The Next 10 Years in Voice Evaluation and Treatment

Julie M. Barkmeier-Kraemer, Ph.D.¹ and Rita R. Patel, Ph.D.²

ABSTRACT

Voice disorders are thought to affect approximately one third of all individuals within the United States during their lifetime. Individuals who require the use of their voice as part of their occupations are at highest risk for developing voice problems. Unfortunately, efficient diagnosis and effective management of voice disorders can be challenged by difficulty accessing professionals with the necessary expertise to diagnose and treat voice problems efficiently. Within the next decade, technological advancements show promise for improving the efficiency and effectiveness of intervention for voice disorders. Exciting developments in laryngeal imaging, modeling of patient-specific vocal patterns, and implementation of smart mobile technology and telehealth will greatly improve the accuracy of diagnosing voice problems and enhance implementation and carryover of effective voice treatment methods to daily communication demands.

KEYWORDS: Voice, voice disorders, evaluation, treatment, technology

Learning Outcomes: As a result of this activity, the reader will be able to (1) identify and discuss current directions in laryngeal imaging technology; (2) identify and discuss current directions in developing patient-specific methods for diagnosis and treatment; and (3) identify and discuss future methods for implementing mobile smart technology and telehealth technology to address current challenges in voice treatment.

Voice disorders affect nearly one-third of individuals in the United States age 20 to 66 years of age during their lifetime, with 39% of older people aged 45 to 64 years of age reporting voice problems.¹ Of those diagnosed with voice disorders, two thirds are female.² Twenty-eight million people in the United States are estimated to require their

¹Division of Otolaryngology, University of Utah, Salt Lake City, Utah; ²Department of Speech and Hearing Sciences, Indiana University, Bloomington, Indiana.

Forecasting the Future: Expert Opinions on the Next Decade for Adult Speech and Language Disorders; Guest Editor, Audrey L. Holland, Ph.D.


ISSN 0734-0478.
voice for job performance. Teachers represent one such voice-intensive occupation with 20% of teachers missing work due to voice problems. Approximately $2.5 billion dollars are spent annually within the United States on teacher sick days and voice treatment expenses. Epidemiology research estimates that individuals with voice disorders miss an average of 39.2 work days/year/person due to a voice problem. One of the larger barriers to all individuals with voice problems is delayed referral to an appropriate health care provider. Delayed referral costs are estimated to be $711 per person when the delay is greater than 3 months compared with $271 per person when appropriately referred within the first 3 months of onset of a voice disorder.

Over one million, or 6 to 9%, of all children in the United States also suffer from dysphonia. The proportion of school-aged children exhibiting a voice disorder ranges from 6 to 23%, depending upon the cohort. Studies have shown that voice disorders have an adverse impact on children academically as well as psychologically. Hence, early identification and treatment of voice disorders is critical.

By 2026, individuals with voice problems will likely enjoy a more accessible, efficient, and convenient method for detection and treatment of voice disorders. Evaluation and treatment of voice disorders will likely involve machine-based diagnostics using technological advances that map abnormal voicing patterns onto underlying pathophysiology (Fig. 1). Furthermore, voice therapy will be available in the comfort of the patient’s home via telehealth or be implemented using smart technology during the course of daily life activities. Voice therapy approaches will be recommended for each individual using artificial intelligence capable of identifying optimal custom-based treatment methods from recorded voice use patterns. Optimal voice therapy outcomes require accurate implementation of specified voicing techniques and exercises that can be applied within typical communication activities. To ensure accurate implementation of therapeutic methods and goals, smart devices and phone apps will utilize patient-specific principles of behavior modification to cue individuals to complete treatment exercises at designated times during the day or use targeted voice therapy techniques during conversation when voice sensors determine aberrant performance. Thus, implementation of new technology can be used outside the clinic setting to facilitate improved implementation of therapeutic voice methods within the context of daily life activities. Technological advances by 2026 will revolutionize the efficiency and accuracy with which various voice disorders are diagnosed and effective voice therapy or medical/surgical treatments are recommended.

**HIGH-SPEED VIDEO IMAGING**

Comprehensive evaluation of voice is multidimensional; however, direct imaging of vocal fold vibratory patterns is currently fundamental for appropriate evaluation of the cause of the dysphonia. High-speed video (HSV) imaging, though not the current gold standard in laryngeal imaging, has transformed voice evaluations in state-of-the-art voice centers. Compared with the current gold standard of laryngeal imaging, videostroboscopy, current HSV imaging systems have the advantage of evaluating severely dysphonic voice qualities and rapid transient fluctuations in voice. HSV imaging offers high temporal resolution of up to 8,000 frames per second compared with stroboscopy, which can only capture images at an average rate of 30 frames per second. Thus, stroboscopy frequently cannot accurately track the signal used to trigger stroboscopic lighting due to severely impaired vocal fold vibratory patterns. HSV imaging does not have this limitation because the high rate of image capture does not require synchronization with the voice signal. That is, valid HSV images without movement artifact can be obtained irrespective of the degree and duration of the voice disturbance during image recordings. Clinical indications for the use of HSV imaging support its value in evaluating severely affected voices, differentiating adductor spasmodic dysphonia from muscle tension dysphonia, comparing voice onset patterns during different types of singing, age-related voice changes, diplophonia, tremor, and various structural abnormalities affecting the vocal folds (e.g., sulcus vocalis, nodules and polyps). HSV
imaging also provides new insights into evaluating the effects of superficial dehydration on vibratory motion, evaluating the growth and development of the vibratory motion in children, monitoring changes in voice physiology following voice therapy, and quantifying outcomes of phonosurgical treatment of early glottic carcinoma. Furthermore, HSV
imaging combined with multiple arrays of lasers provides valuable insights into the basic mechanisms of phonation in adults and children by aiding in quantifying not only the two-dimensional vibratory motion, but also the three-dimensional vocal fold vibration kinematics in vivo, which is unavailable on standard laryngeal imaging.

The current limitation to clinical use of HSV imaging is the significant time required to complete visual-perceptual analysis of the long and detailed recordings. For example, 4 seconds of HSV recorded at 4,000 frames per second requires about 27 minutes to watch the entire recording at 10 frames per second using the playback rate. Moreover, visual-perceptual analysis of HSV is subjective and often results in low inter- and intra-rater reliability. However, recent work addressing these limitations using methods of machine learning, feature extraction, and automated analysis of HSV recordings has recently shown promise for developing simulation approaches based on laryngeal tissue secretion biomarkers that could eventually help clinicians determine optimal timing and treatment approaches for patients with laryngeal tissue injuries (e.g., post-laryngeal surgery, or trauma).

Another method to develop future applications for improved diagnostic accuracy entails the acquisition of physiologic parameters associated with normal and abnormal voice use for simulation followed by empirical testing. Simulation-based modeling is useful for testing the accuracy of predictions about voice output patterns associated with parameters of voice physiology. Such methods are being utilized to develop future diagnostic tools that will be able to detect the presence and characteristics of such voice disorders as vocal tremor, vocal fold paralysis, and functional voice abnormalities.

**DIAGNOSIS USING PATIENT-SPECIFIC VOICE MODELING**

In addition to advances in laryngeal imaging, current efforts are underway to implement the use of bioinformatics and artificial intelligence for diagnosing and treating voice disorders. Artificial intelligence can detect patterns that the humans cannot perceive or predict and has the capacity to synthesize large amounts of data within seconds. Current voice modeling efforts study tissue-based biomarkers or voice physiology parameters to determine patient-specific patterns associated with each voice disorder and predict optimal custom-based treatment approaches for optimal outcomes.

One approach to simulating patient-specific information for treatment purposes is the use of biomarker agent-based models. Li and colleagues studied laryngeal secretion composition associated with onset of inflammatory responses to phonotrauma and mapped recovery patterns in a small group of individuals. Using the agent-based models approach, they determined which biomarkers from laryngeal secretions best predicted the recovery trajectory of speakers across a baseline, vocal loading (i.e., phonotrauma), and recovery period using three different paradigms of voice use. An extension of the initial model showed promise for developing simulation approaches based on laryngeal tissue secretion biomarkers that could eventually help clinicians determine optimal timing and treatment approaches for patients with laryngeal tissue injuries (e.g., post-laryngeal surgery, or trauma).

Another promising approach for evaluation and treatment purposes is the development of mobile technology for recording daily voicing patterns or cueing and facilitating the use of voice therapy techniques during daily life activities. The first form of these tools utilized attachment of an accelerometer to the skin overlying the larynx to record the frequency and intensity of the voice across an 8-hour period. Recordings are stored on a small computer device worn in a waist pack or in a pocket. Early work using these devices was used to compare patterns of voice use in teachers between work and nonwork environments. Voicing patterns of teachers compared with nonteachers elucidated the concept of “vocal dose” levels associated with onset of voice problems. More recent work extended the information gained from the voice sensors to predict subglottal pressures, aerodynamic...
patterns, and muscle activation patterns associated with normal and abnormal voicing. In the near future, these devices could be purchased by consumers as disposable sensors that communicate with smartphones to record voice use patterns and eventually provide biofeedback to the user regarding their accuracy in implementing targeted therapeutic voice use patterns in daily life.

Recent use of smartphones to facilitate accuracy and adherence to voice therapy methods also shows promise for future implementation of media that can be used by clinicians to ensure consistent and accurate patient practice of voice therapy methods. The use of voice therapy smartphone applications should lead to custom-based treatment recommendations such that patients are cued to practice at selected times during the day by their smartphone or mobile device based on patient-specific principles of behavior change to optimize acquisition and carryover of the target therapeutic skills. Partnered with physiologic feedback from the dosimeter or ambulatory phonation monitor, future treatment plans may enable direct cueing to patients to improve voice production techniques within their typical daily environment when aberrant physiology is detected by the sensor placed on their throat.

**TELEHEALTH**

Another advancement changing the face of voice therapy is the increased utilization of telemedicine/telehealth. Although literature on the efficacy and effectiveness of telehealth for voice treatment is not extensive, telehealth appears to be utilized to deliver voice therapy with comparable levels of treatment outcomes. The use of telehealth to deliver voice therapy will be particularly important for individuals with voice disorders living in rural regions without voice specialists or those with transportation barriers. Thus, implementation of telehealth and improvements in telehealth technology will enable improved access to voice treatment in the comfort of an individual’s home or local site with necessary reception specifications (e.g., local library, medical center).

Current common barriers to use of telehealth for voice treatment include the lack of insurance coverage by Medicare and other third-party payers as well as the need for the clinician to be licensed in the state where the patient resides. Although the latter is not a problem for in-state patients, regions where multiple state regions are served by one voice center make it necessary to maintain multiple state licenses and associated requirements. Hopefully, the ability of patients to access qualified providers for voice evaluation and treatment through telehealth will be addressed in upcoming efforts by the American Speech-Language-Hearing Association to initiate changes to current legislative and licensure barriers. We predict that implementation of telehealth for voice evaluation and treatment will improve accessibility to state-of-the-art clinical methods for improving diagnostic precision and treatment planning for those with voice disorders.

In summary, the area of voice and voice disorders faces an exciting time as technological advancements may significantly improve diagnostic precision and treatment planning through the development of improved imaging and computer diagnostic systems that will soon be available to communication science and disorder specialists. With early detection of various conditions and information regarding risk factors at their fingertips, clinicians will be more involved in preventive care. With the aid of technological advances in artificial intelligence, communication science and disorder clinicians will be able to use computers to quickly determine evidenced-based diagnoses and highlight the risk factors, thereby facilitating early detection and treatment of various voice disorders. Finally, mobile technology methods will eventually be available to improve the efficiency with which voice treatment methods can be trained so that patients can quickly carryover improved voicing techniques and exercises into their daily lives. The increased availability and use of technology by voice patients to facilitate improved voicing methods in their daily lives will help reduce overall medical costs associated with inaccurate diagnoses and extended use of ineffective treatment methods. We predict that technological advances within the upcoming decade will have an impact on evaluating and treating voice disorders. This should result in significantly more efficient and effective approaches to voice
disorder management that is more accessible to the elderly and parents of children with voice disorders and to those living in rural regions than currently.

REFERENCES


45. Khosla S, Oren L, Gutmark E. An example of the role of basic science research to inform the treatment of unilateral vocal fold paralysis. SIG 3 Perspectives on Voice and Voice Disorders. ASHA 2014;24:37–50.


