



AMERICAN  
SPEECH-LANGUAGE-  
HEARING  
ASSOCIATION

# Vocal Tract Visualization and Imaging: Technical Report

*ASHA Special Interest Division 3, Working Group on Voice and Voice Disorders*

Reference this material as: American Speech-Language-Hearing Association. (2004). *Vocal Tract Visualization and Imaging: Technical Report* [Technical Report]. Available from [www.asha.org/policy](http://www.asha.org/policy).

Index terms: endoscopy, stroboscopy, voice

DOI: 10.1044/policy.TR2004-00156

---

© Copyright 2004 American Speech-Language-Hearing Association. All rights reserved.

*Disclaimer:* The American Speech-Language-Hearing Association disclaims any liability to any party for the accuracy, completeness, or availability of these documents, or for any damages arising out of the use of the documents and any information they contain.

**About This Document**

This technical report was prepared by Special Interest Division 3, Working Group on Voice and Voice Disorders of the American Speech-Language-Hearing Association (ASHA). Members of the working group were Julie Barkmeier (Chair), Glenn W. Bunting, Douglas M. Hicks, Michael P. Karnell, Stephen C. McFarlane, Robert E. Stone, Shelley Von Berg, and Thomas L. Watterson. Alex F. Johnson served as monitoring vice president. Amy Knapp and Diane R. Paul served as ex officio members. The ASHA Executive Board approved this report in March 2003.

\*\*\*\*

**Introduction**

The purpose of this technical report is to: (a) define the collection of procedures known as vocal tract visualization and imaging, (b) inform speech-language pathologists (SLPs) that performing vocal tract visualization and imaging procedures is within the scope of practice of speech-language pathology provided that appropriate training and experience requirements are met, (c) set the scientific foundation for the topic, (d) review the problems and issues pertinent to the topic, and (e) educate ASHA members and other interested parties about this topic.

**Definition of the Topic**

There are several procedures available for vocal tract visualization and imaging. Rigid fiberoptic oral endoscopy is performed with a rigid tube inserted into the oral cavity or oropharynx. The rigid oral endoscope has a prism optic system that projects high-intensity light at a predetermined angle to illuminate the structures that are to be observed and recorded. The advantages of this technique are high illumination, wide field of view, and excellent image reproduction. The disadvantages are equipment expense, interference with normal speech production, and possible difficulties associated with a gag response in a few patients. Flexible fiberoptic nasendoscopy is performed with a flexible nasendoscope inserted through the nasal passage. High-intensity light, transmitted by a fiberoptic bundle, illuminates structures to be viewed and/or recorded. Advantages are an excellent image of the soft palate and vocal folds during conversation, singing, or deglutition, and the potential for image recording and instant replay. The disadvantages are equipment expense and a possible unpleasant feeling in the nasal passageway or stimulation of the gag response in a few patients. Stroboscopy is performed with any of the above instrumentation, combined with a strobe light that is driven by vocal fold vibration, to permit visualization of vocal tract structures in an apparent slow motion format. Advantages are an extensive body of information relative to the effect of pathology on the process of voicing, and the potential for providing information about the neuromuscular and physiological integrity of the vocal folds and supraglottic structures. A disadvantage is that the image is restricted to isolated vowel production when the strobe light is used in conjunction with a laryngeal mirror or an endoscope.

**Education and Training**

Clinical certification by ASHA ensures that practitioners have met ASHA's education, knowledge, and experience requirements for providing basic clinical services in the profession of speech-language pathology or audiology. ASHA certification in speech-language pathology is necessary but not sufficient for performing the specific clinical procedure(s) discussed in this report. Practitioners are bound by the ASHA Code of Ethics (ASHA, 2003) to maintain high standards

of professional competence. Therefore, practitioners should engage only in those aspects of the professions that are within the scope of their competence, considering their level of education, training, and experience.

Education and training for vocal tract visualization and imaging may be obtained by a variety of means. Some of the training should take place in a clinical setting, allowing the SLP to work with more experienced professionals and a number of patients. SLPs who intend to participate in vocal tract visualization and imaging must ensure that they have acquired the knowledge and skills necessary to provide a continuum of service. These knowledge and skill areas form the basis for assessing clinical competency in this specialized area of practice. An accompanying knowledge and skills document (ASHA, 2004) outlines specific objectives to attain adequate preparation for this procedure, as well as necessary proficiencies and knowledge and skills required to accomplish each objective. This document is intended to provide background information on the application of vocal tract visualization in the clinic. While it is recognized that vocal tract visualization is also used during some research procedures, this document will address only clinical application of vocal tract visualization and imaging.

## Background

Vocal tract visualization and imaging is the collection of procedures for performing a detailed visual examination of the vocal tract and laryngeal and velopharyngeal structures. These procedures enable the SLP to assess the problem as well as to appraise the effect of treatment strategies. Vocal tract visualization and imaging can be an effective tool for evaluating and adjusting treatment of voice, possibly deglutition, and resonance/aeromechanical disorders. These procedures include the use of a constant or stroboscopic light source for indirect laryngoscopy, rigid fiberoptic oral endoscopy, or flexible fiberoptic nasendoscopy. The image produced by any of these techniques can be stored on photographic film, videotape, or digital media. Definitions of these procedures are provided at the end of this report.

Videofluoroscopy, ultrasound, and common motion pictures can also be used to image all or part of the vocal tract and oral structures as a means of assessing function or pathology. However, these procedures are not discussed in this report.

## Scope of Practice

It is the position of ASHA that vocal tract visualization and imaging for the purpose of diagnosing and treating patients with voice, resonance/aeromechanical, or deglutition disorders is within the scope of practice of the SLP (see the position statement on vocal tract visualization and imaging, ASHA, 2004).

The scope of practice (ASHA 2001) grows with advances in technology that enable practitioners to provide new and improved methods of diagnosis and treatment. If practitioners choose to perform these procedures, indicators should be developed, as part of a continuous quality improvement process, to monitor and evaluate the appropriateness, efficacy, and safety of the procedures.

## Education and Training

The procedure of vocal tract visualization and imaging has its roots in the profession of speech-language pathology and in the early efforts of speech/voice scientists, frequently in collaboration with the discipline of medicine required for effective diagnosis. Academic programs have responded by including discussion

and description of endoscopic procedures in courses on clinical voice assessment procedures. Many programs have also acquired endoscopic equipment in order to provide hands-on training for students. While this trend may be limited somewhat by the expense of acquiring and maintaining the equipment, the availability of training may be expected to increase. Post-graduate training seminars will continue to be available for those not associated with academic programs where such training is offered. For many clinicians, it will be necessary to seek training in visualization and imaging after completion of the requirements for the Certificate of Clinical Competence through intensive continuing education, pre-service, or in-service training programs. Education and training may vary for each of these procedures. The training should take place in a clinical setting, allowing the professional to work with more experienced professionals and a number and variety of patients. Each practitioner must determine whether or not he or she has obtained a sufficient degree of education and training to be competent to perform vocal tract visualization and imaging. The safety of the patient is paramount when considering any procedure.

## Precautions

Before undertaking these procedures, practitioners should take the following precautions:

1. Inform institutional and/or regulatory bodies, such as state licensure boards, about these procedures as within the scope of practice;
2. Check with state licensure board(s), where appropriate, to determine whether there are limitations on the scope of SLP practice that restrict the performance of these procedures;
3. Follow universal precautions to prevent the risk of disease transmission from blood/airborne pathogens, such as contained in the Centers for Disease Control Morbidity and Mortality Weekly Report (1988) or ASHA's AIDS/HIV Update (ASHA, 1990);
4. Have available immediate emergency medical assistance when using topical anesthesia or flexible fiberoptic nasendoscopy;
5. Hold a current Basic Life Support Certificate if performing flexible fiberoptic nasendoscopy or using topical anesthesia;
6. Obtain informed consent of the patient and maintain complete and appropriate documentation when performing flexible fiberoptic nasendoscopy or when using topical anesthesia.

## Review of Scientific Foundation

Vocal tract imaging using a continuous light source was first reported by Manuel Garcia (1855). Oertel (1895) applied a stroboscopic light source to a laryngeal mirror in order to examine the slow-motion appearance of the vocal folds throughout the pitch range. More recently, equipment developed for visualizing the vocal tract has incorporated the ability to vary the light source from continuous to stroboscopic light synchronized to the patient's fundamental frequency with permanent recording capabilities. Current methods for visualizing the vocal tract use a flexible or rigid scope. The flexible scope requires placement through the nasal passage, allowing visualization of the velopharynx as well as the laryngopharynx. Since the flexible scope by passes the oral cavity, individuals may be evaluated during both sustained as well as connected speech samples (Karnell, 1994; McFarlane, 1990). The rigid scope is placed transorally for examination of the larynx. Thus, evaluation using this scope is limited to sustained phonation of a vowel. The disadvantage of using a flexible scope is that it contains fewer light

and image conducting fibers, resulting in less light intensity and image resolution than the rigid scope. The rigid scope also offers a higher magnification and less image distortion than the flexible scope (Karnell, 1994).

Research has documented the importance of videoendoscopy for improved diagnosis of voice disorders (Casiano, Zaveri, & Lundy, 1992; Woo, Colton, Casper, & Brewer, 1991). Vocal fold vibratory parameters associated with normal and abnormal voicing have also been proposed (Bless, Hirano, & Feder, 1987). Woo et al. (1991) found that the use of vocal tract imaging played a role in changing the voice disorder diagnosis in 27.2% of 146 patients. Casiano et al. (1992) reported alteration of diagnosis and treatment outcomes in 14% of patients evaluated for a voice disorder. Videostroboscopy was reported as particularly instrumental in determining whether individuals exhibited a vocal fold paralysis or “functional dysphonia” (Casiano et al., 1992). Finally, videostroboscopy allowed clinicians to address assessment of vocal fold vibratory patterns such as glottal configuration, amplitude of vibration, and mucosal wave characteristics (Casiano et al., 1992).

Bless, Hirano, and Feder (1987) described the parameters of vocal fold vibration assessed during videostroboscopic evaluation of phonation. They described the following parameters of vocal fold vibration for perceptual assessment:

1. Symmetry of vocal fold movement, or how well the vocal folds mirror each other, during vibration;
2. Periodicity, or regularity, in the appearance of cycle-to-cycle apparent vocal fold vibration during phonation;
3. Glottic closure, or the appearance of the glottis when the vocal folds are approximated during phonation;
4. Amplitude of vibration, or degree of lateral displacement of the vocal folds from approximation during phonation; and
5. Mucosal wave, or the appearance of a traveling wave along the superior surface of the vocal folds during phonation.

These observations have become part of routine clinical practice, although few data are available about their reliability and validity. Teitler (1995) pointed out the possibility that observer bias is influenced by knowledge of the patients' case history. In this study, however, the rater's degree of bias appeared directly related to experience as a rater of videostroboscopic recordings. In addition, cases associated with case history bias occurred most frequently when judging milder laryngeal pathologies. Bless et al. (1987) addressed possible observer bias related to poor understanding of vocal fold physiology. They concluded that understanding normal vocal fold vibratory alterations with changes in loudness and pitch is important for accurately judging videostroboscopy images of the vocal folds.

Application of imaging techniques to velopharyngeal and oral articulation for speech has revolutionized our understanding of the physiologic bases for oronasal resonance balance. Early studies focusing on imaging techniques (Bjork, 1961; Carrell, 1952) led to later research that has formed the foundation for our current understanding of velopharyngeal movements during speech. Lubker (1968) used cineradiography and electromyography to examine normal levator muscle activity during speech. Subtelny, Kha, and McCormack (1969) provided seminal information about normal bilabial stop and nasal consonant production using

cineradiographic and aerodynamic techniques. Glaser, Skolnick, McWilliams, and Shprintzen (1979) described in detail the activity of Passavant's ridge in normal speakers and speakers with abnormal velopharyngeal mechanisms. Fluoroscopic imaging techniques have been and continue to be a major contributor to basic and clinical speech science. Investigations using magnetic resonance imaging extend this line of research to three dimensions (McGowan et al., 1992; Yamawaki, Nishimura, & Suzuki, 1996).

Studies incorporating endoscopic methods for imaging the velopharyngeal mechanism began to appear in the speech literature more than 30 years ago (Taub, 1966). Since then, dozens of reports have employed endoscopy to describe velopharyngeal physiology in both normal and abnormal speech. For example, Matsuya, Miyazaki, and Yamaoka (1974) used endoscopic images to speculate about the motor innervation of the velopharyngeal mechanism. Bell-Berti and Hirose (1975) used endoscopy and electromyography (EMG) to describe the activity of the palate during normal voiced and unvoiced consonant production. Shelton et al. (1978) and later Siegel-Sadewitz and Shprintzen (1982) described the successful use of videoendoscopic biofeedback therapy designed to improve velopharyngeal closure in hypernasal speakers. Croft, Shprintzen, and Rackoff (1981) described four major patterns of velopharyngeal closure observed during speech. Karnell, Linville, and Edwards (1988) described variations in velar position that normally occur during oral speech using videoendoscopic images. Endoscopic data regarding the details of normal and abnormal velopharyngeal function for speech continue to appear in the speech literature today.

In most of the studies described above, the distinction between applications of vocal tract visualization techniques applied by laryngologists and those applied by speech-language pathologists was not clearly distinguished. A position statement has since been formulated to both affirm and specify the roles of the otolaryngologist and speech-language pathologist in the appropriate use of vocal tract visualization approaches in clinical practice (ASHA, 2000).

## Summary

Vocal tract visualization and imaging techniques represent sophisticated technology that is both currently available and within the scope of practice for speech-language pathologists. Proper clinical use requires specialized training and experience that generally occurs subsequent to graduate education. Equipment cost, patient demographics, and medical support may limit availability to select clinical settings. These imaging techniques provide crucial information for both the differential diagnosis and clinical management of speech, voice, and resonance disorders. Primary benefit relates to our ability to objectively study the physiologic bases of oral communication by revealing the behavioral correlates of phonation, resonance, and articulation. Continued advances in technology will refine the utility of these tools for both research and clinical practice.

## Research Directions

A large body of published research documents the many creative ways visualization and imaging techniques have been applied to basic and clinical science questions regarding speech and swallowing. It is not the purpose of this document to detail these applications. However, there are a variety of technical advances that may be expected to enhance these procedures.

Endoscopic techniques have been successfully integrated into the practices of many SLPs largely because of advances in technology impacting fiberoptic light transmission, light sources, cameras, and recording media. While these advances have been impressive, the field may see even greater developments in the future. For example, we may expect light transmitting fibers to be improved so that reductions in fiber diameter may be achieved without sacrificing light transmission properties or fiber flexibility. The result will likely be smaller diameter flexible endoscopes (<2 mm) that have improved image resolution and light transmission characteristics. A parallel development competing for attention involves the placement of a miniature digital camera at the tip of the insertion tube of the fiberoptic endoscope. This approach has been shown to have merit in that it can produce better quality images than traditionally designed flexible fiberoptic endoscopes that place the camera at the endoscope's eyepiece. As of this writing, this technology is limited to larger diameter endoscopes (>3.5 mm) used primarily in adults. Expected advances in camera technology will eventually lead to similarly designed endoscopes that are small enough (<2 mm) to be used in small children.

An equally exciting area of development involves advances in digital video storage media. Video compression advances coupled with increases in computer processor speeds and network bandwidth expansion all lead to the expectation that digital video records with sound will continue to be produced in less time and will require less digital storage space. Moreover, advances in video streaming technology will permit these records to be easily transmitted and shared over the Internet, ultimately as they are recorded, with professionals, patients, and students. As these records become more accessible, we may expect advances in interpretation and treatment options, not to mention associated advances in education.

Other technological advances that have and will likely continue to lead to improvements in the use of visualization and imaging techniques for speech, swallowing, and laryngeal respiratory function assessment include highspeed video recording techniques, videokymography, application of non-penetrating lasers to quantification techniques, use of air jets for sensory testing, stereoscopic binocular imaging, and synchronized recording with other instruments. Creative advances such as these have already led to improved interpretation and measurement and hint at the inevitability of solutions that can only be imagined today. For example, we may expect a solution to the problems associated with objective, calibrated, reliable measurement of three-dimensional objects with two-dimensional endoscopic images. We may also expect that image distortion due to small lens characteristics will become instantly corrected through digital technology. In the near future we will surely see better color representation that will be valued by speech and medical professionals alike.

There are, no doubt, many other exciting areas where technology will enhance the accessibility and quality of the visualization and imaging techniques we currently use. Investigators and researchers in the speech and hearing sciences have and will continue to be important partners with our colleagues from many other related disciplines who contribute to these advances in science and clinical care.

## Glossary

**Flexible fiberoptic nasendoscopy:** is performed with a flexible nasendoscope inserted through the nasal passage. High-intensity light, transmitted by a fiberoptic bundle, illuminates structures to be viewed and/or recorded. Advantages of this

technique are an excellent image of the vocal folds and velopharyngeal structures during voicing, conversation, or singing, and the potential for image recording and instant replay. Disadvantages are equipment expense and possible patient discomfort.

**Rigid fiberoptic oral endoscopy:** (RFOE) is performed with a rigid tube inserted into the oral or pharyngeal cavity. A prism optic system projects high-intensity light at a predetermined angle to illuminate the structures to be observed and recorded. Advantages are high illumination, wide field of view, and excellent image reproduction. Disadvantages are interference with normal speech production and minor patient discomfort.

**Stroboscopy:** is performed with any of the above instrumentation, combined with a strobe light correlated to vocal fold vibration, permitting vocal tract structures to be visualized in an apparent slow motion format. Advantages are an extensive body of information relative to the effect of pathology on the process of voicing, and the potential for providing information about the neuromuscular and physiological integrity of the vocal folds and supraglottic structures. Disadvantages are patient discomfort related to the use of flexible fiberoptic endoscopy or rigid fiberoptic oral endoscopy, and an image restricted to isolated vowel production when the strobe light is used.

## References

- American Speech-Language-Hearing Association. (1990, December). Report update, AIDS/HIV: Implications for speech-language-pathologists and audiologists. In *Asha* (Vol. 32, pp. 46-48).
- American Speech-Language-Hearing Association. (2000). Roles of the speech-language pathologist and otolaryngologist in the performance and interpretation of endoscopic examinations of swallowing (position statement). *Asha*, 20(Suppl.), 18.
- American Speech-Language-Hearing Association. (2001). *Scope of practice in speech-language pathology*. Rockville, MD: Author.
- American Speech-Language-Hearing Association. (2003). *Code of ethics*. Rockville, MD: Author.
- American Speech-Language-Hearing Association. (2004). *Knowledge and skills for speech-language pathologists with respect to vocal tract visualization and imaging*. Rockville, MD: Author.
- American Speech-Language-Hearing Association. (2004). *Vocal tract visualization and imaging: Position statement*. Rockville, MD: Author.
- Bell-Berti, F., & Hirose, T. (1975). Palatal activity in voicing distinctions: A simultaneous fiberoptic and emg study. *Journal of Phonetics*, 3, 69-74.
- Bjork, L., & Nylen, B.O. (1961). Cineradiography with Synchronized Sound Spectrum Analysis. *Plastic & Reconstructive Surgery*, 27(4), 397-412.
- Bless, D. M., Hirano, M., & Feder, R. J. (1987). Videostroboscopic evaluation of the larynx. *Ear, Nose, and Throat Journal*, 66(7), 289-296.
- Carrell, J. A. (1952). Cinefluorographic technique for the study of velopharyngeal closure. *Journal of Speech and Hearing Disorders*, 17, 224.
- Casiano, R. R., Zaveri, V., & Lundy, D. S. (1992). Efficacy of videostroboscopy in the diagnosis of voice disorders. *Otolaryngology-Head Neck Surgery*, 107(1), 95-100.
- Centers for Disease Control. (1988). Morbidity and mortality weekly report: Perspectives in disease prevention and health promotion. *CDC*, 37, 377-388.
- Croft, C. B., Shprintzen, R. J., & Rakoff, S. J. (1981). Patterns of velopharyngeal valving in normal and cleft palate subjects: A multi-view videofluoroscopic and nasendoscopic study. *Laryngoscope*, 91, 265-271.

- Garcia, M. (1855). Observations of the human voice. *Philosophical Magazine Journal of Science*, 10, 511-513.
- Glaser, E. R., Skolnick, M. L., McWilliams, B. J., & Shprintzen, R. J. (1979). The dynamics of Passavants Ridge in subjects with and without velopharyngeal insufficiency: A multi-view video fluoroscopic study. *Cleft Palate Journal*, 16, 24-33.
- Karnell, M. P. (1994). *Videoendoscopy: From velopharynx to larynx*. San Diego: Singular.
- Karnell, M. P., Linville, R. N., & Edwards, B. A. (1988). Variations in velar position over time: A nasal videoendoscopic study. *Journal of Speech and Hearing Research*, 31, 417-424.
- Lubker, J. F. An electromyographic-cinefluorographic investigation of velar function during normal speech production. *Cleft Palate Journal* 1968 5 1 17
- Matsuya, T., Miyazaki, T., & Yamaoka, H. (1974). Fiberscopic examination of velopharyngeal closure in normal individuals. *Cleft Palate Journal*, 11, 286-291.
- McFarlane, S.C. (1990). Videolaryngoendoscopy and voice disorders. In *Seminars in Speech and Language* (Vol. 11, Issue 1). New York: Thieme Medical.
- McGowan, J. C., Hatabu, H., Yousem, D. M., Randall, P., & Kressel, H. Y. (1992). Evaluation of soft palate function with MRI: Application to the cleft palate patient. *Journal of Computer Assisted Tomography*, 16(6), 877-882.
- Oertel, M. J. (1895). Das Laryngo-Stroboskop und die laryngostroboskopische Untersuchung. *Archiv für Laryngologie und Rhinologie*, 3, 1-5.
- Shelton, R.L., Beaumont, K., Trier, W.C., & Furr, M.L. (1978). Videoendoscopic feedback in training velopharyngeal closure. *Cleft Palate Journal*, 15(1), 6-12.
- Siegel-Sadewitz, V. L., & Shprintzen, R. J. (1982). Nasopharyngoscopy of the normal velopharyngeal sphincter: An experiment of biofeedback. *Cleft Palate Journal*, 19, 194-200.
- Subtelny, J. C., Kho, G. H., & McCormack, R. M. (1969). Multidimensional analysis of bilabial stop and nasal consonants: Cineradiographic and pressured-flow analysis. *Cleft Palate Journal*, 6, 263-289.
- Taub, S. (1966). The Taub oral panendoscope: A new technique. *Cleft Palate Journal*, 3, 328-346.
- Teitler, N. (1995). Examiner bias: Influence of patient history on perceptual ratings of videostroboscopy. *Journal of Voice*, 9(1), 95-105.
- Woo, P., Colton, R., Casper, J., & Brewer, D. (1991). Diagnostic value of stroboscopic examination in hoarse patients. *Journal of Voice*, 5(3), 231-238.
- Yamawaki, Y., Nishimura, Y., & Suzuki, Y. (1996). Velopharyngeal closure and the longus capitis muscle. *Acta Oto-Laryngologica*, 116(5), 774-777.