

Letter to the Editor

Laryngeal High-Speed Videoendoscopy: Rationale and Recommendation for Accurate and Consistent Terminology

Dimitar D. Deliyski,^{a,b} Robert E. Hillman,^{c,d,e} and Daryush D. Mehta^{c,d,e}

Purpose: The authors discuss the rationale behind the term *laryngeal high-speed videoendoscopy* to describe the application of high-speed endoscopic imaging techniques to the visualization of vocal fold vibration.

Method: Commentary on the advantages of using accurate and consistent terminology in the field of voice research is provided. Specific justification is described for each component of the term *high-speed videoendoscopy*, which is compared and contrasted with alternative terminologies in the literature.

Results: In addition to the ubiquitous *high-speed* descriptor, the term *endoscopy* is necessary to specify the appropriate imaging technology and distinguish among modalities such as ultrasound, magnetic resonance imaging, and nonendoscopic optical imaging. Furthermore, the term *video* critically

indicates the electronic recording of a sequence of optical still images representing scenes in motion, in contrast to strobed images using high-speed photography and non-optical high-speed magnetic resonance imaging. *High-speed videoendoscopy* thus concisely describes the technology and can be appended by the desired anatomical nomenclature such as *laryngeal*.

Conclusions: *Laryngeal high-speed videoendoscopy* strikes a balance between conciseness and specificity when referring to the typical high-speed imaging method performed on human participants. Guidance for the creation of future terminology provides clarity and context for current and future experiments and the dissemination of results among researchers.

The accurate and consistent use of terminology is an important trait for any scientific field. Describing technologies with concise and specific language helps to disambiguate nomenclature, provides a common reference across investigators, and prevents the potential for misconceptions to propagate in the future. The purpose of this letter is to discuss the rationale behind the term *laryngeal high-speed videoendoscopy* (laryngeal HSV) to describe the application of high-speed endoscopic imaging techniques to the visualization of vocal fold vibration. For example, the ubiquitous term *high-speed digital imaging* could refer to the general application of several different

kinds of imaging technologies, such as magnetic resonance imaging; optical coherence tomography; and ultrasound used in ex vivo, in vivo, and synthetic material environments. Applying the term HSV helps make the needed distinction among these technologies. Notably, the focus of this letter is on full-frame HSV technology versus high-speed videokymographic techniques (Švec & Schutte, 1996).

Table 1 lists popular terms that have been used in the voice literature to describe high-speed imaging of vocal fold vibration and is approximately ordered by the first appearance of a particular term. The history of laryngeal HSV thus can be traced from its earliest years using high-speed (nonvideo) motion picture cameras (Farnsworth, 1940; Herriott & Farnsworth, 1938) and electronic, video-based high-speed camera technology (Hess & Gross, 1993; Kiritani, Hirose, & Imagawa, 1993; Wittenberg, Moser, Tigges, & Eysholdt, 1995). Significant advances in the development and application of HSV technology have been made by researchers in several countries, including the United States, Japan, Germany, Sweden, and France. Using the citations in Table 1, one may explore the evolution of terminology in regard to laboratories, journals, and investigator type (clinician, engineer, etc.).

^aCommunication Sciences Research Center, Cincinnati Children's Hospital Medical Center, OH

^bUniversity of Cincinnati, OH

^cCenter for Laryngeal Surgery and Voice Rehabilitation, Massachusetts General Hospital, Boston, MA

^dHarvard Medical School, Boston, MA

^eMassachusetts General Hospital Institute of Health Professions, Charlestown, MA

Correspondence to Daryush D. Mehta: daryush.mehta@alum.mit.edu

Editor: Jody Kreiman

Received September 10, 2014

Revision received February 25, 2015

Accepted June 9, 2015

DOI: 10.1044/2015_JSLHR-S-14-0253

Disclosure: The authors have declared that no competing interests existed at the time of publication.

Table 1. Tabulation of terminology in the voice literature used to describe the endoscopic visualization of human vocal fold vibration using high-speed imaging techniques.

Terminology	Selected references
High-speed motion picture	Farnsworth, 1940; Herriott & Farnsworth, 1938
Ultrahigh-speed photography	Harden, 1975
Ultra high-speed film	Yamada et al., 1982
High-speed laryngeal film	Baer et al., 1983; Booth & Childers, 1979
High-speed digital imaging	Kendall, 2012; Kiritani et al., 1993; Patel et al., 2008; Yamauchi et al., 2012, 2013; Yan et al., 2006
High-speed imaging	Chen et al., 2014; Echternach et al., 2010; Kendall et al., 2005; Kimura, Imagawa, et al., 2010; Kimura, Nito, et al., 2010; Kreiman et al., 2012; Larsson & Hertegård, 2004; Larsson et al., 2000; Maurer et al., 1996; Neubauer et al., 2001; Patel et al., 2012; van As et al., 1999
High-speed recording	Eysholdt et al., 1996
High-speed laryngoscopy	Deliyski et al., 2006; Granqvist & Lindestad, 2001
Endoscopic digital high-speed images	Döllinger et al., 2003
High-speed video imaging	Granqvist et al., 2003
High-speed digital videoendoscopy	Hertegård, 2005
High-speed laryngeal imaging	Kendall, 2009; Yan et al., 2005
Endoscopic digital high-speed recording	Schwarz et al., 2006
Endoscopic digital high-speed video	Lohscheller et al., 2007
High-speed movies	Lohscheller et al., 2008
High-speed videoendoscopy ^a	Chen et al., 2013; Deliyski et al., 2008; Ikuma et al., 2012, 2013a, 2013b, 2014; Karakozoglou et al., 2012; Krenmayr et al., 2012; Kunduk et al., 2012; Mehta et al., 2011
High-speed digital endoscopy	Döllinger, 2009
Ultra-rapid cinematography	Chevallier et al., 2010
High-speed endoscopy	Döllinger et al., 2012
High-speed videolaryngoscopy	Bohr et al., 2014; Pinheiro et al., 2014

Note. Ordering is by earliest publication year.

^aRecommended terminology.

What is clear from the list and the selected references is the necessity of including the modifier *high-speed* to clarify that the object of interest (vocal fold tissue) is being captured at a rate faster than the eye can see and faster than conventional video-recording technology because the object itself is moving at such a rapid rate. Although we acknowledge that *high-speed* is a relative adjective (today's high speed is often tomorrow's slow speed), the descriptor indicates to the reader that the recording mode is in the general category of high-frame rate imaging. In addition, it is important to note that the HSV modality is inherently in the category of non-real-time imaging in regard to the necessity of slow-motion video playback; in contrast, although other modalities, such as videostroboscopy, may also be played back in slow motion, these modalities are primarily designed to be displayed to the viewer in real time to allow for the simultaneous auditory evaluation of voice quality (Mehta & Hillman, 2012).

To date, practically all technologies for imaging the temporal details of vocal fold vibration make use of high-speed video cameras coupled to an endoscope that is inserted transorally or transnasally to view the laryngeal anatomy. Thus, it might not seem necessary to specify the term *endoscopy* in describing the technology. However, in all likelihood it will simply be a matter of time before other imaging modalities—optical coherence tomography, ultrasound, magnetic resonance imaging, and the like—will catch up in regard to image sampling rate. When that occurs, general terms such as *digital imaging* will lose their uniqueness

in describing a particular imaging modality. Specifying *endoscopy* in the terminology is thus necessary to take this eventuality into account. Even today, high-speed cameras are being used to image everything from the beating of butterfly wings to airbags deployed during vehicle crash testing, and drawing clear distinctions among these applications is already necessary. Also, high-speed video cameras have been all-digital for decades, and certainly they have been digital since the very first time video cameras were used for laryngeal high-speed endoscopic recordings (Hess & Gross, 1993; Kiritani et al., 1993; Wittenberg et al., 1995). Finally, specifying the term *imaging* does not add information but rather creates ambiguity, because biomedical professionals are accustomed to relating this term to techniques based on non-optical principles—such as gamma ray, ultrasound, or nuclear-magnetic resonance—that also require some method of image reconstruction.

Inclusion of the term *video* in our terminology makes an important distinction in disambiguating the term *high-speed*, which by itself could refer to methods based on electronic versus non-electronic principles or optical versus non-optical principles (e.g. high-speed photography, high-speed magnetic resonance imaging, or high-speed video). Although an ingenious technique of revealing previously invisible characteristics, high-speed photography is inherently designed to freeze the rapid motion of an object and capture a single snapshot in time (think of Harold Edgerton's bullet through an apple). High-speed video, however, specifies that more than a few quick image bursts are captured

because a video is a sequence of images that reflect the natural and observable temporal progression of a moving object. Specifically, the Society of Motion Picture & Television Engineers (White Plains, NY) defines *video* as “the technology of electronically capturing, recording, processing, storing, transmitting, and reconstructing a sequence of optical still images representing scenes in motion” (quoted in Watson & Hill, 2012, p. 314). Thus, animated films or sequences of pictures (as in a slideshow) are not videos. Because the definition also limits a video to having captured a singular event, the concatenation of videos taken at different points in time is not a video anymore; it is then considered a *movie* (motion picture). Note that, in clinical voice assessment, the use of *video* in the term *videostroboscopy* is appropriate because stroboscopic sampling (albeit performed at non-uniform intervals) occurs in time order to capture a sequence of observable images. However, analog film-based high-speed motion picture/photography technology (Brubaker & Holinger, 1947; Farnsworth, 1940; Herriott & Farnsworth, 1938; Moore, White, & Von Leden, 1962) is not considered video because those cameras did not electronically capture video images; instead, they used mechanical and chemical methods to expose film for the capture and saving of images.

Laryngeal HSV thus strikes a balance between conciseness and specificity when referring to the currently typical high-speed imaging method performed on human participants. Terms such as *high-speed videolaryngoscopy* (or with variants such as *videostrobolaryngoscopy* and *strobosvideolaryngoscopy*) may be considered unwieldy. In addition, because the technique itself may be succinctly referred to as *HSV*, the acronym *HSV* remains flexible so it can be qualified by other anatomical terms to describe the body part being imaged (e.g., heart valve HSV, velar HSV, or tongue HSV).

We hope that this brief letter provides clarity and context to the terminology chosen and guidance for the creation of future terminology. The suggested HSV terminology has been adopted by an ad hoc committee of the American Speech-Language-Hearing Association that is developing recommended protocols for instrumental voice assessment and protocols for endoscopic, acoustic, and aerodynamic assessment of voice (see Ad Hoc Committee on Developing Instrumental Voice Assessment Protocols, <http://www.asha.org/About/governance/committees/Active-Ad-Hoc-Committees/>). The committee is composed of research scientists and clinicians with international representation and backgrounds spanning voice science, speech-language pathology, otolaryngology, physics, and engineering.

Acknowledgments

This work was supported by a grant from the National Institute on Deafness and Other Communication Disorders (R01 DC007640), awarded to Dimitar D. Deliyski, and by the Voice Health Institute. The article's contents are solely the responsibility of the authors and do not necessarily represent the official views of the National Institutes of Health.

References

- Baer, T., Löfqvist, A., & McGarr, N. S. (1983). Laryngeal vibrations: A comparison between high-speed filming and glottographic techniques. *The Journal of the Acoustical Society of America*, 73, 1304–1308.
- Bohr, C., Kräck, A., Dubrovskiy, D., Eysholdt, U., Svec, J., Psychogios, G., . . . Döllinger, M. (2014). Spatiotemporal analysis of high-speed videolaryngoscopic imaging of organic pathologies in males. *Journal of Speech, Language, and Hearing Research*, 57, 1148–1161.
- Booth, J. R., & Childers, D. G. (1979). Automated analysis of ultra high-speed laryngeal films. *IEEE Transactions on Bio-medical Engineering*, 26, 185–192.
- Brubaker, J. D., & Holinger, P. H. (1947). An endoscopic motion picture camera for otolaryngology and broncho-esophagology. *Journal of the Biological Photographic Association*, 15, 171–192.
- Chen, G., Kreiman, J., & Alwan, A. (2014). The glottaltopogram: A method of analyzing high-speed images of the vocal folds. *Computer Speech and Language*, 28, 1156–1169.
- Chen, G., Kreiman, J., Gerratt, B. R., Neubauer, J., Shue, Y.-L., & Alwan, A. (2013). Development of a glottal area index that integrates glottal gap size and open quotient. *The Journal of the Acoustical Society of America*, 133, 1656–1666.
- Chevallier, G., Sauvaget, E., Herman, P., & Tran Ba Huy, P. (2010). The ultra-rapid cinematography of the larynx, its contributions in speech pathology [French]. *Revue de Laryngologie Otologie Rhinologie*, 131, 23–29.
- Deliyski, D. D., Cieciva, S., & Zielinski, T. (2006). Fast and robust endoscopic motion estimation in high-speed laryngoscopy. *7th International Conference: Advances in Quantitative Laryngology, Voice and Speech Research*, 7(8), 1–12.
- Deliyski, D. D., Petrushev, P. P., Bonilha, H. S., Gerlach, T. T., Martin-Harris, B., & Hillman, R. E. (2008). Clinical implementation of laryngeal high-speed videoendoscopy: Challenges and evolution. *Folia Phoniatrica et Logopaedica*, 60, 33–44.
- Döllinger, M. (2009). The next step in voice assessment: High-speed digital endoscopy and objective evaluation. *Current Bioinformatics*, 4, 101–111.
- Döllinger, M., Braunschweig, T., Lohscheller, J., Eysholdt, U., & Hoppe, U. (2003). Normal voice production: Computation of driving parameters from endoscopic digital high speed images. *Methods of Information in Medicine*, 42, 271–276.
- Döllinger, M., Kunduk, M., Kaltenbacher, M., Vondenhoff, S., Ziethe, A., Eysholdt, U., & Bohr, C. (2012). Analysis of vocal fold function from acoustic data simultaneously recorded with high-speed endoscopy. *Journal of Voice*, 26, 726–733.
- Echternach, M., Dippold, S., Sundberg, J., Arndt, S., Zander, M. F., & Richter, B. (2010). High-speed imaging and electroglottography measurements of the open quotient in untrained male voices' register transitions. *Journal of Voice*, 24, 644–650.
- Eysholdt, U., Tigges, M., Wittenberg, T., & Proschel, U. (1996). Direct evaluation of high-speed recordings of vocal fold vibrations. *Folia Phoniatrica et Logopaedica*, 48, 163–170.
- Farnsworth, D. W. (1940). High-speed motion pictures of the human vocal cords. *Bell Laboratories Record*, 18, 203–208.
- Granqvist, S., Hertegård, S., Larsson, H., & Sundberg, J. (2003). Simultaneous analysis of vocal fold vibration and transglottal airflow: Exploring a new experimental setup. *Journal of Voice*, 17, 319–330.
- Granqvist, S., & Lindestad, P.-Å. (2001). A method of applying Fourier analysis to high-speed laryngoscopy. *The Journal of the Acoustical Society of America*, 110, 3193–3197.

- Harden, R. J.** (1975). Comparison of glottal area changes as measured from ultrahigh-speed photographs and photoelectric glottographs. *Journal of Speech and Hearing Research*, 18, 728–738.
- Herriott, W., & Farnsworth, D. W.** (1938). High speed motion pictures of the vocal cords. *The Journal of the Acoustical Society of America*, 9, 274.
- Hertegård, S.** (2005). What have we learned about laryngeal physiology from high-speed digital videoendoscopy? *Current Opinion in Otolaryngology & Head and Neck Surgery*, 13, 152–156.
- Hess, M. M., & Gross, M.** (1993). High-speed, light-intensified digital imaging of vocal fold vibrations in high optical resolution via indirect microlaryngoscopy. *Annals of Otolaryngology, Rhinology, and Laryngology*, 102, 502–507.
- Ikuma, T., Kunduk, M., & McWhorter, A. J.** (2012). Mitigation of temporal aliasing via harmonic modeling of laryngeal waveforms in high-speed videoendoscopy. *The Journal of the Acoustical Society of America*, 132, 1636–1645.
- Ikuma, T., Kunduk, M., & McWhorter, A. J.** (2013a). Advanced waveform decomposition for high-speed videoendoscopy analysis. *Journal of Voice*, 27, 369–375.
- Ikuma, T., Kunduk, M., & McWhorter, A. J.** (2013b). Preprocessing techniques for high-speed videoendoscopy analysis. *Journal of Voice*, 27, 500–505.
- Ikuma, T., Kunduk, M., & McWhorter, A. J.** (2014). Objective quantification of pre- and postphonosurgery vocal fold vibratory characteristics using high-speed videoendoscopy and a harmonic waveform model. *Journal of Speech, Language, and Hearing Research*, 57, 743–757.
- Karakozoglou, S.-Z., Henrich, N., d'Alessandro, C., & Stylianou, Y.** (2012). Automatic glottal segmentation using local-based active contours and application to glottovibrography. *Speech Communication*, 54, 641–654.
- Kendall, K. A.** (2009). High-speed laryngeal imaging compared with videostroboscopy in healthy subjects. *Archives of Otolaryngology–Head and Neck Surgery*, 135, 274–281.
- Kendall, K. A.** (2012). High-speed digital imaging of the larynx: Recent advances. *Current Opinion in Otolaryngology & Head and Neck Surgery*, 20, 466–471.
- Kendall, K. A., Browning, M. M., & Skovlund, S. M.** (2005). Introduction to high-speed imaging of the larynx. *Current Opinion in Otolaryngology & Head and Neck Surgery*, 13, 135–137.
- Kimura, M., Imagawa, H., Nito, T., Sakakibara, K.-I., Chan, R. W., & Tayama, N.** (2010). Arytenoid adduction for correcting vocal fold asymmetry: High-speed imaging. *Annals of Otolaryngology, Rhinology, and Laryngology*, 119, 439–446.
- Kimura, M., Nito, T., Imagawa, H., Sakakibara, K.-I., Chan, R. W., & Tayama, N.** (2010). Collagen injection for correcting vocal fold asymmetry: High-speed imaging. *Annals of Otolaryngology, Rhinology, and Laryngology*, 119, 359–368.
- Kiritani, S., Hirose, H., & Imagawa, H.** (1993). High-speed digital image analysis of vocal cord vibration in diplophonia. *Speech Communication*, 13, 23–32.
- Kreiman, J., Shue, Y.-L., Chen, G., Iseli, M., Gerratt, B. R., Neubauer, J., & Alwan, A.** (2012). Variability in the relationships among voice quality, harmonic amplitudes, open quotient, and glottal area waveform shape in sustained phonation. *The Journal of the Acoustical Society of America*, 132, 2625–2632.
- Krenmayr, A., Wöllner, T., Supper, N., & Zorowka, P.** (2012). Visualizing phase relations of the vocal folds by means of high-speed videoendoscopy. *Journal of Voice*, 26, 471–479.
- Kunduk, M., Döllinger, M., McWhorter, A., Švec, J., & Lohscheller, J.** (2012). Vocal fold vibratory behavior changes following surgical treatment of polyps investigated with high-speed videoendoscopy and phonovibrography. *Annals of Otolaryngology, Rhinology, and Laryngology*, 121, 355–363.
- Larsson, H., & Hertegård, S.** (2004). Calibration of high-speed imaging by laser triangulation. *Logopedics Phoniatrics Vocology*, 29, 154–161.
- Larsson, H., Hertegård, S., Lindestad, P. A., & Hammarberg, B.** (2000). Vocal fold vibrations: High-speed imaging, kymography, and acoustic analysis: A preliminary report. *Laryngoscope*, 110, 2117–2122.
- Lohscheller, J., Eysholdt, U., Toy, H., & Döllinger, M.** (2008). Phonovibrography: Mapping high-speed movies of vocal fold vibrations into 2-D diagrams for visualizing and analyzing the underlying laryngeal dynamics. *IEEE Transactions on Medical Imaging*, 27, 300–309.
- Lohscheller, J., Toy, H., Rosanowski, F., Eysholdt, U., & Döllinger, M.** (2007). Clinically evaluated procedure for the reconstruction of vocal fold vibrations from endoscopic digital high-speed videos. *Medical Image Analysis*, 11, 400–413.
- Maurer, D., Hess, M., & Gross, M.** (1996). High-speed imaging of vocal fold vibrations and larynx movements within vocalizations of different vowels. *Annals of Otolaryngology, Rhinology, and Laryngology*, 105, 975–981.
- 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., Zañartu, M., Quatieri, T. F., Deliyski, D. D., & Hillman, R. E.** (2011). Investigating acoustic correlates of human vocal fold vibratory phase asymmetry through modeling and laryngeal high-speed videoendoscopy. *The Journal of the Acoustical Society of America*, 130, 3999–4009.
- Moore, G. P., White, F. D., & Von Leden, H.** (1962). Ultra high speed photography in laryngeal physiology. *Journal of Speech and Hearing Disorders*, 27, 165–171.
- Neubauer, J., Mergell, P., Eysholdt, U., & Herzel, H.** (2001). Spatio-temporal analysis of irregular vocal fold oscillations: Biphonation due to desynchronization of spatial modes. *The Journal of the Acoustical Society of America*, 110, 3179–3192.
- Patel, R., Dailey, S., & Bless, D.** (2008). Comparison of high-speed digital imaging with stroboscopy for laryngeal imaging of glottal disorders. *Annals of Otolaryngology, Rhinology, and Laryngology*, 117, 413–424.
- Patel, R. R., Pickering, J., Stemple, J., & Donohue, K. D.** (2012). A case report in changes in phonatory physiology following voice therapy: Application of high-speed imaging. *Journal of Voice*, 26, 734–741.
- Pinheiro, A. P., Dajer, M. E., Hachiya, A., Montagnoli, A. N., & Tsuji, D.** (2014). Graphical evaluation of vocal fold vibratory patterns by high-speed videolaryngoscopy. *Journal of Voice*, 28, 106–111.
- Schwarz, R., Hoppe, U., Schuster, M., Wurzbacher, T., Eysholdt, U., & Lohscheller, J.** (2006). Classification of unilateral vocal fold paralysis by endoscopic digital high-speed recordings and inversion of a biomechanical model. *IEEE Transactions on Biomedical Engineering*, 53, 1099–1108.
- Švec, J. G., & Schutte, H. K.** (1996). Videokymography: High-speed line scanning of vocal fold vibration. *Journal of Voice*, 10, 201–205.
- van As, C. J., Tigges, M., Wittenberg, T., Op de Coul, B. M., Eysholdt, U., & Hilgers, F. J.** (1999). High-speed digital imaging of neoglottic vibration after total laryngectomy. *Archives of Otolaryngology–Head and Neck Surgery*, 125, 891–897.
- Watson, J., & Hill, A.** (2012). *Dictionary of media and communication studies*. New York, NY: Bloomsbury.

-
- Wittenberg, T., Moser, M., Tigges, M., & Eysholdt, U. (1995). Recording, processing, and analysis of digital high-speed sequences in glottography. *Machine Vision and Applications*, 8, 399–404.
- Yamada, M., Ozawa, S., Ono, H., Fukuda, H., & Saito, S. (1982). Automated analysis of vocal cord vibration from ultra high-speed films using angular bend function [Japanese]. *Japanese Journal of Medical Electronics and Biological Engineering*, 20, 101–108.
- Yamauchi, A., Imagawa, H., Sakakibara, K.-I., Yokonishi, H., Nito, T., Yamasoba, T., & Tayama, N. (2013). Phase difference of vocally healthy subjects in high-speed digital imaging analyzed with laryngotopography. *Journal of Voice*, 27, 39–45.
- Yamauchi, A., Imagawa, H., Yokonishi, H., Nito, T., Yamasoba, T., Goto, T., . . . Tayama, N. (2012). Evaluation of vocal fold vibration with an assessment form for high-speed digital imaging: Comparative study between healthy young and elderly subjects. *Journal of Voice*, 26, 742–750.
- Yan, Y., Chen, X., & Bless, D. (2006). Automatic tracing of vocal-fold motion from high-speed digital images. *IEEE Transactions on Biomedical Engineering*, 53, 1394–1400.
- Yan, Y. L., Ahmad, K., Kunduk, M., & Bless, D. (2005). Analysis of vocal-fold vibrations from high-speed laryngeal images using a Hilbert transform-based methodology. *Journal of Voice*, 19, 161–175.