Assessment of Vocal Cord Function and Voice Disorders

Phillip Song and Ryan J. Song

Contents

Abstract

The assessment of vocal fold function requires multiple elements, including a detailed patient history that considers occupation and environmental exposures, as well as a focused review of systems related to swallowing, breathing, and voice. The physical examination should include evaluation of the larynx, neck, oral cavity, and cranial nerves, along with a perceptual assessment of the voice and laryngoscopic examination. Voice disorders are unique in that their functional impact varies widely depending on a patient's personality,

P. Song (✉)
Division Chief of Laryngology, Department of Otolaryngology and Head and Neck Surgery, Harvard Medical School, Massachusetts Eye and Ear Infirmary, Boston, MA, USA
e-mail: phillip_song@meei.harvard.edu

R. J. Song
Boston University, Boston, MA, USA
e-mail: rjsong@bu.edu

L. Yarmus et al. (eds.), *Principles and Practice of Interventional Pulmonology*,
https://doi.org/10.1007/978-3-031-49583-0_11-1

lifestyle, occupation, and social context. This chapter reviews common assessment tools, including objective acoustic and aerodynamic testing, perceptual rating scales, laryngoscopy, and mucosal wave analysis via videostroboscopy. Additionally, it covers common causes of hoarseness, such as functional and movement-related voice disorders, COVID-19-related issues, and mucosal wave abnormalities.

Keywords

Voice Assessment · Laryngology · Stroboscopy · Perceptual Voice Analysis · Laryngoscopy · Videostroboscopy

1 Introduction

The comprehensive voice assessment is a multidisciplinary examination that includes several different components: a voice-specific history, clinical laryngeal examination with rigid and/or flexible laryngoscopy, videostroboscopy, perceptual voice analysis, acoustic recordings, objective acoustic and aerodynamic data, and physical examination. The voice team may include laryngologists, speech and language pathologists, respiratory therapists, pulmonologists, allergists, and vocal pedagogues. The purpose of this chapter is to introduce the fundamental components of the laryngeal examination including laryngoscopy and stroboscopy as well as discuss common voice disorders. The focus will be on the instrumentation, technique, and discussion of pathology pertinent to airway and interventional pulmonology.

Voice is the product of a complex interplay between the chest, larynx, pharynx, and oral cavity. Our understanding of laryngeal physiology, especially as it pertains to voice production, has evolved from different disciplines including musicology, phonology, endoscopy, and head and neck surgery. Instruments for performing tracheostomy and endoscopy have been found among ancient artifacts of Pompeii and Egypt. The current instruments for the office examination of indirect laryngoscopy using an angled curved mirror and outside light source were developed by Dr. Philip Bozzini in 1807. The first individual to describe voice production as a component of laryngeal function was a baritone opera singer and music professor named Manuel Garcia. Using an angulated dental mirror, he was able to visualize his vocal folds during complex phonatory maneuvers. He observed that the vocal folds closed together to produce sound and opened during respiration. He noted that the tension and length of the vocal folds increased with pitch. He detailed his observations in a manuscript entitled "Physiological Observations on the Human Voice" and presented the paper before the Royal Society of London on May 24, 1855. In 1895, Alfred Kirstein modified a rigid esophagoscope to visualize the vocal folds that eventually gave rise to a generation of laryngoscopes. The laryngoscope enabled direct instrumentation of the larynx, enabling surgeons the ability to remove tumors, treat infectious diseases, and remove foreign bodies. The introduction of fiberoptic technology in the 1960s and 1970s gave rise to flexible endoscopes which allowed for visualization of the larynx in a natural phonatory position. In the mid-1990s, dedicated laryngology fellowships became more widespread, producing another generation of surgeons into the evolving fields of laryngology, professional voice, and neurolaryngology.

2 The Voice History and Patient-Reported Outcome Measures

Referral for a laryngeal examination for dysphonia is determined by risk factors, clinical behavior, and vocal demands of the patient. In general, most hoarseness in the presence of other infectious symptoms referable to the throat spontaneously resolves. The American Academy of Otolaryngology in a clinical practice guideline recommended evaluation of the larynx for hoarseness of greater than 4-week duration regardless of risk assessment and clinical history [1]. However, if there is a risk of malignant or progressive pathology, or if an individual is functionally impaired, assessment should come sooner. Hoarseness of unclear etiology in a high-risk patient with significant smoking and alcohol history requires expedient evaluation of the voice box. Vocal complaints by elite vocal professionals such as actors and singers may require immediate management and evaluation at the onset of symptoms.

The larynx has three primary functions, the regulation of deglutination, respiration, and phonation. The voice history is focused on the nature of the laryngeal dysfunction, difficulty swallowing, aspiration, voice loss, and breathing, and these functions should be seen as the larynx-specific review of systems. Regardless of the nature of the voice complaint, associated symptoms such as dysphagia, odynophagia, respiratory difficulties, stridor, dyspnea on exertion, coughing or choking during meals, or frank aspiration should be elicited. Reflux-related symptoms are also frequently encountered during the voice history, and symptoms such as globus, heartburn, altered taste, and excessive mucus/phlegm should be recorded (Table 1).

The functional impact of hoarseness can vary considerably. The patient's self-perception of the vocal complaints is often not proportional to the perception by an observer. Oftentimes emotional or social factors, such as the diagnosis of throat cancer in a family member, can bring specific awareness for one's voice problem that might otherwise have been neglected. Causes of hoarseness prompting investigation are often inaccurate. Laryngitis is common and is

Table 1 Special topics to include within a voice history

Time course and onset
Profession and vocal demands, including singing
Vocal abuse
Upper respiratory infections
Endotracheal intubation
Surgery in the head and neck, thoracic, cardiac, and esophageal
Laryngeal trauma, external and internal
Tobacco, alcohol, and drug use
Dietary habits
Heartburn and acid reflux
Hydration
Allergies and asthma
Environmental and chemical exposures and reactions
Climate and seasonal changes
Heating and cooling units

frequently attributed to a myriad of things such as colds and other infections (tonsillitis, strep throat, pneumonia, bronchitis, etc.), allergies, overuse, environment, and reflux.

Hoarseness is a general term and there is no physiological or perceptual definition for hoarseness. The examiner should try to direct and focus the patient toward more specific symptoms and problems. Examples include change in the quality of voice, altered pitch, changes in vocal stamina, strain, early vocal fatigue, loss of range, inability to project, inability to be heard, difficulty with articulation, or alteration in clarity. In addition, other components of the voice history should include details regarding the onset of symptoms (gradual or sudden), environmental issues (seasonality, pets, chemical triggers, and sensitivities), precipitating events (upper respiratory infections, stress, travel, screaming or shouting), and exacerbating factors (telephone use, diet, and reflux).

The onset of voice problems and inciting factors should be recorded. An acute onset of voice loss suggests a sudden change in vocal fold pliability or altered vocal behavior. Sudden voice changes occur with vocal hemorrhage, infectious laryngitis, vocal polyp, vocal fold paralysis, and muscle tension dysphonia. More gradual onset of voice changes suggests a more insidious or slow-growing process such as the development of vocal fold nodules, scarring, Reinke's space edema or polypoid corditis (smoker's larynx), presbylarynges, laryngopharyngeal reflux, vocal fold keratosis or leukoplakia, and laryngeal cancer. Most voicing disorders, especially in the early stages, will have fluctuating levels of dysphonia. For instance, patients with vocal fold nodules will note recurrent "laryngitis" with the return of normal voice. Vocal fold paralysis may result in a substantial loss of projection and vocal stamina; however, the degree of dysphonia may vary considerably based on the resting position of the immobile vocal fold, the degree of vocal fold edema, and the pulmonary status.

The quality of life impact of phonatory disorders is highly specific and individual to the patient. An examiner's perception may not correlate with the patient's perspective on the severity of the issue. The magnitude of the clinical problem may not correlate with the perceptual evaluation of hoarseness. The most important factor in assessing the impact of someone's voice dysfunction is usually occupation. A hoarse librarian will not have the same level of impact from dysphonia as a professional singer. The impact of voice dysfunction tends to be highly varied from individual to individual. Some individuals are highly aware of small differences and changes in voice and alterations in voice timbre or frequency. Those with vocally demanding occupations tend to be more aware of minute voice changes than those who do not have significant voice use. Professionals who rely heavily on their voices, such as actors and singers, represent a special class of patients: elite vocal athletes. These individuals produce heavy demands on the vocal folds and test the limits of vocal abilities. Shear trauma or phonotrauma in these individuals is exceedingly high, making them susceptible to vocal injury.

Vocal users can generally be placed into three categories based on voice requirements: the standard user, the vocal professional, and the elite vocal performer. The elite vocal performer is required to use the extreme ranges of his or her phonatory abilities daily for income. This group includes professional actors, singers, and presenters. Vocal professionals are required to speak in public with large groups regularly; this category includes teachers, clergy, politicians, and attorneys. The standard user or nonvocal professional is not significantly impacted by hoarseness or vocal fatigue.

Patient-reported outcome measures (PROMs) are important because phonatory disorders' impact is highly variable and may not correspond measurably to perceptual differences in actual voice. The Voice Handicap Index (VHI) [2] and the Voice Handicap Index-10 (VHI-10) [3] are common patient-reported outcome measurements (see Table 2). It is extremely useful to have patients fill out these self-questionnaires at the initial visit as well as during subsequent follow-up especially if a medical or surgical intervention has taken place (Table 3).

3 Physical Examination of the Larynx

The physical examination should always include a complete examination of the head and neck and may also include relevant inspection of other systems including respiratory, gastrointestinal, and neurological based on history and clinical suspicion. The chapter's focus will be on the laryngeal examination for the diagnosis of voice conditions.

The mucosal tract of the oral cavity, oropharynx, nasopharynx, larynx, and respiratory system all contribute to voice production. These areas should be inspected for mucosal lesions, masses, and movement abnormalities.

Table 2 Voice handicap index-10 (VHI-10)

F1	My voice makes it difficult for people to hear me	0	1	2	3	4
F2	People have difficulty understanding me in a noisy room	0	1	2	3	4
F8	My voice difficulties restrict personal and social life	0	1	2	3	4
F9	I feel left out of conversations because of my voice	0	1	2	3	4
F10	My voice problem cause me to lose income	0	1	2	3	4
P5	I feel as though I have to strain to produce voice	0	1	2	3	4
P6	The clarity of my voice is unpredictable	0	1	2	3	4
E4	My voice problem upsets me	0	1	2	3	4
E6	My voice makes me feel handicapped	0	1	2	3	4
P3	People ask, "What's wrong with your voice?"	0	1	2	3	4

Adapted from source: Rosen et al. [3]

Table 3 Clinical characteristics of common vocal fold pathology

Pathology	Onset	Symptoms (general)
Vocal nodules	Gradual, preceded by recurrent and prolonged laryngitis	Frequent voice loss, roughness, and strain in the voice Easy vocal fatigue and inconsistency
Vocal polyp	Sudden, subacute, generally worse acutely with mild improvement over time	Significant roughness and strain. Difficulty with projection and very inconsistent voice
Vocal hemorrhage	Sudden onset of voice loss with improvement almost back to baseline	In the beginning, severe roughness and strain
Laryngeal granuloma or contact ulcer	Preceding history of intubation or gradual onset	Globus sensation, excessive throat clearing, roughness, and vocal strain
Laryngeal web	Preceding history of surgery or laryngeal trauma, gradual loss of voice	Gradual loss of voice and projection. One of the rare physiological causes of aphonia
Subglottic stenosis	Gradual	Progressive difficulty primarily with projecting the voice and stamina
Bilateral vocal fold immobility	Gradual or sudden	Voice is strained and rough, but often not weak. Breathing symptoms predominate
Unilateral vocal fold immobility	Sudden, subacute	Breathy, weak, and asthenic voice. Easy vocal fatigue and pitch may increase. May complain of breathlessness
Laryngeal cancer	Gradual or subacute	Voice is typically rough and strained with decreased pitch

Examination of the mucosal quality should be carefully performed. Normal, well-hydrated mucosa is important for good sound production, and the presence of abnormal mucosal surfaces, excessive dryness, thick phlegm/mucus, purulent drainage, and fungal changes will result in dysphonia. The tongue and palate should be inspected for good mobility. The nasal exam should include an evaluation for mucosal hypertrophy and allergic changes.

The neck examination should include manual palpation of the neck and thyroid. The larynx should be mobile and minimally tender. During swallowing, the larynx should elevate several centimeters freely and comfortably. However, during comfortable phonation, there should be relatively little vertical elevation of the larynx. The cricothyroid space and the thyrohyoid space should be palpated during phonation. Excessive tension or tenderness of these areas during phonation or inspection would suggest excessive strain or the recruitment of supraglottic muscles to achieve voice. A highly strained and effortful voice may reveal excessive elevation of the larynx, short and tender thyrohyoid and cricothyroid spaces, and tight pharyngeal and strap muscles.

4 Perceptual Voice Examination

Perceptual voice analysis is a physical examination component of the laryngeal exam. The examiner is hearing and analyzing the voice quality to correlate the findings on laryngoscopy. During an exam, the patient may sound stressed or anxious; there may be difficulty with projection or a high degree of strain. The voice may sound flat and disaffected or too emotional. On many levels, all clinicians respond to these differences but often do not act or document these observations. Perceptual voice analysis requires practice systematically and consciously.

Two common perceptual voice scales are the GRBAS (Grade, Roughness, Breathiness, Asthenia, and Strain) scale and the Consensus Auditory-Perceptual Evaluation of Voice (CAPE-V). The GRBAS scale was developed in the 1970s by the Japanese Phoniatric Society [4]. It is widely used and internationally adopted in a wide variety of languages. The GRBAS scale is a perceptual rating scale with five basic parameters. The scale Grade (G) stands for the assessment of the global degree of dysphonia. Roughness (R) is the perceived quality of irregular vibration from the larynx and corresponds to irregular fluctuations of the fundamental frequency or amplitude of the glottal sound source. Breathiness (B) is the voice quality produced by turbulence, air leak, and/or glottal incompetence. Asthenia (A) describes a weakness of spontaneous phonation. Strain (S) or vocal hyperfunction is the auditory impression of excessive effort and tension.

The CAPE-V [5] is a rating scale using parameters of severity, roughness, breathiness, strain, pitch, and loudness, using a 10-cm visual analog scale on a standardized reading passage. The voice evaluation is not limited to these scales but is part of a comprehensive description of the voice and its dysfunction. Oftentimes, patients will be asked to focus and concentrate on a component of their symptoms to better elucidate the problem while the examiner assists in applying relevant descriptors and terminology. The goal of the symptom assessment is to take a global descriptor like hoarseness and break it down into separate components of dysphonia to be able to identify individual aspects of laryngeal dysfunction.

5 Vocal Fold Anatomy and Physiology

Voice is the product of many different structures including lungs, chest wall, neck, pharynx, larynx, oral cavity, nasal cavity, and face. The lungs provide the air and the power to the vocal tract. The vocal folds close to generate subglottal air pressure which is released through the laryngeal valve in the form of a stream of controlled air “puffs” which produce acoustic wave vibration. The vocal folds are known as the “sound generator.” The pharynx, oral cavity, and lips manipulate the acoustic vibration to provide resonance and articulation. The neurophysiological control of the larynx is complex and poorly understood. One of the hallmarks of the voice exam is that form does not always mirror function, meaning that two identical pathological findings on the vocal folds may produce substantially different types of voice. Because of the complexity of the vocal system and the ability of individuals to compensate for a severely disordered larynx, treatment should be tailored to the patient, not the pathology.

The vocal fold comprises three discrete layers: the epithelium, lamina propria, and thyroarytenoid muscle. The true vocal fold is physiologically adapted to produce vocal fold vibration. It has an innate viscoelastic property which allows it to produce vibratory pitch within physiologic limits of subglottal airflow. The medial edge is adapted to sustain repetitive trauma produced by these vibrations. Most of what we know about vocal fold vibration and the function of the vocal cord is attributed to the work of Minoru Hirano in the 1970s [4]. It was his observations that led to the development of the current theory of voice production and vocal fold vibration which is the cover-body theory. The basis of this theory is that there are two distinct structural elements to the vocal folds, the cover which is composed of epithelium and vibratory tissue (superficial layer of the lamina propria) and the body which is the more rigid underlying elements (muscle, ligament, deep lamina propria) (Fig. 1). Examination of the vocal folds evaluates the cover predominantly. The musculomembranous vocal fold refers to the medial edges of the vocal folds which extend from the vocalis process of the arytenoid cartilage to the anterior commissure. The striking zone is an area where the vibration of the vocal folds is felt to be at maximum amplitude and is present at the junction between the anterior and middle third of the musculomembranous vocal fold. A variety of phonotraumatic lesions such as nodules, polyps, and cysts occur in this region. The arytenoid complexes include the arytenoid cartilages and the two sesamoid cartilages, the corniculate and cuneiform. The cricoarytenoid joint is a synovial joint that has rotational as well as translational motion. The arytenoid cartilage has two processes, the medially oriented vocalis process which points toward the airway and provides the attachment to the vocalis muscle (the medial belly of the thyroarytenoid muscle) and vocal folds and the lateral muscular process.

The vocal folds produce acoustic signals by transforming the mechanical energy of vocal fold vibration into acoustic energy with aerodynamic vibration. An ideal vibration can be produced by smooth mucosal edges that close easily by two vocal folds with symmetrical rheological properties and elasticity. This produces an efficient translation of aerodynamic pressure into an acoustic signal.

All methods of physical examination of the vocal cords are from the superior viewpoint. Whether with a rigid or flexible endoscope or with indirect laryngoscopy, views are always from the top looking inferiorly to the base of tongue, epiglottis, superior aspects of the arytenoids, piriform sinuses, true vocal folds, and supraglottis. Areas that are difficult to view from the superior position are the laryngeal ventricles, the infraglottic segment, and the posterior vocal cords, which are sometimes obscured by the arytenoid complex, the pyriform sinuses which are often hidden within folds of tissue, and the vallecula which can be obscured by lymphoid hypertrophy. The postcricoid area of the cervical esophagus is a difficult area to view because the area is usually held closed by the action of the cricopharyngeus muscle. In addition, when the larynx is compromised, the natural adaptive behavior is to contract the false vocal folds and supraglottis, which often obscures visualization of the true vocal cords.

6 Office-Based Examination of the Larynx

6.1 Indirect Laryngoscopy

Office-based assessment of the larynx is performed in several ways. The most common method is indirect laryngoscopy using an angulated dental mirror and headlamp. The angulated mirror, when inserted correctly into the oropharynx

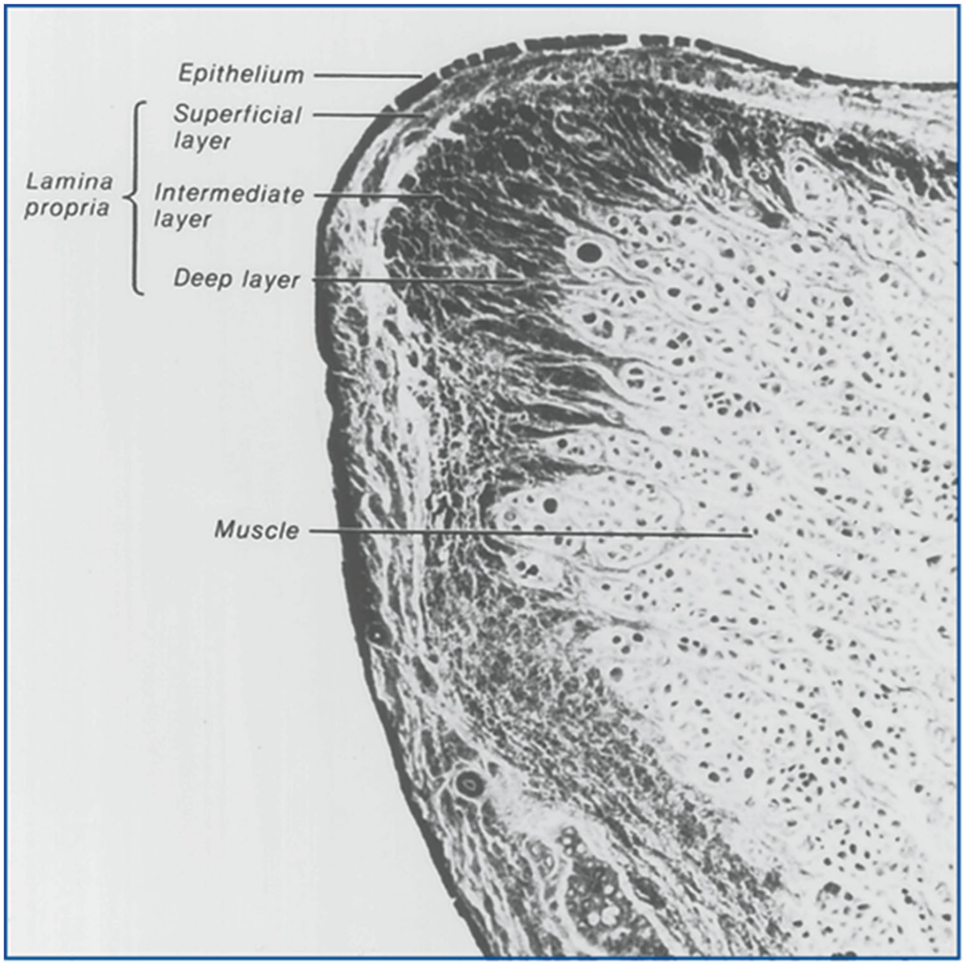

Fig. 1 Vocal fold cross-sectional histology. (Reprinted from Hirano M. Clinical Examination of Voice in Disorders of Human Communication Vol. 5, Springer Verlag; 1981, with kind permission of Springer Science + Business Media)

with a directed light source, can be used to visualize the true vocal fold, false vocal folds, arytenoids, and supraglottic structures quite adequately. The exam is limited by the patient's anatomy and ability to tolerate an instrument deep within the back of the throat. A substantial gag reflex, a large palate, enlarged base of tongue, cervical osteophytes, and hypertrophic tonsils can prevent examination.

6.2 Flexible Transnasal Laryngoscopy

The 1960s and 1970s gave rise to fiber-optic technology and the development of small-diameter fiber-optic endoscopes. Inserted transnasally, these fiber-optic endoscopes could visualize the larynx as the patient performs normal laryngeal tasks. The advent of digital imaging technology has given rise to a new generation of flexible laryngeal scopes or nasopharyngoscopes which has improved visualization further. High-definition imaging with the use of special light filters has also improved the diagnostic sensitivity of the examiner. To perform transnasal laryngoscopy or nasopharyngoscopy, the nasal cavity is topically anesthetized with lidocaine and a nasal decongestant. The endoscope is inserted into the nose toward the nasopharynx. Generally, entrance along the floor of the nasal cavity is preferred, as the sensory innervation is denser in the more superior segments. The endoscope is guided into the oropharynx until the vocal folds are in view.

Once the epiglottis and vocal folds are visualized, the patient is instructed to produce a steady /i/ at a comfortable pitch and volume and again at a high and low register. The flexible endoscope also allows for visualization of the larynx during complex phonatory tasks, such as speaking and singing, and vegetative tasks (coughing, laughing, and throat clearing). Implementing a standardized reading passage and a singing sample (for singers) is an important component of the flexible examinations.

6.3 Rigid Telescope

The angulated rigid telescope is also an excellent method for visualizing the true vocal folds. The degree of angulation is generally 60°, 70°, or 90°. The rigid telescope allows for excellent visualization of the true vocal folds with substantial light and magnification. The rigid telescope is often paired with a stroboscopy unit which is used to determine the pliability or mucosal ways of the vocal folds.

6.4 Videostroboscopy

Videostroboscopy is a powerful tool designed to evaluate the pliability of the vocal fold and the larynx's ability to generate the mucosal wave. The vocal fold is a multilayered structure with several different components responsible for phonatory vibration, and the mucosal wave forms the basis for vocal fold vibration and the generation of an acoustic signal. The stroboscopy unit is essentially a flashing light that is calibrated to the frequency of phonation. The flashing light, or strobe, renders an optical image that seemingly slows down the vibratory movement or "mucosal wave" of the vocal folds. The stroboscopy unit is usually paired with a video system to record and document the examination, and the pairing is referred to as videostroboscopy. Because the viscoelastic properties of the vocal folds determine the signal characteristics of voice production, videostroboscopy is a key examination component. The magnified view of the medial edge of the vocal folds allows for the diagnosis of mass lesions, scars, adynamic components, and loss of tissue. In addition, the depth of a mucosal or submucosal lesion can be better understood by evaluating the impact of the mass on the mucosal wave. A superficial keratosis or leukoplakia will appear to float along the surface of the vocal fold while an invasive carcinoma will stick down to the deeper laryngeal structures.

The instrumentation needed to perform videostroboscopy includes a stroboscopic unit that generates the light and a microphone that captures the frequency of the acoustic signal. An electroglottograph, a device that can measure electrical resistance across the vocal folds using surface electrodes, can also capture the frequency of vocal fold closure to aid the stroboscopic unit. A video capture system that includes a monitor, a rigid or flexible endoscope, and a camera is also necessary. Recording and archiving the examination is important because stroboscopic interpretation often requires repeated rounds of viewing. Videostroboscopy can be performed with a flexible endoscope or rigid telescope.

The performance of rigid videostroboscopy can vary among practitioners. The basic procedure begins by placing the patient in a "sniffing" position. The patient sits forward, with hands or elbows resting lightly on the knees, along with the neck and chin extended. The mouth is opened, and the patient extends the tongue forward out of the mouth. The patient or practitioner can grasp the tongue in the forward position with a cotton pad or 4 × 4 gauze. The rigid telescope is inserted along the floor of the mouth until the larynx or the epiglottis is visualized. Patients with highly sensitive gag reflexes may have difficulty tolerating this portion of the examination. Care should be taken to avoid striking the anterior tonsillar pillars and the posterior pharyngeal wall, as these areas are particularly sensitive. Patients with significant difficulty may benefit from the application of a topical anesthetic such as cetacaine or lidocaine.

Once the epiglottis or vocal folds are visualized, the patient is instructed to produce a steady/i/at their most comfortable pitch and volume. This is the fundamental frequency. The majority of the interpretation should be carried out at this frequency. The patient is then instructed to produce a steady/i/at their upper pitch range and then again at their lowest register. Increasing the pitch will lengthen and tense the vocal folds, thereby highlighting adynamic segments and bringing submucosal lesions more superficial and easier to delineate. The lower pitches can also highlight adynamic segments, especially sulcus deformities. Additional maneuvers to perform during the evaluation include a pitch glide or glissando, where the patient goes from the lower register to the upper in a continuous stream. Brisk nasal inhalation "sniffs" or alternating "sniffs" with/i/can reveal the range of abduction of the vocal folds.

Stroboscopy is a method of assessment rather than an objective test. There are several obstacles in the interpretation of videostroboscopy, the most fundamental of which is the lack of a standardized and reliable methodology of evaluation. There have been several attempts to make a uniform grading system for stroboscopic evaluation; however, no single methodology has been universally accepted. There is inherent variability in the interpretation of the examination that has been well illustrated within the literature.

7 Interpretation of the Laryngeal Exam

Different components of the examination can be broken down into three categories: anatomy, laryngeal motion, and mucosal wave properties. For the sake of consistency, a standardized form should be used so that all components of the examination can be commented on.

Anatomic components include an assessment of supraglottic and laryngeal structures, especially when describing lesions along the medial vocal fold length. Mucosal lesions, trauma, inflammation, and masses along with a description of distribution, location, and size should be noted.

Laryngeal motion includes a functional assessment of gross vocal fold motion, symmetry, and supraglottic function. Special care should be taken to note supraglottic hyperfunction and glottic closure pattern. Hirano and Bless described several different glottic closure patterns [6]. These include complete, incomplete, hourglass, irregular, posterior gap, anterior gap, and spindle gap. Supraglottic compression may be altered between flexible and rigid stroboscopy. In particular, the forced anterior extension of the tongue will alter the hyperfunctional state of the larynx. A description of the supraglottic compression should be included. A common classification system for supraglottic compression divides the types into an open posterior gap-type hyperfunction, predominantly lateral overclosure, predominantly anterior-posterior overclosure, and sphincteric closure. Mucosal wave properties are an assessment of the pliability of the vocal folds and the larynx's ability to generate vibratory energy taken during videostroboscopy. Mucosal wave properties include phase closure, amplitude and phase symmetry, periodicity, adynamic or nonvibratory segments, and duration of closure. Mucosal wave characteristics should be recorded at the patient's comfortable or habitual pitch which reflects the fundamental frequency.

Periodicity describes the cycle-to-cycle regularity of successive acoustic waves. A periodic signal implies a clear, resonant pitch without significant variations between vibratory cycles. An aperiodic or irregular signal implies significant differences in vibratory cycles, producing a rough acoustic signal.

Duration of closure or the closed phase of the mucosal wave is the length of time, relative to the vibratory cycle, that the vocal folds remain in a closed position. The duration of closure can be determined based on a gestalt or approximate measure or more quantitatively by measuring the length of the closed phase with digital photography or videokymography.

Mucosal wave amplitude describes the horizontal excursion of the vocal folds and is an indication of the viscoelastic properties of the vocal fold. It is a somewhat subjective description, as the amplitude can vary depending on pitch, frequency, effort, and volume. Other factors such as angulation of the endoscope and supraglottic compression can influence the degree of vocal fold show.

8 Voice Analysis

Acoustical analysis and airflow studies are instrumental measures of voice and useful adjuncts for vocal analysis. These measures can localize areas of dysfunction within the phonatory mechanism and are quantifiable values that can be used to measure the progress of therapy. Much effort has been expended in acoustic and speech studies toward the search for standardized acoustic parameters that are reliable and reproducible among different subjects and have diagnostic value for various vocal pathologies. Although no perfect acoustic analysis measure has yet to be identified, these measures are useful for following the effectiveness of therapies and are being applied to machine learning models. By building a body of normative data, these parameters may have further applications in the future.

The frequency is the rate of vibration of the vocal folds measured in cycles per second or Hertz. The fundamental frequency is the natural or most comfortable pitch delivered by the larynx. Pitch can be measured by waveform analysis of the voice. Smooth vocalizations when digitized form periodic waves that can be graphically represented. The mean fundamental frequency is an average value that is calculated during sustained vowels or extracted during speech. The phonation range is the range of fundamental frequencies an individual can produce, which is an important characteristic for singers.

Intensity is a measure of loudness. On an acoustic waveform, intensity corresponds to the amplitude height. Maximal and minimal intensity can be measured at different fundamental frequencies to graph a phonetogram. Like an audiogram, the phonetogram is a visual representation of the vocal range in a given individual. Frequency range and intensity range are effort-dependent and can vary when changing testing environments. Standardization is extremely important when eliciting these values.

Jitter and shimmer are two perturbation measures. Perturbation is the cycle-to-cycle variability of the acoustic waveform during a sustained vowel. Shimmer is the perturbation in intensity or loudness, and jitter is the perturbation in frequency. These measures correspond to how smooth a voice sounds. These values are most consistent in the absence of significant dysphonia/dysarthria as it takes a sustained uninterrupted vowel sound to measure waveforms. Perturbation parameters are not useful for patients with severe dysphonia and dysarthria; however, they can be used to evaluate clinical treatment efficacy.

Signal-to-noise ratios are measures comparing harmonic signal energy to aperiodic or noise energy. During smooth, uninterrupted vocalization, the majority of energy is harmonic and forms well-defined periodic waveforms. As the mucosal wave is disrupted, the acoustic waveforms become irregular, forming white noise. Generally, large noise energy with greater random aperiodicity represents an abnormal vocal function.

Acoustic analysis testing in general requires a rigorously standardized testing environment, skilled evaluators, and highly motivated voice users. Accurate waveforms are difficult to elicit in the presence of severe dysphonia or dysarthria, rendering most parameters inconsistent. While clinical correlation to mucosal wave vibration, glottic closure, and power is possible, acoustic parameters do not have diagnostic value for the localization of vocal pathology.

Machine learning applications have recently emerged as powerful tools for disease identification through voice analysis. The standard approach involves extracting acoustic features of the vocal tract, such as pitch, volume, tone, duration, and intensity. These features are automatically extracted using signal processing techniques and compared to healthy individuals to diagnose disease. Various illnesses can alter these vocal characteristics; unilateral vocal fold paralysis typically lowers the voice's volume, duration, and pitch. Machine learning algorithms can analyze these vocal qualities and identify correlations between sound patterns and throat health by converting audio recordings of patients' voices into visual representations (such as spectrograms or waveforms). This process enables faster and more accurate medical diagnosis and treatment planning. The ability of ML models to detect subtle changes in voice features may be valuable in early disease detection and monitoring of treatment progress.

9 Aerodynamic Measurements

Because of the important role of the lungs in vocal production, pulmonary function testing is helpful for the evaluation of weak and asthenic voice and vocal fatigue. The lungs provide the ventilatory support needed to initiate the mucosal wave. Poor pulmonary reserve is associated with weak, breathy, and poor voice quality.

Subglottal pressure can be measured directly through the introduction of a pressure transducer through the trachea; however, most of the subglottic pressure measurements are made indirectly. By holding the transducer in the mouth, pursing the lips tightly, and exhaling, an indirect measure of subglottal pressure via transoral pressure is obtained. Normative testing has shown that the minimal subglottic air pressure needed to support voice is between 3 and 7 cm of water. Average airflow rates between the vocal cords are around 50–200 ml of water. The phonation threshold is the minimal subglottal pressure needed to initiate vocal fold vibration. Taken together, these aerodynamic measurements provide a sample of laryngeal valve function and mucosal wave integrity.

In general, objective measurements of acoustic and aerodynamic voice analysis suffer from similar problems when using them clinically. Instrumental standardization has not yet been fully implemented, so readings taken with different equipment and in different laboratories are not necessarily comparable. Technical expertise is required to administer these tests and to interpret the results. The type of speech sample used also affects results. These tests require a high level of understanding and function for the subject, and a number of these tests are effort dependent. For severely dysphonic patients, there is considerable variability in testing results.

10 Common Vocal Problems

Common voice problems related to laryngeal dysfunction can be broken down into several different categories: nonorganic or functional voice disorders, motion or movement disorders, and anatomic laryngeal pathology which disrupts the mucosal wave. There is a great deal of controversy in the laryngology literature regarding definitions of common vocal fold pathology. This chapter will try to be consistent with descriptions and labels, although in practice, there may be significant differences between definitions from one interpreter to another. Although a comprehensive discussion of the causes of dysphonia is beyond the scope of this chapter, an attempt will be made to describe the most common etiologies in the voice clinic.

11 Functional Voice Disorders

Functional voice disorders are nonorganic vocal problems characterized by dysphonia secondary to inappropriate vocal tension without anatomic, physiological, or neurological basis. Laryngeal anatomy is normal or near normal, and the symptoms are disproportionate to the examination findings. The motion of the vocal folds is intact and symmetric, and there is no discernable laryngeal pathology that corresponds to the hoarseness. Hoarseness is often much more severe than the laryngeal findings. The symptoms can be highly variable. The voice can be strained or breathy, weak or excessively loud, and consistent or inconsistent. There can be pitch alterations. During the laryngeal exam, attention is carefully made to the behavior of the larynx. There are often features of excessive force or traction on the vocal folds, with increased contraction and function of the supraglottic larynx and extralaryngeal tension. Clinically, the patient has a highly strained vocal quality as well as features of excessive work of voicing. The neck is often tight and tender to palpation, especially in the thyrohyoid space and cricothyroid space, and the larynx is in a high, tight position against the hyoid bone.

The hallmark of functional voice disorders is disproportionate hoarseness relative to the laryngeal findings. The most common functional voice disorder is primary muscle tension dysphonia (MTD) or vocal hyperfunction. There is excessive tension and effort placed onto the vocal track producing a strained, strangled voice quality. Critical evaluation of which laryngeal muscles are dysfunctional is very important. Muscle tension dysphonia can be primary or secondary. Primary muscle tension dysphonia is often associated with anxiety or laryngeal irritability. Secondary muscle tension disorder refers to compensation secondary to some glottic abnormality or mucosal wave pathology. Secondary muscle tension dysphonia may be a very normal compensatory response.

There are several ways to categorize types of muscle tension dysphonia. Supraglottic manifestations of hyperfunction include ventricular hyperfunction which is the overclosure of the false vocal folds over the true vocal folds, anterior-posterior compression which refers to overclosure secondary to the epiglottis and arytenoid complex narrowing, and sphincteric closure where both dimensions are involved, closing off the larynx. Muscle tension may also be the product of excessive tension of the intrinsic laryngeal muscles without significant supraglottic activity. There may be overrotation of the lateral cricoarytenoid muscle creating a posterior glottic gap, and excessive vocalis process shows excessive cricothyroid muscle motion with a lengthened and tight vocal fold or hyperfunction of the transverse arytenoid muscles causing a scissoring action of the posterior glottis. Voice therapy is the treatment of choice for vocal hyperfunction.

12 Motion Disorders of the Larynx

Motion disorders of the larynx are a group of disorders that cause hoarseness secondary to alteration in neurological input and control. The motion may be hyperkinetic or hypokinetic. Examples of hyperkinetic motion abnormalities include neurological diseases such as spasmodic dysphonia, laryngeal tics, myoclonus, and vocal tremor. Hypokinetic motion disorders include vocal fold immobility, paresis, or paralysis. These problems may require systemic medications, botulinum toxin injections to the larynx, voice therapy, and phonosurgery.

Vocal fold immobility describes a fixed vocal fold or reduced motion observed during examination. Vocal fold paralysis and paresis are the result of injury to the recurrent or superior laryngeal nerve. The major causes of vocal fold paralysis include iatrogenic injury (intubation, head and neck or cardiac surgery, trauma), neoplasm (lung, head and neck, and brain), and idiopathic. Vocal fold immobility may result from injury at the cricoarytenoid joint. Subluxation or dislocation of the joint can be the consequence of trauma and intubation. There may also be mass effect from an adjacent neoplasm or fibrosis which interferes with the range of motion of the vocal folds.

Unilateral vocal fold immobility causes glottic insufficiency with symptoms of breathy voice, vocal fatigue, inability to project, and excessive strain. Phonatory dyspnea describes running out of breath during speech because of inefficient breath support.

Bilateral vocal fold immobility may be the result of bilateral nerve injury but is more commonly a result of mechanical fixation. The mechanical fixation can be from bilateral cricoarytenoid joint fixation or posterior glottic scarring. The common causes of bilateral vocal fold fixation include prolonged intubation, intralaryngeal trauma, and radiation fibrosis. Laryngopharyngeal reflux and autoimmune diseases such as Wegener's disease can predispose to granulation tissue and posterior glottic scarring. Bilateral paralysis (neurogenic) can be the result of tumors, strokes, and surgery, notably esophagectomy, total thyroidectomy, and tracheal resection.

Bilateral vocal fold immobility results in primarily respiratory symptoms, shortness of breath, difficulty breathing at night, exertional dyspnea, and biphasic stridor, while voice changes are usually minimal. In fact, in those operated on to open the airway, the return of voice is typically followed by the return of breathing symptoms with restricted airflow.

13 Mucosal Wave Pathology

Vocal fold pathology disrupts vocal fold pliability and the mucosal wave. The mucosal wave refers to the vibratory wave of the vocal folds which produces the main sound signal. The viscoelastic properties of the vocal folds may be impaired by a variety of benign and malignant pathologies. In general, lesions are phonotraumatic, which are a consequence of vocal overuse, or nonphonotraumatic lesions.

Phonotraumatic lesions are typically submucosal and cause loss of the superficial lamina propria early in the presentation. This results in stiffness of the vocal fold and difficulty producing the mucosal wave. Common phonotraumatic lesions include vocal nodules (singer's or screamer's nodules), polyps, sulcus deformities, vocal hemorrhages, varices/ectasias, and cysts.

Common phonotraumatic vocal fold pathologies are vocal nodules, polyps, and cysts. The etiology of these lesions is from excessive phonotrauma, the trauma created as the vocal folds strike each other during vibration. These lesions occur in a specific location of the vocal folds, the "striking zone," which is in the junction of the anterior and middle third of the musculomembranous vocal fold. These benign disorders have a characteristic appearance. Vocal fold nodules are typically symmetric with a broad base and significant scarring (see Fig. 2). Vocal fold polyps are typically unilateral raised mass lesions, although a small scar is commonly seen on the contralateral vocal fold (see Fig. 3). Both nodules and polyps are histologically similar and occur in the subepithelial layer or the basement membrane of the epithelium. Vocal cysts are deeper lesions and may be true epithelial-lined cysts or false cysts. The cysts can be tethered to the vocalis process, and the margins can extend a considerable distance beyond the actual mass lesion.

Vocal ectasias and varices are abnormal distended blood vessels on the superficial surface of the vocal folds. The lesions can rupture with excessive pressure and lead to vocal hemorrhage (see Fig. 4).

Nonphonotraumatic lesions are mostly epithelial lesions affecting the mucosal surface. Epithelial changes include laryngeal squamous cell carcinoma, premalignant changes (leukoplakia and erythroplakia), and infections (recurrent respiratory papillomatosis, fungal infection, and bacterial and viral laryngitis). Certain conditions such as mucosal trauma, vocal fold ulcerations, and anterior glottic webbing (see Fig. 5) are also epithelial-based.

Vocal fold pathologies which affect the epithelium include squamous cell carcinoma, leukoplakia, and keratosis. Squamous cell carcinoma is by far the most common malignancy of the vocal folds. Leukoplakia is a white patch on the vocal fold and implies a premalignant clonal proliferation of abnormal tissue folds (see Fig. 6). Keratosis or hyperkeratosis is the

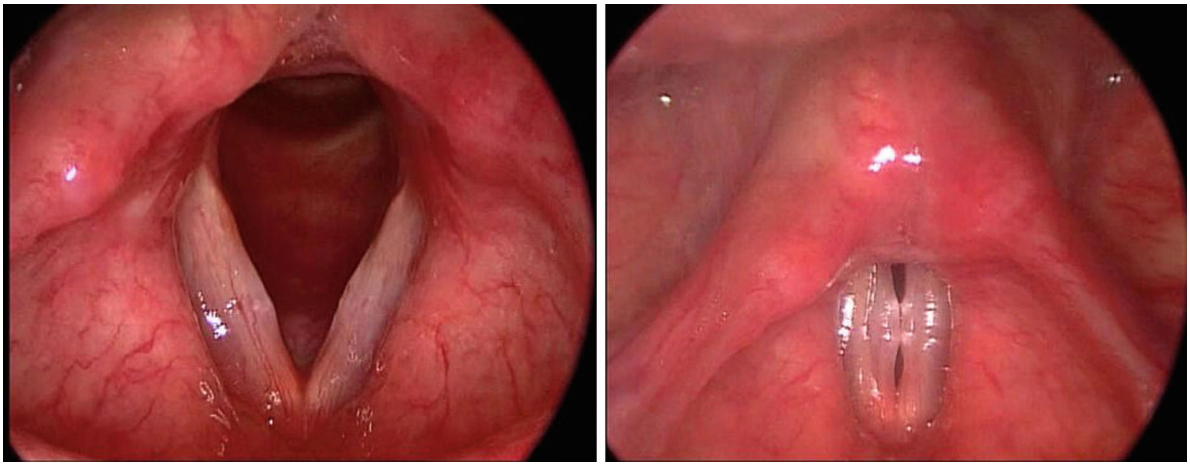

Fig. 2 Vocal fold anatomy and demonstration of vocal fold nodules

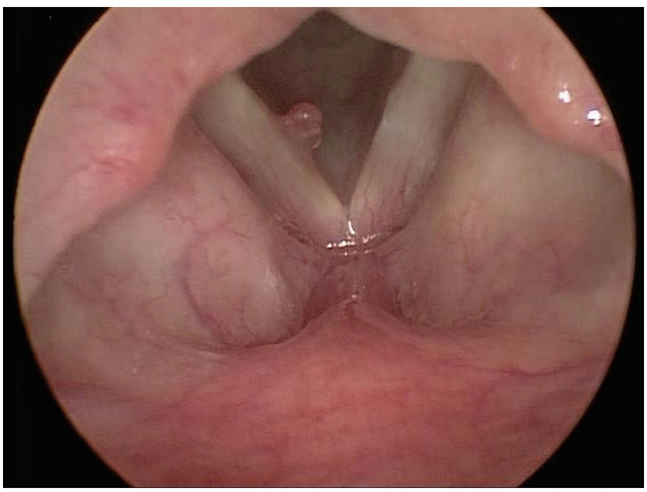

Fig. 3 Vocal fold polyp

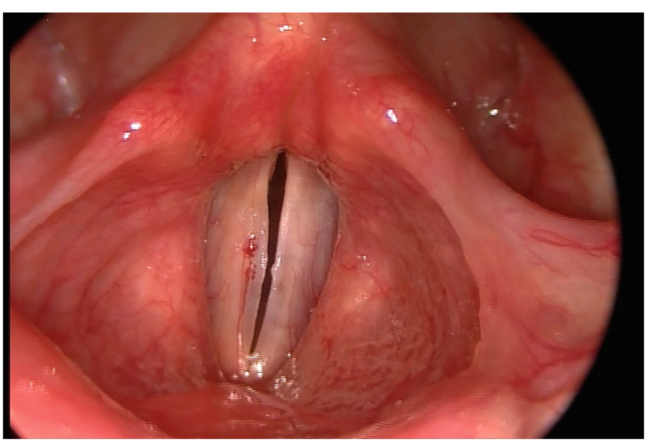

Fig. 4 Varix

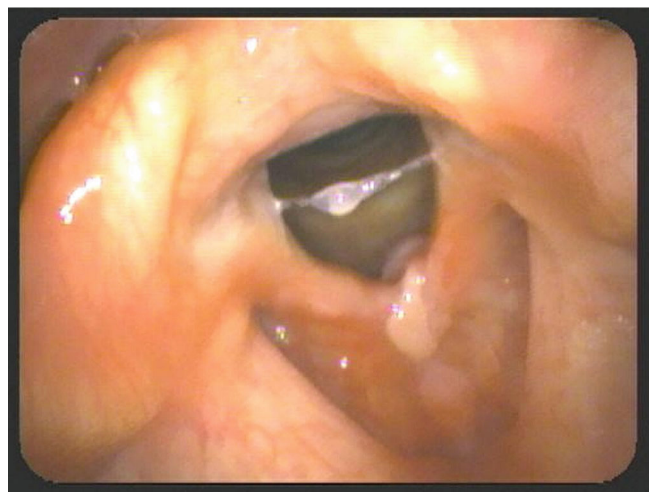

Fig. 5 Anterior glottic web

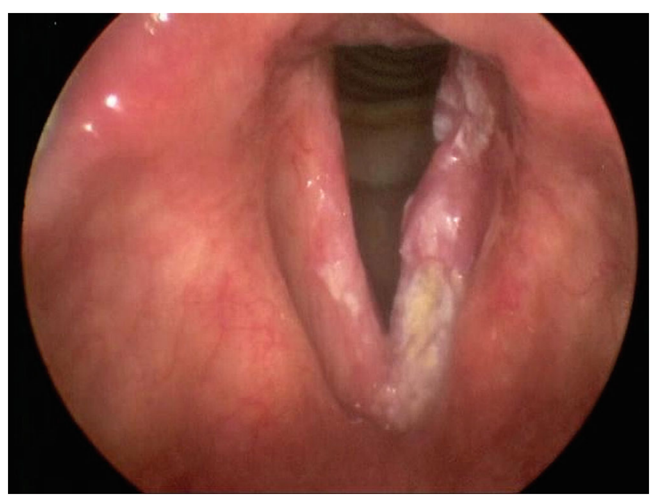

Fig. 6 Leukoplakia

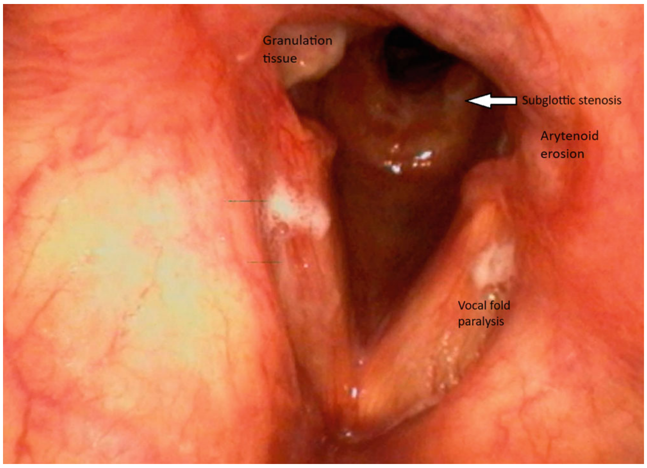

Fig. 7 Impact of prolonged intubation after COVID respiratory failure

development of the white patch from excessive keratin formation on the normally nonkeratinizing epithelium of the vocal. Rates of progression from cellular atypia to squamous cell carcinoma vary from 5% to 18% depending on the degree of atypia. The impact on voice depends on the location of the lesion. Lesions anterior to and at the striking zone will have earlier voice changes than those that occur elsewhere. Lesions in the laryngeal ventricles, false vocal folds, posterior true vocal folds, and arytenoid regions will need to be much larger and more exophytic to cause voice changes.

14 COVID-19 and Voice

The coronavirus disease 2019 (COVID-19) pandemic has led to an increase in the rates of voice disorders [7]. This rise may be attributed to the direct effects of COVID-19 on the larynx, including neuropathy, autonomic dysfunction, inflammation, and respiratory problems. Autonomic dysfunction, such as cough and laryngeal spasm, is commonly associated with long-COVID symptoms. Additionally, the effects of respiratory failure and hospitalization have created a group of patients with laryngeal injuries, including vocal fold paralysis, posterior glottic scarring, granulation tissue, and subglottic stenosis (see Fig. 7). The impact of COVID-19 on voice issues may also be linked to widespread lifestyle and occupational changes, such as increased use of teleconferences and phone meetings, which exacerbate vocal strain. Moreover, advancements in technology, like the proliferation of voice recognition software and virtual assistants that struggle with patients who have voice disorders, may also be drawing more attention to hoarseness and voice changes, leading to increased patient visits [8].

15 Key Points

- The comprehensive voice evaluation involves several components including a voice-specific history, perceptual auditory evaluation, visualization of the larynx in motion, stroboscopic evaluation of the mucosal wave, and aerodynamic and acoustical analysis.
- Members of a multidisciplinary voice care team involve laryngologists and speech and language pathologists with voice-specific training and can include other professionals such as vocal coaches, teachers, pedagogues, and medical specialties.
- Quality of life and self-assessment tools are an important component of the voice history.
- Image quality is important in evaluating vocal anomalies, and the examiner should choose the best available instruments for evaluating anatomic and functional components of the vocal folds.
- Videostroboscopy allows for the assessment of mucosal wave or vibratory function of the vocal folds.
- Flexible nasopharyngeal endoscopy enables visualization of the larynx in a more natural posture during complex phonatory tasks.
- Instrumental analysis of voice including aerodynamic and acoustical analysis provides quantifiable tools to the laryngologist; however, no perfect value or test exists. Careful history taking and clinical evaluation are still the most important tools toward diagnosis.
- Voice problems can be the result of anatomic and physiological disruption of the mucosal wave as well as motion abnormalities of the larynx.
- Functional voice disorders are very common and are usually the product of elevated tension and strain in the larynx.
- Phonotraumatic lesions typically occur on the middle third of the vocal fold called the striking zone. This area corresponds to the location of maximal velocity and amplitude of the mucosal wave. Lesions that occur outside of the striking zone need a higher degree of suspicion.
- Motion disorders of the larynx include vocal fold paralysis and paresis, vocal fold fixation from disruption of the cricoarytenoid joint, and neurological problems.

16 Conclusion

Voice production involves multiple physiological systems, including the neurological, respiratory, and laryngeal systems. The impact of communication loss or hoarseness varies greatly among individuals and may not always correlate with physical examination findings. A thorough voice assessment should include standardized perceptual evaluation tools such as the GRBAS or CAPE-V scales. Patient-reported outcome measures help to better understand the specific functional effects of voice disorders. Although objective measures like acoustic and aerodynamic testing are useful, they may not fully reflect the patient's experience. Direct visualization of the larynx through laryngoscopy and mucosal wave analysis using videostroboscopy remains a key component of comprehensive voice assessment.

Competing Interest Declaration The author(s) has no competing interests to declare that are relevant to the content of this manuscript.

Bibliography

1. Stachler RJ, Francis DO, Schwartz SR, Damask CC, Digoy GP, Krouse HJ, et al. Clinical practice guideline: hoarseness (dysphonia) (update). Otolaryngol Head Neck Surg Off J Am Acad Otolaryngol-Head Neck Surg. 2018;158(1_suppl):S1–42.
2. Jacobson BH, Johnson A, Grywalski C, Silbergleit A, Jacobson G, Benninger MS, et al. The voice handicap index (VHI): development and validation. Am J Speech Lang Pathol [Internet]. 1997 [cited 2024 Aug 19];6(3):66–70. http://pubs.asha.org/doi/10.1044/1058-0360.0603.66
3. Rosen CA, Lee AS, Osborne J, Zullo T, Murry T. Development and validation of the voice handicap index-10. Laryngoscope. 2004;114 (9):1549–56.
4. Hirano M. Clinical examination of voice. New York: Springer; 1981.
5. Zraick RI, Kempster GB, Connor NP, Thibeault S, Klaben BK, Bursac Z, et al. Establishing validity of the consensus auditory-perceptual evaluation of voice (CAPE-V). Am J Speech Lang Pathol [Internet]. 2011 [cited 2024 Aug 18];20(1):14–22. http://pubs.asha.org/doi/10.1044/1058-0360%282010/09-0105%29
6. Hirano M, Bless DM. Videostroboscopic examination of the larynx. San Diego: Singular Publishing Group; 1993. 249 p.
7. Naunheim MR, DeVore EK, Huston MN, Song PC, Franco RA, Bhattacharyya N. Increasing prevalence of voice disorders in the USA: updates in the COVID era. Laryngoscope. 2024;134(8): 3713–8.
8. Naunheim MR, Zhou AS, Puka E, Franco RA, Carroll TL, Teng SE, et al. Laryngeal complications of COVID-19. Laryngosc Investig Otolaryngol. 2020;5(6):1117–24.

Suggested Reading

Bless DM. Measurement of vocal function. Otolaryngol Clin North Am. 1991;24(5):1023–33.

Hirano M. The clinical evaluation of voice. Vienna: Springer; 1981.

Koufman JA, Blalock PD. Functional voice disorders. Otolaryngol Clin North Am. 1991;24(5):1059–73.

Jacobson BH, Johnson A, Grywalsky C, et al. The voice handicap index (VHI): development and validation. Am J Speech Lang Pathol. 1997;6:66–70.

Gray S, Hammond E, Hanson D. Benign pathologic responses of the larynx. Ann Otol Rhinol Laryngol. 1995;104:13–8.