

語音比對

中央警察大學鑑識科學學系教授 徐健民

電話聲響起，你拿起聽筒就可聽到對方的聲音。如果是熟識的人所打來的電話你往往能夠猜中到底是誰！為什麼你有這種能耐？主要是每一個人的聲音都有獨特的性質。當你熟悉這些特性就能夠分辨出到底是誰的聲音。在這個電話、收音機、錄音機及網路普及的年代，聲音很容易就可以傳達出去，也很容易就可錄得語音。罪犯常常利用電話進行恐嚇或詐欺。而恐嚇或詐欺的語音也很容易被錄取下來。當嫌犯被逮捕時，恐嚇電話與嫌犯聲音的比對往往成為有價值的證物。

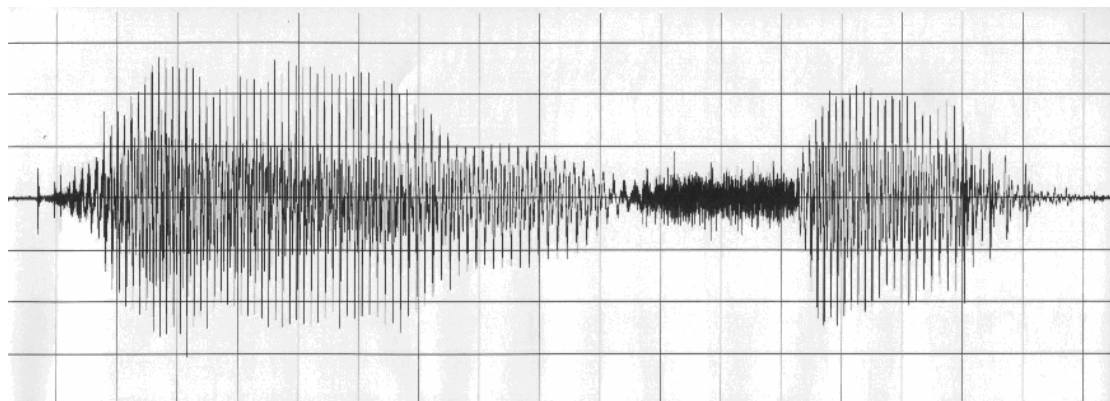
人類想要向外界傳達某些訊息時，必須先將欲傳達的信號在腦中轉換成語言的形式，然後透過神經系統的刺激，氣體由肺部、支氣管、氣管呼出，流經喉頭，通過咽腔、口腔及鼻腔，經由口鼻將聲波送出。當呼氣經過咽腔、口腔及鼻腔，由口鼻送出聲波時，唇、齒、舌等位置的變化改變口腔及鼻腔內部空間的大小及形狀，以及氣體呼出體外的模式，因而影響氣柱的共鳴，發出不同的聲音。當呼氣經過喉頭的聲帶時，如果聲帶有近似週期性的一開一閉動作，則可發出能量較大，特性穩定，具類週期性聲波形狀的「有聲語音」；如果聲帶未振動，氣體在聲腔受阻摩擦及聲腔急速開閉則產生波形不規則的「無聲

語音」。

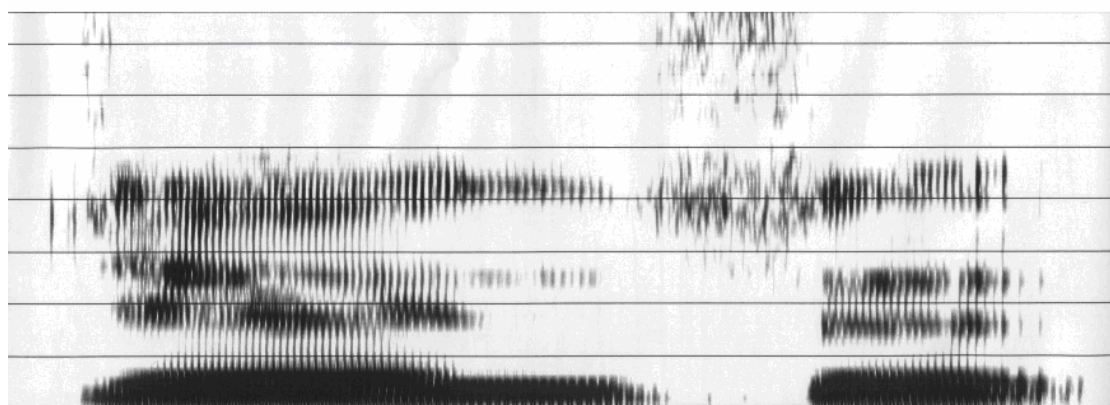
當人類耳朵聽到聲音時，耳朵都會感覺到它的音調、音量和音色。音調與聲帶振動的快慢有關，而聲帶振動頻率稱為基頻。聲帶振動得快音調高，聲帶振動得慢音調低。聲帶振動的快慢與聲帶的長短厚薄及緊張程度有關。聲帶的長短厚薄因人而異，且每個人聲帶的振動有一定的頻率的變化範圍，所以音調也因人而異。音量就是聲音的強弱程度，人類聽覺感受到聲音大小與聲音強弱程度的對數值有關，稱之為分貝。音色是一個人發音時因聲腔共鳴的差異會同時發出某些泛音，其振動頻率為基頻的倍數，並增強某些頻率區域的聲音強度。這些較強的頻率區稱為共振峰。有聲語音通常有三到四條共振峰。

如果將一連續的語音區分成許多連續的短時區，並分析每一時區的語音在不同頻區的聲音強度。以語音時間為橫座標，聲音的頻率為縱座標，並將每一時間之每一頻率的音量分貝值以灰色程度標示出來。此三度空間圖形稱為聲譜或稱為聲紋。圖 1 是『鑑識』語音的波形圖及聲譜圖。在聲譜中較黑的帶狀部分由底部算起又分別稱為第一共振峰、第二共振峰、第三共振峰及第四共振峰。共振峰的形狀變化與語音的內容有關。雖然不是同一個人所言，但是語音內容相同也會具有類似形狀的共振峰。內容不同之語音的共振峰形狀是有相當的差異。不同人講相同的一個字雖然具有類似的共振峰形狀，但是共振峰

的相對位置會有差異。這些共振峰位置的差異程度也是用來辨識語者的重要依據之一。



(a) "鑑識" 語音波形圖



(b) "鑑識" 語音聲譜圖

圖 1. 『鑑識』語音的相關圖形

個人的語音特色包括音色、基頻的高低、音量的大小、語音的清晰程度、說話的快慢、說話的節拍、說話的腔調及詞彙的應用。這些都是人體發音器官複雜的交互作用所產生的結果。語音比對人員錄取語音內容相同的兩句語音經由聲音聆聽與聲譜比對來研判兩句語音的異同。比對後依據語音相符的字數、不符的字數或錄音品質的優劣等情況作綜合性的判斷。