



台灣聽力語言學會電子學報

The Speech-Language-Hearing Association, Taiwan

- 主題文章：人工電子耳的過去、現在及未來 - 編碼策略篇
- 撰 稿 者：柳佩辰



主題文章

人工電子耳的過去、現在及未來 - 編碼策略篇

柳佩辰

聽力師們一般在接觸聽覺輔具的時候，都會自然的把焦點放在目前的技術及未來的發展方向上，但觀今宜鑒古，無古不成今，我們不妨藉由人工電子耳發展的歷程來認識現在技術的應用，並展望未來可能的趨勢。

1985 年美國 FDA 首度批准將人工電子耳用於極重度聽力損失的成人，至今已有 35 年歷史，隨著植入者數量的增加，人工電子耳的安全性和有效性普遍被大家認可，使得人工電子耳早已成為助聽器幫助有限時的最佳聽力解決方案。然而，從語後失聰的成人植入者的回饋中，我們得知，電子耳聽到的聲音有些像機器人的聲音，聽音樂時聽不出旋律。為何會有這些回饋？這和電子耳如何截取、分析聲音信號，再將它轉換成電刺激有關，而編碼策略就是這個截取、分析和刺激的核心。

最理想的編碼策略應該是以類比的方式，完全保留並還原輸入信號聲波中的時間、頻率及響度特性。然而，由於受到直接電刺激聽神經時，聽神經纖維的高同步放電、神經不反應期等生理因素的影響，以及電池及植入體電極耦合技術的硬體限制，因此，40 年前多頻道的人工電子耳系統設計初期，編碼策略就分成以頻率解析為主或以時間解析

為主的兩大主軸。

● 以頻率解析為主的編碼策略

以頻率解析為主的編碼策略又被稱作特徵提取的策略，一直以來 Cochlear 公司都以此編碼策略為主。從最早的 F0/F2 策略，到後來的 F0/F1/F2，再來的 MPEAK、SPEAK 以及目前使用的 ACE、ACE(RE)、MP3000 等，全都屬於特徵提取的編碼策略。早期的 F0/F2、F0/F1/F2 策略，由於僅提取說話者的基頻和母音共振峰，對子音的辨識能力較差，加以改良後，研發出 MPEAK 編碼策略，MPEAK 是在 F0/F1/F2 基礎上，再加上三個高頻帶（2k-6k Hz），以增加更多子音資訊。這三個編碼策略，主要以提取語音信息為主。

自 1994 年，Cochlear 推出 N22 型植入體，便將語音提取技術進一步改為頻譜提取，即 n-of-m 的編碼策略。它會針對麥克風傳入的聲音，進行頻譜採樣分析，並根據聽力師設定的最大數 (Maxima) 挑選出頻譜包絡中最高的那幾個點，然後在對應該頻率的頻道上給出電刺激。最大數就是 n-of-m 當中的 n，而 m 指的是總頻道數。例如 ACE 總共有 22 個頻道，如果最大數設定的是 8，那就是 8-of-22，意思就是 22 個頻道都可用，但每次採樣只會挑選出頻譜包絡上顯示振幅最高的八個頻率，並在對應的八個頻道放電。

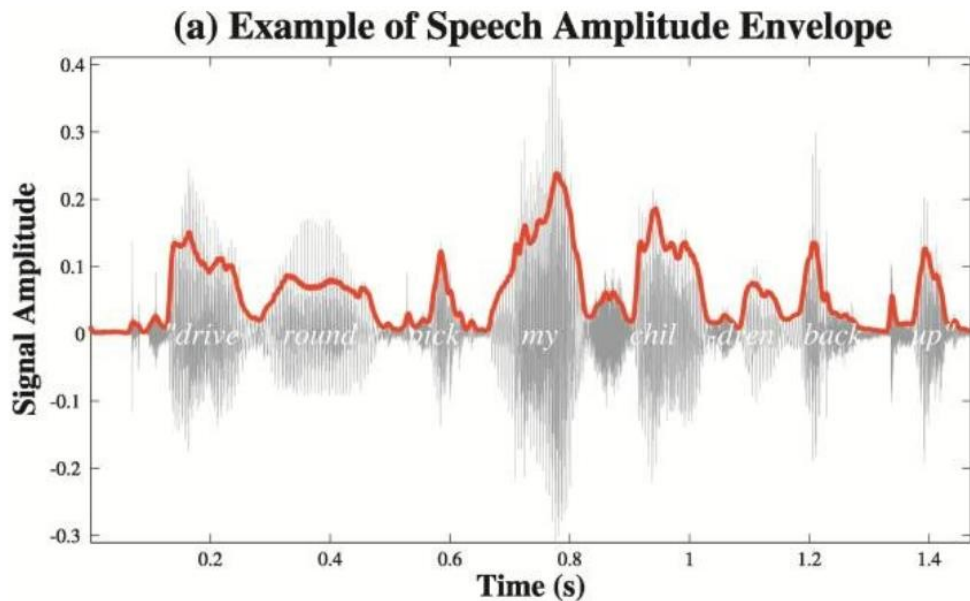
SPEAK、ACE、ACE(RE)都是相同的原理，唯一的區別就在於越後期的編碼策略，刺激速率越高，以期在頻率解析的基礎上還能更快的反應聲音在時間軸上的變化，達到更好的時間解析。而最新的 MP3000，則是考慮了，當原始信號的頻譜能量分佈較寬時，相鄰兩個頻道會同時被選中，但能量強、振幅高的頻道所對應的電極放電也較大，因此可能連相鄰電極下方的神經纖維都會被刺激到，如果這時候兩個相鄰通道都被選中的實際意義就不大了，也造成通道的浪費，MP3000 主要就在解決這個問題。

● 以時間解析為主的編碼策略

以時間解析為主的編碼策略就是 CIS 了，說它以“時間解析”為主，是由於 CIS 所提取的是信號的振幅包絡而不是頻譜包絡。它的信號處理方式是先讓輸入的聲音進入一組帶通濾波器，將聲音區分成數個頻帶，每個帶通濾波器後方都有一個包絡偵測 (envelop detector) 來偵測該通道中信號的振幅變化，之後再將振幅包絡的變化，在對應的電極上，以一連串的調製脈衝刺激 (modulated biphasic pulses) 來呈現。為了避免混疊錯誤 (aliasing error)，CIS 的最低刺激速率至少要是低通濾波器截止頻率 (cutoff frequency) 的四到五倍，若低通濾波器的截止頻率為 400Hz，那麼 CIS 每個通道的刺激速率就至少要 1600 至 2000 pps，因此相較於特徵提取的策略，CIS 通常需要更高的刺激速率，CIS 也因它不間斷的採樣包絡信號，並快速且交錯的在每個電極上呈現脈衝刺激而得名。

CIS 不僅是三個廠牌都有的編碼策略，它還是現在許多新的編碼策略的基礎，如 MED-EL 公司的 FSP、AB 公司的 HiRes、HiRes120 等。這些新的編碼策略主要目的都在

提升所謂的“精細結構”。精細結構指的是在振幅包絡中包含的頻率資訊，如下圖（註一），紅色線條代表振幅包絡，當中的灰色線條則是頻率資訊。



以目前的編碼策略 (ACE, HiRes120, FSP) 來說，多數的植入者在安靜環境下都可以達到 80% 以上的語句識別，有些甚至可以達到 90% 以上，從言語識別的角度來說，現有的技術已經做得很好了。目前最大的缺憾在於音樂的識別。但不論是 CIS 或 n-of-m，都無法提取旋律的資訊，這可能需要一個全新的提取技術才能做到吧？期望在不久的將來，科學家們可以研發出全新的信號提取技術，讓電子耳植入者可以更好的聆聽和識別音樂的旋律。

<註一>

圖片出自google 圖片搜尋，源自“Goswami U, Leong V. 2013. Speech rhythm and temporal structure: Converging perspectives?, *Laboratory Phonology March 2013*”

● 參考文獻

- [1] Wolfe, J. (2018). *Cochlear Implant: Audiologic Management and Considerations for Implantable Hearing Devices*. Plural Publishing Inc.
- [2] Somek B., et al. (2006). Coding Strategies for Cochlear Implant. *Automatika* 47(1-2):69-74
- [3] Wilson, B., Dorman, M. (2008). Cochlear Implants: a remarkable past and a brilliant future. *Hear Res.*, 242(0): 3-21.

關於作者

學歷	墨爾本大學臨床聽力學碩士
經歷	亞洲大學聽力暨語言治療學系 助理教授級專業技術人員 領先仿生醫療器械（上海）有限公司 中國區康復暨研究經理 澳科利耳醫療器械（北京）有限公司 大中華區臨床經理



編輯

發行單位：台灣聽力語言學會

發行人：葉文英

主編：張晏銘

編輯顧問：曾進興

網址：www.slh.org.tw

發行日期：2020.08.02

聽語學報：第九十二期

執行編輯：張矩嫻

助理編輯：陳奕秀