Assessment of Voice and Respiratory Function

2

Philippe H. Dejonckere

Core Messages

- Voice is multidimensional.
- Audio recording is the most important basic requisite for voice quality assessment.
- Once a high-quality complete recording has been attained, it can be stored and remains available.
- Existing research does not support the complete substitution of instrumental measures for auditory perceptual assessment.
- To be valuable, however, perceptual assessment should follow a standard procedure, as does the voice recording. A currently used scale for making perceptual judgments is the GRBAS scale.
- Videolaryngostroboscopy is the main clinical tool for diagnosing the etiology of voice disorders, but it can also be used to assess the quality of vocal fold vibration and thus evaluate the effectiveness of a treatment.
- The simplest aerodynamic parameter of voicing is the maximum phonation time (MPT), in seconds. It consists of the prolongation of an /a:/ for as long as possible after maximal inspiration and at a spontaneous, comfortable pitch and loudness.A reduction of possible bias (e.g., supportive respiratory capabilities compensating for poor membranous vocal fold closure) is possible by computing the ratio or quotient : Averaged phonation airflow or PQ = VC (ml)/ MPT (s).
- Accurate estimation of subglottal pressure can be achieved by measuring the intraoral air pressure produced during the repeated pronunciation of /pVp/ syllables (i.e., a vowel between two plosive consonants).
- Among Voice Range Profile parameters, the highest and lowest frequencies and the softest intensity (decibels, or dBA, at 30 cm) seem most sensitive for changes in voice quality.
- Although subjective by definition, self-evaluation is of great importance in clinical practice. Careful quantification is needed for self-evaluation to be compared and correlated with the objective assessment provided by the voice, an important adjuvant technique, laboratory. The Voice Handicap Index is a largely diffused, validated protocol.
- Electromyography (EMG), an important adjuvant technique, is an electrophysiological investigation of neuromuscular function.

2.1 Introduction

The voice laboratory is considered an essential tool for the assessment and treatment evaluation of voice patients and for clinical research on voice disorders. Several specific questions may be answered from the information obtained in the voice laboratory.

1. Is a given voice or voice function measurement considered normal (within normal limits) or pathological?
2. If the voice or voice function is considered pathological, how severe is the alteration?

P. H. Dejonckere
The Institute of Phoniatrics and Logopedics,
University Medical Center Utrecht, AZU F.02.504,
PO Box 85500, Nl-3508 GA Utrecht, The Netherlands
e-mail: Ph.deJonckere@umcutrecht.nl

M. Remacle, H. E. Eckel (eds.), *Surgery of Larynx and Trachea*,
DOI: 10.1007/978-3-540-79136-2_2, © Springer-Verlag Berlin Heidelberg 2010

3. Which aspects or mechanisms of voice production are involved with the voice disorder? How does the primary (medical) etiology or lesion explain the components of voice production that are perceived or analyzed as deviant (e.g., by limiting vocal fold closure or by eliciting irregular vibrations related to vocal fold asymmetry)? How do they account for the patient's complaints (e.g., voice fatigue or compensation mechanisms)?
4. What is the result of a comparison of voice production two or more times (e.g., before and after therapy), in two or several situations or voicing conditions (spontaneously vs. louder, when doing an Isshiki maneuver, or when applying a defined therapeutic technique)? Have the changes returned the voice to normal function as indicated by voice measurement [1]?

2.2 Prerequisite: Recording a Voice Sample

Audio recording is the most important basic requisite for voice quality assessment. Once a high-quality recording has been performed, it can be stored and remains available—as a document—for performing additional investigations at a later time (e.g., blind perceptual evaluation by a panel or sophisticated acoustical analyses) [2]. A sampling frequency of at least 20,000 Hz is recommended. Ideally, the recordings are made in a sound-treated room, although a quiet room with ambient noise permanently < 45 dB is acceptable. The mouth-to-microphone distance needs to be held constant at 10 cm. A (miniature) head-mounted microphone offers a clear advantage. Off-axis positioning (45°–90° from the mouth axis) reduces aerodynamic noise from the mouth during speech [3, 4].

In regard to voice/speech material, examples of protocol for standard recording are as follows.

- /a:/ at (spontaneous) comfortable pitch/loudness, recorded three times to evaluate variability of quality [5]
- /a:/ slightly louder to evaluate the possible change in quality (plasticity) and the slope of the regression line frequency/sound pressure level [6, 7]
- A single sentence or a short standard passage

Phonetic selection can be useful, such as a short sentence with constant voicing (no voiceless sounds and spoken without interruption) and no fricatives. Such a sentence (e.g., "We mow our lawn all year") can be analyzed by a computer program for sustained vowels; and because it contains no articulation noise, there is no biasing of harmonics-to-noise computations. Computation of the percent voiceless (normal in this case is 100%) is useful for neurological voices or spasmodic dysphonia [9]. Furthermore, it allows easy determination of the mean habitual fundamental speaking frequency.

Another example of a criterion for phonetic selection is a multiplication of voice onsets, as they are critical in disturbed voices [10]. Such criteria are not language-linked.

A standard reading passage should also be recorded whenever possible. Two classic, often used reading passages for English-speaking persons are "The Rainbow Passage" (a phonetically selected passage including all the speech sounds of English) and "Marvin Williams" (an all-voiced passage) [3].

2.2.1 Perception

Existing research does not support the complete substitution of instrumental measures for auditory perceptual assessment. To be valuable, however, perceptual assessment must follow a standard procedure, as does voice recording. A currently used scale for making perceptual judgments is the GRBAS scale, which rates grade, roughness, breathiness, asthenicity, and strain on a scale of 0–3 [11]. The rating is made by assessing current conversational speech or when reading a passage. The severity of hoarseness is quantified under the parameter "grade" (G), which relates to the overall voice quality, integrating all deviant components. There are two main components of hoarseness, as shown by principal component analysis [12].

1. Breathiness (B): an auditive impression of turbulent air leakage through an insufficient glottic closure, including short aphonic moments (unvoiced segments)
2. Roughness or harshness (R): an impression of irregular glottic pulses, abnormal fluctuations in fundamental frequency, and separately perceived acoustic impulses (as in vocal fry), including diplophonia and register breaks. When present, diplophonia can also be noted as "d."

These parameters have shown sufficient reliability (inter- and intrarater reproducibility) [13, 14]. A reliability analysis provided further evidence to support the GRBAS scale as a simple, reliable measure for clinical use [15]. The behavioral parameters asthenicity (A) and strain (S) appear to be less reliable. The remaining simplified scale, GRB, then becomes similar to the RBH scale used in German-speaking countries [16].

For reporting purposes, a four point grading scale is convenient (0 = normal or absence of deviance; 1 = slight deviance; 2 = moderate deviance; 3 = severe deviance). However, it is also possible to score on a visual analogue scale (VAS) of 10 cm, possibly with anchoring points [14, 17].

It is proposed that the term "dysphonia" be used for any kind of perceived voice pathology. The deviation may concern pitch or loudness as well as timbre or rhythmic and prosodic features. "Hoarseness" is limited to deviant voice "quality" (or timbre) and excludes pitch, loudness, and rhythm factors. A limited number of voice pathology categories—such as those related to mutation or transsexuality—are specifically concerned with pitch and register. Rhinophonia is a specific abnormality of resonance and if present needs to be reported separately. Tremor is a characteristic temporal feature and when present must also be reported separately. A special protocol is required for substitution voices [19, 77].

Perceptual evaluation—if averaged among several blinded raters—is very well suited to demonstrate treatment efficacy in voice pathology [21].

2.2.2 *Vocal Fold Imaging*

2.2.2.1 Videolaryngostroboscopy

Videolaryngostroboscopy is the main clinical tool for diagnosing the etiology of voice disorders, but it can also be used to assess the quality of vocal fold vibration and thus evaluate the effectiveness of a treatment. Stroboscopy involves a video-perceptual series of judgments and ratings (e.g., glottic closure, regularity, symmetry, mucosal wave). The pertinence of stroboscopic parameters is based on a combination of reliability (inter- and intraobserver reproducibility), no redundance (from the factor analysis), and clinical sense (relation to physiological concepts) [23].

The basic parameters are the following:

1. Glottal closure. It is recommended that the type of insufficient closure also be recorded and categorized.
   - Longitudinal. It is important to consider that a slight dorsal insufficiency—even reaching into the membranous portion of the glottis—occurs in about 60% of middle-aged healthy women during normal voice effort. Fifty percent of the women close the glottis completely during loud voice.
   - Ventral.
   - Irregular.
   - Oval. It is over the whole length of the glottis but with a dorsal closure.
   - Hour-glass shaped.

Rating glottal closure has been found highly reliable [24, 25]. Objective quantitative measurements are also possible [26]

2. Regularity: quantitative rating of the degree of irregular slow motion, as perceived with stroboscopy [27].
3. Mucosal wave: quantitative rating of the quality of the mucosal wave, accounting for the physiology of the layered structure of the vocal folds [11].
4. Symmetry: quantitative rating of the "mirror" motion of both vocal folds. Usually asymmetry is caused by the limited vibratory quality of a lesion (e.g., diffuse scar, localized cyst, leukoplakia) [28].

For each stroboscopic parameter, a four-point grading scale can be used (0 = no deviance; … 3 = severe deviance), but a VAS may also be useful [23, 28]. Videostroboscopy can be documented on hard copy and thus be archived. Rating *a posteriori* is possible.

It is classically recommended to observe and record videostroboscopic pictures under various voicing conditions. For example, the degree of glottal closure usually increases with increased loudness [24, 25]. However, this basic rating concerns a comfortable pitch and loudness. Laryngostroboscopic ratings and measurements have been found relevant for documenting therapeutic effects [20–22, 29].

2.2.2.2 Digital High-Speed Pictures

With modern technology, it has become possible to capture and store digital vocal fold images at a rate of 2000 (and more) per second with sufficient definition (several hundred pixels) and to display the image sequence at a rate of, for example, 20/s immediately after capture. This procedure does not seem to be appropriate for routine use in the diagnosis of voice problems as a long review time is needed for a short sequence without simultaneous sound. A specific indication for digital high-speed cinematography is to analyze and understand the vibratory characteristics in aperiodic voices, during voice onsets or accidents (breaks), or in case of diplo- or triplophonia [30, 31].

2.2.2.3 High-Speed Single-Line Scanning (Video-Kymography)

High-speed single-line scanning (video-kymography) is an imaging technique for investigating vocal fold vibration, especially when the vibration is irregular and when the focus is on accidents or short events in this vibration, making conventional stroboscopy unsuitable. A modified video-camera selects a single horizontal line from the whole image and monitors it at high speed (8000/s). The displayed image shows successive high-speed line images below each other, thereby demonstrating the vibration of the selected ventrodorsal level of the vocal folds over time. An important practical advantage is that the display is in real time. This type of imaging provides relevant, timely information e.g., for comparing the vibration amplitude of both folds or for understanding diplophonia [32, 33]. It also clearly demonstrates the mucosal wave phenomenon and its absence or asymmetry.

Single-line scanning can also be performed on a high-speed video recording (Fig. 2.1). If several lines are displayed, phase shifts between different ventrodorsal segments of one or both vocal folds—as

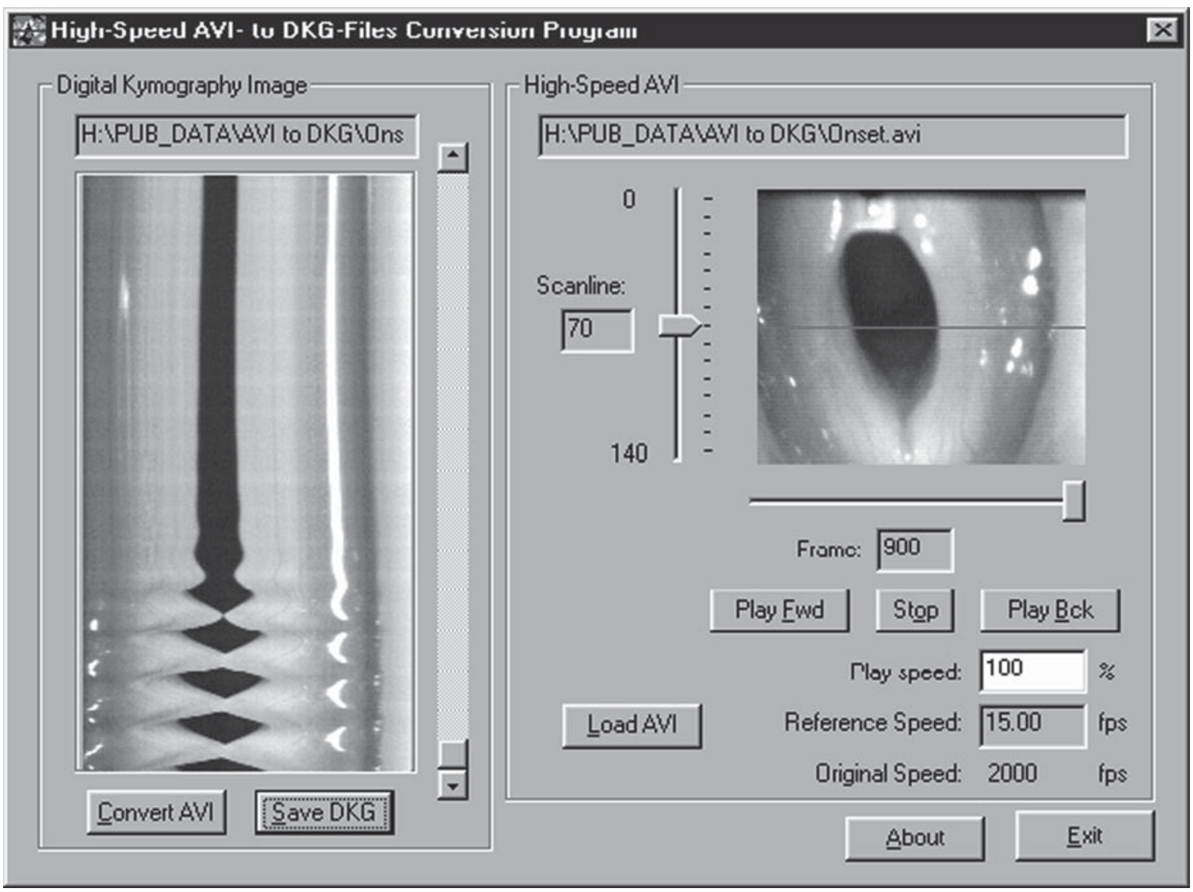

Fig. 2.1. Single-line scan (video-kymography) obtained from a high-speed video recording in a normal subject (Kay System; Kaypentax, Lincoln Park, NJ, USA). *Right* Vibrating vocal folds and the *single line* that was selected. *Left* Oscillation pattern at that specific level during a voice onset

frequently occurs in case of a vocal fold cyst—can be demonstrated [34].

2.2.3 Aerodynamics

Aerodynamic analysis of voice production includes measurement of airflow and air pressure, and their relation during phonation. Using appropriate instrumentation, a number of derived measurements can provide information regarding vocal efficiency, although for certain measurements only a stopwatch is needed.

2.2.3.1 Phonation Airflow

The simplest aerodynamic parameter of voicing is the MPT (in seconds). It consists of the prolongation of an /a:/ for as long as possible after maximum inspiration and at a spontaneous, comfortable pitch and loudness. It is one of the most widely used clinical measures in voice assessment worldwide [35]. A prior demonstration is necessary, and three trials are required, the longest being selected for comparison to the norm [36]. As it concerns an "extreme" performance, it has been shown to be extremely sensitive to learning and fatigue effects. Furthermore, in good voices the duration of "apnea" can become the limiting factor, rather than the available air. Children show significant lower MPT values as their lung volume is smaller [37]. A reduction of possible bias (e.g., supportive respiratory capabilities compensating for poor membranous vocal fold closure) is possible by computing the following ratio: (Phonation Quotient).

Averaged phonation airflow or PQ = VC (ml)/MPT (s)

Vital capacity (VC) is defined as "the volume change at the mouth between the position of full inspiration and complete expiration." It can be measured in a reliable way using a hand-held spirometer [38]. In normal subjects, the VC depends on anthropometric factors and is quite strongly correlated, for example, with height [39]. It is also sensitive to lung disease. As the VC is not directly related to voice quality, it is meaningful to take it into account, especially if a child is being investigated.

The mean airflow rate can also be measured using pneumotachography. This technique directly measures the mean airflow rate (ml/s) for sustained phonation over a comfortable duration, usually 2–3 s, at the habitual pitch and intensity level and following habitual inspiration. Pathophysiological backgrounds and normative values have been reported [11, 35, 40–43].

The variation of averaged phonation airflow varies considerably among normal subjects, and there is a large overlapping range of values in normal and dysphonic subjects, which limits its value for diagnostic purposes [44]. Nevertheless, when comparing glottal function before and after surgical intervention or nonsurgical voice training techniques, airflow measurement may be useful for monitoring therapeutic effects [45] (e.g., in the case of paralytic dysphonia [46–48] or when microlaryngeal phonosurgery is performed) [42]. The method is especially useful for demonstrating changes in a single test subject over time. For comparisons (pretreatment/posttreatment), it is recommended that the same technique (PQ or mean airflow rate measured by pneumotachography) be used for each measurement.

Flow glottography (FLOG) consists of inverse filtering of the oral airflow waveform. The basic tool is a high-frequency pressure transducer incorporated into an airtight Rothenberg mask [49]. The inverse filtering procedure removes the resonant effects of the vocal tract and produces an estimate of the waveform produced at the vocal folds. The special advantage of this technique is that it differentiates, and after calibration quantifies, leakage airflow (the DC component of the air flow) and pulsated airflow (AC component). Leakage airflow is an important concept: It assumes that there is an opening somewhere along the total length of the vocal folds through which air escapes. Calibration is critical for reliable measurements. FLOG can also be used to analyze voice onset.

2.2.3.2 Subglottal Air Pressure

Measurements of subglottal air pressure using esophageal balloons or pressure transducers, transglottal catheters, or tracheal puncture are semi-invasive or invasive and are limited to research situations. Subglottal pressure can be accurately estimated by measuring the intraoral air pressure produced during the repeated pronunciation of /pVp/ syllables (i.e., a vowel between two plosive consonants). A thin catheter is introduced into the mouth through the labial commissure, is sealed by the lips, and is not occluded by the tongue. If there is no closure of the vocal folds, the intraoral air pressure should be similar to the pressure elsewhere in the respiratory tract. During production of a voiceless consonant, the vocal folds are abducted and should not impose any significant obstruction to airflow from the lungs. Thus, the pressure behind the lips is the same everywhere and reflects the pressure available to drive the vocal folds if they were to vibrate [50, 51]. This technique also allows measurement of the phonation threshold pressure (PTP), the minimum pressure required to initiate phonation [52]. Pressure is usually reported in pascal units: 1 Pa = 1 N/m^2; and 1 kPa = 10 cm H_2O.

2.2.3.3 Efficiency of Phonation

Together with airflow and vocal intensity, subglottal air pressure can be used to estimate the efficiency of phonation. Obviously, reduced efficiency is expected to induce voice fatigue. Vocal efficiency—defined as the ratio of acoustical power to aerodynamic power—can be estimated by dividing the acoustical intensity of the utterance by the product of the air pressure and the airflow used to produce the utterance [54].

2.2.3.4 Flow Versus Volume Loops

Spirometry is important for investigating cases in which voice problems are associated with laryngeal obstruction, such as bilateral abduction paralysis, stenosis caused by extensive webs and scars, cancer, or even severe Reinke's edema. The flow–volume loop is generated when measurements of maximum forced expiration and maximum forced inspiration are plotted on a graph, with the flow rate on the ordinate and lung volume on the abscissa (Fig. 2.2). Lack of effort is easy to detect because there is reduced flow at the beginning of the expiratory curve, and the inspiratory curve is abnormal (Fig. 2.2b). Obstructive lesions of the larynx are easily detected and quantified because the morphology of the flow–volume loop is altered. Variable extrathoracic obstruction (as with bilateral vocal fold paralysis) manifests as a decrease in inspiratory flow only (Fig. 2.2c), whereas a fixed obstruction of the upper airway (e.g., extensive laryngeal cancer) is demonstrated and quantified by a symmetrical reduction of inspiratory and expiratory flow (Fig. 2.2d) [55–57].

2.2.4 Acoustics

Acoustical measures provide, in an objective and noninvasive way, a great deal of information about vocal function. Increasingly, these measures have become available at affordable cost and appear to have succeeded well in monitoring changes in voice quality across time (e.g., before and after treatment). Acoustical measures reflect the status of vocal function and do not relate specifically to certain voice disorders because basic biomechanical changes resulting in acoustical differences can be induced by various lesions and dysfunctions.

2.2.4.1 Visible Speech

Acoustical analysis can be used to make the voice and speech visible (e.g., in spectrograms) [50]. This visual representation may be a considerable aid to the perception and description of voice characteristics. Spectrograms are also useful for comparing normal phonation with phonation characterized by excessive noise. Commercially available software packages provide synchronized displays of the microphone signal and the spectrogram, showing the frequency distribution of acoustical energy over time. A choice can be made between narrowband filtering (frequency resolution, mainly demonstrating fundamental frequency, harmonics, interharmonic and high-frequency noise, subharmonics) and broadband filtering (temporal resolution, mainly demonstrating periodicity but also formant location). Voice characteristics such as the sound pressure level (SPL), fundamental frequency, and formant central frequency can also be displayed over time for analysis of the singing voice. Visualizing fast Fourier transform (FFT) graphics (power spectrum) and long-time average spectra (LTAS) is usually possible (Fig. 2.3). When visible speech is provided simultaneously with voice sound, the interrater consistency of the perceptual quality evaluation significantly increases [8, 9]. Martens WMAF, Versnel H, Dejonckere PH (2007) The effect of visible speech on the parceptual rating of pathological voices. Arch Otolaryngol Head Neck Surg 133 : 178–185.

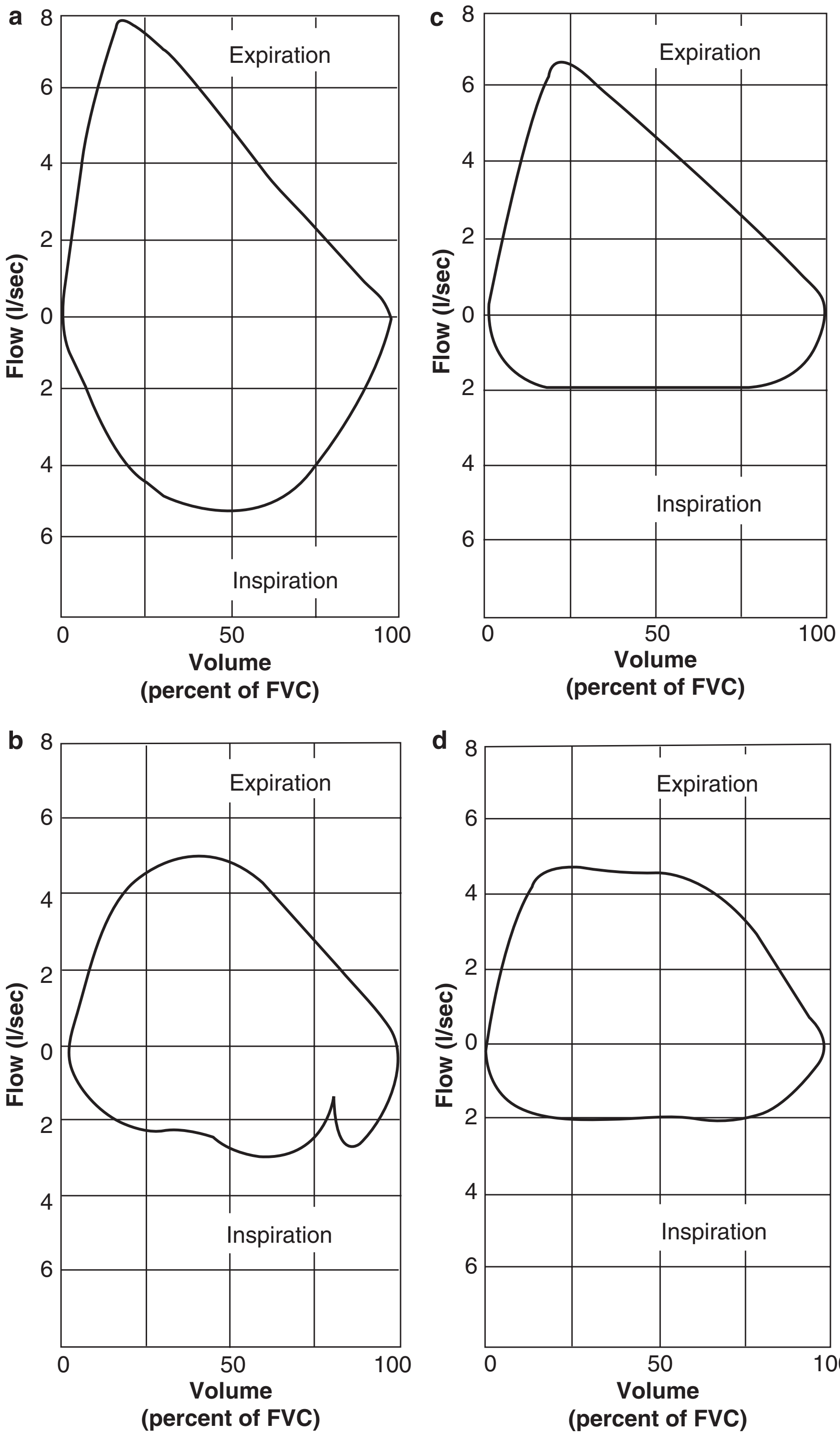

Fig. 2.2. Flow-volume curves. Measurements of a maximum forced expiration and a maximum forced inspiration are plotted on a graph with flow rates on the ordinate and lung volume on the abscissa. (**a**) During normal respiration the expiratory flow curve decays linearly. (**b**) When effort is poor, the initial slope of part of the expiratory curve is decreased, and the inspiratory curve is also abnormal. (**c**) A variable extrathoracic obstruction decreases only the inspiratory flow rate. (**d**) In case of fixed obstruction of the upper airway, inspiratory and expiratory flow rates are both reduced

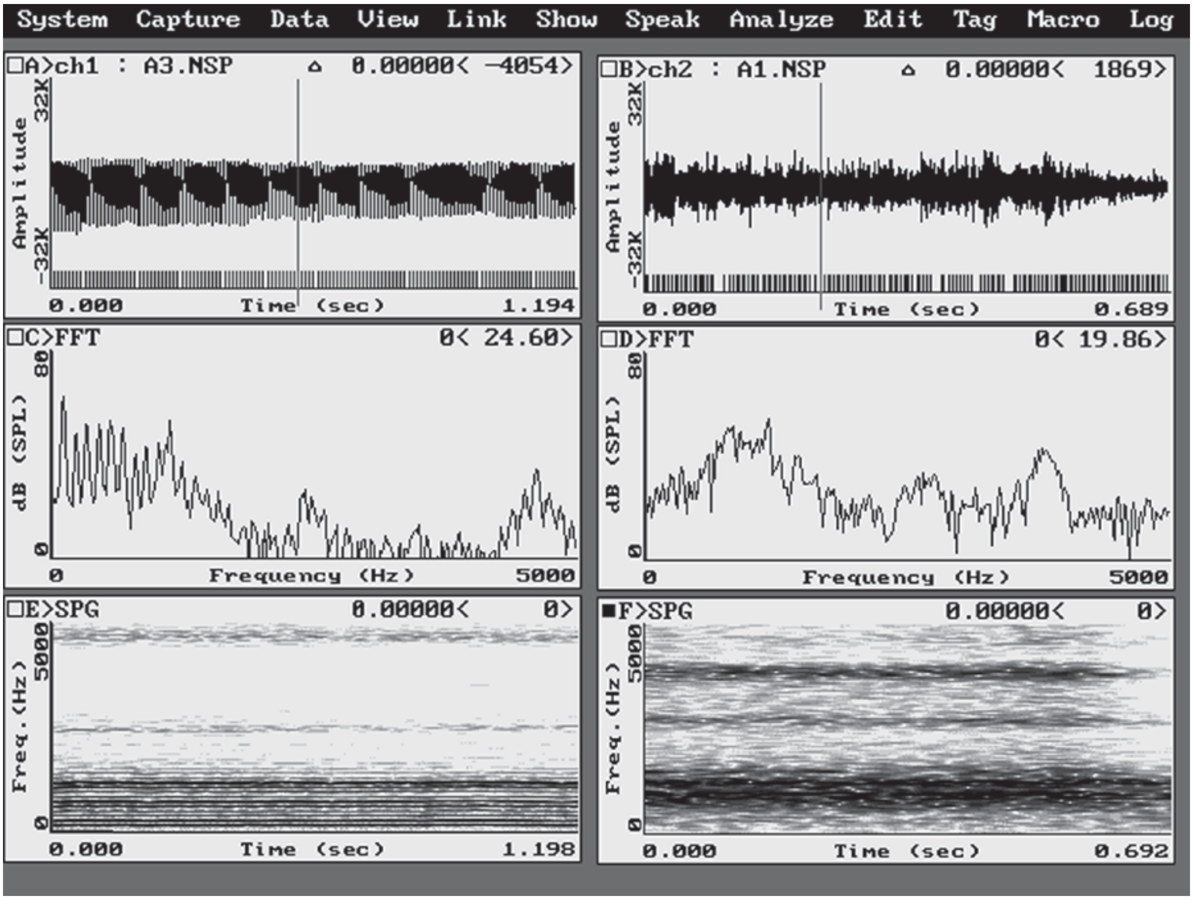

Fig. 2.3. Visible speech or sonagraphy, as displayed by the Computerized Speech Laboratory (Kay Elemetrics, Lincoln Park, NJ, USA). Sustained /a/: on the left by a normal voice and on the *right* by a breathy voice. *From top to bottom* : microphone signal), power spectrum (0–5000 Hz), and spectrogram (sonogram), frequency display 0–5000 Hz over time: about 1.2 s *left* and 0.7 s *right* , narrowband fi ltering 25 Hz with frequency (resolution). *Left panels*: in Power spectrum and spectrogram the harmonics are easy to identify, whereas they are lacking on the *right panels* . Here the power spectrum and spectrogram are replaced by aperiodic acoustical energy (noise). This kind of display also provides information about formant location

2.2.4.2 Acoustical Parameters

Acoustical analysis can also provide precise numerical values for many voice parameters, from averaged fundamental frequency to sophisticated calculations for noise components or tremor features.

Factor analysis allows the large number of acoustical parameters to be reduced to a limited number of clusters [14].

- Short-term fundamental frequency perturbation
- Short- or medium-term amplitude perturbation and voiceless segments
- Harmonics-to-noise ratio
- Long-term frequency and amplitude modulation
- Very long-term amplitude variation
- Subharmonics
- Tremor

Perturbation measures (in period and amplitude) and harmonics-to-noise computations on a sustained vowel (/a:/) at comfortable frequency and intensity appear to be the most robust measures and seem to determine the basic perceptual elements of voice quality: grade, roughness, and breathiness. Nevertheless, correlations with perceptual data remain usually moderate [14, 58]. *Jitter* is computed as the mean difference between the periods of adjacent cycles divided by the mean period. It is thus a fundamental frequency (F_0)-related measurement (Fig. 2.4). For *shimmer*, a similar computation is made on peak-to-peak amplitudes. Voice breaks must always be excluded. For pathological voices, the coefficients of

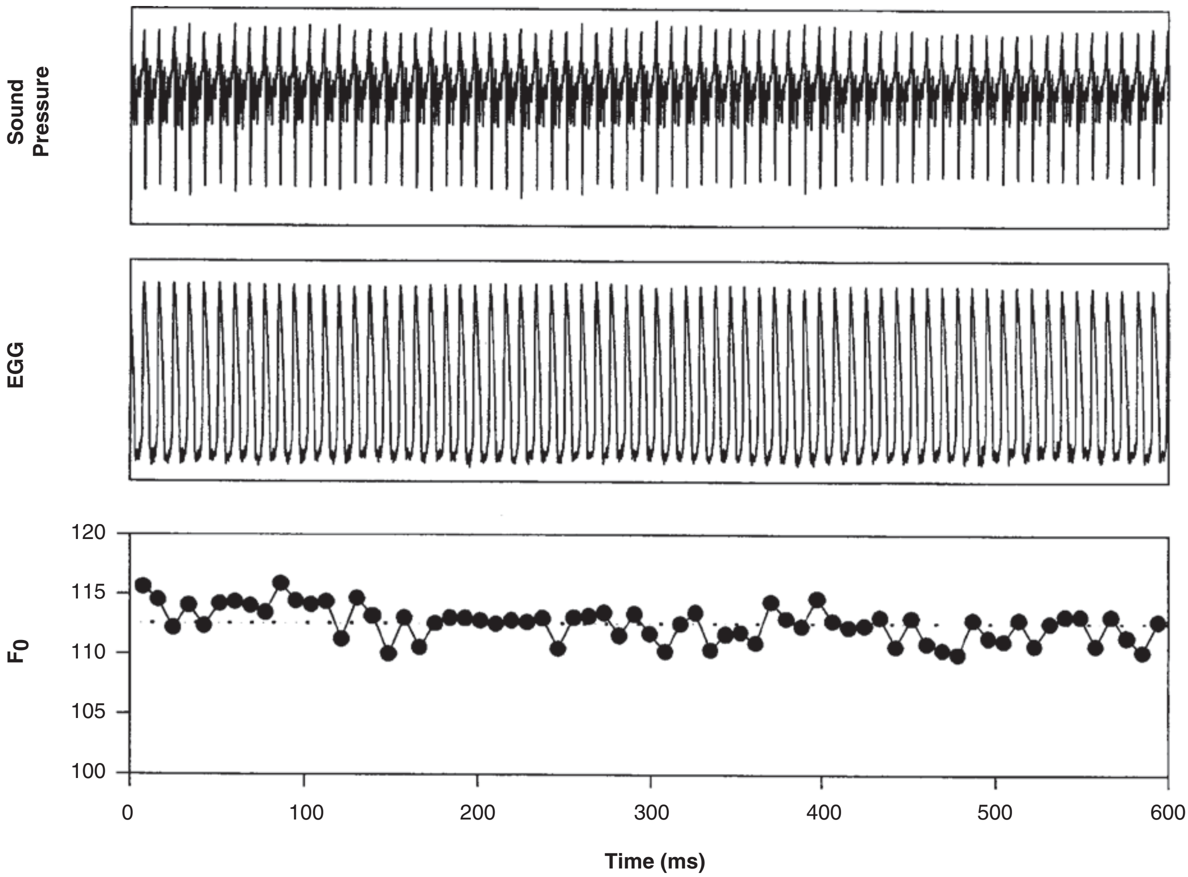

Fig. 2.4. Normal male voice, sustained /a:/. Microphone signal, electroglottogram, and F_0 plot across time. Normal voice is characterized by slight (< 1%) random variation of the fundamental frequency. In most cases of pathology, this aperiodicity (jitter) increases

variation of jitter and shimmer for a sustained /a/ are in the order of 20–30% for successive single trials as well as trials on different days [20–22]. A general limitation is that the systems employed for acoustical analysis cannot (or not in a reliable way) analyze strongly aperiodical acoustical signals. Perturbation measures of less than about 5% have been found to be reliable [59]. Only "quasi-periodic" voices are suited for perturbation analysis. Therefore, visual control of the period definition on the microphone signal is always necessary: Even in regular voices, a strong harmonic or subharmonic may account for erratic values. Alternatives from the field of nonlinear dynamics, such as the coefficient of Lyapunov, have been proposed for analyzing "chaotic" or "bifurcated" signals [60]. Also for substitution voices, special acoustical approaches of frequency perturbation have been proposed [18, 19].

For signal-to-noise ratio computations—e.g., normalized noise energy (NNE), harmonics to noise ratio (HNR), cepstral peak prominence (CPP)—there is currently insufficient standardization of the optimal algorithm(s) and insufficient knowledge about normative values for widespread clinical use. The harmonics-to-noise ratio was also found to be less well suited for demonstrating the effects of therapy [20–22].

Rhinophonia is a particular resonance characteristic of the voice. It may be present without a concomitant articulation disorder. Acoustical nasometry provides objective measurements by (schematically) computing the ratio between nasal and whole voice (nasal + oral) sound pressure levels [61].

2.2.4.3 Phonetography/Voice Range Profile

The phonetogram plots the dynamic range (dBA) as a function of the fundamental frequency range (Hz), thereby documenting the extreme possibilities of voice.

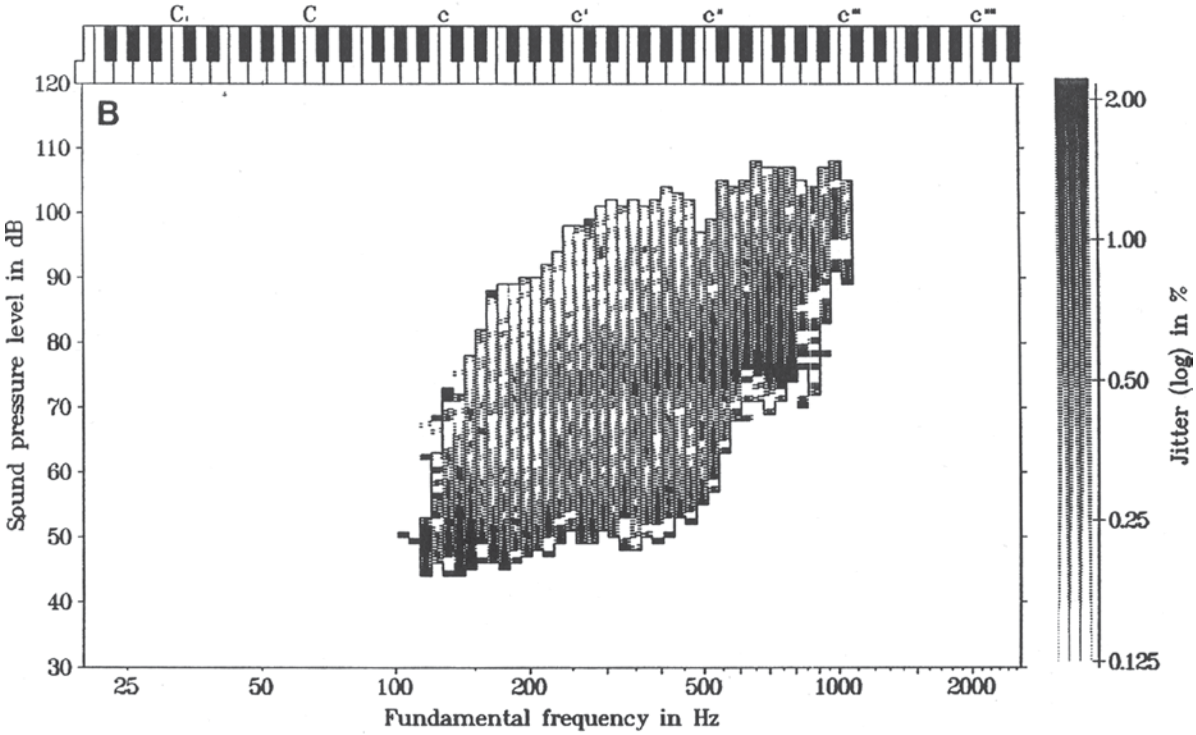

Fig. 2.5. Computerized phonetogram (voice range profile), with a gray scale indicating the amount of jitter (normal female voice). The more jitter, the darker the area. *Horizontal axis*: fundamental frequency in Hz (or musical tones on a keyboard). *Vertical axis*: sound pressure level, measured at 30 cm (dBA). This plot combines information about extreme possibilities of voice as well as an aspect of voice quality

These extremes are of importance for professional voice users, especially singers [62], but they must be interpreted with care [52, 53] because the acoustical energy is related to spectral distribution. Normative values for children and teachers have been defined [63].

Computerized systems make possible real-time measurement and display of fundamental frequency versus SPL and also of quality parameters such as jitter. Jitter results in various color gradations within the voice area, showing specific altered zones, or register boundaries (Fig. 2.5). Such computerized systems can also provide range profiles of current speech, possibly coupled with provocation tests, such as the task of reading at a controlled, louder intensity. These profiles are expected to be relevant for occupational voice users.

The highest and lowest frequencies and the softest intensity (dBA at 30 cm) seem most sensitive for changes in voice quality [5, 64–66], the latter being related to the phonation threshold pressure (PTP) [52, 53]. Measuring the lowest frequency allows one to compute the fundamental frequency range. Such a "three points range profile" can be obtained without completing a (time-consuming) whole voice range profile. However, as these three points represent "extreme" performances, they are (as are the MPT and CV) highly sensitive to learning and fatigue effects.

2.2.5 Self-Evaluation by Patient

Although subjective by definition, self-evaluation is of great importance in clinical practice. Careful quantification is needed for self-evaluation to be compared and correlated with the objective assessment provided by the voice laboratory. The purpose of subjective self-evaluation is to determine the deviance of voice quality and the severity of disability or handicap in daily professional and social life and the possible emotional repercussions of the dysphonia.

The basic aim is to differentiate the deviance of voice quality *stricto sensu* and the severity of the disability/handicap in daily social and/or professional life. A Voice Handicap Index can be computed on the base

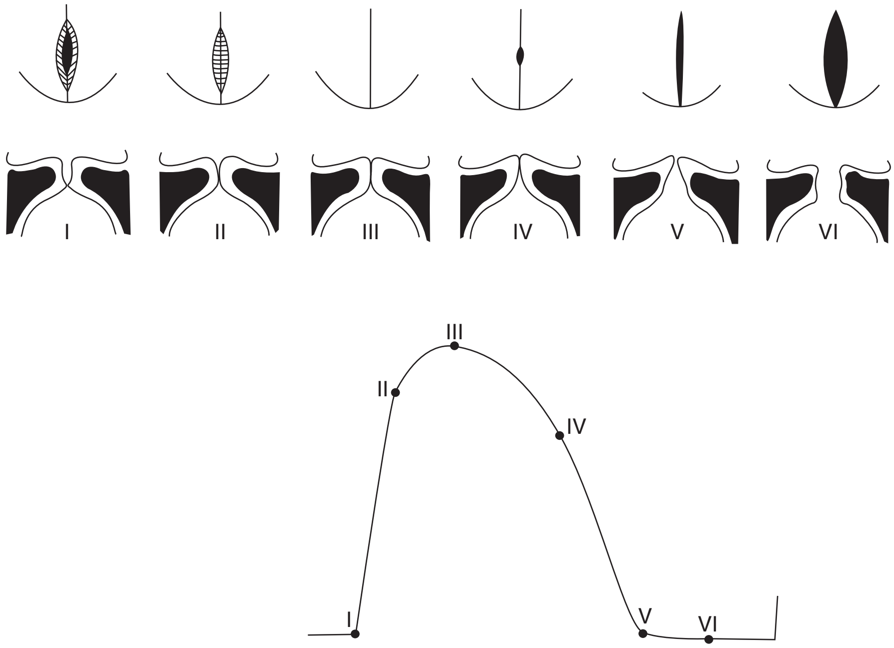

Fig. 2.6. Electroglottographic waveform with the corresponding laryngoscopic view and a frontal section through the midportion of the glottis. Point *III* corresponds to minimum impedance, which is at maximum closure of the glottis

of the patient's responses to a carefully selected list of questions [67]. It also investigates the possible emotional repercussion of the dysphonia. Rosen et al. [68] proposed a validated shortened version of the Voice Handicap Index: the VHI–10. However, for a basic protocol, a minimal subjective evaluation can be provided by patients themselves on a double VAS of 100 mm: their impressions about voice quality *stricto sensu* and about the repercussion of the voice problem regarding every-day social and, if relevant, professional life and activities. A score of "0" (maximum left) means a normal voice on the first scale and no handicap (related to voice) in daily life on the second scale. A score of "100" (maximum right) means extreme voice deviance on the first scale and extreme disability or handicap in daily social (and, when relevant, professional) activities, as rated by the patients themselves. A comparative study does not suggest that the exhaustive questionnaire is more reliable than the simple scales [20–22].

2.2.6 Adjuvant Techniques

2.2.6.1 Electroglottography

Electroglottography (or electrolaryngography) (EGG) is a method for monitoring vocal fold contact, rate of vibration, and perturbation of regularity during voice production (Fig. 2.6). The major advantage of EGG is that it does not interfere with the physiologic processes of speaking or singing. The signal originates from two electrodes lightly placed on the speaker's neck at the level of the thyroid cartilage. Pitch extraction from the EGG waveform is particularly reliable—so long as there is at least partial vocal fold contact during the vibration cycle—because the waveform is unaffected by vocal tract resonances and environment noise [69, 70].

The main applications of EGG are as follows:

- Fundamental frequency computations (e.g., range, regularity, distribution, display across time, cross plots) so long as there is vocal fold contact
- Voice onset time
- Prephonatory and postphonatory laryngeal gestures
- Closed phase information (hyperkinetic vs. hypokinetic adduction)
- Voice range profile of spontaneous speech (falsetto excluded)
- Triggering of a stroboscopic light source

2.2.6.2 Electromyography

Electromyography is an electrophysiological investigation of neuromuscular function. The main indications are mobility disorders (especially reduced mobility). Neuromuscular pathological conditions in laryngeal muscles do not basically differ from neuromuscular pathological conditions in other muscles, so it is recommended that these investigations be performed in cooperation with a general myography specialist [71–74]. With the patient supine and the neck extended, a concentric needle electrode is inserted through the cricothyroid ligament to approach the thyroid muscle. The needle electrode is then angled cranially 45° and laterally 20° to an approximate depth of 1.5–2.0 cm. The cricothyroid muscle is reached by inserting the electrode off the midline close to the inferior border of the thyroid cartilage. EMG is also used for monitoring *Botulinum* injections in vocal muscles. An evidence-based review has been provided by Sataloff et al. [75].

2.2.7 Specific Techniques for Substitution Voices and Spasmodic Dysphonia

A special protocol for perceptual evaluation is required for substitution voices (i.e., voices in which the sound is not generated by two vocal folds) [18, 19]. Substitution voicing cannot be evaluated accurately by the GRBAS perceptual rating scale as most substitution voices are G_2 or G_3 and because an optimal substitution voice is never G_0. Therefore, the IINFVo scale was proposed as an alternative (and was found reliable). It has the following parameters [76, 77].

- I—overall impression of voice quality, acceptability and adequacy for daily communication
- I—impression of intelligibility
- N—unintended additive noise, a parameter that reflects the amount of annoyance caused by the audibility of all sorts of uncontrolled noises (e.g., bubbly noise, air turbulence, clicks) produced during speech
- F—fluency, which reflects the perceived smoothness of the sound production and accounts for all kinds of undesirable interruptions
- Vo—voicing, which means that utterances are heard as voiced or voiceless when they need to be voiced or voiceless

Scoring is similar to the GRBAS scale, with a four-point scale (from no deviance to severe deviance), or using a VAS. Spasmodic dysphonias can also be rated in this way.

Classic acoustical parameters, such as perturbation measures, are not suited for substitution voices or spasmodic dysphonias, as they are not "quasi-periodic" signals but usually show chaotic "bifurcations," such as breaks or diplophonic moments. Nonlinear approaches are promising in this field.

Using a program based on a peripheral auditory model, Moerman et al. [18, 19] analyzed 10-ms frames of the signal and confirmed by objective acoustical analysis that the quality of tracheoesophageal speech is superior to that of esophageal speech but inferior to that of normal speech or speech with the preservation of one vocal fold.

In cases of spasmodic dysphonia (AD type), simple acoustical measures for fluency, such as the total duration of the sentence (ms) and the ratio of the total duration of voiced segments to the total duration of the sentence—both parameters measured in a short 100% voiced sentence for normal speakers—appear to be valuable objective criteria for effectiveness of treatment. They are useful for monitoring the evolution and timing of a new *Botulinum* injection [8, 9].

For patients with substitution voices, a slightly corrected version of the Voice Handicap Index has been proposed [18, 19].

2.2.8 Basic Protocol for Functional Assessment of Voice Pathology Recommended by the European Laryngological Society, Especially for Investigating Efficacy of (Phonosurgical) Treatments and Evaluating New Assessment Techniques

The basic protocol for functional assessment of voice pathology proposed by the European Laryngological Society was an attempt to reach better agreement and uniformity concerning the basic methodology for functional assessment of pathological voices. The purpose was to allow relevant comparisons with the literature when presenting/publishing the results of any kind of voice treatment (e.g., a phonosurgical technique or a new/improved instrument or procedure for investigating the pathological voice) [78]. A few basic principles served as guidelines.

1. Voice function is multidimensional [35].
2. A (minimum) set of basic requirements for presenting (publishing) results of voice treatments is necessary to make comparisons and meta-analyses possible.
3. New and more sophisticated measurement or evaluation techniques and procedures are to be encouraged, but the basic set must be performed in all cases for comparison.
4. The recommendations must be suited to all "common" dysphonias, but a few specific categories of voice pathology need a specific protocol for increasing sensitivity (e.g., substitution voices and spasmodic dysphonia) [8, 9, 18, 19].
5. In the basic set, or "truncus communis," for the assessment of common dysphonias, the following components need to be considered. Each provides quantitative data.
   a. Perception
   b. Videostroboscopy
   c. Acoustics
   d. Aerodynamics/efficiency
   e. Subjective rating by patient

When investigating substitution voices and spasmodic dysphonias, elements such as intelligibility and fluency should be added. Furthermore, scales and algorithms should be adjusted for the dimensions "perception," "videostroboscopy" and "acoustics."

6. Each of the above items has its own specific relevance when reporting results or statistics, as it provides a particular insight (multidimensionality). Combined scores or indexes [66] integrating these or other data into a single value may be useful, but optimal evaluation and understanding of the treatment effect also requires intrinsic comparison of the scores for the different components [79]. An example of such an Index is the Dysphonia Severity Index (DSI): it was constructed by logistic regression (Fisher discriminant analysis) and combines the highest fundamental frequency (F_0) (Hz), softest intensity, MPT, and jitter (%) according to the formula DSI = 0.13 × MPT + 0.0053 × highest F_0 – 0.26 × lowest intensity – 1.18 × jitter (%) + 12.4. For normal voices the DSI is +5, and for severely dysphonic voices it is –5. The first implementation studies pointed out that for some patients treatment effects can vary considerably from one dimension to another [80].
7. When assessing treatment outcomes, maximum objectivity must be constant. However, even objective data, such as audio recordings or videolaryngostroboscopic pictures, may be subjectively rated and interpreted. Nevertheless, for research purposes, it remains possible to improve the validity considerably by (1) averaging the ratings of a panel and (2) rating blindly, which means without knowing the conditions (e.g., before and after treatment).
8. Although the present guideline concerns only basic, nonsophisticated approaches, it is not to be considered as the ultimate way to conduct a basic assessment of voice function. Further implementation studies and more research are necessary and are warmly encouraged [81].
9. Instrumentation is kept to a minimum although considered essential for professionals performing phonosurgery. The ENT surgeon can be assisted in performing this basic set of measurements by a qualified, trained speech therapist.
10. In summary, two of the dimensions are considered objective (so long as the subject is cooperating normally): aerodynamics and acoustics. Although these two dimensions are considered objective, they are rated subjectively by the examiner (ratings can be made blindly by a panel!) via recording a voice sample and videostroboscopy. Finally, one dimension remains totally subjective (self-rating by the patient).

Implementation of this protocol demonstrates the clinical relevance of each item, and the low redundance. When investigating treatment effects, the correlations between the pre-post changes for the different parameters are weak [20–22, 82]. Multidimensional information about changes induced by therapy helps clinicians better understand the way in which a treatment works.

2.2.8.1 Example of a report using the proposed basic protocol

The patient was a 26-year-old woman who was diagnosed with vocal fold nodules. These results were obtained before treatment.

Perception	G34 B52 R18d
Stroboscopy	Clo 40hs Reg10 MW25 Sym0
Aerodynamics	PQ 285 ml/s (MPT 13 s)
Acoustics	Ji 1.2%; Shi 6.1%; F_0 – range c–g1; softest intensity 5 dBA 30 cm
Subjective evaluation	Vo30 Dis50

Explanation

Perception was rated on three VASs of 100 mm: grade, roughness, breathiness. Grade is scored 34/100, 52/100, and 18/100, where 0 = normal (no deviance) and 100 = extremely deviant. Diplophonia is present (d).

Stroboscopy was rated on four VASs of 100 mm: closure, regularity, quality of mucosal wave, symmetry. For closure, if abnormal, a categorical choice is also recommended; in this case, there was an hourglass-shaped pattern. Symmetry was normal.

Aerodynamics: phonation quotient (ml/s) and maximum phonation time (s). VC was 3705 ml.

Acoustics: jitter% and shimmer% on a sustained /a:/, at comfortable pitch and loudness. The “c” corresponds to 131 Hz and “g1” to 392 Hz. As for phonetography, the distance of the microphone must be 30 cm.

Subjective evaluation: provided by the patient himself on a double VAS of 100 mm. The first scale concerns the impression about voice quality *stricto sensu* (i.c. 30/100 = slight to moderate), and the second scale concerns the impression about repercussions of the voice problem during everyday social and, if relevant, professional life and activities (i.c. 50/100 = moderate to severe).

References

1. Dejonckere PH (2000) Perceptual and laboratory assessment of dysphonia. Otolaryngol Clin North Am;33:731–750
2. Titze I (1995) Workshop on Acoustic Voice Analysis: Summary Statement. National Center for Voice and Speech. The University of Iowa, Denver Co, 36 pp
3. Sataloff RT (1997) Professional Voice. Singular Publishing Group, San Diego
4. Watson C (1994) Database management of the voice clinic and laboratory. J Voice;3:99–106
5. Speyer R, Wieneke GH, van Wijck-Warnaar I, Dejonckere PH (2003) Effects of voice therapy on the voice range profiles of dysphonic patients. J Voice;17:544–556
6. Dejonckere PH (1998) Effect of louder voicing on acoustical measurements in dysphonic patients. Logoped Phoniatr Vocol;23:79–84
7. Dejonckere PH, Lebacq J (2001) Plasticity of voice quality: a prognostic factor for outcome of voice therapy? J Voice;15:251–256
8. Dejonckere PH (2007) Améliorer la fiabilité de l’analyse perceptive de la voix pathologique par un support visuel sonagraphique. In: Klein-Dallant C (Éd). Voix parlée et chantée. Editions Ortho, Paris, pp. 35–45
9. Dejonckere PH (2007) Critères acoustiques de fluence pour l’évaluation des dysphonies spasmodiques. In: Klein-Dallant C (Éd). Voix parlée et chantée. Editions Ortho, Paris, pp. 63–73
10. Revis J, Barberis S,Giovanni A (2000) Definition of a new temporal voice onset measurement. Rev Laryngol Otol Rhinol (Bord);121:291–296
11. Hirano M (1981) Clinical Examination of Voice. Springer, New York
12. Dejonckere PH, Lebacq J (1996) Acoustic, perceptual, aerodynamic and anatomical correlations in voice pathology. ORL J Otorhinolaryngol Relat Spec;58:326–332
13. De Bodt, M, Wuyts, F,Van de Heyning, P, Croeckx, C (1997) Test - re-test Study of GRBAS-Scale. J Voice;11:74–80
14. Dejonckere PH, Remacle M, Fresnel-Elbaz E, Woisard V, Crevier Buchman L, Millet B (1996) Differentiated perceptual evaluation of pathological voice quality: reliability and correlations with acoustic measurements. Rev Laryngol Otol Rhinol;117:219–224
15. Webb AL, Carding PN, Deary IJ, MacKenzie K, Steen N, Wilson JA (2004) The reliability of three perceptual evaluation scales for dysphonia. Eur Arch Otorhinolaryngol;261: 429–434
16. Nawka T, Anders LC, Wendler J (1994) Die auditive Beurteilung heiserer Stimmen nach dem RBH-System. Sprache Stimme Gehör;18:130–133
17. Wuyts F, De Bodt M, Van de Heyning PH (1999) Is the reliability of a visual analog scale higher than of an ordinal scale ? An experiment with the GRBAS-scale for the perceptual evaluation of dysphonia. J Voice;13:508–517
18. Moerman M, Martens JP, Dejonckere PH (2004) Application of the voice handicap index in 45 patients with substitution voicing after total laryngectomy. Eur Arch Otolaryngol;261: 423–428
19. Moerman M, Pieters G, Martens JP, Van Der Borgt MJ, Dejonckere PH (2004) Objective evaluation of the quality of substitution voices. Eur Arch Otorhinolaryngol;261: 541–547
20. Speyer R, Wieneke GH, Dejonckere PH (2004) The use of acoustic parameters for the evaluation of voice therapy for dysphonic patients. Acta Acustica United Acustica;90:520–527
21. Speyer R, Wieneke GH, Dejonckere PH (2004) Documentation of progress in voice therapy: perceptual, acoustic and laryngostroboscopic findings pretherapy and postterapy. J Voice;18: 325–339
22. Speyer R, Wieneke GH, Dejonckere PH (2004) Self-assessment of voice therapy for chronic dysphonia. Clin Otolaryngol;29:66–74
23. Dejonckere PH, Crevier L, Elbaz E, Marraco M, Millet B, Remacle M, Woisard V (1999) Quantitative rating of videolaryngostroboscopy: factor analysis of the vibratory characteristics. In: Dejonckere PH, Peters H (Eds). Communication and its Disorders: A Science in Progress. Nijmegen University Press, Nijmegen, pp. 170–171

24. Södersten M, Hertegard S, Hammarberg B (1995) Glottal closure, transglottal air flow and voice quality in healthy middle-aged women. J Voice;9:182–197
25. Sulter AM, Schutte HK, Miller DG (1996) Standardized laryngeal videostroboscopic rating: differences between untrained and trained male and female subjects, and effects of varying sound intensity, fundamental frequency, and age. J Voice;10:175–189
26. Speyer R, Wieneke GH, Hosseini EG, Kempen PA, Kersing W, Dejonckere PH (2002) Effects of voice therapy as objectively evaluated by digitized laryngeal stroboscopic imaging. Ann Otol Rhinol Laryngol;111:902–908
27. Dejonckere PH, Wieneke GH, Lebacq J (1989) Laryngostroboscopy and glottic dysrhythmia's. Acta Otorhinolaryngol Belg;43:19–29
28. Hirano M, Bless DM (1993) Videostroboscopic Examination of the Larynx. Singular Publishing, San Diego
29. Speyer R, Wieneke GH, Kersing W, Dejonckere PH (2005) Accuracy of measurements on videostroboscopic images of the vocal folds. Ann Otol Rhinol Laryngol;114:443–450
30. Eysholdt U, Rosanowski F, Hoppe U (2003) Measurement and interpretation of irregular vocal fold vibrations. HNO;51: 710–716
31. Imagawa H, Kiritani S, Hirose H (1987) High speed digital image recording system for observing vocal fold vibration using an image sensor. J Med Electron Biol Eng;25: 284–290
32. Neubauer, J, Mergell P, Eysholdt U, Herzl, H (2001) Spatio-temporal analysis of irregular vocal fold oscillations: biphonation due to desynchronization of spatial modes. J Acoust Soc Am;110:3179–3192
33. Svec JG, Schutte HK (1996) Videokymography: high-speed linescanning of vocal fold vibration. J Voice;10:201–205
34. Dejonckere PH, Versnel H (2005) High speed imaging of vocal fold vibration: analysis by 4 synchronous line-scans of onset, offset and register breaks. Proceedings IFOS Congress, Rome, 9pp*
35. Hirano M (1989) Objective evaluation of the human voice: clinical aspects. Folia Phoniatr;41:89–144
36. Neiman GS, Edeson B (1981) Procedural aspects of eliciting maximum phonation time. Folia Phoniatr;33:285–293
37. Kent RD, Kent JF, Rosenbek JC (1987) Maximum performance tests of speech production. J Speech Hear Res;52: 367–387
38. Rau D, Beckett RL (1984) Aerodynamic assessment of vocal function using hand-held spirometers. J Speech Hear Dis;49: 183–188
39. Morris S, Jawad MSM, Eccles R (1992) Relationships between vital capacity, height and nasal airway resistance in asymptomatic volunteers. Rhinology;30:259–264
40. Colton RH, Casper JK (1996) Understanding Voice Problems. Williams & Wilkins, Baltimore
41. Verdolini K (1994) Voice disorders. In: Tomblin JB, Morris HL, Spriesterbach OC (Eds). Diagnosis in Speech-Language Pathology. Singular Publishing Group, San Diego, pp. 247–306
42. Woo P, Casper J, Colton R, Brewer D (1994) Aerodynamic and stroboscopic findings before and after microlaryngeal phonosurgery. J Voice;8:186–194
43. Woo P, Colton RH, Shangold L (1987) Phonatory air flow analysis in patients with laryngeal disease. Ann Otol Rhinol Laryngol;96:549–555
44. Schutte HK (1980) The efficiency of voice production, Thesis. University of Groningen. Kemper, Groningen
45. Schutte HK (1992) Integrated aerodynamic measurements. J Voice;6:127–134
46. Fritzell B, Hallen O, Sundberg J (1974) Evaluation of teflon injection procedures for paralytic dysphonia. Folia Phoniatr; 26:414–421
47. Hirano M, Koike Y, von Leden H (1968) Maximum phonation time and air usage during phonation. Folia Phoniatr; 20:185–201
48. Murry T, Bone RC (1978) Aerodynamic relationships associated with normal phonation and paralytic dysphonia. Laryngoscope;88:100–109
49. Rothenberg M (1973) A new inverse filtering technique for deriving the glottal airflow during voicing. J Acoust Soc Am;53:1632–1645
50. Baken RJ, Orlikoff R (2000) Clinical Measurement of Speech and Voice. Singular Thomson Learning, San Diego
51. Rothenberg M (1982) Interpolating subglottal pressure from oral pressure. J Speech Hear Disord;47:218–224
52. Titze IR (1992) Phonation threshold pressure – a missing link in glottal aerodynamics. J Acoust Soc;91:2926–2935
53. Titze IR (1992) Acoustic interpretation of the voice profile (phonetogram). J Speech Hear Res;35:21–34
54. Colton RH, Woo P (1995) Measuring vocal fold function. In: Rubin JS, Sataloff RT, Korovin GS, et al (Eds). Diagnosis and Treatment of Voice Disorders. Igaku - Shoin, New York, pp. 290–315
55. Hyatt RE, Black LF (1973) The flow-volume curve: a current perspective. Am Rev Respir Dis;107:191
56. Robin ED (1982) Respiratory Medicine. In Medicine, Scientific American, New York, 14;III:1–7**
57. Slavit DH (1995) Role of the pulmonary junction in voice assessment. In: Rubin JS, Sataloff RT, Korovin GS, Gould WJ (Eds). Diagnosis and Treatment of Voice Disorders. Igaku - Shoin, New York, pp. 327–340
58. Wolfe V, Fitch J, Martin D (1997) Acoustic measures of dysphonic severity across and within voice types. Folia Phoniatr Logop;49:292–299
59. Titze I, Liang H (1993) Comparison of Fo extraction models for high precision voice perturbation measurements. J Speech Hear Res;36:1120–1133
60. Yu P, Ouaknine M,Giovanni A (2000) Clinical significance of calculating the coefficients of Lyapunov in the objective assessment of dysphonia. Rev Laryngol Otol Rhinol (Bord); 5:301–305
61. Dejonckere PH, van Wijngaarden HA (2001) Retropharyngeal autologous fat transplantation for congenital short palate: a nasometric assessment of functional results. Ann Otol Rhinol Laryngol;110:168–172
62. Schultz-Coulon HJ (1990) Stimmfeldmessung. Berlin, Springer
63. Heylen L, Wuyts FL, Mertens F, De Bodt M, Dan de Heyning PH (2002) Normative voice range profiles of male and female professional voice users. J Voice;16:1–7
64. Heylen L, Wuyts F, Mertens F, De Bodt M, Pattyn J, Croeckx C, Van de Heyning P (1998) Evaluation of the vocal performance of children using a voice range profile index. J Speech Lang Hear Res;41:232–238
65. Van de Heyning PH, et al (1996) Research work of the Belgian study group on voice disorders. Acta Otorhinolaryngol Belg;50:321–386

66. Wuyts F, De Bodt M, Molenberghs G, Remacle M, Heylen L, Millet B, Van Lierde K, Raes J, Van de Heyning PH (2000) The Dysphonia Severity Index: an objective measure of quality based on a multiparameter approach. J Speech Lang Hear Res;43:796–809
67. Jacobson BH, Johnson A, Grywalski C, Silbergleit A, Jacobson G, Benninger MS, Newman CW (1997) The Voice Handicap Index (VHI): development and validation. Am J Speech Lang Pathol;6:66–70
68. Rosen CA, Lee AS, Osborne J, Zullo T, Murry T (2004) Development and validation of the voice handicap index – 10. Laryngoscope;114:1549–1556
69. Dejonckere PH (1996) Electroglottography: a useful method in voice investigation. In: Pais-Clemente M (Ed). Voice Update. Excerpta Medica. Elsevier, Amsterdam, pp. 29–33
70. Fourcin A, Abberton E, Miller D, et al (1995) Laryngography. Eur J Disord Commun;30:101–115
71. Blitzer A (1995) Laryngeal electromyography. In: Rubin JS, Sataloff RT, Korovin GS, Gould WJ (Eds). Diagnosis and Treatment of Voice Disorders. Igaku - Shoin, New York, pp. 316–326
72. Dejonckere PH (1987) EMG of the Larynx. Press Productions, Liège
73. Dejonckere PH, Knoops P, Lebacq J (1988) Evoked muscular potentials in laryngeal muscles. Acta Otolaryngol Belg;42:494–501
74. Munin MC, Murry T, Rosen CA (2000) Laryngeal electromyography. Otolaryngol clin North Am;33:759–770
75. Sataloff RT, Mandel S, Mann EA, Ludlow CL (2004) Practice parameter: laryngeal electromyography (an evidence-based review). J Voice;18:261–274
76. Moerman M, Martens JP, Crevier-Buchman L, de Haan E, Grand S, Tessier C, Woisard V, Dejonckere PH (2006) The INFVo perceptual rating scale for substitution voicing: development and reliability. Eur Arch Otorhinolaryngol; 263(5):435–439
77. Moerman M, Martens, JP,Van Der Borgt MJ, Pelemans M, Gillis M, Dejonckere PH (2006) Perceptual evaluation of substitution voices: development and evaluation of the (I)INFVo rating scale. Eur Arch Otorhinolaryngol;263:183–187
78. Dejonckere PH, Bradley P, Clemente P, Cornut G, Crevier-Buchmann L, Friedrich G, Van De Heyning P, Remacle M, Woisard V (2001) A basic protocol for functional assessment of voice pathology. Eur Arch Otorhinolaryngol;258:77–82
79. Verdonck-de Leeuw IM, Mahieu HF (2000) Multidimensional assessment of voice characteristics following radiotherapy for early glottic cancer and the DSI. Eur Arch Otolaryngol; 257(Suppl 1):S20
80. Dejonckere PH, Crevier L, Elbaz E, Marraco M, Millet B, Remacle M, Woisard V (2000) Clinical implementation of a multidimensional basic protocol for assessing functional results of voice therapy. In: Jahnke K, Fischer M (Eds). Proceedings EUFOS Congress. Monduzzi editore, Berlin, pp. 561–565
81. Friedrich G, Dejonckere PH (2005) The voice evaluation protocol of the European Laryngological Society. First results of a multicenter study. Laryngorhinootologie;84:744–752
82. Dejonckere PH, Crevier-Buchman L, Marie JP, Moerman M, Remacle M, Woisard V (2003) Implementation of the European Laryngological Society (ELS) - basic protocol for assessing voice treatment effect. Rev Laryngol Otol Rhinol; 124:279–283