Glissando: Laryngeal Motorics and Acoustics

Ulrich Hoppe, Frank Rosanowski, Michael Döllinger, Jörg Lohscheller, Maria Schuster, and Ulrich Eysholdt

Erlangen, Germany

Summary: The objective of this study was to investigate the laryngeal mechanisms and the acoustical signal during a glissando. In particular, glottal length, maximum glottal area, and vibratory amplitudes during a glissando maneuver of a healthy male adult were measured. An endoscopic high-speed system combined with a laser projection device was used to obtain quantitative data both in the time and spatial domains. Simultaneously to the endoscopic investigation, the acoustic signal was recorded. Fundamental frequency and sound pressure level derived from the acoustic recordings were compared to vocal fold length and glottis area derived from the high-speed recordings. Results were used for interpretation of the phonation mechanism during glissando by means of laryngeal and acoustic parameters. The transition between the chest register and the falsetto register was identified by the absence of vocal fold contact. A rather early onset of the falsetto register was observed at 160 Hz. Although fundamental frequency of the vocal folds increased linearly even at the transition point, sound pressure level dropped down.

These data represent the first ever quantitative description and interpretation of the glissando based on both voice properties and laryngeal motorics. In the presented example of an untrained singer, the falsetto sets in at comparatively low frequencies. Although the chest-falsetto transition is rather smooth for laryngeal motorics and voice pitch, a sudden drop of voice intensity was observed.

Key Words: Glissando—High speed camera—Vocal fold length—Pitch—Quantitative laryngoscopy.

INTRODUCTION

The ability to change voice pitch is an important precondition for oral communication and is used for producing prosodic features in speech and pitch variation in singing. Although the pitch of comfortable phonation is primarily determined by the actual size of individual anatomical structures, the frequency range can be substantially extended by voice training. An important aspect of voice training is to enable the singer to stabilize their voice
Phonation accompanied by a monotonic increase (or decrease) of pitch within a time range of several seconds is called glissando. Essentially, the pitch can be regarded as the perceptual analog on of the fundamental frequency of the vocal fold vibrations (ie, 1/duration of time needed for one oscillation cycle). From a physical point of view, this corresponds to a change of the eigenfrequency of the vocal folds, which is primarily determined by the length, tension, and mass of the individual vocal folds. Subglottal pressure is just a marginal parameter influencing the pitch. As neither mass, length, nor tension can be varied separately, the details of the underlying mechanisms for pitch variation remain unknown. The study of glissando maneuvers is particularly interesting for clinical purposes because it allows the investigation of the nonstationary voice, which is closer to the everyday condition.

Widespread stroboscopy is a useful tool for investigation of regular vibrations. However, it does not allow for the examination of nonstationary vocal fold movements occurring, for example, during a glissando. In principle, this problem can only be solved with real-time methods, the most advanced being digital high-speed glottography. Using current systems, up to 10,000 images per second can be recorded. They are particularly helpful for nonstationary effects occurring in a large variety of physiologic voice phenomena and voice disorders.

High-speed systems allow the determination of voice parameters in the time domain. Until recently, the derivation of absolute metric units within the endoscopic recordings was impossible. However, our group has developed a laser projection system that enables determination of the actual size of objects within the endoscopic images. Combined with a digital high-speed camera, it allows the fully quantitative investigation of the phonation process even when no stationarity is present.

This paper describes a method that allows the identification of laryngeal mechanisms responsible for the pitch variation during a glissando. As an example of application, a single healthy male subject was examined.

**MATERIALS AND METHODS**

**Laser Projection System**

To determine absolute measures within the endoscopic recordings, a laser projection system (LPS) consisting of a portable laser unit and a projection device was designed by our group. The laser unit was a battery-driven semiconductor laser diode with a wavelength of 633 nm (GaAs). The laser beam was channeled through a light pipe to the projection device shown in Figure 1. A mirror of 50% reflectivity split the laser beam: The first part of the beam propagated through the glass plate mounted on the bottom of the LPS, and the second part of the beam was reflected at a mirror with 100% reflectivity. The mirrors were adjusted in parallel. Hence, the two
FIGURE 2. Sequence of eight images of a high-speed recording during sustained phonation. For clarity, only every fourth frame is shown. Laser spots are visible on the subject’s left vocal fold as marked within the first frame. The scale shown in the upper right figure marks a length of 5 mm.

Endoscopic Investigation

For the endoscopic investigation, the LPS was mounted on the tip of a rigid 90°-endoscope (Wolf Corp., Germany) and attached to a digital high-speed video camera as shown in Figure 1. The spatial resolution of the camera was set to 128 × 64 pixel (128 from left to right, 64 from dorsal to ventral). The temporal resolution was 0.27 ms, which corresponded to a frame rate of 3704 frames per second. The high-speed camera allowed continuous recording for up to 8 seconds. A 250-Watt Xe-light lamp was used as a light source.

A 31-year-old healthy male volunteer with a perceptual normal voice and a mean speaking fundamental frequency of 102 Hz served as experimentee. He was used to endoscopic investigations and exhibited no clinical signs of voice disturbances. However, he had not undergone any kind of vocal education. For the test, he was asked to sustain a vowel [i:] (as in “beat”) at a comfortable pitch. Then he was asked to perform a glissando from low to high pitch after the endoscope had been inserted into the oropharynx. He was asked to start phonation within the chest register in a comfortable way and to increase frequency until phonation becomes impossible.

Acoustical Recording

Simultaneously to the high-speed recording, the acoustical signal was recorded by a condenser microphone (type B&K 4129, Bruel & Kjaer Corp.) mounted on the top of the camera at a distance of approximately 20 cm in front of the experimentee. The acoustic signal was digitized (44.4-kHz sample rate, 16-bit resolution) and saved to the hard disk. Fundamental frequency (F0) and sound pressure level (SPL) were determined using a commercially available speech analysis software (Dr. Speech, Tiger Electronics, Inc., USA).14,15

Image Processing

Figure 2 shows eight frames of a high-speed sequence covering a time interval of 7.56 ms. In this figure, only every fourth frame of the sequence is shown. The sequence corresponds to one complete glottal cycle starting and finishing with the maximum open glottis during the sustained phonation with a fundamental frequency of approximately 125 Hz. The primary aim of image processing was to automatically extract the glottis as a function of time. This was performed by a knowledge-based...
algorithm enhancing, which enhances image contrast and allows automatic recognition of the vibratory pattern of the vocal folds. This algorithm was shown to be stable even when the images of the high-speed sequence were dark. For the purpose of this study, the previously described algorithm was modified in order to focus on extraction of the glottis for a predefined high-speed sequence. Three different parameters were evaluated:

1. Glottal length
2. Maximum glottal area ($A_{\text{max}}$)
3. Vibratory amplitudes

For the most open glottis condition, the glottal length was determined from the digital images in time-steps of 50 ms. $A_{\text{max}}$ was derived automatically as the glottal area for this condition. Vibratory amplitudes were determined as the distance of two adjacent points on the medial line (50% from anterior to posterior).

RESULTS

Figure 3 shows the microphone signal of the acoustical recording for the sustained phonation over a time period of 30 ms. Arrows in the upper left corner mark the time instants corresponding to the frames shown in Figure 2. They were corrected for different velocities of light and sound in air. The period duration of 8 ms corresponds to a fundamental frequency of 125 Hz.

Figure 4A (top) shows the fundamental frequency ($F_0$) as a function of time during the glissando. $F_0$ increases monotonically within a time range of 1.3 s from 125 Hz to 244 Hz. Despite the fact that no instructions were given to the subject about the amount of frequency shift, the frequency shift corresponds nearly to one octave. Between 580 ms and 1500 ms, a nearly linear pitch increase, with a slope of approximately 0.1 Hz per millisecond, can be observed. Linear regression is shown as a dashed line. The horizontal line within Figure 4A marks the frequency of the nearly constant frequency of 125 Hz prior to the glissando. Two vertical reference lines are introduced indicating particular instants: (1) The first vertical reference line, at 350 ms, marks the intersection of the constant frequency line with the extrapolated line of the linear frequency increase. (2) The second vertical line at approximately 580 ms marks the beginning of the linear pitch increase.

The sound pressure level (SPL) as a function of time is shown in Figure 4B. During the first 350 ms, the sound pressure level tends to increase. Between 350 and 600 ms (corresponding to $F_0$ of 130 Hz
and 150 Hz, respectively), a decay of about 5 dB is observable. Above 600 ms, SPL tends to increase monotonically. However, larger short-term fluctuations occur.

During the complete high-speed sequence, the whole glissando maneuver consists of more than 5900 single frames, wherein about 300 vibratory cycles can be observed. Figure 5 shows the glottal length (A), maximum glottal area (A<sub>max</sub>) (B), and vibratory amplitudes (C) as a function of time. Using the LPS, derivation of metric units (mm, mm²) was possible. The glottal length remains nearly constant during the first 85 ms and increases steeply up to 1400 ms. For A<sub>max</sub>, a tendency to increase can be observed during the first 580 ms. This is followed by a nearly linear decrease with a slope of 1.7 mm² per 100 ms. Vibratory amplitudes exhibited a similar behavior: Until 600 ms, amplitudes remain nearly constant; afterward, a linear decrease with a slope of about 0.14 mm per 100 ms can be observed.

Additionally to the above-described quantitative investigation, the vibratory pattern of the vocal folds was visually analyzed. Up to approximately 600 ms, the vocal folds vibrate with clearly detectable mucosal waves and glottal closure is complete within each glottal cycle. Compared to sustained phonation, mucosal waves are reduced during the glissando. At about 600 ms, the glottal closure becomes incomplete along the whole glottal axis and a resting area of approximately 10 mm² is observed. This resting area remains nearly constant until the end of the recording.

**DISCUSSION**

In this paper, we presented the investigation of an upward glissando performed by a male subject from B to b during a time of 1700 ms. Acoustical and endoscopic parameters were measured simultaneously and described fully quantitatively. Generally, during the glissando, glottal length increased of about 3 mm and both the maximum glottal area and vibratory amplitudes decreased. The sound pressure level exhibited a nonmonotonic behavior and varied between 66 and 74 dB(a).

Considering the pitch curve, three time intervals as separated by the reference lines in Figures 4 and 5 were identified:

(1) During the first 350 ms, the fundamental frequency F₀ remained more or less constant. A slight increase of the vibratory amplitude, maximum glottal area, and sound pressure level was observed. Glottal length remained constant. This interval can be regarded as the preparation phase of the glissando where the subglottal pressure is deliberately increased. Consequently, vocal fold amplitudes and sound pressure level increases.

(2) Between 350 ms and 580 ms glottal length, vibratory amplitudes and maximum glottal area remained constant, pitch increased linearly, and sound pressure level dropped down. This corresponds to the initial phase of the glissando where the vocal fold tension begins to increase. Consequently, the pitch goes up on the one side and sound pressure level drops down on the other side.
The third phase of the glissando above 580 ms was characterized by a linear pitch increase. As vibratory amplitude and $A_{\text{max}}$ decreased, SPL was observed to increase. Large fluctuations of SPL are indicative of an instable voice in this phase. A considerable glottal length increase could not be observed until 850 ms.

The above-described classification of the glissando is initially based on the fundamental frequency curve of the acoustical signal. It is also reflected in the SPL curve and in most of the laryngeal parameters. However, there is one important exception: The glottal length remains constant over a broad pitch range and does not increases until fundamental frequencies of 170 Hz are reached. Hence, glottal length is the most stable of the observed laryngeal parameters. This behavior confirms that it is not the vocal fold length that contributes primarily to the pitch, but the vocal fold tension. Obviously, vocal fold tension has to be increased during a glissando. In the preparation phase, this increase contributes little to the pitch. However, beginning with the initial phase, the isometric increase of vocal fold tension involves a substantial increase of the pitch. At 170 Hz (i.e., 850 ms), the tension leaves the isometric range and glottal length is increased. Although the pitch of the voice is not affected by this laryngeal change, a sudden decrease of the sound pressure level of 3 dB is present.

Surprisingly, the increase of fundamental frequency during the glissando corresponds to almost exactly to one octave, even though the subject was not instructed to cover a certain frequency range. This phenomenon was observed in other subjects too and may be explained by the octave dominance of the ear physiology.

As the experimentee is an untrained singer, only the chest (or modal) register and the falsetto register must be considered. The transition between both registers is clearly audible. From the different register definitions, the clinician is in favor of the laryngeal one: The chest register is characterized by large vibratory amplitudes of the vocal folds and clearly observable mucosal waves. The falsetto register shows small vibratory amplitudes and complete missing of glottal closure. Vocal folds do not entirely vibrate along the anterior–posterior axis but only along a small fraction. For the presented glissando, the glottal contact vanishes at about 600 ms (i.e., 150 Hz). At the same instant, the register transition is observed when the voice SPL drops down.

A second surprising finding is observed in both registers: The glottal length does not change neither in the chest nor in the basal falsetto register. A measurable length increase can be observed only six semitones above the register transition. This is in contrast to previous studies in trained singers by Hollien and Rubin. This observation may have been missed in other studies where the superior digital high-speed system was not used. Alternatively, this may be explained by physiology because the subject as an untrained voice user exhibits a rather early onset of the falsetto register and may use another kind of motoric laryngeal control. Consequently, as indicated by large short-term fluctuations of the SPL function above 600 ms, the voice sounds “thin” and “instable” within the falsetto register.

This work is the first fully quantitative description of a glissando maneuver based on evidence from both endoscopic and acoustic data. Although the simultaneous analysis allowed the interpretation of most of the topics concerning the laryngeal motorics for the investigated subject, the transferability of the results is highly limited. Firstly, due to different phonation mechanisms of trained and untrained singers, a great variability between subjects has to be assumed. Secondly, there are numerous studies that give evidence for different phonation processes in males and females. For example, mechanical adjustment in moving from chest to falsetto voice source is considerably greater in the male voice. Additionally, the subglottal pressure, which contributes substantially to the kind of phonation, was not investigated. Therefore, studies including measurements of the subglottal pressure in larger groups of subjects are in progress.

**Conclusion**

For the first time, the simultaneous recording of both high-speed endoscopic measurements and acoustical parameters during a glissando was demonstrated. The use of a laser projection system...
enabled the derivation of absolute spatial units (ie, vocal fold length, vibratory amplitudes etc.). The presented data demonstrate that glissando studies provide important information about laryngeal motorics. The simultaneous recording of acoustic and endoscopic parameters during a glissando is particularly interesting for clinical investigation of the laryngeal motorics.

Acknowledgments: The authors would like to thank Dr. L. A. McKenna for her constructive suggestions and comments in preparing this manuscript. The work was supported by grants from the “Deutsche Forschungs Gemeinschaft” (Project-Nr.: DFG EY 15/8 and DFG EY 15/10).

REFERENCES