VOCAL CAPABILITIES

Subglottic Pressure
(Volume)

High

Low

Pitch

Typical speaking pitch

Typical speaking range

Low

High

Current sp. pitch
Clear sound
Air leak
Rough
No sound

Flutter
Diplophonia
Onset delay
Pitch break
Vocal capabilities pattern matching is a group of voice elicitations that the examiner may use to evaluate and document phonatory ability and impairment of the human voice (hoarseness). Pitch and volume are varied while the examiner records the voice and notes when sound (the voice) is clear (harmonic) and at what pitch/volume combinations the voice is impaired. Secondly, the examiner notes how it is impaired – that is, which tasks generate breathiness, which generate roughness and which generate both qualities. Third, for further precision, the tones on which impairment occurs are noted. Every type of voice disorder will thus have a vocal signature – specific findings of types and degrees of roughness and/or breathiness occurring at various pitch/volume combinations – which are elicited and recorded as a vocal capabilities pattern.

When vocal capabilities pattern matching is utilized in conjunction with the patient history, the differential diagnosis of hoarseness becomes more refined before proceeding with visual examinations of endoscopy and stroboscopy. Audible findings will both predict visual findings as well as direct the endoscopic examination. During endoscopy, the vibratory edges of the vocal cords will then be visualized and recorded during appropriate pitch and volume to visually identify and document impaired sound production. The endoscopic exam also documents air leak and multiple sound sources as the etiology for the perceived breathiness and roughness, which in turn, represent the patient’s complaint of hoarseness. In addition to focusing the exam, the examiner is more confident in the subsequent visual diagnosis.

Audio recording of vocal capabilities is essential documentation of vocal impairment. This recording should be part of the minimum examination performed for the diagnosis of hoarseness as well as recorded before and after any treatment or surgery that has the potential to impact the voice directly (e.g. microlaryngoscopy) or indirectly (e.g. thyroid surgery).

KEYWORDS

Voice measurement; hoarseness; dysphonia; voice pathology; laryngoscopy; stroboscopy; roughness; breathiness; larynx; vocal nodules; vocal polyps; vocal cords; vocal fold bowing; hoarseness; behavior; vocal trauma; vocal abuse; vocal overuse; Voice disorders/diagnosis; Dysphonia/diagnosis; Dysphonia/etiology; Predictive Value of Tests; Laryngeal Diseases/diagnosis; voice quality; Larynx/physiopathology
The problem: inadequate or no voice assessment

The most common puzzle-solving algorithm for hoarseness is a two-part method. A history is taken followed by visual examination of the throat, quite often with endoscopy or for some examiners utilizing a mirror, looking for a lesion on the larynx. Although this technique will tend to identify gross problems, it exposes the examiner to uncertainty, specifically; is a visual finding on the larynx actually related to the patient’s hoarse voice? If I didn’t see anything, is nothing truly there? An examiner may feel confident about making a correct diagnosis and still be incorrect. Too commonly patients hear, “I don’t see anything wrong,” or “It looks red.” This two-part examination misses less-than-obvious problems and without audible correlation of vocal cord function, may even lead the examiner down inappropriate diagnostic and treatment pathways.

What is hoarseness?

Accurate troubleshooting of the impaired human voice requires a basic understanding of the physics of sound production. For most individuals, a non-impaired, functional voice is one that produces a clear tone and can be manipulated rather freely in terms of volume and pitch. We can also manipulate clarity, which can be thought of as the intentional introduction of air leak or irregular vibrations. When clarity is impaired unintentionally, a patient typically complains of hoarseness. Two primary sound production impairments are 1) unwanted breathiness (white noise) and 2) unwanted roughness (simultaneous production of more than one pitch or polyphony). Both represent essentially non-harmonic passage of air through the sound generating system. Both

While “white noise” is considered a technical term and “breathiness” a non-technical term, breathiness has clearly been used in voice literature for at least 5 decades. White noise has a flat spectrum over the audible frequency range and is perceived as a /sh/ sound. Because breathiness is such a common term in laryngology literature, I will consider white noise and breathiness as synonymous in this article.

Nearly the same can be said for roughness, a consumer-type term that has been used in laryngology literature for decades. It represents the perceived quality of two or more tones, which are not multiples of each other. When two non-harmonic tones interact, the sound waves cancel and multiply with each other in terms of volume, altering our perception of the sound. Voice loses clarity. We often use the term diplophonia when there are two almost distinct tones, but depending on the spectral distance from each other, it may be difficult to perceive whether there are two tones or more and we could actually be hearing a triphonic sound, or more generally a polyphonic sound.

In day-to-day musical terms though, polyphony is perceived as beautiful, as when an orchestra is in tune and more than one note harmonically blends with related notes. For this article though, the terms roughness, diplophonia and polyphony are used as synonyms.
terms have been described as significant components of vocal impairment throughout at least half a century[3-5].

Unwanted breathiness is created when air passes through the vocal cord aperture during intended phonation, without entrainment. The most common source of breathiness occurs when the vocal cords fail to close between oscillations. Continuous air leak through a gap results in non-laminar flow at or beyond the edges of the vocal cords, generating white noise via turbulence; essentially the sound produced intentionally by a whisper. Stiff vocal cords require greater airflow to produce oscillations; consequently, some airflow passing between stiff vocal cords is converted to turbulence. Unwanted breathiness can be thought of as the inefficient conversion of subglottic pressure to sound production by a volume leak (large gap) or a pressure leak (stiffness).

Pitch is determined by tension, mass and length of the vibrating source. Unwanted roughness may be created anytime there is more than one non-harmonic vibratory source; that is to say, two (or more) different pitches generated simultaneously. When two vocal cords are uneven in terms of tension, mass or length, each will tend to vibrate at a different pitch. A single vocal cord may also oscillate with more than one vibratory segment when there is a non-linear density along the vocal cord length (e.g. nodule, polyp, scar, sulcus). Consequently, asynchronous oscillations of each vocal cord or multiple segmental oscillations of one vocal cord may generate multiple non-harmonic pitches when a difference in tension, length or mass exists. With two or more sound sources, competing sound waves that have no simple mathematical relationship cancel and augment each other, resulting in sound with irregular pitch and volume which tends to be displeasing to the ear - roughness. A single cord that is extremely lax, when driven with enough air pressure, may also oscillate in a temporal non-harmonic manner producing more than a single pitch and is often audibly perceived as flutter, a severe roughness.

Both of these irregularities in sound production, breathiness and roughness, are generated along the vibratory margin or edge of the vocal cords with rare exceptions. Only in the unusual case is sound produced by some other structure or portion of the larynx, such as in false cord phonation or aryepiglottic fold phonation and then the quality is substantially different, mostly lower in pitch given the larger mass of these structures, mostly monotonal given the relative inability to alter tension of these structures.
We are hoping to “tune your ear” with this course. Since patients are visiting you for hoarseness or some other impairment or limitation of their voice, your ability to hear the problem they are complaining about puts you way ahead of someone who can only look in the throat, or even another professional who has a camera they can pass to look at the vocal cords. The voice will guide you to the pathology.

We sincerely hope you enjoy this course and take home some sonic tidbits.
Existing voice evaluations

A consensus review of vocal assessments used to document thyroid surgery vocal impairments[12] doubles as a reasonable review of current common voice evaluations and comes to the conclusion that a surgeon should document assessment of the patient’s voice before surgery, yet evidence for voice assessment providing clinical value is only a Grade C. This low level of scientific support for assessing the voice is not because it shouldn’t be assessed, but because current methods of evaluation fall short in several respects.

Various existing voice assessment protocols tend to describe the voice from a given perspective. VHI and V-RQOL describe the sense of degree of impairment of the voice from the patient’s perspective. GRBAS and Cape-V describe the degrees of vocal impairment from the physician or therapist’s perspective, and is often called “expert perceptual rating” if performed by a team of various professionals. Phonotograms, aerodynamic measures and other computer voice measurement protocols are seemingly objective, unbiased and precise measurements and could be termed, the computer’s perspective. A basic audio recording, without any interpretation, is also recommended as a potential clinical practice with a perhaps neutral perspective. The important question is, do any of these assessments orient the examiner to the physical location (where), the temporal location (when), and the type of vibration impairment (what) present? These where, when and what questions should be answered by a complete voice assessment of a vocal impairment.

The patient’s view: Several surveys have been developed to assess the patient’s perception of the degree of vocal impairment. Voice Handicap Index[13] (VHI) and Voice Related Quality Of Life[14] (V-RQOL) surveys are frequently used tools. However, surveys do not direct the examiner towards pathology. They also do not direct the examiner towards any particular location on the vocal cords where pathologic vibration might be present. They do not orient the examiner to areas of the vocal range most impaired nor do they measure any characteristics about the voice. They do not record the voice for later comparison. In summary, they do not orient the examiner to the where, when and what questions of vibration impairment.

The examiner’s view: Various perceptual factors such as loudness, pitch, clarity, roughness along with other terms have been considered for measurement[15]. In 1969, Isshiki[3] utilized roughness and breathiness as two primary descriptive audible impairments in the hoarse human voice, as he tried to correlate the ratio of these components directly with disease processes. His difficulty seemed to be the mixing of these two parameters (and perhaps more) in systems trying to infer directly from audible impairment the disease diagnosis by listening to the voice.

In 1981, Hirano published an overview of a consensus group’s findings from the Society of Logopedics and Phoniatrics in Japan. In Chapter 6, the now frequently used GRBAS scale[4,16] (Grade, Roughness, Breathiness, Asthenia, and Strain) was proposed. The GRBAS scale rates several vocal features using an ordinal 4-point rating scale. G represents severity. The R and B represent roughness and breathiness. A relates to power (willingness or strength and also to fullness of upper harmonics). S relates to hyperfunction. Another auditory perceptual scale, the Consensus Auditory-Perceptual Evaluation of Voice (CAPE-V)[17] utilizes visual analog scaling for rating the parameters of Overall Severity, Strain, Roughness, Breathiness, Pitch, and Loudness.

The attribution of a degree of impairment for these parameters is useful, perhaps over time, for intra-individual comparison. However, in its present form, at least three of the ratings are not utilized to direct the examiner towards pathology. The “G”, representing severity, does not direct the examiner. “A” seems to relate more to the pharyngeal tract, for example the pharynx which, when tuned properly, provides resonance filling in upper harmonics. “A” might also refer to breath support or pulmonary or diaphragmatic function. It is difficult to discern precisely what the authors were listening for with the “S” ratings’ “psycho-acoustic impression of a hyperfunctional state.” Hyperfunction seems to carry different meanings with different authors, ranging from an impression of non-organic psychological vocal alteration, to supraglottic compensation for impaired glottic closure.

Two of these ratings though, roughness and breathiness are necessary components for localizing laryngeal pathology, but not sufficient. The insufficiencies of the GRBAS scale are how roughness and breathiness correlate with two other parameters, pitch and volume. In almost all voice disorders, roughness will change between low and high pitch and roughness will also change between low and high volume. Breathiness has a similar variability over the ranges of pitch and volume. The GRBAS scale does not take this relationship into account.

Roughness and breathiness need to be correlated with two other parameters, pitch and volume.
Hoarseness is all too frequently misattributed to reflux\textsuperscript{[2,6,7]} at least in part because of a lack of diagnostic rigor. The terms laryngopharyngeal reflux (LPR), gastroesophageal reflux (GERD), extraesophageal reflux and silent reflux seemingly imply a causal relationship between the stomach’s secretions and the symptoms of hoarseness. On the scientific publication side, lumping together multiple symptoms (e.g. lump in the throat, pain, hoarseness, throat clearing, dysphagia, cough etc)\textsuperscript{[8-10]} and searching for a potential correlation is fraught with the perils of leniency. On the diagnostic side, despite a Cochrane meta-analysis calling into question the reliability of peer reviewed data on the theory of vocal impairment by gastroesophageal reflux\textsuperscript{[9]}, the concept of reflux induced hoarseness has spread. It has trickled down to the point where reflux laryngitis has become a default diagnosis for otolaryngologists and primary care physicians, where on a quotidian basis, hoarseness is empirically treated with anti-reflux medication\textsuperscript{[11]} without a vocal assessment or all too frequently, without even a laryngeal examination.

With vocal capabilities pattern matching, we maintain diagnostic rigor by searching only for the cause of hoarseness. Other complaints potentially coming from the throat might be related to the larynx or even the vocal cords, however, by focusing on the single complaint of hoarseness, we are much more likely to identify the cause of sound impairment. This focused approach is much stronger than a simple correlation, where simple correlation is often misconstrued to represent cause.

A complete voice assessment should have three parts, rather than only two. There should be a history, a voice evaluation and a visual examination of the larynx. During the third part of the laryngeal examination (endoscopy with stroboscopy) we know that sound is produced almost exclusively by the mucosal edges of the vocal cords. With vocal capabilities findings as a guide, visual attention during endoscopy can be directed to the edges of the vocal cords during an appropriate pitch and volume that optimally elicits inappropriate vocal cord vibration. Then the type of vibratory impairment can be visibly noted and recorded.
INTRODUCTION

Complexity, compensation & capabilities

If each muscle in the larynx can produce only a single action, why is there so much complexity? How do we take all these variables into account? There is complexity in the larynx, first because, as a symmetric, midline organ there is mirrored duplication of essentially all of the components of the larynx. Then there is an additional level of redundancy since more than one muscle can contribute to a single function. A simple rating of roughness and breathiness fails to orient the examiner because of this redundancy built into the larynx.

Starting with symmetry, if one lateral cricoarytenoid muscle is mildly injured, the opposite lateral cricoarytenoid muscle has some capacity to move the vocal process on the healthy side past the midline to meet the incompletely adducted vocal process, resulting in greater closure than expected, perhaps even visually complete closure. There is an injury, but there is compensation, at least for basic vocal tasks. However, when the system is sufficiently stressed, the stronger muscle may not be able to compensate for the weaker muscle and the vocal impairment will become more audible. An individual with this injury will have a different level of audible hoarseness at the end of a long day of phonation because of fatigue and loss of compensation from the overworked healthy side.

A second example of redundancy is that pitch will increase with thyroarytenoid muscle activation by increasing intrinsic vocal cord tension. Pitch will also increase with cricothyroid muscle activation by lengthening the vocal cord. Together, in the fully functioning larynx, these two muscles offer the ability for a larger vocal range than either muscle alone. In the injured larynx, if one muscle is impaired, the other will substitute to some degree. However, the overall range will be reduced. Again, there is an injury, but there is compensation, and the compensation can be complete for comfortable, mid-range vocal tasks.

There are other types of compensation, which increase diagnostic complexity. When a vocal cord marginal swelling impairs vibration by touching and dampening the opposite vocal cord’s vibration, increased airflow (volume) can maintain entrained vibration even after touching, by blowing the dampened cords apart. Essentially the diaphragm muscle compensates for the vibratory impairment. Alternatively the vocal cords can be moved further apart to avoid touching and dampening from the lesion by tensing muscles opposing glottic closure, the posterior cricoarytenoid muscles. This tension trades increased air leak (breathiness) for reduced irregular vibration (roughness).

Examining the sound produced at any single pitch/volume combination provides an incomplete picture of roughness and breathiness because of various forms of compensation. However, the astute examiner can, to some degree, isolate the function of the intrinsic laryngeal muscles and reduce or alter compensation by examining the voice over a range of pitch at high volume, then over a range of pitch at low volume. The overall pattern of impairment for roughness and breathiness will yield a vocal signature for a particular voice disorder. For example, the gap of lateral cricoarytenoid paresis which becomes audible after vocal fatigue can also be heard and visualized by altering pitch and volume, because the stronger healthy side cannot compensate with equal efficacy at low and high pitch nor can it compensate with equal efficacy at low and high volumes.

Generally speaking, unwanted vocal gaps and vocal weakness impair low pitch at soft volume more than high pitch at loud volume. Swellings and scarring tend to impair high pitch at soft volume more than low pitch at high volume. Other disorders may generate clarity in one portion of the vocal range, breathiness in another portion of the vocal range and roughness in a different portion of the vocal range, vice versa or in some other combination. Psychogenic voice disorders fail to consistently follow a pattern.

Degrees of roughness and breathiness do not then correlate with a specific pathology. Neither does the ratio of roughness to breathiness correlate with a specific pathology. Rather roughness and breathiness match the physical manifestation of the disease and with the compensation attempted by the individual with the voice pathology. We can utilize a vocal capabilities pattern matching examination to define where, when and what types of vibration impairment are present. This orientation will focus a subsequent visual examination of the impaired vocal cords. “Where” to look is determined primarily by the laryngeal complaint. Since in this monograph, we are concerned only with laryngeal disorders of sound production (not the related laryngeal functions of swallowing and breathing), all voice disorders will be manifest on the vocal cord’s membranous vibratory margin, where sound is produced. Roughness and breathiness assessment will direct the examiner temporarily “when” to look for impairment, that is, during which pitch and which volume combinations to observe vocal cord vibration. “What” to look for indicates whether a gap or multiple sound sources will be identified.

Existing voice assessment protocols fail because they do not correlate presence of roughness and breathiness with pitch and volume. They also fail in orienting a subsequent visual examination to one or more aspects the “Where”, “When” and “What” questions.
Vocal capabilities pattern matching is the mid-portion of a three-part technique for identifying the cause of hoarseness. Vocal capabilities testing, first described by Robert Bastian, evolved from audibly identifying vocal swellings. A standard battery of vocal capabilities proves to be useful in all voice disorders.

Although many of the parameters in this technique could also be measured with various types of hardware and software under the heading of vocal outcome assessments, the dynamic and interactive nature of eliciting vocal capabilities leads to immediate decision-making, which allows the examiner to probe the voice and identify the etiology for the complaint of hoarseness. Rather than a goal of extremely precise measurement, vocal capabilities pattern matching is used initially for recognition of an overall pattern, which then orients the following visual examination of the larynx.

Ideally in Part I of a patient interaction, the patient offers a history. In Part II vocal capabilities are assessed. Then in Part III of a laryngology exam, vocal cords are examined with endoscope and stroboscope, oriented by the findings from Part I & II.
Vocal capabilities pattern matching is not only a sensitive method for assessing changes in the voice; recording vocal capabilities before and after any intervention that has the potential for altering vocal cord function has significant benefits and few negatives. For example, recording vocal capabilities before surgery that could directly alter the vocal cords (microsurgery) or indirectly (surgery near the nerve supply of the larynx in the brain, skull base, neck or chest) would provide necessary and even sufficient documentation for later comparison. It might not be a preposterous idea to have such a recording even before general anesthesia, if one wished to learn the true incidence of significant vocal injury following intubation. An audio recording of vocal capabilities essentially documents the vocal functional status of the larynx, including the motor nerves, muscles and mucosal covering of the vocal cords and outlines vocal limitations.

On the upside, first, the only method to go back in time is to have already made a recording. Second, optimal legal evidence that no unintentional change has occurred during an intervention or that change had occurred before the intervention is from a recording. Third, physicians who operate near the recurrent and superior laryngeal nerves would have a much better sense of how often the nerves are injured both temporarily and permanently and could offer their patients reasonably accurate estimates during a pre-surgery conference as well as alter their future surgical techniques based on this feedback. Fourth, an audio recording is a far more accurate record for comparison than a physician’s memory or written notes or even a phonetogram without sound. Fifth, recording from a microphone attached to a laptop computer takes little effort, less than 5 minutes of time and costs are minimal.

This contrasts with the most current position paper on thyroid surgery by the American Academy of Otolaryngology – Head and Neck Surgery Foundation. There is not even a strong recommendation to record the voice. The minimum recommendation is that the surgeon subjectively assesses the voice. If neither the physician nor the patient feels there is anything obviously wrong with the voice, then not even a recording is recommended. If a recording is performed, then a 3 - 5 second recording of “ah” or “ee” and a 30 second conversation and reading passage are deemed adequate. None of these vocal tasks documents much of the individual’s vocal capabilities and certainly doesn’t adequately assess the full motor capabilities of the superior and recurrent laryngeal nerves. The reported incidence of nerve injury in thyroid surgery may be low not because there are so few injuries, but rather because the capabilities of the laryngeal nerves are essentially rarely measured pre- and post-operatively.

A suggested method for recording an individual’s vocal capabilities are via a microphone held a set distance in front of the mouth. The following tasks comprise the examination for pitch and volume variation: reading aloud, maximum phonation time, vocal range (lowest and highest pitch), maximum volume, vegetative sound, and vocal swelling tests.
The patient reads aloud a paragraph using a comfortable voice. Using the same passage for every exam provides for easy future comparison. While reading is a mixture of voice and speech, reading aloud provides a rough measure of “comfortable speaking pitch.”

Reading often relaxes the patient and takes the focus away from the examination – many patients start out with a great deal of anxiety during an exam, anticipating foul tasting medicine placed in the nose and throat, worried how big the tube is that goes in the nose and how much it will hurt. These fears are not irrational, as previously examined patients complain of terrible tasting sprays, uncomfortable or even painful endoscopic exams and they may have gagged terribly.

Second, by listening, the approximate average speaking pitch is noted (perhaps clinically by matching the voice with a tone on a piano). It is not necessary to know the precise pitch, as previously examined patients complained of terrible tasting sprays, uncomfortable or even painful endoscopic exams and they may have gagged terribly.

Third, the reading task allows time to hear any speech issues. Problems with the rate of speaking or poor enunciation become audible during this task. Involvement of muscles innervated by other branches of cranial nerve X (e.g. palate) and other cranial nerves (e.g. XII, IX, VII) involved in articulation may be audible.

Fourth, severe hoarseness apparent during this task cues the astute examiner to severe breathiness or roughness present at the comfortable speaking pitch.

Fifth, if the comfortable speaking pitch is elevated above the typical range for the patient’s gender, such as in obligate falsetto, atypical recruitment of the cricothyroid muscle may be deduced.

When the same person performing endoscopy performs these vocal elicitations, the process of differential diagnosis formulation begins during this task and progresses during further vocal capabilities testing. The examiner begins a visual thinking process about where to look for the sound impairment.
**FINDING PITCH**

Comfortable Pitch
Maximum phonation time

Using the /i/ sound, ask the patient to see how long they can say /i/ on one breath, at their comfortable speaking pitch and comfortable volume. MPT (maximum phonation time) is recorded as the number of seconds a single phonation is maintained at a specific pitch. MPT typically increases with higher pitch, as less air is utilized for the shorter oscillation intervals and the lower amplitude of oscillation releases less air. An increase in volume leads to a shorter MPT as more air passes between the cords. One method of standardization is to try and record the MPT at the same pitch and volume as the comfortable speaking pitch determined during the reading task. While not controlling pitch and volume as precisely as a researcher might with computerized testing equipment in a soundproof booth, this test, the maximum phonation time (MPT) at the comfortable speaking pitch, is a rough measure of the degree of vocal cord approximation.

The more closure, the less air is wasted and the longer sound can be maintained. As a rough guide, with an MPT of less than 10 seconds duration at the comfortable speaking pitch, most people will complain of being out of breath with talking. Healthy young people can typically go beyond 20 to 30 seconds on MPT testing. There are many variables that affect this test, including lung capacity as well as vocal strategies used to produce sound, but the more that the pitch and volume are kept constant; the more the test represents vocal cord approximation. This is an especially helpful measurement for one individual over time. For example, after implementation of some treatment to the voice, change in MPT after the intervention is often secondary to the intervention.

Near the end of the maximum phonation time, when there is reduced breath support, vocal impairments will be more noticeable. Essentially, the compensation provided by high subglottic pressure diminishes and impairments such as stiffness and glottic gaps become more audible.
Next, the patient attempts to produce sound at their lowest pitch, at any volume, defining the vocal floor of their voice. Sometimes the person has excellent vocal rapport, capable of matching their voice to notes played on a piano. Some people are not so talented and the examiner may ask them to slide down in pitch and then by ear try to determine the lowest note produced. Moving lower in pitch removes any compensation from the cricothyroid muscle. For example, a vocal paresis will become more apparent as increased air leak at lower pitches.

A similar investigation is performed moving up in pitch until the patient produces the highest note they are capable of. Typically this vocal ceiling is reached where the vocal cords, placed on a stretch, reach the limit of their ability to vibrate given their mass and stiffness, as well as the limit of energy that subglottic airflow can impart. When the uppermost notes have a tight quality, we could term this a muscle-quality vocal ceiling. There are other qualities possible for a vocal ceiling. The individual may reach a note where the voice suddenly cuts out and this can be suggestive of a swelling-quality vocal ceiling caused by a sudden dampening of vibrations when a swelling touches the opposite vocal cord or they might leak air at the highest note suggestive of a gap-quality vocal ceiling caused by a lack of closure.
robust vocalization, not a scream, but a well-supported yell on the word “Hey” assesses the ability of the patient to maintain or recruit additional closure with increased subglottic pressure. The additional energy from increased pressure beneath the vocal cords can cause weak vocal cords to flutter. The task may allow stiff vocal cords to actually produce sound, when quiet sounds were almost impossible. Psychogenic problems often show up on this test when the patient hesitates or exaggerates performing this task, perhaps subconsciously worrying that the voice will create sound in an unexpected way where there previously was no sound.

Vocal effort and vocal quality during high vocal intensity (volume) may be assessed at both low and high pitch. Notation is made whether the volume seems normal, reduced (typical of paralysis) or better than expected (typical of bowing and termed vocal recruitment). Notation is made of volume relative to the pitch. A loud sound that can only be produced at high pitch is suggestive of weakness such as a recurrent laryngeal nerve paresis. Notation is made of quality. A loud phonation that is clear at high pitch and causes flutter at low pitch is suggestive of an anterior branch, recurrent laryngeal nerve paresis. Notation is made of patient effort. A patient with a nonorganic voice issue will typically defer on this task or be surprised when their voice is suddenly normal or there may be facial signs such as la belle indifférence.

Ask the patient to cough, followed by a clearing of their throat. These tasks can be helpful, like yelling, in sorting out weakness of the glottis or psychogenic / nonorganic vocal problems. For instance, if a patient could only whisper up to this point in the exam, but can produce a robust cough, then the vocal cords or some portions of the glottis have the capacity to come together, hold back air and then on release, generate sound. Notation is made whether the patient can or cannot produce sound on this task and whether it is normal, soft, seal-like bark quality or strained.
Perhaps the most detailed task, the examiner reassesses the upper and lower ends of the vocal range at the very softest volume the patient can produce, for comparison with the previously recorded maximum vocal range. Quite often, this requires some coaching. There are a number of disorders that impair soft voicing and despite the patient’s interest in solving their problem, no one likes to “fail” at a test, not even a patient. This is especially pronounced in professional voice users. Even when an individual's chief complaint is that they are missing notes, they utilize significant effort to avoid sounding “bad” on these notes during an exam. Coaching the patient to sing softer and softer and emphasizing the importance of hearing on what tone the vocal cords stop vibrating and when the voice sounds bad, can improve patient compliance with the test. Emphasis is placed on hearing and discovering the hoarse or impaired voice.

Generally, a healthy larynx should be able to produce similar tones at both soft and loud volumes at both the upper and lower ends of their vocal pitch range. When one cannot reach almost the same note softly that one can reach loudly, there is probably a vibratory impairment. The greater the difference between the soft-volume vocal pitch range and the high-volume vocal pitch range, the more significant the vocal cord vibratory problem.

One of the easiest ways to determine the upper limit of the soft vocal range is to have the patient sing the first four words of the nearly universally known song, “Happy Birthday to You.” When singing the words, “happy birthday to you,” between the words “day” and “to” is a melodic interval of a fourth (5 semitones). If no sound comes out on the word “to,” or if there is a significant onset delay to the start of vocal cord vibration on that word, then there is likely some mechanical vibratory limitation of the vocal cords commencing within this interval of a fourth. This test can be repeated at a lower or higher tone and the tone where the voice cuts out more precisely determined. This denotes the soft cutoff point.

Robert Bastian has termed this test for the soft, upper vocal ceiling, the “vocal swelling test”[19] and the exam is very sensitive for vocal cord vibratory margin swellings (nodules and polyps). In general, the point at which there is an onset delay or soft sound cutoff point signifies the tension at which a swelling on one vocal cord touches the other vocal cord and stops cord vibration. It is just like putting your finger lightly on a vibrating guitar string, dampening or stopping the vibration and sound.

It is also possible to learn to hear a central glottic gap with this test. The point at which the patient cannot start the vocal cords vibrating (because all the air leaks out between the cords without entraining them) does not occur at as precise a pitch as when a swelling stops the vibrations. But there will be a general pitch range over which the vocal cords cannot be entrained during low air flow because of the width of the glottic gap.

Occasionally vibratory impairment may be secondary to compression from a mass above the vocal cord. A dilated saccular cyst may dampen the vocal cord’s vibrations at certain pitches if compressed against the true vocal cord. Compression from the saccular cyst may also shorten the effective vibratory length of one vocal cord creating diplophonia.
Complaint

**Hoarseness**

Vocal capabilities pattern testing

- **breathiness**
  - Qualities listened for:
  - vary volume
  - vary pitch

- **roughness**

Vocal signature

- Pitch breaks & air leak at high pitch, low pressure phonation.
- Clear at low pitch, high pressure phonation

Endoscopy Stroboscopy

High pitch, low pressure view of vocal cords reveals pathology:
- swelling touch
- pitch breaks
- 2 segments
- diplophonia
- gaps
- air leak
If we have a moderate sized swelling (~2mm wide and ~1mm tall – e.g. nodule) located in the mid-portion of the medial vibratory margin of the membranous vocal cord, it will tend have the following acoustic effects.

Since the comfortable speaking pitch is typically in the bottom quarter of the vocal range, the vocal cords are relatively lax during speaking. At low pitch, with either loud or soft volume, a small mass will not impair closure of the vocal cords because the cords are short and loose and the swelling can compress into the vocal cords during the closed phase of vibration. There is no air leak. The asymmetric mass difference between the two vocal cords with a small nodule is insufficient to break the phase lock and so diplophonia does not occur during speaking and the voice sounds clear.

For the same reason, maximum phonation time at comfortable speaking pitch is typical, since a small swelling does not promote air leak or diplophonia.

During vocal range testing, we increase pitch initially by tightening the thyroarytenoid muscle, which stiffens the vocal cord. With further pitch elevation, the cricothyroid muscle lengthens the vocal cord, indirectly adding additional stiffness while raising pitch. The oscillatory amplitude is reduced with increasing tension. The swelling begins to stand proud of the margin of the vocal cord and it no longer compresses into the deeper layers of the vocal cord. At some pitch, the swelling begins touching the opposite vocal cord during every closing phase of the cycle. Air begins to leak from anterior and posterior to the swelling. If airflow is high enough, i.e. high-volume, the phase lock will remain. However, if volume is low and there is just enough air passing through the vocal cords to maintain entrainment, as the swelling begins to touch the opposite vocal cord with sufficient pressure, it may stop entrainment, the pure tone will cease and air will leak. When occurring at phonatory onset, this can be termed an onset delay. A delay, because the individual being examined usually quickly makes a laryngeal adjustment to break the loss of sound production, either increasing the airflow or partially opening the posterior commissure which pulls the swelling away from the opposite vocal cord and allows entrainment to resume.

Secondly, at some point the compression of the swelling against the opposite vocal cord will create an acoustic node and separate the anterior and posterior aspects of each vocal cord into two separate sound sources. As our example swelling is in the exact central portion, the length of the anterior and posterior segments are exactly 1/2 of the original length and the vocal pitch would suddenly double, jumping about one octave, with two short segments generating an identical pitch. We could call this jump a pitch break. If the swelling were not in the exact mid-portion, the length of the anterior segment would be different than the length of the posterior segment and after the pitch break, two separate tones or diplophonia occurs.

We can rerun the scenario with a larger mass. The findings of onset delay and diplophonia will now tend to occur at a lower pitch. If the mass is large enough, such as in the case of a unilateral smoker’s polyp, the phase lock may be broken even at very low pitches and diplophonia will occur even in the low range.

Many individuals will have developed compensation to avoid a rough voice, especially if the mass has enlarged slowly. The typical compensation is to open the posterior commissure, trading increased breathiness for less roughness. Breathiness will occur initially in the upper vocal range, then later the mid-vocal range, as a swelling enlarges. The farther apart the vocal cords are held, the less likely they are to be entrained at higher pitches with low subglottic pressure, since tension increases stiffness of the vocal cord. Increased tension requires higher subglottic pressure to entrain the stiffer cords.

These acoustic changes are not intrinsically dependent upon the composition of the mass or swelling. A benign nodule of the same density and mass as a carcinoma, located in the same position, will have the exact same acoustic effects. Swellings that differ in density or mass may alter the acoustic impairment.
somewhat. A polyp and a nodule of the same size will impair vibration at slightly different pitches because of differing densities, but the overall vocal capabilities pattern will be essentially the same. A hemorrhagic polyp may enlarge during phonatory use if it fills with blood during oscillation, and so a vocal capabilities pattern may be different in degree after extensive vocal use than after vocal rest.

Consequently, vocal capabilities pattern matching which has these characteristics (diplopohnia at high, soft sound production, onset delays during swelling testing in the upper range, pitch breaks with a sudden jump upward) will direct the endoscopic examiner to look at the vocal cord’s medial vibratory margin for a mass. The smaller the mass, the more this test clues the examiner where to look. A very small vocal nodule on the inferior vibratory lip will be essentially invisible to an examiner using a fiber-optic endoscope, looking from far away (tip of scope above the epiglottis), without a stroboscope.

The same swelling will be visualized by a second examiner whose ears are tuned to vocal capabilities pattern matching, by topically anesthetizing the vocal cords if necessary, moving the endoscope very close to the vocal cords (perhaps 1 mm), utilizing a stroboscope and examining and recording the vocal cords while vibrating at high pitch and low airflow, when the small swelling is most likely to be protuberant and interrupting the vibrations of the vocal cords. Then perhaps reviewing the video frame by frame, the small, nearly hidden swelling is identified as the source of harmonic sound impairment.

Even if you assume that the first examiner hasn’t missed anything important (presupposing that a benign nodule is only important to a world-class singer - possibly a false assumption), the first examiner will be predisposed to an erroneous diagnosis, perhaps ordering inappropriate diagnostic testing, perhaps pursuing pH probe studies, esophageal manometry, prolonged prescribing of anti-reflux medication with the potential for side effects, consultations with gastroenterologists, endoscopy of the G.I. tract, fundoplication… consuming the individuals time and money, creating discomfort and wasting a healthcare system’s limited resources.

The second examiner notes the small swelling, reviews a video recording of the swelling, explains the mechanics of vocal impairment to the patient and decides on an appropriate plan of action. Plans might range from an informative discussion where the patient is willing to live with the vocal impairment, to appropriate voice therapy, to surgical excision of the lesion. Although of differing expense, all three of these treatments are appropriate and cost effective and the patient has the data and information to make a personal yet educated decision.
Let’s consider a left, distal recurrent laryngeal nerve injury - a different, yet still common vocal impairment. Various injuries to the nerve supply of the larynx weaken some of the muscles and in this case the nerve injury involves the anterior branch of the recurrent laryngeal nerve. The thyroarytenoid muscle and the lateral cricoarytenoid muscle on the left side have only partial innervation. We might hear the following acoustic effects. Beginning with comfortable speaking pitch, we notice that the patient is speaking at a higher pitch than typical. Whenever the vocal cords cannot approximate tightly from thyroarytenoid muscle lack of tension or from lack of rotation by the lateral cricoarytenoid muscle, the cricothyroid muscle tends to tighten, adding compensatory tension to the vocal cords. Adding tension increases the comfortable speaking pitch. If the unilateral paresis is significant enough, he may be speaking in falsetto full-time.

The voice quality is soft and the maximum phonation time is less than 10 seconds when producing sound at his comfortable speaking pitch. The softness in the voice is secondary to air leak. When the vocal cords cannot close tightly, air leaks through an incomplete closure. This may occur centrally between the membranous vocal cords and posteriorly between the vocal processes. Even when the lateral cricoarytenoid muscle remains functional, air leak occurs centrally through the paretic vocal cord, bowed from a lack of tension in the thyroarytenoid muscle as well as from atrophy and lack of mass within the vocal cord. The shortened maximum phonation time is secondary to air leak. A larger quantity of air is needed to produce vocal cord entrainment, so there is air wasting.

The lowest pitch that can be achieved is only produced softly and at a higher pitch than typical. By listening to a lower pitch, the examiner is removing the compensation provided by the cricothroid muscle and the weak vocal cord will rest in a more lateral position as well as a more concave configuration, allowing more air leak. The highest pitch that can be produced would typically be normal or slightly less than normal and should be fairly clear. The weaker vocal cord still receives tension from the uninjured cricothyroid muscle but lacks intrinsic tension. The patient might be able to yell with moderate volume at a high pitch, but, at a low pitch, the yell will be either weak or there will be obvious flutter from the paretic vocal cord’s intrinsic lack of tension.

In the case of complete paralysis, the unilateral gap will be obvious during endoscopy, however for the astute examiner, even a mild paresis of the anterior branch of the recurrent laryngeal nerve will also be visible. A stroboscopic examination of the vocal cords during attempted adduction at low pitch and soft volume will magnify the impairment. If necessary, even with supraglottic compensation (false cord squeeze), topical anesthesia of the larynx will allow placement of the endoscope between the false vocal cords during phonation and any gap between the membranous cords may be visualized. The endoscope may be angled beneath the arytenoids to visualize the asymmetric angles of closure between the vocal processes in a unilateral lateral cricoarytenoid muscle weakness, one vocal process remaining lateral while the normal side’s vocal process hyperextends past midline trying to reach the weakened vocal cord.
A man notes end-of-day hoarseness, which over time progresses to hoarseness earlier and earlier in the day. An otolaryngologist treats him with a proton pump inhibitor and subsequently excises leukoplakia from the left vocal cord, which is benign appearing on histopathologic exam. His voice becomes worse although his vocal cord looks better. His otolaryngologist feels this ongoing hoarseness may be something he just needs to tolerate. Ultimately though, after assessing his vocal capabilities and endoscopy we find that he has a malignancy within the false vocal cord which is compressing the true vocal cord during phonation.

In this case, vocal capabilities pattern matching yields the following. His voice is gravelly during reading, gravelly being another descriptive term for roughness. On maximum phonation time testing the duration is 15 seconds and the quality is again rough. While maximum phonation time testing can be utilized to determine approximately how much air is converted to sound, the degree of harmonic sound production can also be noted. In this case, roughness suggests that there is an asymmetry and likely two sound sources even at his comfortable speaking pitch. During phonation, the unilateral false vocal cord mass is compressing the ipsilateral vocal cord, but not the opposite cord. This compression from the false cord tumor on the superior surface, near the anterior part of the left cord, both tensions the ipsilateral vocal cord and shortens the effective vibrating length on that side so that the left vocal cord tends to vibrate at a higher pitch than the right. We may attribute his moderate phonation time to partial loss of vibratory energy secondary to a damping effect from the false vocal cord mass pressing on the true vocal cord. This pressure also stiffens the left vocal cord.

When we test his upper vocal range, going higher in pitch, we hear more air leak and less roughness. As his pitch rises, the cricothyroid muscle stretching the vocal cords, supraglottic squeeze pushes the mass more firmly onto the ipsilateral vocal cord, such that the left true cord stops vibrating entirely, essentially becoming extremely stiff from compression. Even though there is a complete closed phase, additional airflow is required to keep the remaining vocal cord vibrating against the compressed left cord and we hear breathiness.

Vocal capabilities pattern matching directed our endoscopic exam to take place at both low and high pitch and seek an explanation for the roughness in the low range and the breathiness in the high range. The compression from the false cord mass though becomes the obvious source of his vocal impairment because it accounts for both these audible findings. Ultimately an excisional biopsy of the false vocal cord reveals the carcinoma enlarging beneath the mucosa. Similar findings may occur with supraglottic masses such as a dilated saccular cyst or a laryngocoele. This compression may be present in only a portion of the vocal range.

It is easy enough to focus one's selective attention on the physical surface characteristics of the vocal cord and miss an even larger mass of the false vocal cord, particularly if the false cord's surface is smooth. In this case, the leukoplakia attracted the examiner's visual attention so much that even during surgery, the false vocal cord was moved out of sight by the operative laryngoscope. Often only in hindsight is such a mass so obvious, unless one listens to where the voice is impaired and then visualizes the vocal cord's vibrations during that impairment. Even if the leukoplakia was initially impairing vibration, failure to improve after excision would warrant a reevaluation of vocal cord vibration to determine the etiology of the ongoing impairment. The most obvious visual finding may distract the examiner, when not correlating hoarseness with vocal cord vibration impairment.
Learning

Three examples are not enough to cover all the audible patterns potentially created by laryngeal disorders which cause hoarseness and by definition impair harmonic voice production. However, whatever pattern one hears, the vocal pattern should be explainable by the subsequent visual examination.

In fact, this technique is self-teaching, in that when a new vocal impairment pattern is heard, the examiner who views the vocal cords endoscopically using the same vocal maneuvers that elicit an impaired voice on vocal capabilities testing, will often discover the reason for the vibratory impairment, even if it is new to the examiner. The examiner who records audio of vocal capabilities pattern matching will also have feedback to discover and learn when the phonatory system is altered during surgery on the vocal cords or in the vicinity of the motor nerve supply of the larynx, whether intentionally or unintentionally.

Conclusion

Vocal impairments can be described in terms of roughness and breathiness, the “R” and “B” of the GRBAS system. Roughness is typically diplophonia, although other quantities of multiple simultaneous pitches can be produced, all creating the perceived quality of roughness. Breathiness is unwanted air leak or air escape between vocal cords that do not completely approximate or are stiff. We can be more precise than simple grading of the amount of roughness and breathiness. A more accurate descriptive method is noting the onset of roughness and/or breathiness as present at high pitch, low pitch or at both. We can be even more precise and note the specific pitch at which diplophonia begins to be produced or breathiness become significantly noticeable, and then whether or not this condition is present from this onset pitch upward or this onset pitch downward. The most accurate record is to have dated audio and video recordings maintained for future review or comparison.

Utilizing the following parameters for vocal capabilities pattern matching, comfortable speaking pitch, maximum phonation time at comfortable speaking pitch, vocal range (lowest pitch, highest pitch), loudness capability, vegetative sound capability and vocal swelling test we can then define or describe the vocal signature of each patient with a complaint of hoarseness. This vocal signature orients the examiner to the where (vocal cord margins), when (pitch and volume) and what to observe for (gap or diplophonia) during recording of laryngoscopy and stroboscopy.

If each physician were to record the vocal capabilities of every patient before and after interventions to the vocal cords, and before and after interventions in the region of the recurrent laryngeal nerve, we would learn more about vocal injuries. We would more precisely learn when we are successful in altering the voice, since harmonic sound production is a successful outcome, not vocal cord appearance. Vocal capabilities pattern matching essentially tests the status of the laryngeal muscles, the status of the closure of the margins of the vocal cords, the flexibility of the vocal cord mucosa, as well as the status of the symmetry of the vocal cords.
Bibliography


Further reading

Thomas, James P., "Why is there a frog in my throat?" 2012