Resonance properties of the vocal folds: In vivo laryngoscopic investigation of the externally excited laryngeal vibrations

Jan G. Švec a)  
Center for Communication Disorders, Medical Healthcom, Ltd., Řeřovská 10/491, 181 00 Prague 8, the Czech Republic

Jaromír Horáček  
Institute of Thermomechanics Academy of Sciences of the Czech Republic, Dolejškova 5, 182 00 Prague 8, the Czech Republic

František Šram  
Center for Communication Disorders, Medical Healthcom, Ltd., Řeřovská 10/491, 181 00 Prague 8, the Czech Republic

Jan Veselý  
Institute of Thermomechanics Academy of Sciences of the Czech Republic, Dolejškova 5, 182 00 Prague 8, the Czech Republic

(Received 26 July 1999; accepted for publication 20 June 2000)

The study presents the first attempt to investigate resonance properties of the living vocal folds by means of laryngoscopy. Laryngeal vibrations were excited via a shaker placed on the neck of a male subject and observed by means of videostroboscopy and videokymography (VKG). When the vocal folds were tuned to the phonation frequency of 110 Hz and sinusoidal vibration with sweeping frequency (in the range 50–400 Hz) was delivered to the larynx, three clearly pronounced resonance peaks at frequencies around 110, 170, and 240 Hz were identified in the vocal fold tissues. Different modes of vibration of the vocal folds, observed as distinct lateral-medial oscillations with one, two, and three half-wavelengths along the glottal length, respectively, were associated with these resonance frequencies. At the external excitation frequencies below 100 Hz, vibrations of the ventricular folds, aryepiglottic folds and arytenoid cartilages were dominant in the larynx. © 2000 Acoustical Society of America. [S0001-4966(00)00210-1]

PACS numbers: 43.70.Aj, 43.40.At [AL]

I. INTRODUCTION

From the theory of vibration it is known that vibration of a structure can generally be decomposed into a set of independent characteristic vibration patterns, called eigenmodes. Like any other vibrating structure, vocal folds have inherent eigenmodes which are crucial in determining their possible vibration behavior. Each of the eigenmodes is associated with a specific eigenfrequency and exhibits certain damping. The eigenmodes, eigenfrequencies, and damping are called the “dynamic characteristics” of the vibrating system and are independent of the excitation mechanism of the vibration. These characteristics can be used to describe inherent vibration properties of the vocal folds. Information on the dynamic characteristics of the true vocal folds has been rather limited, however, mostly due to difficulties related to measurement of these characteristics in the delicate and hardly accessible vocal fold tissues.

Titze and Strong (1975) were the first who theoretically studied the eigenmodes of the vocal folds. The theory predicts generally infinite number of eigenmodes in the vocal fold tissues (e.g., Titze and Strong, 1975; Titze, 1994). Only a few dominant (lowest) modes, however, are assumed to play a substantial role in the actual vibration of the vocal folds. In a study with a finite-element model of the vocal folds, Berry et al. (1994) have found that combination of only two eigenmodes captures more than 95% of the variance of the vocal fold vibration in normal phonation and more than 70% of irregular vocal fold vibration. Prevalence of two dominant eigenmodes has also been recently found experimentally in vibration of the vocal folds in excised canine larynges (Berry, in press).

The two dominant eigenmodes have been theoretically known and designated as x-10 and x-11. The x-10 mode presents a lateral-medial oscillation which is responsible for opening and closing of the vocal folds in a vibratory cycle. The x-11 mode presents an out-of-phase motion of the upper and lower margins of the vocal folds, which plays an important role in the transfer of the aerodynamic energy into the motion of the vocal fold tissues (Stevens, 1977; Ishizaka, 1988; Titze, 1988). Higher-order modes, such as, e.g., x-20, x-21 or x-30, x-31 (lateral-medial oscillations encompassing two or three half-wavelengths along the glottal length) are assumed to partially contribute to production of more complex vocal fold vibration patterns, especially those related to pathologic voice quality (Titze and Strong, 1975; Titze, 1994; Berry et al., 1994).

During phonation, the oscillation of the vocal folds is significantly influenced by phonatory airflow. Aerodynamic coupling leads, along with inherent nonlinearity of the vocal
folds, to phenomena such as entrainment of the eigenmodes (also known as “mode locking”), which cause rearrangement of the eigenfrequencies of the vocal folds (Berry et al., 1994; Fletcher, 1996; Berry, in press). A well known example can be found in the behavior of a simple two-mass model of the vocal folds (Ishizaka and Flanagan, 1972): under typical conditions the eigenmodes of the model, which correspond to the modes $x$-10 and $x$-11, are tuned to 120 and 201 Hz (e.g., Titze, 1976; Ishizaka, 1988). Under the influence of the airflow these two eigenmodes are entrained to vibrate at identical frequency around 130–150 Hz (depending on the subglottal pressure; see Ishizaka and Flanagan, 1972). This effect illustrates a need for studies of vocal fold behavior not influenced by the phonatory airflow revealing the dynamic characteristics more accurately.

Only a few experimental studies have been devoted to the vibrations of the vocal folds without the airflow. In a study with excised human larynges, Tanabe et al. (1979) displaced a vocal fold via a metal plate attached to a stretched elastic rod. When the rod was cut, the vocal fold was released and exhibited a damped oscillation which was monitored using a high-speed (cinematographic) camera. These experiments were performed for obtaining information on the damping properties of the vocal folds. The results could, in principle, have been used also for studying the eigenfrequencies of the vocal folds, but these authors did not determine these values.

The only studies to date, which provide information on eigenfrequencies as well as damping characteristics of the true vocal folds, were published by Kaneko et al. (1981, 1983, 1987). Here, the vibration of the vocal folds was excited externally using a shaker placed on the thyroid cartilage. The response of the vocal folds was registered by means of modified ultrasonic equipment. Kaneko designed a special phonatory maneuver in order to investigate the eigenfrequencies of the vocal folds in living subjects: the subject phonated at a given pitch and stopped delivering the air from the lungs while keeping the vocal folds in the phonatory position (the so-called “neutral phonatory position”). At that moment the shaker was switched on and the response measurement was done. Kaneko identified two distinct eigenfrequencies of the vocal folds, which changed with the phonation frequency. The lower of the eigenfrequencies was found to be close to the frequency of phonation. Similar results were obtained by Kaneko et al. in a number of living subjects as well as in excised human larynges.

The eigenfrequencies of the true vocal folds measured by Kaneko et al. correspond well to the eigenfrequencies of the two-mass model of the vocal folds designed by Ishizaka and Flanagan (1972). On the basis of this correlation it has been hypothesized (Ishizaka, 1988) that these two eigenfrequencies are related to the eigenmodes $x$-10 and $x$-11. Recently, however, this hypothesis has been challenged by Berry and Titze (1996) who have, on the basis of a theoretical analysis of a continuum model of the vocal folds, predicted the eigenfrequencies of the $x$-10 and $x$-11 modes to be nearly identical. A much smaller difference (0–25 Hz) than that measured by Kaneko et al. (ca. 50–100 Hz) was also found between the eigenfrequencies of the $x$-10 and $x$-11 modes in the finite-element model of the vocal folds by Dedouch et al. (1999). The eigenfrequencies of eigenmodes found in the finite-element model of Jiang et al. (1998) were, on the other hand, quite far apart (>100 Hz). Unfortunately, Kaneko’s experiments did not bring any information on mode shapes of vibration of the vocal folds and thus it has not been clear whether, indeed, the measured eigenfrequencies belong to $x$-10 and $x$-11, or to some other eigenmodes. This uncertainty has called for new, more specific measurements of the dynamic characteristics of the true vocal folds.

This study presents the first attempt to investigate the dynamic characteristics of the vocal folds in vivo laryngoscopically. The basic question of the present study is the following one: Is it possible to externally induce the vibration of the vocal folds to such extent that they can be monitored laryngoscopically?

Certain positive evidence can be found in the studies of Fukuda et al. (1987) and Fukuda (1999). Here, a laryngeal shaker, principally similar to the one used by Kaneko, was used to excite the vocal fold vibrations in patients undergoing surgery under general anesthesia. The vibration of the vocal fold mucosa could be observed laryngoscopically under stroboscopic light. The studies of Fukuda were, however, clinically oriented and did not pay attention to the dynamic characteristics of the vocal folds.

The aim of the present study is to employ the laryngoscopic observation (1) to identify the eigenfrequencies of the vocal folds and (2) to relate these eigenfrequencies to the specific mode shapes of vibration.

II. MATERIALS AND METHODS

A resonance approach, well known from technical practice (e.g., Anderson, 1967; Richardson, 1997) and previously adapted by Kaneko et al. (1981), was modified for the measurement of the dynamic characteristics of the vocal folds. Laryngeal vibrations were excited externally via a shaker placed anteriorly on the thyroid cartilage. The vibration response of the vocal folds was monitored laryngoscopically. Resonance frequencies (i.e., frequencies at which the vocal folds exhibit maximal amplitudes of vibration) and resonance modes of vibration (i.e., vibration patterns of the vocal folds at the resonance frequencies) were examined. The resonance frequencies and resonance modes can be seen as practical approximations of the eigenfrequencies and eigenmodes of the vocal folds. (More detailed information on the relationship between the eigenfrequencies/modes and resonance frequencies/modes can be found in literature on dynamics of vibrating systems, e.g., Anderson, 1967.)

A. Experimental setup

The experimental setup, shown in detail in Fig. 1, consisted of three parts: (1) generation and monitoring of the vibration of the external shaker; (2) monitoring the force by which the shaker is pressed against the neck; and (3) laryngoscopic observation of the excited vocal fold vibrations.

1. Excitation of the laryngeal vibrations

Signal (sinusoidal and impulse, see Sec. II.B for details) from a generator (HP type 3324A Synthesized Function
Sweep Generator) was amplified by a VEB Metra Verstärker (type LV 103) and fed into a shaker (Brüel & Kjaer, type 4810). A specially designed plexiglass cylindrical head was firmly attached (screwed) to a vibrating element of the shaker. This plexiglass head served as an electrically isolated contact element which was placed on the neck of the subject. Acceleration of the cylindrical head was registered via an attached accelerometer (Brüel & Kjaer, type 4344). Signal from the accelerometer was amplified by means of a vibration meter (Brüel & Kjaer, type 4511) and recorded on the first channel of an FFT analyzer (Brüel & Kjaer, type 2034 Dual Channel Signal Analyzer). The vibration force was registered by means of a force transducer (Brüel & Kjaer, type 8200) which was placed between the plexiglass cylinder and the vibrating element of the shaker. Signal from the force transducer was amplified by a Bruel & Kjaer (type 2626) amplifier and recorded on second channel of the FFT analyzer.

Signals from the generator, accelerometer, and force transducer were used to obtain transfer functions of the excitation system. The transfer functions (force/generator as well as acceleration/generator) were found to be essentially flat in the frequency range 100–400 Hz. Below 100 Hz, in the relevant range 50–100 Hz, the transfer functions monotonically decreased with descending frequency. Maximum displacement amplitude of the shaker was observed at 50 Hz and measured to be ca. 3 mm.

2. Contact force measurement

The body of the shaker was firmly attached onto a metal rod (Fig. 1). The rod bent when the shaker was pressed against the neck. The amount of bending was measured by means of a semiconductor strain gauge designed at the Institute of Thermomechanics of the Academy of Sciences of the Czech Republic (Vaněk and Cibulka, 1994). The strain gauge was powered from a stabilized power supply of 10 V (TESLA, type BS 525). Change of voltage in the sensory circuit of the semiconductor strain gauge caused by the bending was monitored by means of a voltmeter (TESLA, type DU 20). The system was calibrated, making it possible to convert the measured voltage to the force applied to the shaker.

3. Laryngoscopic observation of the excited laryngeal vibrations

Two methods of monitoring the vocal fold vibration were used: laryngostroboscopy and high-speed videokymography. Laryngostroboscopy is a well-known technique routinely used in laryngology and a more detailed description can be found elsewhere (e.g., Hirano and Bless, 1993). Videokymography is a newly developed method for high-speed optical investigation of vibrations (Švec et al., 1997, 1999; Schutte et al., 1998).

In Videokymography (VKG) a modified video camera is used. The camera can function in two different modes, standard and high speed. In the standard mode, it works as a standard commercial video camera, monitoring the vibration
of the vocal folds with a speed of 25 frames (respectively, 50 interlaced fields) per second (CCIR/PAL standard was used here). An example of a standard laryngoscopic image is shown in Fig. 2(a). In the high-speed mode the camera delivers images from a single selected line with a speed of 7812.5 line images per second. These line images are put below each other and together create a new, videokymographic image monitoring vibration of the selected part of the vocal folds in time [Fig. 2(b)]. A mechanical switch enables to change between the standard and high-speed modes instantly. Both the normal as well as the high-speed images are transmitted in a standard TV format and can be recorded and monitored using a standard video recorder and a TV-compatible monitor. More detailed information on videokymography can be found elsewhere (Švec and Schutte, 1996; Švec et al., 1997, 1999).

The laryngoscopic setup is shown in Fig. 1. For stroboscopy, the following equipment was used: Light source (Rhino-Laryngeal Stroboscope Kay Elemetrics, model 9100), rigid endoscope (70° Kay Elemetrics, type 9106), and 3CCD color camera (Panasonic GP-US502 with a Control Unit) with a C-mount 35mm lens/adapter (Kay Elemetrics, model 9116). Xenon light source (Richard Wolf Auto LP/FLASH 5135), Lupenlaryngoskop 90° Richard Wolf, model 4450.47), and videokymographic CCD black and white camera (Lambert Instruments) with a C-mount objective/adapter (ATMOS) were employed for videokymography. The laryngeal image, registered by either (standard or videokymographic) video camera attached to the endoscope, was presented simultaneously on two video monitors (in Fig. 1, only one monitor is depicted, for simplicity). The examiner used the first monitor; the second monitor provided feedback to the examined subject. The images were recorded using an s-VHS video tape recorder (Panasonic, model AG 7355). An audio signal was registered by means of an electret microphone (Kay Elemetrics lapel microphone, type 7175-6000) and recorded on audio track of the videotape.

B. Experimental procedure

One of the authors (JGS, male, age 32, an amateur jazz-singer) served as a subject for the study. The phonation frequency of 110 Hz was chosen as a reference for the investigation. During the experiment the subject placed the plexiglass cylindrical head of the shaker anteriorly on the prominence of the thyroid cartilage and pushed the neck against the shaker with a force of ca. 3–5 Newton (higher forces were subjectively judged as uncomfortable). The shaker was not firmly fixed to the neck, for safety reasons. The subject at comfortable intensity in chest register reproduced the reference frequency, given by means of a tuning fork. Next, the endoscope was inserted into the oral cavity and the subject repeated the phonation. After this, the subject took a breath and produced the Kaneko maneuver: a short phonation which was interrupted while keeping the vocal folds in the neutral phonatory position. Attempt was made to avoid any movement of the larynx during the Kaneko maneuver (see also the Appendix). At this time, the external vibrations were delivered to the larynx. Three different approaches were used for investigating the externally excited laryngeal vibrations:

1. Videostroboscopy

Sinusoidal excitation signal with a constant frequency (50, 75, 100, 110, 125, 150, 175, 200, 225, 250, 275, 300, 325, 350, 375, and 400 Hz) was used and the laryngeal vibrations were monitored videostroboscopically.

2. Frequency sweep

A sinusoidal signal of constant input power was delivered to the shaker and the frequency was linearly increased. The vibration amplitude of the shaker was dependent on the frequency of oscillation (see the Appendix). Two frequency sweeps in the frequency range of 100–400 Hz and 50–200 Hz were used, each sweep 5 s in duration. The measurement position was aimed at the middle of the membranous part of the vocal folds, transversally to the glottis [Fig. 2(a)], using a standard mode of the VKG camera. The measurement position was adjusted manually by moving the endoscope to the desirable position during the phonation preceding the Kaneko maneuver. Before the start of the frequency sweep, the camera was switched into the high-speed mode and the position of the endoscope was held still during the sweep. After the end of the sweep, the VKG camera was switched back into the standard mode in order to confirm that the measurement position had not changed during the VKG measurement. The measurement was repeated several times for each sweep.

3. Impulse excitation

A periodic rectangular impulse excitation signal of 1-ms duration was delivered from the generator to the shaker. The vibration response of the laryngeal tissues was observed by means of videokymography. Vocal folds responded with damped oscillations; their amplitude was found rather small, however, which appeared problematic for a detailed analysis. The VKG data from this particular condition were thus not analyzed in detail and are not treated in this study. The signals from the force transducer and accelerometer were used to obtain transfer functions of the excitation system (discussed in Sec. II A 1).

C. Analysis of the video data

1. Frequency sweep

The best VKG samples were selected from the videotape. The basic selection criteria were: (a) correctly produced Kaneko maneuver (see the Appendix); (b) well focused image; (c) clearly visible externally induced vibrations of the vocal folds with highest possible amplitude; (d) VKG examination covering the whole frequency sweep (no interruptions of the VKG image by the standard image); (e) approximately constant VKG measurement position during the sweep. A single sample for each frequency sweep was chosen for analysis. That sample was digitized and fed into a PC using a videoboard (Miro PCTV) and saved as an AVI file. A CorelScript program under CorelPhotoPaint 8 software was
written in order to extract successive video fields from the AVI files and save them into a sequence of bitmaps (250 bitmap images per 1 sweep of 5 s in duration). Each 20 successive bitmaps then were concatenated (concatenating more than 20 bitmaps into a single image appeared unpractical for the software used), forming a “train” of video fields representing altogether the vocal fold vibration pattern within the time sequence of 400 ms, and saved as a new image [a section of such an image with three concatenated video fields is given in Fig. 2(b)]. From these images, after an adjustment of an optimal contrast, amplitudes of the externally induced vocal fold vibration were extracted.

Extremes of the displacements of the vocal folds and ventricular folds were read manually for every period [Fig. 2(d)] using a cursor in SigmaScan (Jandel Scientific) software. Each value was represented by a pair of pixel values \(x, y\) (\(x\) pixel corresponding to the position of the vocal folds, \(y\) pixel corresponding to a specific time). The data were processed in SigmaPlot (Jandel Scientific) software. Here, the \(y\) pixel values were converted to time values in milliseconds, amplitudes were calculated as a half of the difference between the displacement extremes for each vibration period, duration of periods was calculated and converted to frequency values in Hz, best fit of the measured frequency values was done using the known parameters of the sweep (accuracy better than \(\pm 1\) Hz), and a graph of the frequency response was plotted. Further on, the data were interpolated, smoothed via the Kernel smoothing algorithm, and plotted.

Resonance frequencies \((F_r)\) and half-power (3-dB) bandwidths \((\Delta F_r)\) were measured from the smoothed response curves. \(F_r\) were found as frequencies at which the amplitude reached maximal value \((A_{\text{max}})\), \(\Delta F_r\) was measured as difference of two frequencies around \(F_r\) at which the amplitude was equal to \(A_{\text{max}}/\sqrt{2}\) (Anderson, 1967; Herlufsen, 1984).

2. Videostroboscopy

Videostroboscopic records representative of selected discrete frequencies were digitized and saved as AVI files. Four phases of a vibration cycle were selected (maximal and minimal displacements of the vocal folds, or other laryngeal structures, and two intermediate states) and composed into an image using Corel PhotoPaint 8 software. Sketches outlining borders of the laryngeal structures in their extreme positions were drawn carefully by hand using CorelDraw 8 software. No quantitative analysis of the stroboscopic images was done.

III. RESULTS

A. 100–400 Hz frequency sweep

Frequency response function of the left vocal fold for the \(100–400\) Hz sweep is shown in Fig. 3(a). Three resonance peaks with maxima at 114, 171, and 241 Hz are clearly pronounced here. The first resonance maximum is \(4\) Hz higher than the intended phonation frequency \(110\) Hz. The bandwidth of the first resonance peak was impossible to measure since the maximum was too close to the lower limit frequency of the sweep. The second and third resonance maxima were identified at 171 and 241 Hz with bandwidths of 44 and 45 Hz, respectively.

The response of the right vocal fold is given in Fig. 3(b). It reveals two resonance peaks with central frequencies of 164 and 238 Hz and bandwidths of 37 and 41 Hz, respectively. An increase in the vibration amplitude is evident also around \(100\) Hz; this resonance falls, however, partially below the low frequency limit of \(100\) Hz and therefore its central frequency and bandwidth cannot be identified. Smoothed frequency response functions of the vocal folds are set side by side in Fig. 3(c). The measured resonance frequencies \((F_r)\) and bandwidths \((\Delta F_r)\) of both the vocal folds are summarized in Table I.

B. 50–200 Hz frequency sweep

In order to find out the dynamic characteristics of the vocal folds below \(100\) Hz, the data from the \(50–200\) Hz
frequency sweep were analyzed. Smoothed frequency response functions of both the vocal folds for this sweep are shown in Fig. 4. Large left–right asymmetry is evident in the graph. The left vocal fold exhibits two distinct resonance peaks around the central frequencies of 77 and 104 Hz. The right vocal fold shows a weak maximum at 58 Hz and a strong resonance peak with two local maxima at 92 and 100 Hz. Bandwidths of these resonance peaks were not measured due to their nontrivial shape. A local minimum, suggesting an antiresonance, is found in the right vocal fold at the frequency of 75 Hz.

The second resonance peak in both the vocal folds observed in Fig. 4 corresponds to the first resonance peak found in Fig. 3 in the previously described 100–400 Hz sweep. The resonance frequencies from these two measurements are close, but not identical (the observed difference in the left vocal fold is 10 Hz), which may suggest that the vocal folds were adjusted slightly differently during the two measurements.

Besides the vocal folds also ventricular folds clearly responded to the external excitation (Fig. 5). Maximal vibration amplitudes of the ventricular folds were more than twice as large as those of the vocal folds within this frequency range. Maxima were identified at 67 and 72 Hz for the left and right ventricular fold, respectively. Figure 5 reveals local maxima also at 80, 105, and 125 Hz (left) and 94, 108, and 129 Hz (right); these resonance peaks are relatively weak, however. The resonance frequencies of the vocal folds and ventricular folds identified from the 50–200 Hz sweep are summarized in Table II.

### C. Laryngostroboscopy

Laryngostroboscopic investigation was used to find out the vibration shapes of the vocal folds during the external excitation at distinct frequencies in order to distinguish the resonance modes of vibration. A laryngostroboscopic image of the studied vocal folds in the neutral phonatory position is presented in Fig. 2(a); its sketch is given in Fig. 2(c). The larynx appeared generally normal, it differed slightly from an ideal laryngeal outlook in two features: (1) there was a slight left–right asymmetry, especially in the position of the arytenoid cartilages [this finding may not be considered unusual, however, since some degree of asymmetry is observed practically in all larynges (Hirano et al., 1989; Lind-....]
the vocal folds were slightly bowed, thus the glottis remained slightly open in the neutral phonatory position. For the purpose of this study, the bowing was not considered as an impediment but rather an advantage since it eliminated the collision between the vocal folds that would otherwise perturb their oscillations. Sequences of laryngostroboscopic images at frequencies 50, 75, and 100 Hz are presented in Figs. 6 and 8, sketches extracting the oscillation of the laryngeal tissues from these images are given in Figs. 7 and 9.

Excitation at 50 Hz: At this frequency, large oscillations of the laryngeal collar, especially the aryepiglottic folds and arytenoid cartilages, were visually dominant in the stroboscopic view [Figs. 6(A)–(F), Fig. 7 top]. Vocal folds oscillated as a unit with other laryngeal structures. Left–right and anterior–posterior phase differences in oscillation were visible among the laryngeal structures; no analysis of these phase shifts was done, however.

Excitation at 75 Hz: Oscillations of the arytenoid and aryepiglottic folds were dominant at this frequency, their amplitudes were, however, smaller compared to the frequency of 50 Hz [Figs. 6(G)–(L), Fig. 7 middle]. Ventricular folds or, more accurately, the spatial distance between the medial borders of the ventricular folds exhibited large vibrations [this is partially obscured in the stroboscopic images; VKG image L reveals also large oscillations of the ventricular folds at 75 Hz. Opening–closing response of the vocal folds is apparent at 100 Hz (M–O). (See the sketches in Fig. 7.)

Excitation at 100 Hz: At this frequency only 10 Hz lower than the reference phonation frequency the vocal folds responded by a clear opening–closing movement—can be seen at 100 Hz.

Excitation at 110 Hz: Excitation frequency matched the reference phonation frequency here. A clear opening–closing

FIG. 6. Series of laryngostroboscopic images of the externally excited laryngeal vibrations at the frequencies of 50, 75, and 100 Hz. Five successive phases of the vibration cycle are shown; images A and E, G and K, and M and Q represent the same phase at the beginning and the end of the stroboscopic cycle. Videokymographic images at the bottom (F,L,R) show vibration of the laryngeal structures at the positions marked in the images E, K, and Q. Oscillations of the aryepiglottic folds and arytenoid cartilages are dominant at 50 (A–E) and 75 Hz (G–K). The VKG image L reveals also large oscillations of the ventricular folds at 75 Hz. Opening–closing response of the vocal folds is apparent at 100 Hz (M–O). (See the sketches in Fig. 7.)

FIG. 7. Sketches showing positions of the laryngeal structures at two opposite phases of the vibratory cycle at 50, 75, and 100 Hz (solid versus dashed lines, extracted from images A and C, G and I, M and O in Fig. 6, respectively). Large amplitudes of vibration of the aryepiglottic folds and arytenoid cartilages are evident at 50 Hz (top), these amplitudes successively decrease with increasing the external driving frequency to 75 (middle) and 100 Hz (bottom). Large amplitude of the left ventricular fold can be observed at 75 Hz. An evident response of the vocal folds—the opening–closing motion—can be seen at 100 Hz.
Excitation at 175 Hz: Here, the external frequency was very close to the second resonance frequency of the vocal folds (in accordance with Fig. 3). The vocal folds responded with lateral-medial oscillations encompassing two half-wavelengths along the glottal length–anterior and posterior parts of the glottis oscillated with opposite phases [Figs. 8(G)–(L), Fig. 9 middle].

Excitation at 250 Hz: Here, the external frequency was close to the third resonance frequency of the vocal folds (in accordance with Fig. 3). Despite the small amplitude of the vibration it was possible to identify response in the vocal folds showing lateral-medial oscillations encompassing three half-wavelengths along the glottal length–middle part of the glottis oscillated in an opposite phase to the anterior and posterior parts [Figs. 8(M)–(Q), Fig. 9 right].

IV. DISCUSSION

The results bring an encouraging message: principally, it is possible to use laryngoscopy for obtaining more detailed information on dynamic characteristics of the vocal folds. Three distinct resonance frequencies of the vocal folds were found around 110, 170, and 240 Hz. The first resonance frequency around 110 Hz corresponded to the frequency of phonation. The resonance frequencies were found to be associated with different modes of vibration of the vocal folds. Figure 9 reveals the associated resonance mode shapes as seen in the laryngoscopic view; these three mode shapes can be designated as modes \( x \)-1, \( x \)-2, and \( x \)-3, respectively (1, 2, 3 meaning the number of the half-wavelengths along the longitudinal axis of the vocal folds).

The resonance frequencies show an interesting relationship: it can be seen from the results in Table I, particularly in the case of the left vocal fold, that the relationship of the resonance frequencies \( F_{11}:F_{r2} (114:171 \text{ Hz}) \) is exactly 2:3, the relationship \( F_{11}:F_{r2}:F_{r3} (114:171:241 \text{ Hz}) \) is close to 2:3:4. Theoretically, if all the modes associated with these resonance frequencies would be excited simultaneously during phonation, a complex vocal fold vibration with a resulting subharmonic frequency of \( F_{11}/2 \) (57 Hz) would be produced. This finding might be related to an \( F_{0}/2 \) subharmonic phonation which was found in the same subject when phonating with slightly abducted vocal folds at high airflow volume velocities. The complex vibration pattern of the vocal folds typical for this phonation was described in detail in our previous study (Švec et al., 1996). The 2:3:4 relationship is, however, suspected to be not a general but rather specific feature of the vocal folds investigated here, since the phonatory maneuver of Švec et al. (1996) was found to lead to the \( F_{0}/2 \) subharmonic phonation not in every subject.

The laryngoscopic findings do not support the hypothesis of Ishizaka (1988) that the \( x \)-11 mode is related to the second resonance frequency of the vocal folds; it was rather the \( x \)-2 mode which was found to play the role here. The out-of-phase oscillations of the upper and lower margins of the vocal folds, typical of the \( x \)-11 mode, were not distinguished in the present study since the lower margin remained hidden in the laryngoscopic view and it was impossible to detect its movement. For the same reason, it was impossible to clearly specify whether the \( x \)-10 or the \( x \)-11 mode (or
their combination) is responsible for the first resonance peak. The recent analysis of Berry (in press) suggests that all the modes from the $x$-1 class ($x$=10, $x$=11, $x$=12, etc.) cluster into a joint resonance peak which makes them practically undistinguishable in the laryngoscopic view. This result would correspond with our difficulties with distinguishing these modes.

Certain discrepancies were found in our data between the 50–200 Hz and 100–400 Hz sweeps. The responses of the vocal folds were expected to be similar in the overlapping range 100–200 Hz. Comparison of Fig. 3 with the vocal folds were expected to be similar in the overlap-

The recent analysis of Berry (in press) suggests that all the modes from the $x$-1 class ($x$=10, $x$=11, $x$=12, etc.) cluster into a joint resonance peak which makes them practically undistinguishable in the laryngoscopic view. This result would correspond with our difficulties with distinguishing these modes.

Another and more serious discrepancy is that the resonance peak around 170 Hz, which is clearly pronounced in both the vocal folds in the 100–400 Hz response curve [Fig. 3(c)], is not present in the 50–200 Hz response of the vocal folds (Fig. 4). This discrepancy might be attributed to at least two factors. First, the amplitude of the vibrations of the shaker, which was reduced in the 50–200 Hz sweep, might have been too small to excite the vibration of the vocal folds at this resonance frequency to a laryngoscopically detectable level. Second, an even more plausible origin of this discrepancy might be due to slightly different measurement position used for the two VKG investigations. Whereas in the 50–200 Hz sweep the VKG measurement position was close to the middle of the glottis (Fig. 6, images E, K, Q), which is the nodal point of the $x$-2 mode at which the amplitude is minimal, in the 100–400 Hz sweep the measurement line was placed in a more anterior part of the glottis (Fig. 8, images E,K,Q) where the amplitude of the $x$-2 mode is maximal. Such sensitivity of the results to the VKG measurement position is one of the pitfalls of the VKG method used here. A more complete list of the potential complications and pitfalls is given in the Appendix.

Not only the vocal folds but also other laryngeal structures apparently responded to the vibrations applied externally on the neck. Below 100 Hz, the amplitude of vibration of the vocal folds was much smaller than that of the ventricular folds (their resonance frequency was identified to be close to 70 Hz; see Table II) as well as the aryepiglottic folds and arytenoid cartilages (their resonance frequency is suspected to be close to 50 Hz in this experiment). Different parts of the larynx thus appear to be tuned to different resonance frequencies. The resonance and antiresonance peaks of the vocal folds found below 100 Hz (Fig. 4) are suspected to have the origin in an interaction of the vocal-fold vibration with the vibration of the adjacent laryngeal structures. The large left–right asymmetry in the vocal fold response below 100 Hz (Fig. 4) is assumed to be related to the clear asymmetry of the laryngeal structures shown in Fig. 2(a).

In future studies it is important to investigate the dynamic characteristics of the vocal folds at various frequencies of phonation and in more subjects. More detailed information on dynamic characteristics of the other laryngeal structures, like, e.g., the ventricular folds, could also be helpful since these structures contribute to phonation in certain singing styles (Fuks et al., 1998; Lindestad and Södersten, 1999) or in patients with voice disorders (Kruse, 1981; von Doersten et al., 1992; Schutte et al., 1998; Švec et al., 1997, 1999). For a more extensive analysis, however, it is desirable to employ automated or semi-automated image detection (Wittenberg, 1997, 1998; Saadah et al., 1998; Larsson et al., 1999) instead of manual analysis of the images which is exceedingly time consuming. Certain other questions remain to be answered, e.g., whether the direction of the frequency sweep does not influence the observed resonance properties. In general, however, the method presented here appears useful and promising for studying the dynamic characteristics of the larynx.

ACKNOWLEDGMENTS

The paper was presented to the 2nd International Conference on Voice Physiology and Biomechanics, March 12–14, 1999, Berlin, Germany. The authors are grateful to the Sound Studio of the Academy of Performing Arts in Prague for providing us with a part of the equipment needed for the experiment. The research has been supported by the Grant Agency of the Czech Republic (GA CR), project no. 109/98/K019.

APPENDIX: GENERAL COMPLICATIONS AND PITFALLS OF THE METHOD PRESENTED

A. Vocal folds and the Kaneko maneuver

(1) Vocal fold tuning: There is no control of the vocal fold tuning in the neutral phonatory position, thus it is not certain that the tension of the vocal folds remains the same as compared to the actual phonation. The difference in the tension may result in a change of the dynamic characteristics. The accuracy of the Kaneko maneuver in living subjects remains to be specified in this respect.

(2) Degree of adduction: Investigated subject might tend either to abduct (open) or to hyperadduct (press together) the vocal folds during the neutral phonatory position while holding the breath. In order to avoid this tendency, laryngoscopic view was monitored and provided as a feedback to the experimental subject as well as to the examiner. The Kaneko maneuver and the neutral phonatory position were judged correct if the vocal folds were kept essentially in the same position as they were during the preceding phonation.

(3) Airflow: Even a small amount of glottal flow might influence behavior of the vocal folds, thus the glottal flow shall be avoided when the vocal folds are in the neutral phonatory position. Small airflow velocities are, however, difficult to perceive by the examined subject when the shaker vibrations are applied onto the neck. In our case, the far-from-sinusoidal vibration pattern of the vocal folds seen in the videokymogram R in Fig. 6 with an indication of a shear movement of the vocal folds and the occurrence of mucosal waves leads us to suspect that some glottal flow might have...
distorted the externally (sinusoidally) excited vibrations at that particular moment during the VKG examination.

B. Laryngoscopy

An important aspect is the constant measurement position during VKG examination. The difference in the VKG measurement position is suspected to be the main origin of the discrepancies found between the two sweeps. A full-image high-speed video recording system would be an alternative which would allow to select the proper measurement position after the examination as well as to correct the measuring position in case of some unpredictable motion (of the endoscope or of the examined subject) during the examination (Wittenberg et al., 1995; Wittenberg, 1998; Larsson et al., 1999).

Another problem is that the laryngoscopic view does not allow to reliably identify vertical movements of the vocal folds and distinguish, e.g., the x-10 versus x-11 mode, or the theoretically described z-modes of the vocal folds (Berry et al., 1994; Berry and Titze, 1996). Also, a good laryngoscopic view of the larynx and tolerance of the endoscope might be problematic in some subjects.

C. Shaker

For safety reasons, it is not recommended to fix the shaker to the living larynx. Due to this, the contact force as well as the position of the shaker on the neck may vary slightly (laterally or vertically) which could alter the excitation force acting on the vocal folds. The variation of the contact force was suppressed here by monitoring its value and using it as a feedback. The shaker placement is assumed not critical for investigating the dynamic characteristics of the vocal folds: Kaneko et al. (1981) reported no significant differences in the resonance properties of the vocal folds when the position of the shaker on the larynx was varied.

At constant excitation force, the amplitude of vibrations of the shaker is inversely proportional to the second power of the length change on modal properties of the vocal fold tissues (Berry et al., 1994; Berry and Titze, 1996). This leads to excessive vibration amplitudes at low frequencies and small displacement amplitudes at high frequencies. Therefore two different sweeps were used in the present study. The highest possible excitation force was used for the 100–400 Hz sweep in order to achieve maximal oscillatory response in the laryngeal tissues. For the 50–200 Hz sweep, the input power of the shaker was reduced otherwise the excessive shaker amplitudes at low frequencies (excursions of ca. ±3 mm at 50 Hz) caused uncomfortable sensations in the subject. It appears desirable to compensate for this phenomenon in future studies.

1In the x-i/j notation, “x” designates oscillations in the lateral-medial direction and the i/j indices give number of oscillatory half-wavelengths occurring along the horizontal and vertical dimensions of the vocal folds (i.e., length and thickness), respectively. For a more detailed description and examples see, e.g., Titze and Strong (1975); Titze (1994); Berry et al. (1994); Berry and Titze (1996); Berry (in press).

acoustic spectra and inverse filtering.’’ Karolinska Institute, Huddinge University Hospital Phoniatri and Logopedic Progress Report 11, 17–26.