

嗓音声学分析在嗓音疾病中的研究

殷亚杰¹ 蔡晓²

1. 青海大学研究生院 青海西宁 810000

2. 青海省人民医院耳鼻咽喉科 青海西宁 810000

摘要: 基于数字时代的来临,使得人们可以借助短视频分享生活、自我表达、更有越来越多的职业用声者,比如教师、歌唱者、主播行业等社会背景下,加上不正确用嗓、吸烟、饮酒、熬夜等不良生活习惯,使得深受嗓音疾病影响的人们正在逐年增加。关于嗓音疾病的评估也越来越成为嗓音学者们关注的重点,而作为无创性的嗓音声学分析技术可为嗓音疾病的评估提供更客观的指标,可用于各种不同目的的研究,为筛查、精确诊断、鉴别诊断、监测嗓音疾病的治疗过程的提供可靠的客观依据。本文旨在梳理嗓音声学分析在嗓音疾病中的特点和价值,作一综述。

关键词: 嗓音疾病; 嗓音声学分析; 声学参数; 嗓音障碍指数

嗓音是人类交流和表达的重要方式,其声学特性反映了发声器官的生理和病理状态。嗓音疾病是由于各种器质性或功能性原因引起的发声障碍性疾病的统称,轻者嗓音间歇性嘶哑,发音易疲劳,音色低沉、粗糙;重者持续性嘶哑,甚至失音,导致患者产生焦虑等情绪影响其正常社交及心理健康。长期声音嘶哑可能会长声带肿物,甚至出现恶性肿瘤。发声涉及生理,听觉感知,空气动力学,声学和情感方面,这意味着需要一个多学科和跨学科的方法在评估,诊断和治疗的声音障碍^[1]。2001年,欧洲喉科学会(ELS)提出了一项评估嗓音相关疾病的基本协议,推荐的嗓音评估包含五个方面:听觉-感知评估、声学分析、空气动力学、患者自我评价和频闪喉镜。其中建议使用嗓音声学分析作为诊断工具^[2]。嗓音的客观评估主要是借助仪器设备,而嗓音声学分析因其客观描述嗓音并能发现嗓音问题,且提供支持证据及判断疾病的严重程度,评估和监测声音特点,也助于评估对治疗的反应并得出结果,因此越来越受嗓音学者们的关注^[3]。

1 什么是嗓音声学分析

嗓音声学分析是涉及呼吸,发声,音域等多方面语音内容的重要内容,它是旨在根据人们发声时记录嗓音波形自动识别发声的客观特征。嗓音声学分析不仅可用于检测和分析病理性嗓音也可用于检测和分析正常健康人的嗓音及艺术家的嗓音的声学特征,它是一种操作简单、定量输出、价格便宜且安全可靠,可重复的嗓音质量的评估方法。随着现代医学、物理学、统计学和人工智能技术的发展,近年来嗓

音声学分析的研究,特别是在声学参数的开发和适用性上有了新的进展^[4]。嗓音声学分析背后的理论之一是声带质量或声带张力的变化导致声音样本中的声音增加和可测量的不规则性或噪音。它可能是声带附近或声带处的噪声源(即声带之间的气流),或者声带声门波的显著非周期性^[5]。在嗓音疾病的背景下,嗓音声学分析越来越是临床医生工作者对嗓音障碍患者的首选评估方法,它通过对嗓音信号的测量和处理,提取相关参数,以定量和客观的方式描述嗓音特征,为嗓音医学的研究和临床应用提供了有价值的信息。

2 嗓音分析的方法

2.1 时域分析

主要是对声音信号随时间变化的特征进行研究,包括对嗓音信号的持续时间、起因时间、嗓音中断情况及振幅变化等的测量,如振幅微扰、周期微扰等参数,用于评估嗓音的稳定性。时域分析的优点在于直观、易于理解和计算。然而,它在反映声音的频率特性方面存在一定的局限性。通常,时域分析会与频域分析等其他分析方法结合使用,以更全面地评估嗓音特征。

2.2 频域分析

在嗓音声学分析中,频域分析是一种重要的方法,主要用于研究声音信号中不同频率成分的分布和特征。通过傅里叶变换等方法将时域信号转换为频域,获取基频、谐波、共振峰等信息,以了解嗓音的频率特征。频域分析有助于深入了解嗓音的声学特性和发声机制,对于嗓音障碍的诊断、

评估和治疗，以及语音合成、声学研究等领域都具有重要的应用价值。

2.3 非线性分析

嗓音声学分析中的非线性分析是一种用于研究嗓音复杂动态特性的方法。其中，较为常见的非线性分析方法包括：

1. 分形分析：用于描述嗓音信号的复杂性和不规则性。通过计算分形维数，可以定量评估嗓音的不规则程度。分形维数的变化可能反映出声带振动模式的改变或喉部疾病的存在。
2. 混沌分析：例如通过计算最大 Lyapunov 指数来判断嗓音系统是否具有混沌特性。如果存在混沌特征的改变，可能提示发声系统的异常。
3. 相空间重构：将嗓音信号转换到相空间中，以揭示其潜在的动态结构和模式。非线性分析方法能够捕捉到嗓音产生过程中的非线性和复杂性特征，为嗓音疾病的诊断和发声机制的研究提供了新的视角和更深入的信息。

3 主要声学参数在常见嗓音障碍中的定义及特点

(1) 基频 (fundamental frequency, F0) 及基频标准差 (Standard deviation of fundamental frequency, SDF0)：基频即声门开闭使得声带振动而产生的能量集中的最低固有频率，它的大小与声带的长度、厚薄、质量及张力有关^[5]。基频受性别、年龄、情绪、发声方式等多种因素影响。例如，成年男性的基频通常低于成年女性。

(2) 微扰 (perturbation)：分为基频微扰 jitter 和振幅微扰 shimmer，二者分别反映嗓音的嘶哑程度及粗糙程度，是反映声带不规则性振动的有效指标^[6, 13]。

(3) 最长发声时间 (Maximum phonatory time, MPT)：是一个空气动力学参数，描述了元音可以连续发声的最大长度，并以秒表示^[5]。连续发声的能力可间接反映声门闭合程度及喉部肌肉的状态。Tom Karlsen 等认为 MPT、声学分析和嗓音障碍之间可能存在的关联^[5]。

(4) 发声障碍严重程度指数 (Dysphonia Severity Index, DSI) 主要与最低强度、最高频率、最长发声时间及基频微扰有关，是综合呼吸和发声参数评价发声障碍的通用参数。

(5) 谐噪比 (harmonic-to-noise ratio, HNR)：也称为信噪比，基频整数倍的频率成分称为谐波。谐波的分布和强度能够反映声带振动的规则性和嗓音的音质。谐波与噪声的比值即为 HNR，描述了信号的周期性程度^[7]。

(6) 共振峰：是声音频谱中的能量集中区域，其频率和强度对于语音的清晰度和可懂度具有重要意义。在元音发声中，共振峰的特征较为明显。

(7) 倒频谱：是在频谱的基础上再次进行傅里叶转换得到的，具有高尖峰值的图形，即为倒频谱峰值 (cepstral peak prominence, CPP) 是倒频谱相关的一个重要参数，其测量的是倒谱峰距倒谱回归线的距离^[8]。2018年，美国语音-语言-听力协会 (ASHA) 的指导建议将 CPP 作为“测量语音信号中噪音的整体水平”和“测量幻听的一般方法”^[9]。

(8) 声压级 (Sound Pressure Level, SPL)：是基于声波的压力变化用来描述声音强度的物理量，单位为分贝 (dB)，表示声音强弱大小。

嗓音的产生是一个复杂的过程，涉及到气流通过声门时引起声带的振动，进而在声道中产生共鸣和滤波。文献中报道了许多嗓音声学参数，但基频 (F0)、声压级 (SPL) 和微扰测量 (如 jitter 和 shimmer) 已被报道为语音评估的可靠测量。如谐波测量 (如谐波噪声比 (HNR))、基于倒频谱的测量 (如倒频谱峰值突出平滑 (CPPS))、频谱测量 (如 α 比率) 和测量 (如基音周期熵 (PPE)) 已被添加到研究和临床环境中评估嗓音功能的选项中^[7, 9-11]。基于倒频谱的测量已成为识别病理性嗓音并确定嗓音疾病的严重程度更普遍和更有价值的指标之一^[12]。此外，频谱测量可有效区分健康声音与粗糙、呼吸急促和紧张的声音^[13]。基于倒频谱的测量似乎更适合于定义嗓音问题的严重程度，频谱测量更适合于区分不同的语音障碍，谐波水平更有效的区分具有类似语音障碍的人群^[10]。值得注意的是：PPE、shimmer、jitter、CPPS 和 HNR 五个参数的组合被认为嗓音障碍筛查中的强预测因子^[10]。具体来说，与没有嗓音障碍的人相比，嗓音障碍的患者表现出显著较低的 α 比率、CPPS、HNR 和 F0 值^[10]。相比之下，与没有语音障碍的人相比，有嗓音障碍的患者表现出更高的 shimmer, jitter, PPE, 声压级标准差 (SPL SD) 和 SDF0 值。CPPS、HNR、shimmer 和 jitter 是在区分发声障碍和非发声障碍组中显示出更好的预测能力的参数，将这些发现整合到临床方案中可以促进更准确和个性化的治疗计划，最终改善这些特定嗓音障碍患者的整体管理^[10]。当然嗓音声学分析的参数广泛应用于喉部疾病、病理嗓音、痉挛性发音障碍等病理嗓音的研究。用 F0、SDF0、Jitter、Shimmer 评估声带息肉患者、早期喉癌患者、

突变假声患者和健康组的声音质量，证明上述参数是评估持续发声过程中音高变化的重要指标^[14]。声音是一个多维的声学现象，单个声学参数测量（例如 CPPS、F0、Jitter、Shimmer、HNR、最大发声时间）与患者对嗓音障碍的自我评估存在弱相关性^[5]，故而嗓音评估需要多参数的多个维度以及了解每个维度有助于提供更好的评估的是非常重要性。因此，声学多参数指标与其他方面的声音评估的探索可以为临床实践带来补充信息^[15]。

4 嗓音声学分析在常见嗓音障碍中的价值

嗓音疾病患者的声学分析与健康人之间的声学指标存在诸多差异，因其可描述的嗓音参数众多，现今并未有明确嗓音疾病间的确切参数，尤其是嗓音疾病并不能依靠单一的声学参数来准确诊断，而是需要多参数协同。Patrizia Vizza 等^[16]通过 39 个健康受试者与 60 名帕金森病 (PD) 患者和 54 名多发性硬化患者的嗓音障碍评估结果示：F0、jitter、shimmer、NHR 在统计学中有显著的特征。PD 患者由于声门控制能力下降，进而气流紊乱形成噪音，从而导致了 HNR 下降，因此它们可以用作诊断构音障碍相关疾病（如 PD）的指标^[17]。董周威等^[18]通过对 150 例不良发声行为性嗓音疾病患者和 98 例健康对照组进行客观评估，嗓音分析结果示：疾病组与健康组在平均言语 F0、SDF0、Jitter、Shimmer、S/Z、MPT 方面差异有统计学意义 ($P < 0.05$)，并且疾病组治疗前和治疗后的值有显著的差别意义。除用于嗓音疾病评估外，上述参数还被用于研究手术治疗对嗓音的影响、术后嗓音的恢复情况及并发症预测、言语治疗的效果评估等方面，如许倩等^[19]通过对 45 例女性声带小结手术组与 45 例正常对照组对比。运用嗓音声学分析得出：Jitter、Shimmer 及 DSI 可以作为声带小结手术疗效评价的敏感指标。Chhetri SS, Gautam R. 等^[3]在不同的喉部疾病中，以双侧声带小结最常见，其次为声带息肉。发声治疗前后的声学分析表明，声学分析客观地量化了结果，为嗓音质量、嗓音状态和功能的评价提供了一种有效的方法。Rita Alegria 等^[6]在双侧声带小结患者经嗓音治疗后嗓音声学参数是否有改善？Meta 分析得出结论声学分析在双侧声带小结的患者中发挥临床作用，通过客观评估和提供嗓音治疗对嗓音声学参数的影响数据。利用声学分析对双侧声带小结患者的嗓音声学参数进行客观评估，这项分析表明，声带小结患者在接受嗓音治疗后改善了他们的声学语音参数，为嗓音治疗效果提供数

据，具有临床应用价值。Tom Karlsen 等^[5]得出在癌症和退行性或炎性的喉部疾病患者中，VHI 评分与 Jitter、Shimmer 和 NHR 评分相关，在组水平上，由特定喉部疾病研究的嗓音声学分析反映了嗓音病理是否存在。此外，在某种程度上反映了这种病理学的严重程度。因此，这些分析可能是重要的替代变量，尤其是对于有关的潜在治疗效果。目前尤其是 Jitter、Shimmer 和 NHR 的情况，Gillespie 等人是支持的^[15]。为了研究差异功效是否可以在个体水平上使用，进一步研究喉部患者声学分析于与对照组的差异将是未来的目标。在不同的患者之间，目前声学分析似乎只在组水平上最有用。^[5]如果以多中心方式进行新的研究，人们可获得将声学测量添加到个体喉部患者的诊断检查中的更准确的数据^[5]。在癌症和退行性或炎性的喉部疾病患者中，声学分析潜在地增加了对喉科患者研究的有用信息^[5]。尽管声音的客观评估不能代替主观评价，但声学分析已在临床上用于区分正常和异常声音，帮助鉴别诊断并评估不同治疗方法的相对有效性，以及在整个治疗过程中监测患者声音的变化情况^[20]。

5 小结及展望

嗓音声学分析作为一种非侵入性、客观有效的手段，在嗓音相关的研究和实践中发挥着关键作用。近些年，随着计算机技术和信号处理算法的不断发展，嗓音声学分析的精度和可靠性得到了提高。但，仍面临一些挑战，如复杂嗓音疾病的准确诊断、个体差异的影响、环境噪声的干扰等。后续的研究还需要集中在扩大样本量和涵盖更广泛的嗓音疾病，以确保在临床实践中确定的诊断参数的稳健性和适用性^[10]。未来，随着技术的进一步创新和多学科融合，有望为嗓音健康和发声艺术的发展提供更强大的支持。综上所述，关于嗓音评估还需要更科学、更全面的参数协助诊断且能矫治嗓音障碍，但可以确定嗓音声学分析的参数会成为嗓音疾病诊疗过程中的重要依据之一，且具有广泛的应用前景和较高的临床价值。关于嗓音疾病的声学评估还需要嗓音学者们的进一步探讨。

参考文献：

[1] Lopes LW, Batista Simões L, Delfino da Silva J, da Silva Evangelista D, da Nóbrega E Ugulino AC, Oliveira Costa Silva P, Jefferson Dias Vieira V. Accuracy of Acoustic Analysis Measurements in the Evaluation of Patients With Different Laryngeal Diagnoses. *J Voice*. 2017 May;31(3):382.e15–382.e26.

doi:10.1016/j.jvoice.2016.08.015. Epub 2016 Oct 11. PMID: 27742492.

[2]Dejonckere PH, Bradley P, Clemente P, Cornut G, Crevier-Buchman L, Friedrich G, Van De Heyning P, Remacle M, Woisard V; Committee on Phoniatics of the European Laryngological Society (ELS). A basic protocol for functional assessment of voice pathology, especially for investigating the efficacy of (phonosurgical) treatments and evaluating new assessment techniques. Guideline elaborated by the Committee on Phoniatics of the European Laryngological Society (ELS). *Eur Arch Otorhinolaryngol*. 2001 Feb;258(2):77–82. doi: 10.1007/s004050000299. PMID: 11307610.

[3]Chhetri SS, Gautam R. Acoustic Analysis Before and After Voice Therapy for Laryngeal Pathology. *Kathmandu Univ Med J (KUMJ)*. 2015 Oct–Dec;13(52):323–7. doi: 10.3126/kumj.v13i4.16831. PMID: 27423282.

[4] 刘思维, 邵骏. 嗓音声学分析方法的研究现状 [J]. *临床耳鼻咽喉头颈外科杂志*, 2022,36(12):966–970+976. DOI:10.13201/j.issn.2096–7993.2022.12.016.

[5]Karlsen T, Sandvik L, Heimdal JH, Aarstad HJ. Acoustic Voice Analysis and Maximum Phonation Time in Relation to Voice Handicap Index Score and Larynx Disease. *J Voice*. 2020 Jan;34(1):161.e27–161.e35. doi: 10.1016/j.jvoice.2018.07.002. Epub 2018 Aug 6. PMID: 30093166.

[6]Alegria R, Freitas SV, Manso MC. Is there an improvement on acoustic voice parameters in patients with bilateral vocal fold nodules after voice therapy? a meta-analysis. *Eur Arch Otorhinolaryngol*. 2020 Aug;277(8):2163–2172. doi: 10.1007/s00405–020–05956–2. Epub 2020 Apr 15. PMID: 32296978.

[7]Viswanathan R, Arjunan SP, Bingham A, Jelfs B, Kempster P, Raghav S, Kumar DK. Complexity Measures of Voice Recordings as a Discriminative Tool for Parkinson's Disease. *Biosensors (Basel)*. 2019 Dec 20;10(1):1. doi: 10.3390/bios10010001. PMID: 31861890; PMCID: PMC7168233.

[8] 陈爽, 吕士良. 嗓音声学分析在临床医学领域的应用研究 [J]. *西北民族大学学报 (自然科学版)*, 2022,43(03):15–20. DOI:10.14084/j.cnki.cn62–1188/n.2022.03.010.

[9]Murton O, Hillman R, Mehta D. Cepstral Peak

Prominence Values for Clinical Voice Evaluation. *Am J Speech Lang Pathol*. 2020 Aug 4;29(3):1596–1607. doi: 10.1044/2020_AJSLP–20–00001. Epub 2020 Jul 13. PMID: 32658592; PMCID: PMC7893528.

[10]Cantor–Cutiva LC, Ramani SA, Walden PR, Hunter EJ. Screening of Voice Pathologies: Identifying the Predictive Value of Voice Acoustic Parameters for Common Voice Pathologies. *J Voice*. 2023 Dec 23:S0892–1997(23)00390–9. doi: 10.1016/j.jvoice.2023.12.005. Epub ahead of print. PMID: 38143203.

[11]Lopes LW, Sousa ESDS, Silva ACFD, Silva IMD, Paiva MAA, Vieira VJD, Almeida AA. Cepstral measures in the assessment of severity of voice disorders. *Codas*. 2019 Aug 15;31(4):e20180175. Portuguese, English. doi: 10.1590/2317–1782/20182018175. PMID: 31433040.

[12]S V N, K V. Spectral Measures of Hoarseness in Persons with Hyperfunctional Voice Disorder. *J Voice*. 2017 Jan;31(1):57–61. doi: 10.1016/j.jvoice.2016.03.005. Epub 2016 Apr 11. PMID: 27080591.

[13]de Abreu SR, Sousa ESDS, de Moraes RM, Lopes LW. Performance of Acoustic Measures for the Discrimination Among Healthy, Rough, Breathily, and Strained Voices Using the Feedforward Neural Network. *J Voice*. 2022 Aug 23:S0892–1997(22)00203–X.

[14]Li G, Hou Q, Zhang C, Jiang Z, Gong S. Acoustic parameters for the evaluation of voice quality in patients with voice disorders. *Ann Palliat Med*. 2021 Jan;10(1):130–136. doi: 10.21037/apm–20–2102. Epub 2020 Dec 31. PMID: 33440977.

[15]Englert M, Latoszek BBV, Behlau M. Exploring The Validity of Acoustic Measurements and Other Voice Assessments. *J Voice*. 2024 May;38(3):567–571. doi: 10.1016/j.jvoice.2021.12.014. Epub 2022 Jan 10. PMID: 35022153.

[16]Vizza P, Tradigo G, Mirarchi D, Bossio RB, Lombardo N, Arabia G, Quattrone A, Veltri P. Methodologies of speech analysis for neurodegenerative diseases evaluation. *Int J Med Inform*. 2019 Feb;122:45–54. doi: 10.1016/j.ijmedinf.2018.11.008. Epub 2018 Nov 26. PMID: 30623783.

[17]Rusz J, Tykalova T, Ramig LO, Tripoliti E. Guidelines for Speech Recording and Acoustic Analyses in Dysarthrias of

Movement Disorders. *Mov Disord.* 2021 Apr;36(4):803-814. doi: 10.1002/mds.28465. Epub 2020 Dec 29. PMID: 33373483.

[18] 董周威, 成人不良发声行为性嗓音疾病患者嗓音障碍分析及矫治策略. 黑龙江省, 哈尔滨市第四医院, 2018-06-14.

[19] 许倩, 陈磊, 赵运华, 等. 声带小结患者显微手术前后嗓音声学分析[J]. 听力学及言语疾病杂志, 2012, 20(02): 113-

115.

[20] Hillenbrand JM (2011) Acoustic analysis of voice: a tutorial. *Perspect Speech Sci Orofac Disord* 21:31 - 43

作者简介:

殷亚杰(1993—), 女, 汉族, 籍贯: 河南周口, 学历: 在读硕士研究生, 单位: 青海省人民医院, 职称: 住院医师, 研究方向: 耳鼻咽喉头颈外科。