時，搭配 1 聲圖片出現的另一張圖片是代表 2 聲。每情境的得分是根據 240 次嘗試結果 (1 標的 × 5 組 × 2 次 × 2 聲版本 × 12 位受試)。各聲調得分範圍如下：1 聲是 62% - 65% (差 3%)，2 聲是 48% - 66% (差 18%)，3 聲是 58% - 68% (差 10%)，4 聲是 30% - 72% (差 42%)。

以福里曼單因子關聯樣本變異數分析，顯示 2 聲 ($\chi^2 (2, N = 12) = 8.79, p = .02$) 和 4 聲 ($\chi^2 (2, N = 12) = 11.02, p < .00$) 內的三情境得分差異顯著，1、3 聲內的情境則無顯著差異 ($ps = 0.91、0.28$)。魏克森符號等級檢定事後分析， p 值以 Bonferroni 法修正，顯示 2 聲的差異存在於 T2 (3) 和 T2 (4) 間 ($p = .005, d = 3.15$)，而 4 聲的差異存在於 T4 (1) 和 T4 (3) 間 ($p = .005, d = 3.15$)。值得注意的是，同屬 4 聲的兩情境 T4 (1) 和 T4 (3) 得分差距很大，分別是 72% 和 30%，也是 12 情境中得分最高和最低的情境。

12 情境除可根據聲調標的分組外，還可

以視覺刺激(圖片)分組,如 T1(2)、T2(1)等六個聽覺刺激不同、但視覺刺激相同的組別。以魏克森符號檢定比較各組內兩情境的得分, p 值以 Bonferroni 法修正,結果是六組內的差異皆未達顯著($ps = .015 - .35$)。

二、聲調表達

聲調表達評量是根據常人評審對電子耳兒童聲調指認(分類)的結果。如某次嘗試中,電腦播放的是某位兒童所說的2聲語詞,若評審指認為2聲,則指認正確,也視為兒童該次的表達正確;若評審誤認為3聲,則視兒童該次的表達錯誤。簡言之,評審指認的對或錯即為兒童聲調表達的對或錯。此計分方式的正當性在於所有評審都經嚴格的聲調辨識篩選。

(一) 團體表現

本實驗共得評審 14400 次四選一的聲調指認反應(20 評審 \times 12 位語者 \times 4 聲調 \times

15 單音節語詞)。表 3 是根據各聲調分類的百分比和標準差所排列的混淆矩陣(confusion matrix)。每標的聲調的嘗試總數(即橫排)是 3600 次。團體 1 至 4 聲的平均得分是 79.47%、59.17%、59.36%、70.69%(對角線粗體數字),四聲合併得分是 67.17%(標準差 = 9.8%)。福里曼單因子關聯樣本變異數分析顯示聲調類別效果顯著, $\chi^2(3, N=20) = 50.28, p = .0000$ 。魏克森符號等級檢定事後分析, p 值以 Bonferroni 法修正,結果除 2、3 聲得分未達顯著差異外($p = .79, d = 0.12$),其餘任何兩聲調皆差異顯著,且 1 聲得分最高(1 聲 $>$ 2 聲: $p = .0001, d = 3.60$; 1 聲 $>$ 3 聲: $p = .0001, d = 3.49$; 1 聲 $>$ 4 聲: $p = .0001, d = 3.42$),4 聲次之(4 聲 $>$ 2 聲: $p = .0001, d = 3.60$; 4 聲 $>$ 3 聲: $p = .0001, d = 3.60$),2、3 聲則同是得分最低的聲調。

表3

評審對電子耳語者團體聲調指認的混淆矩陣,含各類指認的百分比(M)和標準差(SD)(%)

	「1」	「2」	「3」	「4」
	M (SD)	M (SD)	M (SD)	M (SD)
1聲	79.47 (5.02)	9.11 (3.39)	6.78 (2.59)	4.69 (1.61)
2聲	21.44 (5.52)	59.17 (4.95)	14.58 (4.25)	4.75 (1.95)
3聲	12.06 (4.23)	22.00 (4.95)	59.36 (4.90)	6.67 (2.32)
4聲	12.28 (3.37)	5.44 (2.03)	11.50 (3.49)	70.69 (2.48)

(二) 評審對兒童聲調的誤認狀況

為了解聲調表達的混淆狀況,將評審指認的狀況分析如下。以卡方(chi-square test)分別檢驗各標的之四種指認的百分比(即表 4 各橫排所列),結果皆為不均等($\neq 25\%$)分配,1 到 4 聲的 $\chi^2(3, N=3600)$ 值分別是 5711.31、2443.49、2441.18、4049.32, p 皆小

於.0001。以魏克森等級檢定配對比較各標的之三種錯誤指認的百分比, p 值以 Bonferroni 法修正,顯示評審對四聲調的誤認有不同的集中趨勢: 1 聲最常被誤認成 2 聲,大於次常的 3 或 4 聲(2 聲 $>$ 3 聲: $p = .0048, d = 1.62$; 2 聲 $>$ 4 聲: $p = .0004, d = 2.53$); 2 聲最常被誤認成 1 聲,大於次常的 3 聲(1 聲

> 3 聲： $p = .0024, d = 1.85$ ）；3 聲最常被誤認 2 聲，大於次常的 1 聲（2 聲 > 1 聲： $p = .0006, d = 2.39$ ）；4 聲最常被誤認為 1 或 3 聲（1 聲 = 3 聲： $p = .81, d = 0.11$ ），但最少被誤認成 2 聲（2 聲 < 1 聲： $p = .0001, d = 3.64$ ；2 聲 < 3 聲： $p = .0002, d = 2.95$ ）。總結以上分析，評審的誤認狀況並非文獻所載集中於 1 聲的現象。

（三）兒童個別的聲調表現

為檢視個別兒童的困難所在，以表 4 列出（評審指認）兒童各聲調的錯誤百分比。此百分比的計算是以全體評審 300 次（每聲調 15 題 × 20 評審）嘗試總數為分母，以非標的指認總數為分子。根據二項示檢驗，大於

或等於 74% 即是高於機率（ $N = 300, p = 3/4, \alpha = .05$ ）（ $p < .05$ ），以數字右上方加星號（*）表示，而每位兒童錯誤最多的聲調，則以粗體標示。除四聲合併錯誤呈大幅差距外（最右欄），尚有下列細部觀察：（1）跨四聲的表現普遍不平均。12 位中有 8 位的差距在 40% 以上，而在 10% 以下的僅有兩位，S9 和 S12（這兩位的表現已達常人水準）。（2）個別和團體錯誤的分布情況大略一致。如單一聲調錯誤率高於 40% 的都出現在 2 或 3 聲，符合 2、3 聲表現不如 1、4 聲的團體結果。（3）除 S2 和 S5 外，無 1、4 聲錯誤高於機率者，但 S5 的 4 聲錯誤率高達 99.67%。

表 4

評審對個別電子耳語者各聲調的團體錯誤百分比（%）

	1 聲	2 聲	3 聲	4 聲	四聲合併
S1	12.67	46.33	33.00	4.67	24.17
S2	87.67*	90.67*	53.67	36.33	67.08
S3	5.33	41.33	65.00	18.00	32.42
S4	19.67	38.67	34.67	39.00	33.00
S5	34.00	44.00	73.33	99.67*	62.50
S6	7.33	38.67	71.67	53.00	42.67
S7	2.33	48.00	10.33	1.67	15.58
S8	33.33	7.33	17.67	25.67	21.00
S9	7.00	7.00	3.67	11.67	7.33
S10	2.67	43.67	49.67	1.33	24.25
S11	33.00	83.33*	72.00	59.67	62.00
S12	1.33	2.00	3.00	1.00	1.83

* $p < .05$ ；粗體字代表各受試兒童錯誤最多的聲調。

個人四聲合併的表達錯誤百分比在 1.83%（S12）- 67.08%（S2），差異達 65%。值得注意的是，表達錯誤百分比最低的 S12，

其指認得分也最高（89.17%），而指認得分最低（31.67% - 33.33%）的四位，其中三位（S2、S5、S11）的表達錯誤百分比也最高（> 60%）。

經相關分析，得分與年齡 ($r = .08, p = .81$)、電子耳植入年齡 ($r = -.08, p = .68$) 或電子耳經驗 ($r = .13, p = .91$) 都未達統計相關。

(四) 評審信度

根據評審共 14400 次的指認反應，無論就聲調或語者分析，評審間的反應都呈顯著一致。跨語者 1-4 聲的評審間信度 Kendall's W 值分別是 0.71、0.80、0.88、0.91， p 皆小於 .0001，四聲合併的 W 值是 0.93 ($p < .0001$)。

個別語者的評審間信度(跨聲調共 16 種選項反應)也達一致，其中 S2 和 S11 的 Kendall's W 值分別是 0.66 和 0.77 ($p < .0001$)，其餘 10 位的 W 值皆在 0.83 - 0.90 ($p < .0001$)。

三、聲調指認與表達的相關分析

個人平均聲調的指認和表達得分，以皮爾森積差相關係數 (Pearson product-moment correlation coefficient) 分析，呈現高度正相關 ($r = .813, p = .001$)，如圖 4 所示。

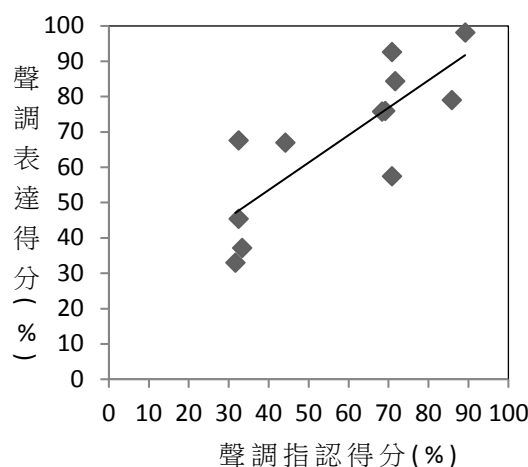


圖 4 12 位電子耳受試的國語聲調指認和表達測驗得分的散佈圖

伍、討論

一、聲調指認

(一) 二選一反應的計分標準與團體得分

在兩次嘗試都指認正確才計分的標準下，團體指認的平均得分為 58.33%，明顯低於過去電子耳兒童研究 67% - 82% 的所得 (Peng et al., 2004; Han et al., 2009; Xu et al., 2011; Zhou et al., 2013)。巧合的是，若將本研究每次嘗試恢復為獨立計分，則平均得分立即增加為 72.36%，與 Peng 等 (2004) 所得相同。由此可見，此前後研究 10% - 20% 的平均差異，應來自於計分標準的差異。另外，根據四選一的聲調評量研究，無論受試是習語前失聰兒童或是習語後失聰的成人，平均指

認水準都低於 70% (Chang et al., 2016; Huang, Wang, & Liu, 1995; Wei et al., 2004)。習語前、後失聰者的差異在於只有後者具備語言經驗 (Boothroyd, Geers, & Moog, 1991; Ciocca et al., 2002)，以致於植入電子耳後的表現常優於前者。雖然本研究二選一聲調指認的表現水準，仍待未來以四選一任務評量習語前失聰兒童才能確認，但可確定的是，二選一的指認評量若忽視測驗信度，將高估電子耳兒童的表現。相形之下，本研所得近 60% 的國語電子耳兒童的聲調指認水準，應較具真實性。

(二) 以「前上」為 3 聲刺激形式的相關指認結果

為與過去研究的指認結果比較，先將 12

情境的平均得分依照六聲調對比合併，如將 T1 (2) 和 T2 (1) 的得分合併為 1-2 聲對比得分。合併後得分最高的三對比是 1-4 聲 (66.67%)、1-3 聲 (65%)、2-4 聲 (62.5%)，次高的兩對比是 1-2 聲 (57%) 和 2-3 聲 (55%)，最低的則是 3-4 聲 (43.75%)。此六對比得分排序除 3-4 聲外，大致符合文獻資料的整體趨勢，即“包含 4 聲的對比分數皆顯著高於不含 4 聲的對比”。

上述 3-4 聲對比的得分最低應與本研究以「前上」為 3 聲的刺激形式有關。本研究將 3 聲從朗誦式的「全上」改為一般對話式的「前上」，不僅截去 $F0$ 曲線後半段的低升

(low rising) 部分，也連帶移除了後半段長於其他聲調 (Shen, 1990) 的時間。這兩項改變都減少了電子耳兒童辨識 3 聲的線索 (Peng et al., 2017)。另一方面，採用「前上」卻大幅提高了 3、4 聲 $F0$ 曲線的相似度，差別僅在於起始頻率的高低。本研究 3 聲的起始頻率約 117-195 Hz，4 聲則是 168-258 Hz，但兩聲調的終點頻率，都在 20 Hz 左右。以圖 5 呈現的女聲版四聲調平均的音域變化為例，雖然 3、4 兩聲調 $F0$ 的起始頻率差異大於 50 Hz，但接下來的走向和頻率卻逐漸合一。由此可推知，3-4 聲的聽覺可區辨性應明顯低於其他聲調對比。

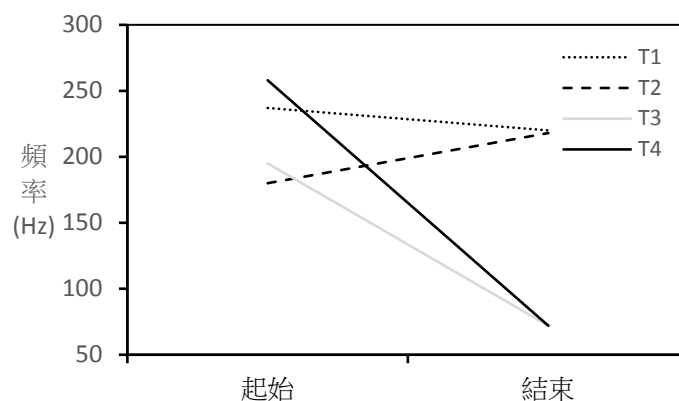


圖 5 聲調指認刺激女聲版四聲調的平均起始和結束頻率

「前上」式 3 聲除改變文獻 3-4 聲對比的得分排序外，也間接讓 1、4 兩聲的平均指認出現未見的顯著差異。如 Peng 等 (2004)、Xu 等 (2011) 皆以含 4 聲的對比得分高於其他對比，但本研究卻以 4 聲的平均得分最低。另外，12 位受試中有多達 10 位的 1 聲得分高於 4 聲。檢視圖 3 即可看出，4 聲的對比指認不再是最容易，且得分的變化也比 1 聲大。具體的說，1 聲內三情境 (空白組) 的表現最穩定，差距範圍僅 3%，完全不受搭配的圖片影響，平均正確率也最高。相對的，4 聲內的三情境 (黑色組) 卻有戲劇性差異：搭配

1 聲圖片時的得分 (72%) 為 12 情境中最高，但搭配 3 聲圖片時的得分卻跌至最低 (30%)。此差異經平均後，即讓 4 聲的正確率低於其他三聲調。Peng 等人 (2004) 因將 T3 (4) 和 T4 (3) 的得分合併為 3-4 聲對比得分，又將 T1 (4) 和 T4 (1) 合併為 1-4 聲對比得分，所以模糊了搭配圖片所能帶動的指認變化。至於為何 3 聲形式的改變並未影響 3 聲的平均得分，或更具體的問，為何影響的是 T4 (3)，而非 T3 (4) 的得分，將在下節分析。

(三) 測試情境效果—搭配圖片的影響

本研究將過去研究的六聲調對比細分為

12 測試情境(4 聲調×3 搭配聲調圖片)進行分析，結果 2、4 聲得分出現情境效果：2 聲內的差異來自 T2 (4) 的得分高於 T2 (3) (差 18%)，4 聲內的差異來自 T4 (1) 得分高於 T4 (3) (差 42%)。

一如前言，聲調辨識最主要的聲學線索是 F_0 。研究顯示，習語後失聰成人在接近基頻附近 (150 Hz) 的區辨閾值只需 6-20 Hz (Geurts & Wouters, 2001)，而國語四選一的聲調指認研究又明確指出，電子耳聽者和常人的差距約是 9.3 比 0.4 個半音 (He, Deroche, Doong, Jiradejvong, & Limb, 2016)。雖然後項數據未必可套用於二選一的結果解讀，但可見電子耳聽者區辨聲調，需要大於常人的頻率差距。

參考圖 5 可大略理解為何 2 聲搭配 4 聲的得分顯著高於搭配 3 聲，及為何 4 聲搭配 1 聲的得分顯著高於搭配 3 聲。雖然就 F_0 終點頻率差距而言，T2 (4) 情境內 2、4 兩聲的差距及 T2 (3) 內的 2、3 兩聲的差距差不多，但就起始頻率而言，T2 (4) 的差距則遠大於 T2 (3)。以此可以解釋為何 T2 (4) 的得分高於 T2 (3)。

再看 T4 (1) 和 T4 (3)。T4 (1) 所涉 4、1 兩聲的 F_0 曲線不僅無重疊，終點頻率又分別是最和最低，所以應區辨容易，而團體得分 (72%) 也一如預期，為所有情境中最高。相對的，T4 (3) 所涉的 4、3 兩聲從始到終頻率走向都是下降，且距離逐漸縮小，團體得分也最低 (30%)。

接下來的問題是：為何 1、3 聲內情境的 F_0 曲線對比無明顯差異？首先，1 聲的 F_0 曲線平穩，因此無論搭配的圖片聲調為何，都是相對簡易的任務，以致於三情境的得分幾乎相同。再看 3 聲的指認，雖然其曲線和 4 聲一樣都是下降，但畢竟下降幅度不如 4 聲，因此可以解釋何以 3 聲的三電腦化評量情境

得分差距遠不如 4 聲的三情境。至於 T3 (4) 和 T4 (3) 相比，兩情境雖圖片組合相同，但 T3 (4) 比 T4 (3) 的得分高了 27%。換言之，把 4 聲誤聽成 3 聲，比把 3 聲誤聽成 4 聲的比率多了 27%。若從頻率角度解析，此差異意味著當兩聲調曲線接近時，電子耳兒童傾向把音域變化較大的 4 聲 (高降) 指認成變化較小的 3 聲 (低降)。此現象反映的正是兒童頻率解析的缺陷 (He et al., 2016)。以上情境效果的分析是根據本實驗兩位語者的平均聲調數據所推論。未來研究可以電腦模擬聲調刺激，操控對比間的頻率差距等方式檢驗。

綜合上述，可得下列結論：一、聽覺標的和視覺圖片混淆的組合，可影響電子耳兒童聲調指認的難易；因此「二選一」的指認所得未必可經由合併、計算平均的方式來估算一組「四選一」的指認結果；二、基於各聲調指認結果皆明顯高於機率，顯然電子耳可以提供國語電子耳兒童部分功能性的頻率解析；三、圖片的混淆雖呈現在視覺領域，但卻可影響聲調的聽覺辨識，顯示部分受試已對聲調發展出不同程度的內在表徵能力；四、本研究六歲以上電子耳兒童的平均聲調指認表現不及加測的三位 6-7 歲正常兒童的表現，甚至未超過 Wong 等 (2012) 的 3 歲正常兒童的表現。個人表現方面，本研究 and Zhou 等 (2013) 的大型研究結果一致，都有近半數受試的表現低於機率。

二、聲調表達

本研究以評審四選一聲調指認反應的對錯為電子耳兒童聲調表達的對錯，所以評審間指認反應的一致性是評量信度的關鍵。20 位評審無論對個別聲調或個別受試的聲調指認，都呈現高評審間信度。整體評審間信度 ($r = .93$) 與 Zhou 等 (2013) 四位評審的指認信度接近 ($r = .95$)，但高於 Peng 等 (2004) 三位專家的聲調等級評分信度 ($r = .84$)。

本研究兒童聲調表達的電腦化評量結果與 Peng 等人(2004)的人工等級評量結果一致，卻與 Han 等三電腦化評量的結果不同。本研究團體得分是 67.17%，明顯優於主要文獻所得（47%-53%）（Peng et al., 2004; Han et al., 2007; Zhou & Xu, 2008; Zhou et al., 2013; Xu et al., 2011），但個別聲調的得分排序差異與其中 Peng 等的等級評量結果一致。具體而言，Peng 等人和本研究的結果是四聲調得分都高於機率，其中 1、4 聲得分近似，2、3 聲得分近似，且前兩聲的得分顯著高於後兩聲，然而 Han 等三電腦化評量的結果卻是四聲調得分互異，以 1 聲最高，2 聲最低，且水準僅及機率。

在評審最常誤認的選項上，本研究和三電腦化評量的差異更大。三電腦化評量不論標的，最常的誤認選項一律是 *F0* 平坦的 1 聲，佔錯誤反應比例的 32.6% - 43.7%。對此結果，研究者認為與聲學分析研究的結論一致，都表示電子耳兒童不會操控他們的聲調，或電子耳提供的訊息阻礙了他們的聲調發展（Zhou & Xu, 2008）。然而此說法卻得不到本研究結果的支持。本研究最常的誤認選項不僅隨表達的標的而改變，且有兩項特徵：（1）與標的聲調的音域接近或 *F0* 曲線走向相似，如 1、2 聲互為最常誤認選項，但 4 聲絕少被誤認為是曲線走向相反的 2 聲；（2）誤認選項的 *F0* 曲線變化幅度常小於標的聲調，如 4 聲雖常被誤認為 3 聲，但 3 聲卻最不常被誤認為 4 聲。（此點和聲調指認 T4（3）與 T3（4）的正確率差異一致。）以上兩特徵顯示，本研究電子耳兒童的聲調控制雖未達常人標準、頻率變化範圍也比正常限縮，但不容否認的是，他們已可製造出可供區辨的變化。

本研究 and 過去三電腦化評量的多項差異，很可能反映的是受試特徵的差異。本研

究雖然亦採電腦化評量，但受試選擇條件，包括年齡和電子耳使用經驗的最低要求，都不同於該三研究。相對的，本研所得聲調得分的排序卻和 Peng 等人（2004）的專家評量結果相似。巧合的是，該研究受試年齡分佈與電子耳經驗皆與本研究相似，受試成熟度也都足以勝任測驗任務。因此，本研究 and 過去研所得的異同應與評量方式無關，而與受試群年齡及連帶的電子耳經驗有關。

隨著電子耳植入年齡普遍逼近 1 歲（Gifford, 2013），年齡愈小的使用者多半使用經驗也愈少，因此年齡和電子耳經驗已成為密不可分的兩項因素。有關年齡對聲調表達的影響，前已述及，即便是正常兒童也要在滿三歲之後才能發展出穩定的聲調表達。電子耳兒童的音韻表達同樣需要時間發展（Peng et al., 2007; Nakata et al., 2012）。至於電子耳經驗是否影響聲調表現，現有文獻雖無定論，但 Reiss、Turner、Erenberg 與 Gantz（2007）的結論卻值得重視。該研究測得電子耳使用者在植入的三至五年內，對音高（pitch）的感知經常改變，且變化幅度可往低頻延伸兩個音程。此低頻感知的擴增應直接關係著聲調的學習。該研究另根據此低頻感知的改變與口語接收的高度相關，指出電子耳經驗對音高感知的影響很可能大於過去研究所認知。事實上，Han 等（2007）已發現，電子耳經驗愈久的兒童，聲調表達的得分愈高。此相關性雖未見於本研究 and Peng 等（2004）的結果，但可能是因受試使用電子耳已超過三至五年，聲調表現進展相對趨緩，使得相關性減弱。是否如此則有待未來研究探討。

三、指認和表達的相關性

指認和表達的相關分析結果呈現高度相關（ $r=.813$ ），與 Xu 等（2011）所得（ $r=.805$ ）接近，但高於 Zhou 等（2013）所得（ $r=.56$ ）。

本研究較高的相關結果應與評分標準修改，即以兩次嘗試都答對才計分有關。如此修改的後果有二：一、擴大指認得分的變動範圍。過往相關研究因採每次嘗試獨立計分，在二選一的機率表現是 0.5 的條件下，個人得分幾乎都在 50%以上。反觀本研究兩次嘗試都答對的機率只有 0.25，因此受試的得分可低至 33.67%。二、信度的增加。此實際效應就是提升指認與表達評量(四選一)結果的相關性。因此，本研究相關係數的提高應與受試人數的多寡無關。

陸、結論

本研究以電腦化測驗評量 12 位習語前失聰電子耳兒童的國語聲調指認和表達。參加兒童皆在六歲以上，電子耳經驗則在兩年以上。結果顯示，當評量信度被納入評分標準後，團體聲調指認表現接近 60%，比過去二選一指認所得低了 10%以上。另外，指認結果可受指認時搭配的图片及 3 聲的刺激形式影響。聲調表達方面，團體得分接近 70%，普遍超出過去研究結果約 20%，個別聲調的得分排序則僅和受試條件相似的非電腦化等級評量一致。綜合以上結果，電子耳兒童的聲調表現可隨評量的信度及受試條件而不同。另外，雖然團體聲調表達結果略優於指認，但因兩項評量的任務與機率不同，所以不宜貿然解讀為聲調表達優於聲調指認，亦不宜推論何項聲調表現進展較速。

個人表現方面，聲調的指認和表達呈高度相關，但多數受試的兩項表現都未臻理想。檢視兩項得分未及 50%者，發現皆具有 3-9 年電子耳經驗，顯然欠缺聲調聽與說訓練的專業介入。在此方面，本研究所得，包括同聲調指認與搭配圖片的難易資料及聲調表達的混淆模式，都可應用於特殊需求課程內容的設計和規劃。

本研究和絕大多數兒童電子耳研究文獻一樣，都受限於參加兒童人數有限，更因實驗地點及資源不足，所以無法隨機取樣。為盡力彌補上述限制，本研究以保守的統計分析檢驗結果的顯著性。另外，由於聲調的指認與表達都採隔音室內的電腦化評量，所以研究結果可與主流文獻直接比較。隨著兒童電子耳人口增加和年齡的成長，未來研究除需注重評量方式的信效度外，更須根據評量任務，篩選成熟度足以勝任的兒童，並持續追蹤聲調發展的特性，例如在正常持續使用電子耳條件下的學習曲線等。另外，有鑑於本研究與部分文獻皆顯示口語接收與表達的高度相關，未來研究可透過教學或臨床介入，檢驗聲調接收或表達的介入效益。

參考文獻

一、中文部分

韓福榮、曹光文、江以文 (2015)。聽覺障礙者接受人工電子耳植入術後溝通成效追蹤研究。台北：中華溝通障礙教育學會。

二、英文部分

- Barry, J., & Blamey, P. (2004). The acoustic analysis of tone differentiation as a means for assessing tone production in speakers of Cantonese. *Journal of the Acoustical Society of America*, 116, 1739-1748.
- Boothroyd, A., Geers, A., & Moog, J. (1991). Practical implications of cochlear implants in children. *Ear and Hearing*, 12, 81-89.
- Boyd, P. (2011). Potential benefits from deeply inserted cochlear implant electrodes. *Ear and Hearing*, 32, 411-427.
- Chang, Y., Chang, R., Lin, C., & Luo, X. (2016). Mandarin tone and vowel recognition in cochlear implant users: Effects of talker

- variability and bimodal hearing. *Ear and Hearing*, 37, 271-281.
- Chen, F., Wong, L., & Hu, Y. (2014). Effects of lexical tone contour on Mandarin sentence intelligibility. *Journal of Speech, Language and Hearing Research*, 57, 338-345.
- Ching, T. Y. C. (1984). Lexical tone pattern learning in Cantonese children. *Language Learning and Communication*, 3, 317-334.
- Ciocca, V., Francis, A., Aisha, R., & Wong, L. (2002). The perception of Cantonese lexical tones by early-deafened cochlear implantees. *Journal of the Acoustical Society of America*, 111, 2250-2255.
- Clumeck, H. (1977). *Studies in the Acquisition of Mandarin Phonology* (Unpublished doctoral dissertation). University of California, Berkeley, California.
- Fu, Q., & Zeng, F. (2000). Identification of temporal envelope cues in Chinese tone recognition. *Asia Pacific Journal of Speech, Language and Hearing*, 5, 45-57.
- Geurts, L., & Wouters, J. (2001). Coding of the fundamental frequency in continuous interleaved sampling processors for cochlear implants. *Journal of Acoustical Society of America*, 109, 713-726.
- Gifford, R. (2013). *Cochlear Implant Assessment: Evaluation of Candidacy, Performance, and Outcomes*. San Diego: Plural Publishing.
- Greenwood, D. (1990). A cochlear frequency-position function for several species— 29 years later. *Journal of Acoustical Society of America*, 87, 2592-2605.
- Han, D., Liu, B., Zhou, N., Chen, X., Kong, Y., Liu, H., Zheng, Y., & Xu, L. (2009). Lexical tone perception with HiResolution and HiResolution 120 sound-processing strategies in pediatric Mandarin-speaking cochlear implant users. *Ear and Hearing*, 30, 169-177.
- Han, D., Zhou, N., Li, Y., Chen, X., Zhao, X., & Xu, L. (2007). Tone production of Mandarin Chinese speaking children with cochlear implants. *International Journal of Pediatric Otorhinolaryngology*, 71, 875-880.
- He, A., Deroche, M., Doong, J., Jiradejvong, P., & Limb, C. (2016). Mandarin tone identification in cochlear implant users using exaggerated pitch contours. *Otology & Neurotology*, 37, 324-331.
- Huang, T.-S., Wang, N.-M., & Liu, S.-Y. (1995). Tone perception of Mandarin-speaking postlingually deaf implantees using the nucleus 22-channel cochlear mini system. *Annals of Otology, Rhinology, and Laryngology Supplement*, 166, 294-298.
- Landsberger, D., Mertens, G., Punte, A., & Heyning, P. (2014). Perceptual changes in place of stimulation with long cochlear implant electrode arrays. *Journal of the Acoustical Society of America*, 135, 75-81.
- Massaro, D., Cohen, M., & Tseng, C. (1985). The evaluation and integration of pitch height and pitch contour in lexical tone perception in Mandarin Chinese. *Journal of Chinese Linguistics*, 13, 267-290.
- Nakata, T., Trehub, S., & Kanda, Y. (2012). Effect of cochlear implants on children's perception and production of speech prosody. *The Journal of the Acoustical*

- Society of America*, 131, 1307-1304.
- Pearsons, K., Bennett, R., & Fidell, S. (1977). Speech levels in various noise environments. Technical Report EPA-600/1-77-025. Washington, D.C.: U.S. Environmental Protection Agency.
- Peng, S., Tomblin, J., Cheung, H., Lin, Y.-S., & Wang, L.-S. (2004). Perception and production of Mandarin tones in prelingually deaf children with cochlear implants. *Ear and Hearing*, 25, 251-264.
- Peng, S., Tomblin, J., Spencer, L., & Hurtig, R. (2007). Imitative production of rising speech intonation in pediatric cochlear implant recipients. *Journal of the Speech, Language, and Hearing Research*, 50, 1210-1227.
- Peng, S., Tomblin, J., & Turner, C. (2008). Production and perception of speech intonation in pediatric cochlear implant recipients and individuals with normal hearing. *Ear and Hearing*, 29, 336-351.
- Peng, S., Lu, H., Lu, N., Lin, Y., Deroche, M., & Chatterjee, M. (2017). Processing of acoustic cues in lexical-tone identification by pediatric cochlear-implant recipients. *Journal of Speech, Language, and Hearing Research*, 60, 1223-1235.
- Reiss, L., Turner, C., Erenberg, S., & Gantz, B. (2007). Changes in pitch with a cochlear implant over time. *Journal of the Association for Research in Otolaryngology*, 8, 241-257.
- Shen, X.-N. S. (1990). Tonal coarticulation in Mandarin. *Journal of Phonetics*, 18, 281-295.
- Su, A. (1985). *The Acquisition of Mandarin Phonology by Taiwanese Children* (Unpublished master's thesis). Fu-Jen Catholic University, New Taipei City, Taiwan.
- Thornton, A., & Raffin, M. (1978). Speech discrimination scores modeled as a binomial variable. *Journal of Speech, Language, and Hearing Research*, 21, 507-518.
- Tye-Murray, N., Spencer, L., & Woodworth, G. (1995). Acquisition of speech by children who have prolonged cochlear implant experience. *Journal of the Speech, Language and Hearing Research*, 38, 327-337.
- Tye-Murray, N., & Kirk, K. (1993). Vowel and diphthong production by young users of cochlear implants and the relationship between the phonetic level evaluation and spontaneous speech. *Journal of the Speech, Language, and Hearing Research*, 36, 488-502.
- Wang, J., Shu, H., Zhang, L., Liu, Z., & Zhang, Y. (2013). The roles of fundamental frequency contours and sentence context in Mandarin Chinese speech intelligibility. *Journal of the Acoustical Society of America*, 134, EL91.
- Wei, C., Cao, K., & Zeng, F.-G. (2004). Mandarin tone recognition in cochlear-implant listeners. *Hearing Research*, 197, 87-95.
- Wong, P., Schwartz, R., & Jenkins, J. (2005). Perception and production of lexical tones by 3-year-old, Mandarin-speaking children. *Journal of the Speech Language and Hearing Research*, 48, 1065-1079.

- Xu, L., Chen, X., Lu, H., Zhou, N., Wang, S., Liu, Q., Li, Y., Zhao, X., & Han, D. (2011). Tone perception and production in pediatric cochlear implants users. *Acta Oto-Laryngologica*, 131, 395–398.
- Zhou, N., Huang, J., Chen, X., & Xu, L. (2013). Relationship between tone perception and production in prelingually-deafened children with cochlear implants. *Otology & Neurotology*, 34, 499–506.
- Zhou, N., Zhang, W., Lee, C., & Xu, L. (2008). Lexical tone recognition with an artificial neural network. *Ear and Hearing*, 29, 326–335.
- Zhu, H. (2002). *Phonological development in specific contexts: Studies of Chinese-speaking children*. Clevedon, England: Multilingual Matters.

Mandarin Tone Identification and Production by Early-Deafened Pediatric Cochlear Implant Users : A Report Based on Performance by Children Older Than Six Years When Assessed Using Computerized Perceptual Tests

Yuan-Chuan Chiang

Department of Speech Language Pathology and Audiology,
National Taipei University of Nursing and Health Sciences

Abstract

Previous studies of Mandarin tone performance by prelingually deafened cochlear implant (CI) users contained identification and production results from subjects aged below three years old or with only months of device experience. This study revisited this dual topic with twelve CI users, aged 7 to 15 years, at an average postoperative period of six years. In the identification experiment, participants completed a 2-Alternative Forced Choice (2AFC) picture identification task in which they identified the four Mandarin contrastive tones of 60 monosyllabic Chinese words under 12 picture-pairing conditions. All the words were pre-recorded in conversational styles. Scores were given only to words that were identified correctly in both trials. In the production experiment, the same 60 words used in the identification experiment were elicited and recorded from each of the above CI users and the tone of each word was judged by 20 normal hearing adults in a 4AFC identification task. The group mean identification score was 58%, compared to 67-73% as previously reported with a 50% response chance level. Identification of tones 2 and 4 was found to vary significantly with the picture-pairing condition. The group mean production score was 67%, compared to 47-53% as previously reported. Individual tone score rankings and tone confusion patterns were inconsistent with results from computerized investigations, while similar to those from a study using a 5-point evaluation that only tested school-aged subjects. The results suggest that tonal performance of the pediatric CI users may vary with the age range, CI experience and test reliability. Individual identification and production scores were highly correlated. The less than satisfactory performance levels in both measures also revealed special needs for auditory training even among the CI users with extended device experience.

Keywords: cochlear implant, Mandarin, tone perception, tone production, prelingual deafness