Perspectives of the ASHA Special Interest Groups
SIG 3, Vol. 2(Part 1), 2017, Copyright © 2017 American Speech-Language-Hearing Association

Recent Innovations in Voice Assessment Expected to Impact the Clinical Management of Voice Disorders

Jarrad H. Van Stan

Center for Laryngeal Surgery & Voice Rehabilitation, Massachusetts General Hospital
Communication Sciences and Disorders, Massachusetts General Hospital
Boston, MA
Institute of Health Professions
Charlestown, MA

Daryush D. Mehta

Center for Laryngeal Surgery & Voice Rehabilitation, Massachusetts General Hospital
Boston, MA
Department of Surgery, Harvard Medical School
Boston, MA
Communication Sciences and Disorders, MGH Institute of Health Professions
Charlestown, MA

Robert E. Hillman

Center for Laryngeal Surgery & Voice Rehabilitation, Massachusetts General Hospital
Boston, MA
Department of Surgery, Harvard Medical School
Boston, MA
Communication Sciences and Disorders, MGH Institute of Health Professions
Charlestown, MA

Disclosures

Financial: Robert E. Hillman has a financial interest in the Ambulatory Phonation Monitor based on a contractual agreement between Sensimetrics, Inc. (R&D for the initial version of the Ambulatory Phonation Monitor) and KayPENTAX, Inc. (manufacturer of the Ambulatory Phonation Monitor). Jarrad H. Van Stan and Daryush D. Mehta have no financial interests to disclose.

Nonfinancial: Jarrad H. Van Stan has no relevant nonfinancial interests to disclose. Daryush D. Mehta and Robert E. Hillman have filed a patent application for the methodology of subglottal impedance-based inverse filtering: Zañartu, M., Ho J. C., Mehta, D. D., Wodicka G. R., & Hillman, R. E.. *System and methods for evaluating vocal function using an impedance-based inverse filtering of neck surface acceleration.* International Patent Publication Number WO 2012/112985. Published August 23, 2012.

Abstract

This article provides a summary of some recent innovations in voice assessment expected to have an impact in the next 5–10 years on how patients with voice disorders are clinically managed by speech-language pathologists. Specific innovations discussed are in the areas of laryngeal imaging, ambulatory voice monitoring, and "big data" analysis using machine learning to produce new metrics for vocal health. Also discussed is the potential for using voice analysis to detect and monitor other health conditions.

Introduction

For speech-language pathologists (SLPs), the purpose of a clinical voice assessment is to fully characterize the impact of a voice disorder on a patient's vocal function and to determine the effects of the disorder on the patient's ability to function in daily life. This information is then used as a basis for determining treatment, prognosis, and the need for additional referrals (Roy et al., 2013). Qualitative judgments of vocal impairment and voice-related disability are typically obtained using instruments such as the Consensus Auditory-Perceptual Evaluation of Voice (Kempster, Gerratt, Verdolini Abbott, Barkmeier-Kraemer & Hillman, 2009) and Voice-Related Quality of Life (Hogikyan & Sethuraman, 1999). Although these rating schemes yield crucial outcome metrics, such patient- or clinician-based perceptual measures are prone to unreliability inherent to human judgment (Kreiman, Gerratt, Kempster, Erman, & Berke, 1993). Therefore, to increase measurement reliability, clinicians have historically searched for quantitative/objective measures of voice production that correlate with perceptual constructs such as degree of dysphonia (Awan, Roy, Jetté, Meltzner, & Hillman, 2010), pathological conditions such as vocal hyperfunction (Stepp, Merchant, Heaton, & Hillman, 2011), and mechanisms such as vocal fold collision (Titze & Hunter, 2015; Zañartu, Galindo, et al., 2014). The purpose of this paper is to highlight some recent advancements in the development of measures for assessing vocal function, especially those expected to impact the clinical management of patients with voice disorders in the next 5–10 years.

Innovations in Laryngeal Imaging

Laryngeal videostroboscopy remains the primary method for clinical assessment of voice disorders. However, videostroboscopy has some well-known limitations that have helped spawn advances in imaging technology. Such limitations include a reliance on periodic vocal fold vibration to obtain a stable videostroboscopic recording and the fact that the resulting images represent averages of multiple cycles of vibration (Mehta & Hillman, 2012). Thus, it is not possible to observe aperiodic vibration or other transient phenomenon that impact individual cycles of vibration (e.g., voice onset and offset, voice breaks, etc.), which could be helpful in better determining the pathophysiology of some voice disorders. These limitations have driven the development of high-speed imaging capabilities such as videokymography (Švec & Schutte, 1996) and high-speed videoendoscopy that record over 4,000 pictures per second (Deliyski & Hillman, 2010). Although these techniques provide the potential to accurately study clinically meaningful concepts, such as gaining a better understanding of the impact of aperiodic vocal fold vibration on laryngeal sound production (Mehta, Zeitels, et al., 2012), high-speed imaging technologies have not gained traction in everyday clinical practice due to high cost, limited availability of systems designed for clinical use (e.g., cameras tend to be large and cumbersome), and challenges associated with managing/analyzing the large amount of imaging data that are produced.

Technological advances continue to reduce the size and cost of digital high-speed video cameras, and it has become technically feasible to perform videostroboscopy and high-speed imaging with the same camera. These advances provide the basis for the future development of hybrid clinical systems that default to performing videostroboscopy but also enable triggering of high-speed imaging to capture selected short phonatory events of interest (Deliyski & Hillman, 2010). There have also been advances in digital image processing techniques for automating and increasing the efficiency of reviewing recordings from laryngeal high-speed imaging, as well as methods for enhancing/optimizing the visual inspection of the recorded images (Bonilha, Deliyski, Whiteside, & Gerlach, 2012)—all of which should be implemented in future clinical systems.

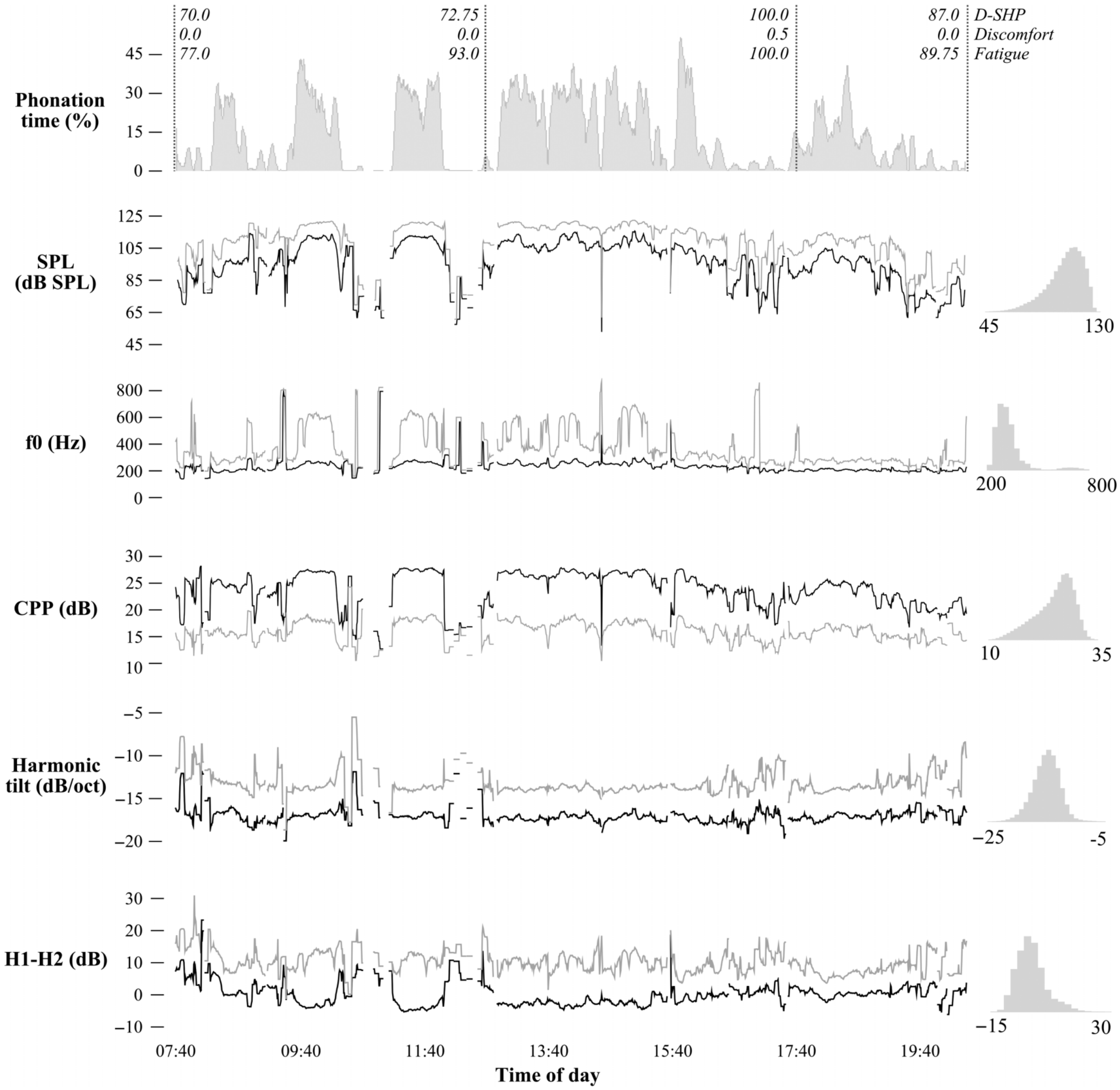
Another limitation of standard laryngeal videostroboscopy is that it only images the superior and some of the medial surface of the vocal folds during vibration, which forces clinicians to make "educated guesses" about the health of the vocal fold sub-epithelial layers based on surface tissue behavior. This limitation has inspired efforts to apply/adapt a technique called optical coherence tomography (OCT) to image the sub-epithelial layers of the vocal folds during static and dynamic (vibratory) conditions. For imaging vocal folds, OCT uses the differential interaction/reflection of laser light to produce pictures of surface and sub-surface (~2 mm) tissue structure (Burns et al., 2005). Initial studies using OCT have demonstrated that these pictures of tissue depth/detail may improve an otolaryngologist's ability to discriminate various types of vocal fold lesions, more accurately describe the depth of the lesion, and more clearly identify areas of scar tissue in the superficial lamina propria (Burns, 2012). This kind of information could have a significant impact on the team (otolaryngologist and SLP) management of voice disorders. For example, better information about vocal fold pathology such as lesion depth and/or amount/location of scar tissue could have profound influences on treatment planning and rehabilitative prognoses, including critical decisions about whether or not a patient should undergo surgical management prior to voice therapy. Efforts are underway to bring OCT technology from the operating room into the office for real-time, dynamic, three-dimensional vocal fold imaging (Coughlan et al., 2016; Iftimia et al., 2016).

Innovations in Ambulatory Voice Monitoring

It is believed that many voice disorders are caused by (or associated with) aberrant or abusive vocal behavior in daily life (Hillman, Holmberg, Perkell, Walsh, & Vaughan, 1989); therefore, ambulatory voice monitoring has been developed to objectively measure vocal function during a patient's activities of daily living. There have previously been three commercially available ambulatory voice monitors developed for clinical use: the Ambulatory Phonation Monitor (APM; 2009), the VoxLog (firmware, 2.2.3, Sonvox AB, Umeå, Sweden), and the VocaLog (Griffin Laboratories, Temecula, CA). These monitors implement real-time processing that produces estimates of vocal sound pressure level (SPL), fundamental frequency (f_o), and vocal dose measures (Švec, Popolo, & Titze, 2003). For specific technical information on all three voice monitors, see Van Stan, Gustafsson, Schalling, and Hillman (2014). Though the ability to track general voice use has been a critical research advancement for the voice field, the clinical impact of these devices has been small. This is due to the prohibitive costs of the devices (approximately $1,000 to more than $5,000 per device), difficulty interpreting the large amount of data collected (which could be over 1 million data points per day), and a lack of clear evidence-based support for clinical use of the measurements provided. For example, group-based differences between patients with voice disorders and healthy controls have been difficult to consistently demonstrate using traditional ambulatory estimates of SPL, f_o, and vocal dose measures (Boudreaux, 2011; Ghassemi et al., 2016; Ghassemi et al., 2014; Masuda, Ikeda, Manako, & Komiyama, 1993; Mehta et al., 2015; Nacci et al., 2013; Van Stan, Mehta, Zeitels, et al., 2015). However, on an individual basis, ambulatory voice biofeedback using SPL or f_o has been successfully used to supplement therapy for patients with Parkinson's disease (Schalling, Gustafsson, Ternstrom, Bulukin Wilen, & Sodersten, 2013), muscle tension dysphonia (Stadelman-Cohen, Van Stan & Hillman, 2014), and vocal fold nodules (Van Stan et al., in press).

Even though clinical adoption of ambulatory voice monitoring has thus far been limited, recent technological advances and ongoing research in this area continue to support its future potential. The cost issue has potentially been addressed by using an application on a smartphone as the data collection platform (Mehta, Zañartu, Feng, Cheyne, & Hillman, 2012). Many investigations using ambulatory voice monitors have used a neck-placed accelerometer (ACC) as the preferred phonation sensor because of its inherent advantages over a microphone for this application (e.g., less obtrusive, maintains confidentiality, less sensitive to acoustic environmental noise). Recent efforts have focused on extracting additional and/or new measures from the ACC signal that are potentially more sensitive to pathological conditions (see Figure 1)—e.g., cepstral peak prominence (CPP), spectral tilt, and periodicity (Mehta, Van Stan, & Hillman, 2016)—and/or better reflect underlying physiological/pathophysiological phonatory mechanisms—for example, aerodynamic measures of subglottal pressure, maximum flow declination rate, and AC flow (Fryd, Van Stan, Hillman & Mehta, 2016; Zañartu, Galindo, et al., 2014; Zañartu, Ho, Mehta, Hillman & Wodicka, 2013) and relative fundamental frequency (Stepp et al., 2011).

Figure 1. Illustration of a Daily Voice Use Profile for an Adult Female Diagnosed With Bilateral Vocal Fold Nodules. Percent phonation time displayed for 5-minute sliding windows with three self-reported ratings of vocal status throughout the day. Other voice measures are shown as 5-minute moving averages (black) or 5th/95th percentiles (gray). Daylong histograms are shown to the right of each time series. D-SHP: difficulty producing soft, high-pitched phonation. dB/oct: decibels per octave



Besides providing access to more clinically salient measures for voice assessment, newer voice monitoring systems have the potential to play a role in the treatment of voice disorders by providing ambulatory biofeedback to reinforce voice therapy tasks or concepts outside of the clinic (Schalling et al., 2013; Van Stan, Mehta, & Hillman, 2015; Van Stan et al., in press), which addresses the frequently noted difficulty of carrying over newly learned vocal behaviors into daily life (Ziegler, Dastolfo, Hersan, Rosen, & Gartner-Schmidt, 2014). Other capabilities could include remote monitoring and transfer of patient voicing data to the clinic (to potentially reduce the number of clinic visits), real-time viewing of current voice use statistics by patients on their smartphone/smartwatch, the incorporation of persuasive technology to help patients accomplish their behavioral change goals (Oinas-Kukkonen & Harjumaa, 2008), and the periodic collection of patient self-assessments such as vocal fatigue or vocal effort throughout the course of a day (Mehta et al., 2015).

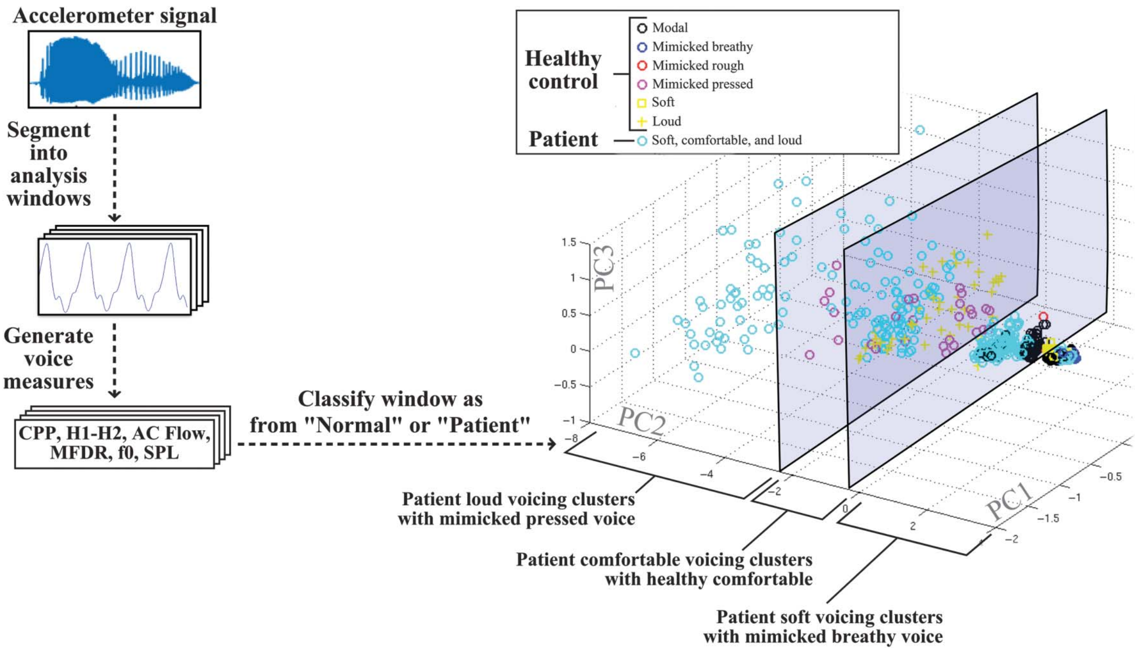
Using Machine Learning on "Big Data" for Voice Assessment

As technology advances throughout the field of health care, more data (and measurements derived from that data) will be collected per patient to improve assessment and treatment procedures. These resulting databases are often too large and complex to enable traditional statistical analysis approaches to be used and could inhibit the efficiency of clinical assessment and decision making. For example, ambulatory voice monitoring can result in more than 1 gigabyte of data per day, which is impossible for a clinician to visually review and interpret in a reasonable amount of time. Also, adoption of electronic health records by most medical practitioners has resulted in the compilation of large amounts (terabytes) of health-care data that can contain unstructured quantitative and qualitative records (Murdoch & Detsky, 2013; Raghupathi & Raghupathi, 2014). Computational algorithms from the field of machine learning (Domingos, 2012) have the potential to find patterns contained in such complex data sets and subsequently provide suggestions or support for clinical decision making. Machine learning algorithms could theoretically take in large amounts of information and "learn" complicated multidimensional relationships that aim to answer clinical questions related to presence/absence of a condition, differential diagnosis (classification) among multiple conditions, and disease severity.

Recent work using machine learning approaches combined with ambulatory voice monitoring has shown potential to improve voice assessment in patients with vocal hyperfunction. More specifically, machine learning has been used on ambulatory voice monitoring data to search for multidimensional biomarkers capable of discriminating between people with vocal hyperfunction and normal voices. For example, Ghassemi et al. (2014) differentiated between patients with vocal fold nodules or polyps and matched normal controls using "supervised" classification techniques to identify salient f_o and SPL-based "features" extracted from day-long recordings of the ACC signal. Also, Ghassemi et al. (2016) differentiated between patients with muscle tension dysphonia and matched normal controls when clustering the shape of individual glottal pulses from the neck-surface ACC waveform (an "unsupervised" algorithm) and also demonstrated potential for measuring the strength of treatment effects within a patient.

In the near future, the clinician could provide various types of patient information for a machine learning software program (e.g., past medical history, ambulatory monitoring data, and in-clinic voice measurements) and obtain an output describing how deviant the patient is from a multidimensional normative database. Therefore, the algorithm could help the voice therapist with clinical decision making through identifying vocally difficult daily situations, salient biofeedback measures, improvement (or lack thereof) from week to week, and discharge readiness, among others. Figure 2 illustrates an example of how machine learning can help improve voice disorder and voice quality classification.

Figure 2. Example of a Machine Learning Framework for the Classification of Voice Modes From a Patient and a Healthy Control. Data points in the three-dimensional space are from sustained vowels.
PC: principal component



Using Voice Analysis to Detect and Monitor Other Health Conditions

The larynx is prominently integrated into many bodily functions that are vital to basic survival (e.g., swallowing, breathing, and coughing), has direct connections to subcortical neurological systems associated with emotion (e.g., limbic system), and is crucial for countless learned behaviors (e.g., verbal communication, acting, singing, and culturally-important behavior; Simonyan & Horwitz, 2011). Combined with the fact that habituated coordination of laryngeal and articulatory muscles in speech and voicing occurs at rates ranging from 110 to 200 words per minute (Laver, 1994), it is reasonable to hypothesize that vocal function could begin to deteriorate at the early stages of various diseases or dysfunctions. Therefore, numerous research groups with diverse focuses are attempting to improve their assessment and treatment programs through investigating vocal biomarkers; which are then used to attempt early-stage disease diagnosis and monitor disease progression/status (e.g., response to medication). More specifically, voice measures such as amplitude, f_o, perturbation, and voice quality have been used to correlate with cognitive load (Quatieri et al., 2015), depression severity (Williamson, Quatieri, Helfer, Ciccarelli, & Mehta, 2014), mild traumatic brain injury (Helfer et al., 2014), Parkinson's disease (Williamson, Quatieri, Helfer, Ciccarelli, & Mehta, 2015), amyotrophic lateral sclerosis (Horwitz-Martin et al., 2016), and congestive heart failure (Murton et al., 2016).

Investigations into neurocognitive status are potentially complementary and may help to improve clinical voice assessment. For example, patients with primary muscle tension dysphonia have been associated with introverted personality profiles, higher degrees of neuroticism, and aberrant physiological stress mechanisms (Roy & Bless, 2000). Therefore, if cognitive or affective disturbances could be detected from the voice signal, it would certainly improve the voice therapist's diagnostic and therapeutic capabilities with this patient population. Also, if Parkinson's disease medication effects could be tracked via vocal biomarkers, voice therapists could account for the effectiveness of their behavioral therapy as the therapy interacts with medicated and non-medicated time periods.

Conclusions

Recent technological advancements are driving the development of new tools for clinical voice assessment. These tools include novel laryngeal imaging techniques, smartphone-based monitoring of patients' vocal behaviors in daily life, machine learning algorithms to aid in clinical decision-making, and novel voice-based health/disease biomarkers. Together, these advancements have great potential to improve the prevention, diagnosis, and treatment of voice disorders. Ambulatory monitoring is already becoming a prevalent part of many people's daily life, as demonstrated by the increasing number of commercial products from technology companies such as Fitbit, Apple, Samsung, Google, etc. And machine learning has been a routine aspect of daily life in recent years (internet search engines, movie recommendations, news feeds, speech recognition software, etc.). Therefore, we envision that the innovations described in this paper will have a high impact on the clinical management of patients with voice disorders in the years to come.

Acknowledgments

We would like to thank Marzyeh Ghassemi for providing images for Figure 2. This work was supported by the Voice Health Institute and the National Institutes of Health (NIH) National Institute on Deafness and Other Communication Disorders under Grant R33 DC011588 and F31 DC014412. Contents are solely the responsibility of the authors and do not necessarily represent the official views of the NIH.

References

Ambulatory Phonation Monitor [Apparatus and software]. (2009). Lincoln Park, NJ: KayPENTAX.

Awan, S. N., Roy, N., Jetté, M. E., Meltzner, G. S., & Hillman, R. E. (2010). Quantifying dysphonia severity using a spectral/cepstral-based acoustic index: Comparisons with auditory-perceptual judgements from the CAPE-V. *Clinical Linguistics & Phonetics, 24,* 742–758.

Bonilha, H. S., Deliyski, D. D., Whiteside, J. P., & Gerlach, T. T. (2012). Vocal fold phase asymmetries in patients with voice disorders: A study across visualization techniques. *American Journal of Speech-Language Pathology, 21,* 3–15.

Boudreaux, D. M. (2011). Using the Ambulatory Phonation Monitor to measure the vocal parameters of older people with and without Parkinson's disease (Unpublished master's thesis). Department of Communication Sciences and Disorders by Danielle Marie Boudreaux BA, Louisiana State University.

Burns, J. A. (2012). Optical coherence tomography: Imaging the larynx. *Current Opinion in Otolaryngology & Head and Neck Surgery, 20,* 477–481.

Burns, J. A., Zeitels, S. M., Anderson, R. R., Kobler, J. B., Pierce, M. C., & de Boer, J. F. (2005). Imaging the mucosa of the human vocal fold with optical coherence tomography. *Annals of Otology, Rhinology, and Laryngology, 114,* 671–676.

Coughlan, C. A., Chou, L. D., Jing, J. C., Chen, J. J., Rangarajan, S., Chang, T. H., . . . Wong, B. J. (2016). In vivo cross-sectional imaging of the phonating larynx using long-range Doppler optical coherence tomography. *Scientific Reports, 6,* 22792.

Deliyski, D. D., & Hillman, R. E. (2010). State of the art laryngeal imaging: Research and clinical implications. *Current Opinion in Otolaryngology & Head and Neck Surgery, 18*(3), 147–152.

Domingos, P. (2012). A few useful things to know about machine learning. *Communications of the ACM, 55*(10), 78–87.

Fryd, A. F., Van Stan, J. H., Hillman, R. E., & Mehta, D. D. (2016). Estimating subglottal pressure from neck-surface acceleration during normal voice production. *Journal of Speech Language and Hearing Research, 59,* 1335–1345.

Ghassemi, M., Syed, Z., Mehta, D. D., Van Stan, J. H., Hillman, R. E., & Guttag, J. V. (2016). Uncovering voice misuse using symbolic mismatch. *JMLR Workshop and Conference Proceedings, 56,* 1–14.

Ghassemi, M., Van Stan, J. H., Mehta, D. D., Zañartu, M., Cheyne, H. A., II, Hillman, R. E., & Guttag, J. V. (2014). Learning to detect vocal hyperfunction from ambulatory neck-surface acceleration features: Initial results for vocal fold nodules. *IEEE Transactions on Biomedical Engineering, 61,* 1668–1675.

Helfer, B. S., Quatieri, T. F., Williamson, J. R., Keyes, L., Evans, B., Greene, W. N., . . . Talavage, T. M. (2014). Articulatory dynamics and coordination in classifying cognitive change with preclinical mTBI. In H. Li & P. Ching (Eds.), *INTERSPEECH 2014: Celebrating the diversity of spoken languages* (pp. 485–489). Red Hook, NY: Curran Associates.

Hillman, R. E., Holmberg, E. B., Perkell, J. S., Walsh, M., & Vaughan, C. (1989). Objective assessment of vocal hyperfunction: An experimental framework and initial results. *Journal of Speech and Hearing Research, 32,* 373–392.

Hogikyan, N. D., & Sethuraman, G. (1999). Validation of an instrument to measure voice-related quality of life (V-RQOL). *Journal of Voice, 13,* 557–569.

Horwitz-Martin, R. L., Quatieri, T. F., Lammert, A. C., Williamson, J. R., Yunusova, Y., Godoy, E., . . . Green, J. R. (2016). Relation of automatically extracted formant trajectories with intelligibility loss and speaking rate decline in amyotrophic lateral sclerosis. In N. Morgan (Ed.), *INTERSPEECH 2015,* (pp. 1205–1219). Retrieved from http://www.isca-speech.org/archive/Interspeech_2016/

Iftimia, N., Maguluri, G., Chang, E., Park, J., Kobler, J., & Mehta, D. (2016). Dynamic vocal fold imaging with combined optical coherence tomography/high-speed video endoscopy. *Proceedings of the 10th International Conference on Voice Physiology and Biomechanics,* 1–2. Retrieved from http://scholar.harvard.edu/files/dmehta/files/iftimiaicvpb2016_dynamic_vocal_fold_imaging_with_combined_optical_coherence_tomography_high-speed_video_endoscopy.pdf

Kempster, G. B., Gerratt, B. R., Verdolini Abbott, K., Barkmeier-Kraemer, J., & Hillman, R. E. (2009). Consensus auditory-perceptual evaluation of voice: Development of a standardized clinical protocol. *American Journal of Speech-Language Pathology, 18*(2), 124–132.

Kreiman, J., Gerratt, B. R., Kempster, G. B., Erman, A., & Berke, G. S. (1993). Perceptual evaluation of voice quality: Review, tutorial, and a framework for future research. *Journal of Speech and Hearing Research, 36*(1), 21–40.

Laver, J. (1994). *Principles of phonetics. Cambridge textbooks in linguistics.* Cambridge, England: Cambridge University Press.

Masuda, T., Ikeda, Y., Manako, H., & Komiyama, S. (1993). Analysis of vocal abuse: Fluctuations in phonation time and intensity in 4 groups of speakers. *Acta Oto-Laryngologica, 113,* 547–552.

Mehta, D. D., & Hillman, R. E. (2012). Current role of stroboscopy in laryngeal imaging. *Current Opinion in Otolaryngology & Head and Neck Surgery, 20,* 429–436.

Mehta, D. D., Van Stan, J. H., & Hillman, R. E. (2016). Relationships between vocal function measures derived from an acoustic microphone and a subglottal neck-surface accelerometer. *IEEE/ACM Transactions on Audio, Speech, and Language Processing, 24,* 659–668.

Mehta, D. D., Van Stan, J. H., Zañartu, M., Ghassemi, M., Guttag, J. V., Espinoza, V. M., . . . Hillman, R. E. (2015). Using ambulatory voice monitoring to investigate common voice disorders: Research update. *Frontiers in Bioengineering and Biotechnology, 3*(155), 1–14.

Mehta, D. D., Zañartu, M., Feng, S. W., Cheyne, H. A., II, & Hillman, R. E. (2012). Mobile voice health monitoring using a wearable accelerometer sensor and a smartphone platform. *IEEE Transactions on Biomedical Engineering, 59,* 3090–3096.

Mehta, D. D., Zeitels, S. M., Burns, J. A., Friedman, A. D., Deliyski, D. D., & Hillman, R. E. (2012). High-speed videoendoscopic analysis of relationships between cepstral-based acoustic measures and voice production mechanisms in patients undergoing phonomicrosurgery. *Annals of Otology, Rhinology, and Laryngology, 121,* 341–347.

Murdoch, T. B., & Detsky, A. S. (2013). The inevitable application of big data to health care. *Journal of the American Medical Association, 309,* 1351–1352.

Murton, O., Mehta, D., Daher, M., Verkouw, K., Tabtabai, S., Steiner, J., . . . Semigran, M. (2016, September). *Impact of congestive heart failure on voice and speech production: A pilot study.* Presented at the the Annual Scientific Meeting of the Heart Failure Society of America, Orlando, FL.

Nacci, A., Fattori, B., Mancini, V., Panicucci, E., Ursino, F., Cartaino, F., & Berrettini, S. (2013). The use and role of the Ambulatory Phonation Monitor (APM) in voice assessment. *ACTA Otorhinolaryngologica Italica, 33,* 49.

Oinas-Kukkonen, H., & Harjumaa, M. (2008). Towards deeper understanding of persuasion in software and information systems. In *Advances in Computer-Human Interaction, 2008 First International Conference on IEEE* (pp. 200–205). Washington, DC: IEEE Computer Society.

Quatieri, T. F., Williamson, J. R., Smalt, C. J., Patel, T., Perricone, J., Mehta, D. D., . . . Moran, J. (2015). Vocal biomarkers to discriminate cognitive load in a working memory task. *INTERSPEECH 2015* (pp. 2684–2688). Red Hook, NY: Curran Associates.

Raghupathi, W., & Raghupathi, V. (2014). Big data analytics in healthcare: Promise and potential. *Health Information Science and Systems, 2*(3), 1–10.

Roy, N., Barkmeier-Kraemer, J., Eadie, T., Sivasankar, M. P., Mehta, D., Paul, D., & Hillman, R. (2013). Evidence-based clinical voice assessment: A systematic review. *American Journal of Speech-Language Pathology, 22,* 212–226.

Roy, N., & Bless, D. M. (2000). Personality traits and psychological factors in voice pathology: A foundation for future research. *Journal of Speech, Language, and Hearing Research, 43,* 737–748.

Schalling, E., Gustafsson, J., Ternstrom, S., Bulukin Wilen, F., & Sodersten, M. (2013). Effects of tactile biofeedback by a portable voice accumulator on voice sound level in speakers with Parkinson's disease. *Journal of Voice, 27,* 729–737.

Simonyan, K., & Horwitz, B. (2011). Laryngeal motor cortex and control of speech in humans. *The Neuroscientist, 17,* 197–208.

Stadelman-Cohen, T., Van Stan, J. H., & Hillman, R. E. (2014). Use of ambulatory biofeedback to supplement traditional voice therapy for treating primary muscle tension dysphonia in an adult female. In J. Stemple & E. Hapner (Eds.), *Voice therapy: Clinical case studies* (4th ed). San Diego, CA: Plural Publishing.

Stepp, C. E., Merchant, G. R., Heaton, J. T., & Hillman, R. E. (2011). Effects of voice therapy on relative fundamental frequency during voicing offset and onset in patients with vocal hyperfunction. *Journal of Speech, Language, and Hearing Research, 54,* 1260–1266.

Švec, J. G., Popolo, P. S., & Titze, I. R. (2003). Measurement of vocal doses in speech: Experimental procedure and signal processing. *Logopedics, Phoniatrics, Vocology, 28,* 181–192.

Švec, J. G., & Schutte, H. K. (1996). Videokymography: High-speed line scanning of vocal fold vibration. *Journal of Voice, 10,* 201–205.

Titze, I. R., & Hunter, E. J. (2015). Comparison of vocal vibration-dose measures for potential-damage risk criteria. *Journal of Speech, Language, and Hearing Research, 58,* 1425–1439.

Van Stan, J. H., Gustafsson, J., Schalling, E., & Hillman, R. E. (2014). Direct comparison of three commercially available devices for voice ambulatory monitoring and biofeedback. *Perspectives on Voice and Voice Disorders, 24,* 80–86.

Van Stan, J. H., Mehta, D. D., & Hillman, R. E. (2015). The effect of voice ambulatory biofeedback on the daily performance and retention of a modified vocal motor behavior in participants with normal voices. *Journal of Speech, Language, and Hearing Research, 58,* 713–721.

Van Stan, J. H., Mehta, D. D., Petit, R., Sternad, D., Muise, J., Burns, J. A., & Hillman, R. E. (in press). Integration of motor learning principles into real-time ambulatory voice biofeedback and example implementation via a clinical case study with vocal fold nodules. *American Journal of Speech-Language Pathology.*

Van Stan, J. H., Mehta, D. D., Zeitels, S. M., Burns, J. A., Barbu, A. M., & Hillman, R. E. (2015). Average ambulatory measures of sound pressure level, fundamental frequency, and vocal dose do not differ between adult females with phonotraumatic lesions and matched control subjects. *Annals of Otology, Rhinology, and Laryngology, 124,* 864–874.

Williamson, J. R., Quatieri, T. F., Helfer, B. S., Ciccarelli, G., & Mehta, D. D. (2014). Vocal and facial biomarkers of depression based on motor incoordination and timing. In *Proceedings of the 4th International Workshop on Audio/Visual Emotion Challenge* (pp. 65–72). New York, NY: Association for Computing Machinery.

Williamson, J. R., Quatieri, T. F., Helfer, B. S., Ciccarelli, G., & Mehta, D. D. (2015). Segment-dependent dynamics in predicting Parkinson's disease. *INTERSPEECH 2015* (pp. 518–522). Red Hook, NY: Curran Associates.

Zañartu, M., Galindo, G. E., Erath, B. D., Peterson, S. D., Wodicka, G. R., & Hillman, R. E. (2014). Modeling the effects of a posterior glottal opening on vocal fold dynamics with implications for vocal hyperfunction. *The Journal of the Acoustical Society of America, 136,* 3262–3271.

Zañartu, M., Ho, J. C., Mehta, D. D., Hillman, R. E., & Wodicka, G. R. (2013). Subglottal impedance-based inverse filtering of voiced sounds using neck surface acceleration. *IEEE Transactions on Audio, Speech, and Language Processing, 21,* 1929–1939.

Ziegler, A., Dastolfo, C., Hersan, R., Rosen, C. A., & Gartner-Schmidt, J. (2014). Perceptions of voice therapy from patients diagnosed with primary muscle tension dysphonia and benign mid-membranous vocal fold lesions. *Journal of Voice, 28,* 742–752.

History:
Received September 15, 2016
Revised September 26, 2016
Accepted September 28, 2016
https://doi.org/10.1044/persp2.SIG3.4