REVIEW

Current role of stroboscopy in laryngeal imaging

Daryush D. Mehta[a,b] and Robert E. Hillman[a,c,d]

Purpose of review
To summarize recent technological advancements and insight into the role of stroboscopy in laryngeal imaging.

Recent findings
Although stroboscopic technology has not undergone major technological improvements, recent clarifications have been made to the application of stroboscopic principles to video-based laryngeal imaging. Also recent advances in coupling stroboscopy with high-definition video cameras provide higher spatial resolution of vocal fold vibratory function during phonation. Studies indicate that the interrater reliability of visual stroboscopic assessment varies depending on the laryngeal feature being rated and that only a subset of features may be needed to be representative of an entire assessment. High-speed videoendoscopy (HSV) judgments have been shown to be more sensitive than stroboscopy for evaluating vocal fold phase asymmetry, pointing to the future potential of complementing stroboscopy with alternative imaging modalities in hybrid systems. Laryngeal videostroboscopy alone continues to play a central role in clinical voice assessment. Even though HSV may provide more detailed information about phonatory function, its eventual clinical adoption will depend on how remaining practical, technical, and methodological challenges will be met.

Summary
Laryngeal videostroboscopy continues to be the modality of choice for imaging vocal fold vibration, but technological advancements in HSV and associated research findings are driving increased interest in the clinical adoption of HSV to complement videostroboscopic assessment.

Keywords
clinical voice assessment, imaging, larynx, stroboscopy, vocal folds

INTRODUCTION

Stroboscopic imaging of vocal fold vibratory function during phonation continues to play a central role in diagnostic, therapeutic, and surgical decisions during the management and treatment of voice disorders. Although sampling rate limitations prevent stroboscopic imaging from capturing cycle-to-cycle details of vocal fold vibratory characteristics, clinicians are able to observe many salient features that cannot be perceived at standard video frame rates. Although newer laryngeal imaging technologies – such as high-speed videoendoscopy (HSV), MRI, and optical coherence tomography [1] – continue to enhance our ability to better define and quantify complex phonatory mechanisms, the cost effectiveness, ease of use, and synchronized audio and visual feedback provided by videostroboscopic assessment serve to maintain its predominant clinical role in laryngeal imaging. This review provides commentary on recent advances and insight into the application of stroboscopic imaging in clinical voice assessment and voice research.

TECHNOLOGICAL ADVANCEMENTS

Imaging of rapid vocal fold motion has a long and storied history. Oertel [2] published the earliest application of stroboscopic principles to observe vocal fold vibrations using a revolving disk with equally spaced holes to mechanically shutter a light source [3,4]. Speakers had to match their pitch to the frequency of the rotating disk to enable the production of a sequence of images that was perceived

[a]Department of Surgery, Center for Laryngeal Surgery and Voice Rehabilitation, Massachusetts General Hospital – Harvard Medical School, Boston, [b]School of Engineering and Applied Sciences, Harvard University, Cambridge, [c]Surgery & Health Sciences and Technology, Harvard Medical School and [d]Institute of Health Professions, Massachusetts General Hospital, Boston, Massachusetts, USA

Correspondence to Daryush D. Mehta, PhD, Center for Laryngeal Surgery and Voice Rehabilitation, Massachusetts General Hospital, One Bowdoin Square, 11th Floor, Boston, MA 02114, USA. Tel: +1 617 643 2466; e-mail: daryush.mehta@alum.mit.edu

Curr Opin Otolaryngol Head Neck Surg 2012, 20:429–436

DOI:10.1097/MOO.0b013e3283585f04

KEY POINTS

- Laryngeal videostroboscopy provides real-time audiovisual feedback and continues to be the imaging modality of choice by voice clinicians.
- Recent advances in coupling stroboscopy with high-definition video cameras provide higher spatial resolution of vocal fold vibratory function during phonation.
- Interrater reliability of visual stroboscopic assessment varies depending on the laryngeal feature being rated, and only a subset of features may be necessary to represent an entire assessment.
- Data suggest that although stroboscopy may be sensitive to certain vocal fold vibratory features such as phase asymmetry, the modality may lack the specificity exhibited by HSV that is required for adequate judgments to be made.
- The potential integration of laryngeal HSV and other complementary imaging modalities into hybrid systems could help improve the management of voice disorders in routine clinical practice.

as a slow-motion representation of the vocal fold vibratory cycle. Today, patients are free to phonate over a wide range of fundamental frequencies that are typically tracked using signals from neck-mounted contact microphones or electroglottograph electrodes.

Principles of stroboscopy

The scientific principles of stroboscopy are well known, but much of the classic voice literature mistakenly attributes the strobe effect to Talbot's law and the persistence of vision. Mehta *et al.* [5] recently debunked these misconceptions in a commentary that explained that two different visual perception phenomena actually play critical roles in laryngeal stroboscopy: the perception of a flicker-free, uniformly illuminated image (satisfied at strobe rates above 50 Hz) and the perception of apparent motion from sampled images when no real motion exists (satisfied at display rates above 17 Hz) [5]. These two requirements are satisfied in modern videostroboscopic systems [6–8], which integrate stroboscopic principles with video-based technologies.

Videostroboscopy: coupling stroboscopic principles with video camera technology

The video recording process in the United States typically follows the National Television System Committee standard that sets the video capture rate to approximately 30 interlaced frames per second, with each frame comprising two half frames, called fields, that are captured at approximately 60 Hz (actual frame and field rates are 30/1.001 Hz and 60/1.001 Hz, respectively) [9]. In 1992, Kay Elemetrics (now KayPENTAX) introduced laryngeal stroboscopy systems that precisely controlled the triggering of light sources so that only one strobe occurred per video field, thereby eliminating artifacts that were previously present due to multiple exposures within each video field [6]. A detailed discussion of the interaction among strobe rate, video camera rate, and phonatory fundamental frequency is provided in Hillman and Mehta [10].

No major technical advancements have been made in recent years regarding stroboscopic imaging. Videostroboscopic technologies typically enable two views of periodic vocal fold vibration. The systems can appear to freeze tissue motion at a selected phase in the periodic vibratory pattern, or they can create an apparent slow-motion view of the periodic vibratory cycles [10]. The specific implementation of sampling the motion of the vocal folds varies by manufacturer. For example, KayPENTAX systems trigger a Xenon light source to illuminate the larynx with flash durations of 5 microseconds [6], whereas the ATMOS system flashes a light-emitting diode light source [8]. An alternative method, employed by JEDMED, applies a constant light source but performs stroboscopic sampling by electronically shuttering the image sensor of the camera [7]. Regardless of method, the flash or shutter durations are sufficiently short to prevent motion blur artifacts in images that may arise due to rapid vocal fold tissue movements that can approach velocities of 1 m/s [11].

High-definition videostroboscopy

Recent advances in coupling stroboscopic systems with High-Definition video camera sensors provide unprecedented spatial resolution of the vocal fold structures involved in phonatory vibration (e.g., mucosa, superficial vasculature, etc.). The High-Definition Digital Stroboscopy System by KayPENTAX, for example, records interlaced video frames with a spatial resolution of 1920×1080 pixels. This wide-format resolution is in contrast to standard-definition video resolution of 720×480 pixels. A formal evaluation of high-definition versus standard-definition video for laryngeal imaging remains to be undertaken, but the significant improvements in image quality associated with high definition are expected to enhance clinical diagnostic capabilities.

As with all imaging modalities, though, the extra resolution afforded is only beneficial if the image target fills up a large portion of the video frame. Figure 1 displays a side-by-side comparison of still frames obtained from standard-definition and high-definition videostroboscopy recordings during sustained phonation. High-definition systems provide added spatial resolution as compared with standard-definition systems, which exhibit pixelation at high levels of magnification.

VISUAL JUDGMENTS OF LARYNGEAL STROBOSCOPY

Efforts to standardize the assessment of laryngeal stroboscopic recordings have produced rating systems that primarily require judgments of various vocal fold vibratory characteristics/parameters and some additional observations [12–14]. Figure 2 [13,15■■] displays one such rating system, the Stroboscopy Evaluation Rating Form, which assesses the following laryngeal properties during phonation:

(1) amplitude: extent of lateral vocal fold displacement
(2) mucosal wave: extent of vocal fold tissue deformation
(3) vibratory behavior: presence or absence of vibration in particular locations
(4) supraglottic activity: extent of laryngeal compression
(5) edge: ratings of smoothness and straightness
(6) vertical level: on-plane versus off-plane vocal fold contact
(7) phase closure: rating of open/closed phase duration
(8) phase symmetry: rating of left–right vibratory phase symmetry
(9) regularity: rating of periodicity

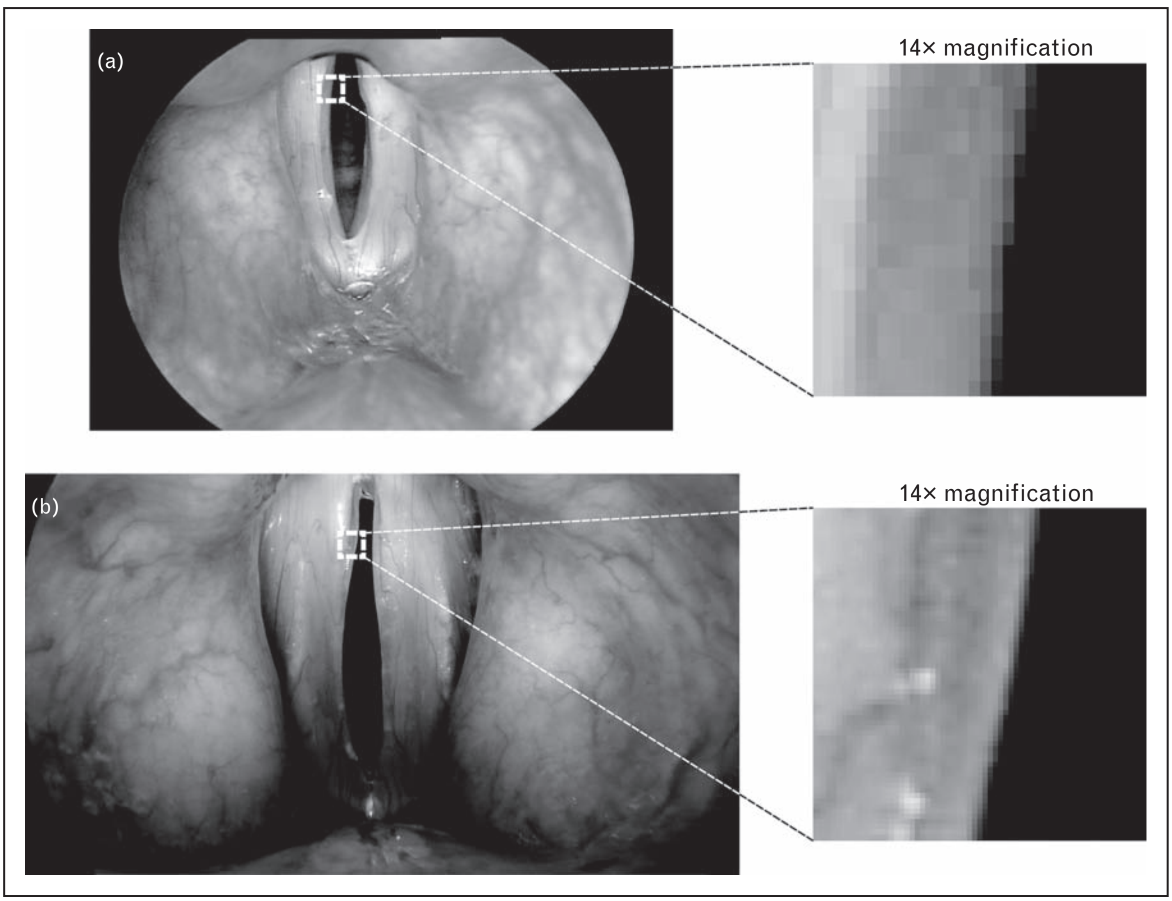

FIGURE 1. Comparison of the spatial resolution of still frames. Spatial resolution of still frames was captured using (a) standard-definition (720 × 480 pixels) videostroboscopy and (b) high-definition (1920 × 1080 pixels) videostroboscopy obtained with rigid endoscopy of normal adult men. A selected segment of the vocal fold edge in each exam is magnified (14x) to illustrate the increased pixelation that occurs in the standard-definition image. High-definition frame courtesy of KayPENTAX Corporation.

(10) glottal closure: categorical rating of the shape of the glottis at closure

Interrater reliability

The interrater reliability of judging the 10 parameters above during stroboscopic imaging has been investigated in a recent study [15▪▪]. Although most of the interval-scaled parameters yielded adequate interrater reliability, the judgments of phase closure, phase symmetry, and regularity exhibited the poorest reliability and call into question the overall validity of obtaining these parameters. The two categorically scaled parameters of vertical level and glottal closure were judged so unreliably that it was suggested that their assessment might hold little information [15▪▪]. Interestingly, parameters exhibiting the most reliable judgments – amplitude, vibratory behavior, and edge – were found by Kelley *et al.* [16▪] to form a minimal subset of parameters that accounted for most of the variance of all the laryngeal stroboscopic characteristics. Although it is unclear whether clinicians are ready to completely dispense with making judgments of vocal fold phonatory parameters that have questionable reliability, it is hoped that ongoing efforts to assess the validity and reliability of measures will continue to inform the refinement and application of such rating schemes.

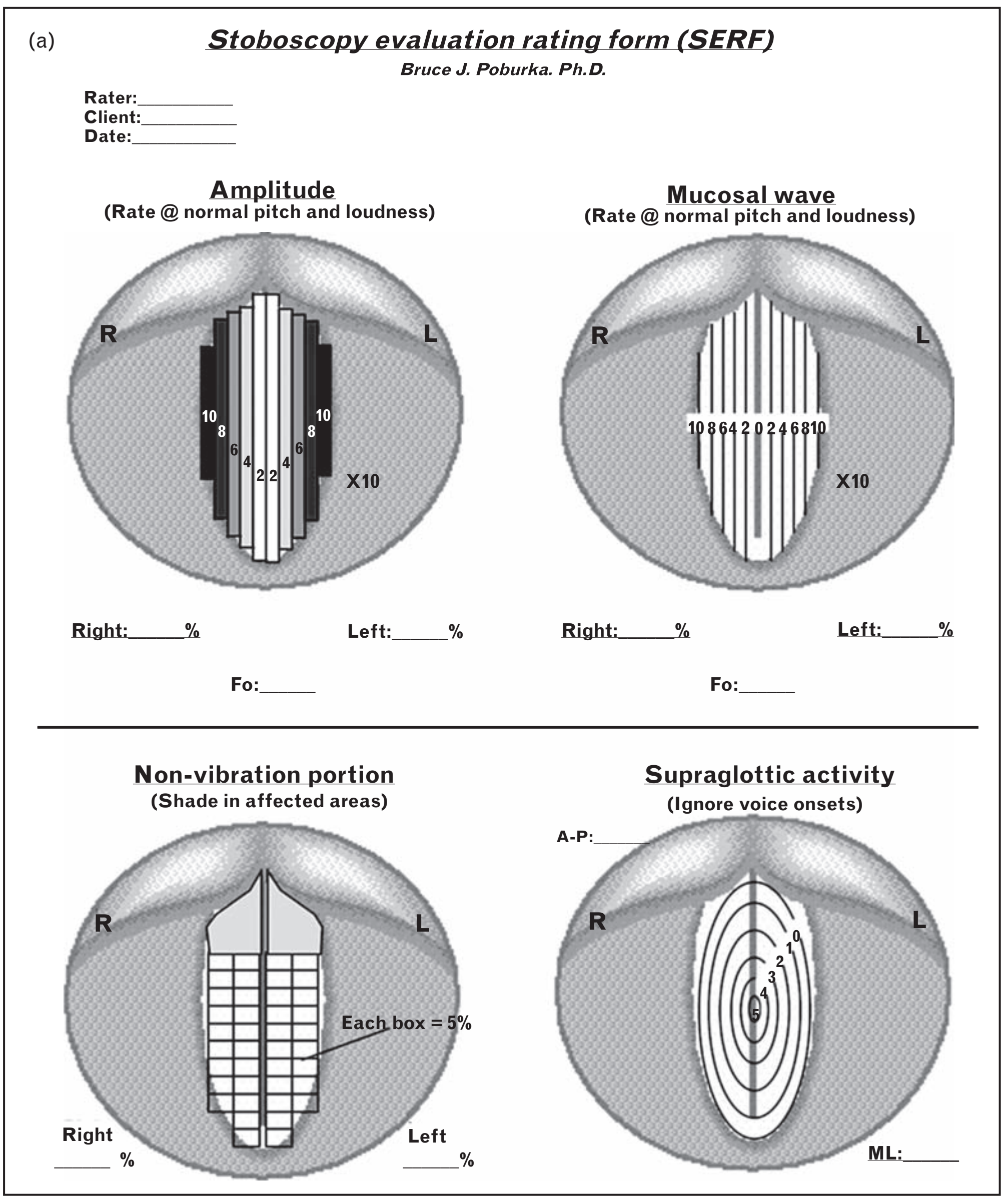

FIGURE 2. Stroboscopy Evaluation Rating Form developed by Poburka [13], whose interrater reliability was evaluated by Nawka and Konerding [15▪▪]. Reproduced with permission from [15▪▪].

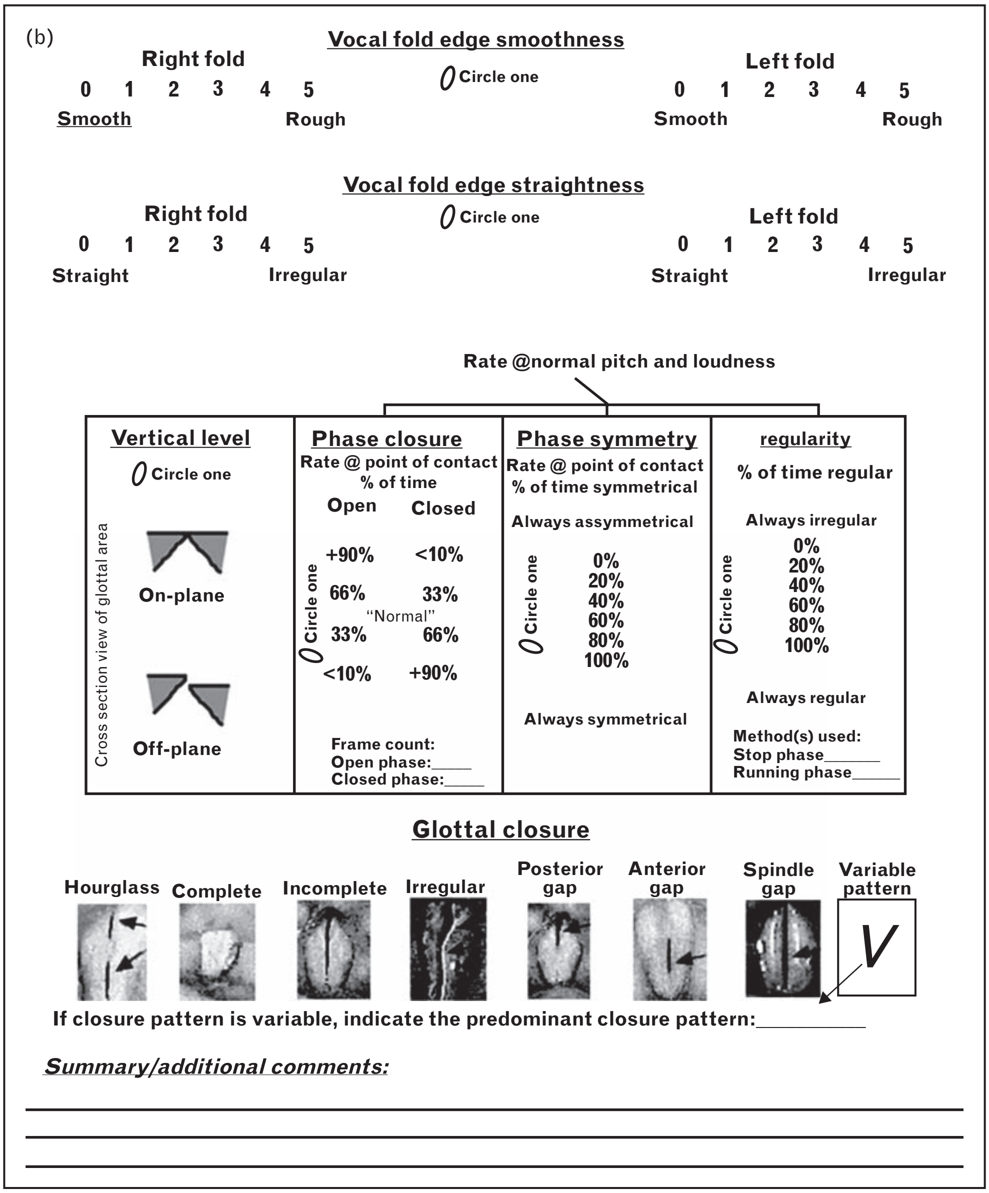

FIGURE 2. (*Continued*)

Comparison of videostroboscopy and high-speed videoendoscopy

As is well known, stroboscopic imaging has inherent limitations due to its sampling technique. The strobe effect can only be produced if the motion being observed is adequately periodic; thus, stroboscopy is typically incapable of revealing vocal fold vibratory patterns once dysphonia exceeds a moderate level [17]. Even when successful, stroboscopy can only provide a highly averaged visualization of periodic motion that is not sensitive enough to capture cycle-to-cycle variations in vocal fold vibration that have been linked to the degradation in acoustic voice quality measures [18]. HSV systems overcome these limitations by recording at frame rates much higher than, and not dependent on, a speaker's fundamental frequency.

Figure 3 [10] illustrates the differences between HSV and stroboscopic sampling. A dual endoscopy was performed on a vocally healthy speaker and simultaneously captured HSV data at 6250 frames per second along with stroboscopic flashes of light triggered once per video field. With the speaker's fundamental frequency at 236 Hz, the HSV recording yielded an average of 26.5 frames per glottal cycle. In contrast, only one videostroboscopic frame (comprised of interlacing two consecutive video fields) is captured every eight glottal cycles.

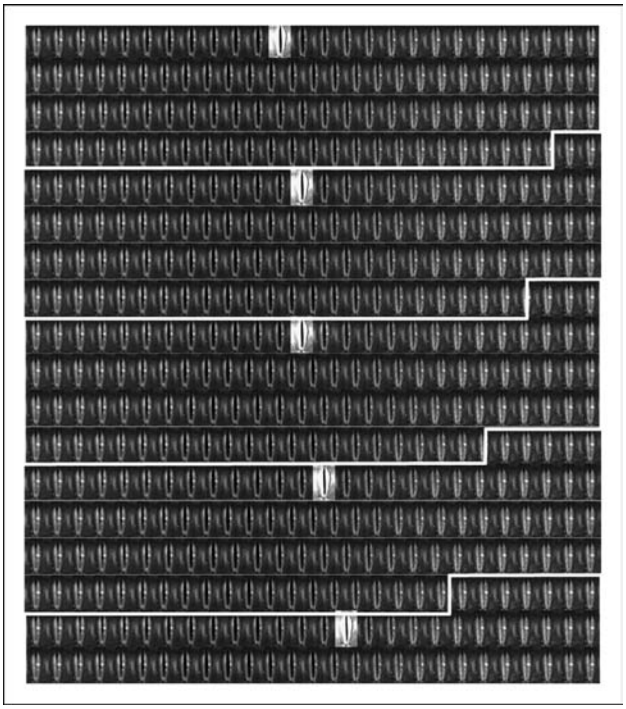

FIGURE 3. Dual rigid endoscopy of a vocally normal speaker sustaining a vowel. Each row of high-speed videoendoscopy frames (6250 frames per second) depicts one cycle of vocal fold vibration. White boundaries indicate the durations of National Television System Committee video fields, during which one strobe is flashed to capture advancing phases in the glottal cycle. With the phonatory fundamental frequency at 236 Hz, one videostroboscopic frame (comprising two fields) would be composed every eight cycles. The figure displays 477 frames, representing 76.32 ms of time. Adapted with permission from [10].

Which visualization is better?

Does the loss of information using stroboscopic imaging matter? One way to answer this question is to ask judges to rate laryngeal features from HSV and videostroboscopy recordings of sustained phonation and compare the ability of each modality to reliably reveal certain features. In a group of healthy speakers, it was found that the reliability of stroboscopic ratings was comparable with similar ratings made on HSV recordings, except for visual judgments of symmetry [19]. Other investigators found high intrarater and interrater reliability of phase asymmetry using stroboscopy in vocally healthy individuals [20] and speakers with voice disorders [21■■]; however, the validity of stroboscopy-based judgments of phase asymmetry was called into question due to lower correlations with an objective measure of phase asymmetry as compared with HSV-based modalities [20,21■■]. A case study also points to the need for HSV-based imaging to describe more detailed vocal fold tissue motion during pretherapy and posttherapy assessment [22■]. These results suggest that although stroboscopy may be sensitive to certain visual features, the modality may lack the specificity required for adequate judgments to be made.

In a study utilizing HSV of glottic cancer patients, variations in levels of acoustic jitter and shimmer were found to be unrelated to average measures of asymmetry; instead, a significant amount of the variation in acoustic jitter was accounted for by the standard deviation in the symmetry of phase and amplitude across the vibratory cycles [18]. This result has implications in terms of stroboscopic imaging because the apparently critical cycle-to-cycle variations in tissue vibratory behavior that were shown to be correlated with the degradation of the acoustic signal would not be reliably revealed using videostroboscopy. Further, as stroboscopic video only captures periodic vocal fold motion, it would be capable of only imaging the kind of highly repetitive asymmetries that do not appear to make a major contribution to disruptions in acoustic sound generation. Research efforts continue to determine optimal visualizations of voice production mechanisms [1,23].

CLINICAL ROLE OF STROBOSCOPY

In a survey of 273 members of the American Academy of Otolaryngology – Head and Neck Surgery (AAO-HNS), 84% of respondents reported that they perform videostroboscopy [24]. This result demonstrates the current routine use of stroboscopy by general otolaryngologists. In the pediatric population, stroboscopy continues to form an integral part of diagnostic voice assessment, even though obtaining high-quality rigid or flexible endoscopic recordings may be challenging in children [25]. Moreover, a pediatric vocal fold nodule rating scale has been developed based on videostroboscopic recordings of sustained vowel production [26].

Diagnostic value of stroboscopy

Recent publications advocate the use of laryngeal stroboscopic assessment to diagnose general hoarseness [27,28], as well as specific pathological conditions, such as organic lesions [29] and vocal fold scarring [30]. Stroboscopic imaging also permits the clinician to simultaneously listen to a patient's voice quality while observing the motion of the vocal folds. A clinical practice guideline published by an AAO-HNS-sponsored committee recently reiterated that stroboscopy is advisable to evaluate vocal function related to hoarseness [28]. In particular, the committee notes that if auditory-perceptual judgments of an individual's dysphonia seem out of proportion with the results of a (nonstroboscopic) laryngoscopic examination, then stroboscopic assessment affords the ability to gain additional information regarding vocal fold tissue pliability that could help in explaining the hoarseness symptoms [28].

Adoption of high-speed videoendoscopy

Even though HSV provides more detailed temporal information about vocal fold kinematics than stroboscopy, the eventual adoption of HSV into clinical practice will depend on the extent to which remaining practical, technical, and methodological challenges can be met. Such HSV-specific challenges include the relatively high cost of current systems, management and processing of large data files, limitations on memory size, potential thermal effects on tissue due to the intense light sources that are required, and a paucity of solid clinical research that demonstrates that HSV significantly improves the diagnosis and management of voice disorders (e.g., controlled clinical trials). Hybrid HSV-stroboscopy systems could take advantage and complement the outputs of each imaging modality – for example, by recording in stroboscopic mode by default to provide simultaneous audiovisual feedback and possessing the flexibility to switch to high-speed mode for specific, short-duration phonatory segments of interest to the clinician.

CONCLUSION

Laryngeal videostroboscopy continues to be the imaging modality of choice by voice clinicians due to its historical use and ability to efficiently capture many salient vocal fold vibratory characteristics. Because visual and objective assessments of certain laryngeal features can be unreliable using stroboscopic imaging, further research is warranted into the integration of laryngeal HSV and other alternate imaging modalities into routine clinical practice to improve the management of voice disorders.

Acknowledgements

The authors would like to thank Stephen Crump and Robert McClurkin of KayPENTAX for technical discussion and for providing examples of high-definition videostroboscopic imaging.

Conflicts of interest

The authors report no conflicts of interest. Support provided by the Institute of Laryngology and Voice Restoration and the National Institutes of Health (grants R01 DC007640 and R33 DC011588).

REFERENCES AND RECOMMENDED READING

Papers of particular interest, published within the annual period of review, have been highlighted as:
■ of special interest
■■ of outstanding interest
Additional references related to this topic can also be found in the Current World Literature section in this issue (p. 540).

1. Deliyski DD, Hillman RE. State of the art laryngeal imaging: research and clinical implications. Curr Opin Otolaryngol Head Neck Surg 2010; 18:147–152.
2. Oertel M. Das Laryngo-stroboskop und die laryngostroboskopische Untersuchung [The laryngostroboscope and laryngostroboscopic examination]. Arch Laryng Rhinol 1895; 3:1–16.
3. Zeitels SM. Premalignant epithelium and microinvasive cancer of the vocal fold: the evolution of phonomicrosurgical management. Laryngoscope 1995; 105:1–51.
4. Wendler J. Stroboscopy. J Voice 1992; 6:149–154.
5. Mehta DD, Deliyski DD, Hillman RE. Commentary on why laryngeal stroboscopy really works: clarifying misconceptions surrounding Talbot's law and the persistence of vision. J Speech Lang Hear Res 2010; 53:1263–1267.
6. KayPENTAX. Instruction manual: Stroboscopy systems and components 2008. Montvale, New Jersey, USA.
7. JEDMED. StroboCAM II specifications 2009. St. Louis, Missouri, USA.
8. ATMOS MedizinTechnik GmbH & Co. KG. Videostroboscopy products 2012. Allentown, Pennsylvania, USA.
9. The Society of Motion Picture and Television Engineers. SMPTE 170M-2004. Television – Composite analog video signal – NTSC for studio applications (Revision of SMPTE 170M-1999). SMPTE Standards 2004.
10. Hillman RE, Mehta DD. The science of stroboscopic imaging. In: Kendall KA, Leonard RJ, editors. Laryngeal evaluation: indirect laryngoscopy to high-speed digital imaging. New York: Thieme Medical Publishers, Inc.; 2010:101–109.
11. Schuster M, Lohscheller J, Kummer P, *et al.* Laser projection in high-speed glottography for high-precision measurements of laryngeal dimensions and dynamics. Eur Arch Otorhinolaryngol 2005; 262:477–481.
12. Bless DM, Hirano M, Feder RJ. Videostroboscopic evaluation of the larynx. Ear Nose Throat J 1987; 66:289–296.
13. Poburka BJ. A new stroboscopy rating form. J Voice 1999; 13:403–413.
14. Poburka BJ, Bless DM. A multimedia, computer-based method for stroboscopy rating training. J Voice 1998; 12:513–526.
15. Nawka T, Konerding U. The interrater reliability of stroboscopy evaluations. J ■■ Voice 2012. [Epub ahead of print]
The authors address a fundamental issue – interrater reliability – that plagues any perceptual rating scales. In their study, laryngeal stroboscopic features exhibited a wide range of interrater reliability, indicating that ratings for certain features should be interpreted with caution.
16. Kelley RT, Colton RH, Casper J, *et al.* Evaluation of stroboscopic signs. ■ J Voice 2011; 25:490–495.
This study offers evidence that visual stroboscopic assessment could be reduced to a small number of parameters to potentially increase efficiency in rating stroboscopic recordings.
17. Patel R, Dailey S, Bless D. Comparison of high-speed digital imaging with stroboscopy for laryngeal imaging of glottal disorders. Ann Otol Rhinol Laryngol 2008; 117:413–424.
18. Mehta DD, Deliyski DD, Zeitels SM, *et al.* Voice production mechanisms following phonosurgical treatment of early glottic cancer. Ann Otol Rhinol Laryngol 2010; 119:1–9.
19. Kendall KA. High-speed laryngeal imaging compared with videostroboscopy in healthy subjects. Arch Otolaryngol Head Neck Surg 2009; 135:274–281.
20. Bonilha HS, Deliyski DD, Gerlach TT. Phase asymmetries in normophonic speakers: visual judgments and objective findings. Am J Speech Lang Pathol 2008; 17:367–376.
21. Bonilha HS, Deliyski DD, Whiteside JP, Gerlach TT. Vocal fold phase ■■ asymmetries in patients with voice disorders: a study across visualization techniques. Am J Speech Lang Pathol 2012; 21:3–15.
This study investigates the ability of raters to visually describe vocal fold phase asymmetries using various stroboscopic and high-speed imaging modalities in a group of patients with voice disorders. More of these types of side-by-side comparisons are necessary to understand the possible inclusion of new imaging modalities into the voice clinic.
22. Patel RR, Pickering J, Stemple J, Donohue KD. A case report in changes in ■ phonatory physiology following voice therapy: application of high-speed imaging. J Voice 2012. [Epub ahead of print]
This case study points to the need for alternative imaging modalities in addition to stroboscopy to assess the progress of voice therapy in a patient with contact granuloma.

23. Krausert CR, Olszewski AE, Taylor LN, *et al.* Mucosal wave measurement and visualization techniques. J Voice 2011; 25:395–405.
24. Cohen SM, Pitman MJ, Noordzij JP, Courey M. Evaluation of dysphonic patients by general otolaryngologists. J Voice 2012. [Epub ahead of print]
25. Kelchner LN, Brehm SB, de Alarcon A, Weinrich B. Update on pediatric voice and airway disorders: assessment and care. Curr Opin Otolaryngol Head Neck Surg 2012; 20:160–164.
26. Nuss R, Ward J, Recko T, *et al.* Validation of a pediatric vocal fold nodule rating scale based on digital video images. Ann Otol Rhinol Laryngol 2012; 121:1–6.
27. Sulica L. Hoarseness. Arch Otolaryngol Head Neck Surg 2011; 137:616–619.
28. Schwartz SR, Cohen SM, Dailey SH, *et al.* Clinical practice guideline: hoarseness (dysphonia). Otolaryngol Head Neck Surg 2009; 141:S1–S31.
29. Rosen CA, Gartner-Schmidt J, Hathaway B, *et al.* A nomenclature paradigm for benign midmembranous vocal fold lesions. Laryngoscope 2012; 122:1335–1341.
30. Welham N, Choi S, Dailey S, *et al.* Prospective multiarm evaluation of surgical treatments for vocal fold scar and pathologic sulcus vocalis. Laryngoscope 2011; 121:1252–1260.