



台灣聽力語言學會電子學報

The Speech-Language-Hearing Association, Taiwan

- 主題文章：淺談頻率
- 撰 稿 者：許家甄博士 語言治療師



主題文章

淺談頻率



許家甄博士/語言治療師

黃自強語言治療所/思比語言治療所

大多數的語言治療師和學生對聲學(acoustics)的第一印象是「一堆數字」、「複雜難懂」、「基礎科學」，所以常常把它定位在「理論」類別，鮮少把其相關知識應用在臨床上。本文章將透過介紹「頻率」這個常用於嗓音和語音的聲學參數，讓大家回顧並了解如何使用聲學解釋不同的嗓音和言語現象。

頻率的定義

「頻率」(frequency; f)泛指某事件(event)在一段時間裡重複的次數，因此常用的頻率單位「赫茲」(hertz; Hz)，表示每秒鐘事件發生的次數。而在聲學的領域裡，頻率是正弦波形(sinusoidal waveform)在一個時間單位裡重複的次數，正弦波形在一個「循環」(cycle)內重複一次，一個循環的時長則稱為「週期」(period)。頻率(f)和週期(T)的關係為 $f = 1/T$ ，在無法直接取得頻率時，可把週期套用在這個公式上，例如：透過觀察發現有個正弦波形每 0.5 秒內重複一次，因此該正弦波形的週期為 0.5 秒($T = 0.5s$)，使用公式 $f = 1/T$ 後，得到 $f = 1/0.5$ ，所以 $f = 2 \text{ Hz}$ ，表示該正弦波形每秒鐘發生 2 次。關於嗓音，所謂的循環為聲門的一開一關，而這一開一關所耗費的時間則為週期。如果一個人的聲門開關一次需要花上 0.01 秒($T = 0.01s$)，按照公式 $f = 1/T$ ，所以 $f = 100 \text{ Hz}$ ，這也就是此人的嗓音基頻。(Johnson, 2012)

頻率、基頻與音高

上面簡單描述了頻率與正弦波的關係。但是人類透過聲帶震動所發出的嗓音(voice)卻不是正弦波，而是週期性複雜聲波(periodic complex sound wave)，此類聲波的波形除了會重複以外，還能透過傅立葉分析(Fourier Analysis)分解成一系列的正弦波。這系列中，頻率最低的正弦波為基頻(fundamental frequency; F_0)，而其他頻率較高的正弦波為諧波(harmonics; H)。根據傅立葉分析的原理， H 頻率為 F_0 頻率的整倍數，例如：透過傅立葉分析分解出一系列的正弦波，其 F_0 為 100 Hz，第一次諧波(H_1)為 $F_0 \times 1 = 100 \times 1 = 100$ ，所以 $H_1 = F_0$ ；第二次諧波(H_2)為 $F_0 \times 2 = 100 \times 2 = 200$ ；第三次諧波(H_3)為 $F_0 \times 3 = 100 \times 3 = 300$ ；以此類推。相較於 H ， F_0 的振幅(amplitude)比較大，其音強(intensity)和音量(loudness)亦較大，因此一個人的嗓音音高(pitch)聽起來較高或較低，取決於 F_0 的高低，雖然較高的 F_0 通常聽起來有較高的音高，而反之亦然，但 F_0 與音高之間的關係並非線性。過去有項研究要求受試者將聽到的音高調至其雙倍的頻率，結果發現大致上受試者能正確地將 500 至 1000 Hz 的聲音調至其雙倍的頻率，但是在較高的音高，例如 2000 Hz 的純音，則需要調至 15000 Hz，才有相同的效果，而較低的音高，例如 125 Hz，只要增至 187 Hz，就有相同的效果。(Ville & Karjalainen, 2015)

頻率、語音與共振峰

嗓音是一種聲音(sound)，是聲帶震動所產生的週期性複雜聲波，由 F_0 和 H 所構成。語音則是嗓音和氣流通過聲道(vocal tract)從嘴發出來的聲音。所謂的聲道就是從鼻腔、口腔、咽腔到聲門上的區域，裡面包含嘴唇、牙齒、硬顎、軟顎、舌頭、下巴、聲門等構音器官(articulators)。構音時會移動構音器官使聲道的型態(configuration)，也就是其長度、大小與形狀，有所變化，而這些聲道的變化會讓一些特別的頻率的震幅增強。簡單來說，嗓音和氣流都是聲源(sound source)，當它們通過聲道時會使在聲道裡的空氣粒子震動，產生共振(resonance)，不同的構音動作會導致不同的聲道型態，其共振特性也有所改變，使得某些頻率帶的 H 的震幅變大。這

些經由聲道共振增強後的 H 最後產生共振峰(formant)，然而聲道共振不會影響 F0。(Gick et al., 2013)

聲學語音學(acoustic phonetics)通常會使用第一共振峰(F1)和第二共振峰(F2)來描述母音(vowel)的聲學特徵。F1 頻率與舌頭的高度(height)成反相關(inversely related)，例如高母音(high vowel)/i/的 F1 頻率比低母音(low vowel)/a/低；F2 頻率則與舌頭的後縮(backness)程度成反相關，例如後母音(back vowel)/u/的 F2 頻率比前母音(front vowel)/i/的低。大致上，相較於高母音，低母音的 F1 有較高的頻率，而相較於前母音，後母音的 F2 有較低的頻率，以及 F2 與 F1 頻率間的差異(F2 – F1)也較小。以構音語音學(articulatory phonetics)而言，F1 頻率反映構音時咽腔的共振特徵，例如說前母音/i/時舌頭往前使得咽腔寬敞，因此共振出較低的頻率；F2 頻率反映構音時口腔的共振特徵，例如說低母音/a/時舌頭往下使得口腔寬敞，因此共振出較低的頻率。(Kent & Read, 2002; Raphael et al., 2011)

僅使用 F1 頻率和 F2 頻率的數據，很難快速做比較或下結論，因此有學者提出使用圖表(graph)或值(value)來描述、對比、評估說話者的母音準確度(accuracy)。目前有三種常用的方式：母音聲學空間面積(vowel space area, VSA, 亦稱作 vowel working space)、母音中央化比值(formant centralization ratio, FCR)，以及母音構音指數(vowel articulation index, VAI)。

VSA 是利用定位母音(cardinal vowels),也就是/a/, /i/, /u/的 F2 – F1 算出面積值並繪圖(可使用 Microsoft Excel)，其公式為：

$ABS\{[F1i*(F2a-F2u)+F1a*(F2u-F2i)+F1u*(F2i-F2a)]/2\}$ (絕對值 absolute value, ABS)。以構音語音學而言，定位母音/a/, /i/, /u/的構音動作較為極端、舌頭的動作範圍較大，因此能代表說話者的母音構音動作的最大範圍，也與說話者的言語清晰度有關聯性，例如當 VSA 比較小，言語清晰度亦較差。有臨床研究發現顱內受傷(close head trauma)、腦性麻痺、巴金森症、肌萎縮性脊髓側索硬化症(amyotrophic lateral sclerosis, ALS)、氣切、舌切除、唐氏症等類型的患者相較於一般說話者有較小的 VSA。(Ball & Gibbon, 2013)

FCR 是把定位母音/a/, /i/, /u/的 F1 和 F2 帶入公式：

$(F2u + F2a + F1i + F1u)/(F2i + F1a)$ 。這個比值的最大值為 2，因此越接近 2 表示說話者的母音中央化越嚴重。以構音語音學的觀點，當說話者的母音越中央化表示其母音構音動作範圍越小，因此言語清晰度也較差。而 VAI 則是 FCR 的倒數，因此其正常數值非常接近 1.0。FCR 與 VAI 都曾在臨床實驗中被用於分析巴金森症患者的言語清晰度。(Ball & Gibbon, 2013; Sapir et al., 2011)

音高與聲調

所有的語言都使用母音、子音(consonant)以及這兩者的搭配來表達不同的字詞與含義。然而聲調語言(tone language)除了上述方法，還可同時搭配使用不同的音高來區分字詞，而這些有語義區分作用的音高稱為聲調(lexical tone)。國語(Mandarin Chinese, MC)有四個聲調外加輕聲，這些聲調的本質不太一樣，簡單來說，聲調可分為兩種：平調(level tone)與曲折聲調(contour tone)。如果聲調從開始到結束都沒有音

高變化就是平調，例如 MC 的一聲(tone 1 or high tone)；反之，如果聲調有音高變化就是曲折聲調，例如 MC 的二聲(tone 2 or rising tone)、三聲(tone 3 or low tone)和四聲(tone 4 or falling tone)。以聽辨而言，MC 的四個聲調，除了一聲較高的音高以外，其餘的聲調則因為都是曲折聲調，所以聽見音高轉折成為判斷的依據，例如二聲的音高上揚、四聲的音高下降，還有三聲的音高下降後上升的轉折。相較於上述的四個聲調，語音學專家們對輕聲(neutral tone)的定義有著不同的看法，一派學者將它視為 MC 的「第五個聲調」，認為其音強(intensity)較弱、其音長(duration)較短、其音高的型態(shape)以及輪廓(contour)會被其前面的音節所影響，也就是說輕聲沒有獨立且固定的音高，常見於語氣助詞(modal particle)例如嗎、呢、吧、結構助詞(structural particle)例如的、地、得，以及時態助詞(tense particle)例如著、了；另一派學者則認為輕聲是變調(tone sandhi)，只會發生在特定的一些語音情境(phonetic context)，常出現於雙字詞而非單獨出現，例如媽媽、看看、包子。(Duanmu, 2004; J. Shen, 2022; Yip, 2002)

音高與語調

除了聲調語言外，還有所謂的非聲調語言，例如語調語言(intonation language)，此類語言使用音高的變化來表達不同的語調(intonation)。在語言學裡，語調有三種主要的功能：語法、語用、情緒。在語法上，語調變化能區分肯定句-疑問句；在語用上，語調變化能讓交談者知道何時說話輪替(turn-taking)、訊息為新或舊；在情緒上，語調變化能表達與辨別說話者的情感和態度，這也是語調的類語言功能(paralinguistic function)。過去有很多學者相信聲調語言沒有語調，所以 MC 母語人士說話時只有聲調沒有語氣(tone of voice)，原因是音高已經被用來表達語義，說話者就無法再透過同樣的方式操弄聲帶震動速率，以表達上述語調的功能。不過研究結果卻發現 MC 是有語調的，原因有三：第一，在口語中，MC 有聲調變調(tone sandhi)，是個只會在連續說話時發生的現象，最著名的就是連續說三聲的字詞時頭幾個三聲的字會被說成二聲，例如我姓許，大家稱我為三聲的「許」老師，但在語音實現(phonetic realization)上卻是二聲的「徐」老師，這也是為何大多學生都稱我為許博士或家甄老師；第二，MC 的疑問句雖然不像英文的疑問句在句尾有音高上升的現象，卻是整個句子的音高都比肯定句來得高；第三，日常談話中，MC 母語人士也是能透過語調傳達與了解情感、態度、重點(focus)、說話輪替、訊息新舊。研究發現，MC 母語人士通常使用較高的音高來表達高興和生氣，而且能透過音高變化聽辨這兩種情緒。(Gussenhoven, 2004; Hsu, 2014; Hsu & Xu, 2014; Ladd, 2008; F. Liu & Xu, 2005; X. Liu & Xu, 2014; Shen, 1990; Vaissière, 2008; Yip, 2002)

音高、發聲與音質

發聲(phonation) 是聲帶震動的產物，其震動方式會因不同的聲帶生理狀態(physiological status of the vocal folds) 與聲門活動型態(glottal configuration) 而異。在聲學語音學中，說話時會呈現兩種聲門活動型態：聲門閉合且震動而產生有聲語音(voiced speech sounds)，以及聲門打開且不震動而產生無聲語音(voiceless speech sounds)。而在嗓音音質(vocal quality) 研究中，通常會把常態聲(modal voice)、氣息

聲(breathy voice)和喉嘎聲(glottal fry or creaky voice)做比較。常態聲是指一般正常說話和唱歌時的音域，發聲時聲帶往中線(midline)靠攏，不會過度用力也沒有閉合不全的現象。音域(vocal register)是由聲帶以特定的震動模式所產生的音高範圍(pitch range)。當說話者使用氣息聲，其聲門閉合不全，導致「漏氣」而產生非週期性噪音(aperiodic noise)。臨床上氣息聲與女性嗓音、嗓音濫用、聲帶麻痺、聲帶創傷、等現象與疾病息息相關。在聲學語音學中，氣息聲亦能做為音韻區別性特徵(phonological distinctive feature)，例如 MC 的一聲除了音高比其他聲調高，還有較多的氣息聲，而研究也發現氣息聲有助於聽者判斷 MC 的一聲。喉嘎聲則是說話者最低的音域，其發聲方式是讓聲帶鬆弛而緩慢地震動，因此使用這樣的發聲方式讓說話者的音高偏低，其特徵是類似氣泡的聲音，也就是所謂的 fry。喉嘎聲通常被視為導致嗓音濫用的一種行為，還有研究發現，越來越多年輕女性說話時使用喉嘎聲，但是她們通常會被評斷為較不能幹、有較低教育水平、比較不值得信賴、缺乏吸引力，以及不可被雇用。而在聲學語音學中，與氣息聲一樣，喉嘎聲也能做為音韻區別性特徵，例如 MC 的三聲除了有降升(falling-rising)的音高轉折外，外加喉嘎聲可有效地將其與二聲做區分。(Anderson et al., 2014; Hillenbrand et al., 1994; Kuang, 2017; Surendran & Levow, 2008)

以嗓音異常而言，主要的臨床治療目的包括促進個案在日常生活中使用常態聲。因此，大多臨床研究著重於測量、評估並了解氣息聲的生理成因和聲學特徵，以有效診斷、改善以及預防嗓音濫用。雖然不能直接用音高來區分或定義不同的音質，但卻能透過基頻和諧音之間的數學關係描述氣息聲的嚴重度。目前常用與氣息聲有關的聲學度量(acoustic measure)有三種：訊號週期性(signal periodicity)、H1 的振幅，以及頻譜斜率(spectral tilt)。訊號週期性預測氣息聲的效果最佳，與聽辨者相比，能達到 80% 的正確率，其中又以倒頻譜峰值顯著性(cepstral prominence peak, CPP)最有希望成為診斷嗓音異常嚴重度的主要聲學度量，美國聽語學會(ASHA)在其 2018 年的「嗓音功能儀器檢查指引」中建議使用 CPP 來取代頻率擾動度(jitter)、振幅擾動度(shimmer)和諧噪比(harmonic-to-noise ratio, HNR)，而這三項傳統的聲學度量需要偵測以計算基頻，但是當個案的嗓音異常達到中度以上則非常難取得其基頻，因此也無法有效計算出這些聲學度量。簡單來說，CPP 是一種以倒頻譜的方式所計算出的諧噪比，可從連續口語(connected speech)與持續母音(sustained vowel)提取，不需要直接計算基頻。近期研究發現，在連續口語中 CPP 低於 6.11 dB 以及持續母音中 CPP 低於 11.46 dB 皆表示有極高的嗓音異常風險。另一項常用於研究的聲學度量是 H1-H2，其演算法是將第一諧波與第二諧波的振幅強度(以分貝 dB 為單位)相減，如果 H1-H2 的數值越大，代表頻譜斜率越大，氣息聲則越明顯。(Hillenbrand et al., 1994; Murton et al., 2020; Patel et al., 2018)

結論

本文章透過聲學參數「頻率」介紹了與其相關的語音與嗓音科學的主要概念：基頻與諧波、音高、共振峰、聲調、語調與嗓音音質。雖然都只是淺談，畢竟每個概念都有更多更深的學問，不過我希望透過文章中簡明的描述，加深各位對聲學的

印象及喜好，也期待著未來在台灣能看見更多以國語為主的聲學相關研究。

參考文獻

- Anderson, R. C., Klofstad, C. A., Mayew, W. J., & Venkatachalam, M. (2014). Vocal fry may undermine the success of young women in the labor market. *PLoS ONE*, 9(5), 1–8. <https://doi.org/10.1371/journal.pone.0097506>
- Ball, M. J., & Gibbon, F. E. (Eds.). (2013). *Handbook of Vowels and Vowel Disorders*. Psychology Press.
- Duanmu, S. (2004). Tone and non-tone languages: An alternative to language typology and parameters. *Language and Linguistics*, 5(4), 891–924.
- Gick, B., Wlison, I., & Derrick, D. (2013). *Articulatory Phonetics*. Blackwell Publishing.
- Gussenhoven, C. (2004). *The Phonology of Tone and Intonation*. Cambridge University Press. <https://doi.org/10.1017/CBO9780511616983>
- Hillenbrand, J., Cleveland, R. A., & Erickson, R. L. (1994). Acoustic correlates of breathy vocal quality. *Journal of Speech and Hearing Research*, 37(4), 769–778. <https://doi.org/10.1044/jshr.3704.769>
- Hillenbrand, J., & Houde, R. A. (1996). Acoustic correlates of breathy vocal quality: Dysphonic voices and continuous speech. *Journal of Speech and Hearing Research*, 39, 311–321.
- Hsu, C. (2014). Emotional prosody production in adolescents with autism. *ASHA Convention*.
- Hsu, C., & Xu, Y. (2014). Can adolescents with autism perceive emotional prosody? *INTERSPEECH*, 1924–1928.
- Johnson, K. (2012). *Acoustic and Auditory Phonetics (3rd Ed.)*. Blackwell Publishing.
- Kent, R. D., & Read, C. (2002). *The Acoustic Analysis of Speech (Second Ed.)*. Singular Thompson Learning.
- Kuang, J. (2017). Covariation between voice quality and pitch: Revisiting the case of Mandarin creaky voice. *The Journal of the Acoustical Society of America*, 142(3), 1693–1706. <https://doi.org/10.1121/1.5003649>
- Ladd, D. R. (2008). *Intonational Phonology (2nd Ed.)*. Cambridge University Press.
- Liu, F., & Xu, Y. (2005). Parallel encoding of focus and interrogative meaning in Mandarin intonation. *Phonetica*, 62(2–4), 70–87. <https://doi.org/10.1159/000090090>
- Liu, X., & Xu, Y. (2014). Body size projection by voice quality in emotional speech-Evidence from Mandarin Chinese. *Proceedings of the International Conference on Speech Prosody*, 974–977. <https://doi.org/10.21437/speechprosody.2014-183>
- Murton, O., Hillman, R., & Mehta, D. (2020). Cepstral peak prominence values for clinical voice evaluation. *American Journal of Speech-Language Pathology*, 29(3),

- 1596–1607. https://doi.org/10.1044/2020_AJSLP-20-00001
- Patel, R. R., Awan, S. N., Barkmeier-Kraemer, J., Courey, M., Deliyiski, D., Eadie, T., Paul, D., Švec, J. G., & Hillman, R. (2018). *Recommended Protocols for Instrumental Assessment of Voice: American Speech- Language-Hearing Association Expert Panel to Develop a Protocol for Instrumental Assessment of Vocal Function*. *American Journal of Speech-Language Pathology*, 27, 887–905.
 - Raphael, L. J., Borden, G. J., & Harris, K. S. (2011). *Speech Science Primer: Physiology, acoustics, and perception of speech* (6th Ed.). Lippincott Williams & Wilkins.
 - Sapir, S., Ramig, L., Spielman, J., & Fox, C. (2011). *Acoustic metrics of vowel articulation in Parkinson's disease: Vowel space area (VSA) vs. Vowel articulation index (VAI)*. *Models and Analysis of Vocal Emissions for Biomedical Applications - 7th International Workshop, MAVEBA 2011*, 9, 173–175.
 - Shen, X. S. (1990). *The prosody of Mandarin Chinese*. University of California Press.
 - Surendran, D., & Levow, G.-A. (2008). *Can voice quality improve Mandarin tone recognition? ICASSP*, 58–61. <https://doi.org/10.1109/ICASSP.2008.4518575>
 - Vaissière, J. (2008). *Perception of Intonation*. In D. B. Pisoni & R. E. Remez (Eds.), *The Handbook of Speech Perception* (pp. 236–263). Blackwell Publishing. <https://doi.org/10.1002/9780470757024.ch10>
 - Ville, P., & Karjalainen, M. (2015). *Communication Acoustics : An Introduction to Speech, Audio and Psychoacoustics*. Blackwell Publishing.
 - Yip, M. (2002). *Tone*. Cambridge University Press.

關於作者

現任	黃自強語言治療所 思比語言治療所
學歷	英國倫敦大學學院言語聽力科學暨語言學博士 國立台北護理健康大學語言治療與聽力學系所碩士 英國里茲大學語言學碩士
經歷	馬偕醫學院 聽力暨語言學系 助理教授 亞洲大學聽力暨語言治療學系 助理教授 亞洲大學附設醫院復健科 語言治療師



編 輯

發行單位：台灣聽力語言學會

發行人：葉文英

主 編：簡欣瑜

編輯顧問：洪右真

網 址：www.slh.org.tw

發行日期：2022.12.07

聽語學報：第 105 期

編輯群：郭婉倫、黃友琳、林峯全、

席 芸、張偉倩、陳昱彤

助理編輯：陳德容