Vocal Fold Collision Speed *in vivo*: The Effect of Loudness

*Philippe Henri DeJonckere, and †Jean Lebacq, *†*Brussels, Belgium*

Summary: Mechanical impact stress on the vocal fold surface, particularly when excessive, has been postulated to cause the so-called phonotraumatic tissue lesions, such as nodules and polyps. The collision stress between the vocal folds depends on the vocal fold velocity at the time of impact. Hence this vocal fold collision speed is a relevant parameter when considering biomechanical economy of phonation, especially in voice professionals needing a louder voice than normal. Combining a precise photometric measurement of glottal area and simultaneous measurements of translaryngeal impedance (electroglottogram) for identifying the time of the maximum rate of increase of vocal fold contact allows computing the vocal fold collision speed in a wide range of loudnesses. The vocal fold collision speed is - for modal voicing - always smaller than the maximum vocal fold velocity during the closing phase, but it strongly increases with intensity. Moreover, this increase shows a biphasic pattern, with a significant enhancement from a certain value of dB on. Understanding physiological variables that influence vocal fold collision forces provides relevant insight into the pathophysiology and the prevention of voice disorders associated with phonotraumatic vocal hyperfunction.

Key Words: Vocal fold collision−Collision speed−Voice intensity−Photoglottography−Glottal area−EGG.

INTRODUCTION

Mechanical impact stress on the vocal fold (VF) surface, particularly when excessive, has been postulated to cause the so-called phonotraumatic tissue lesions.[1] VF nodules and polyps are the best known examples.[2,3] The maximum area declination rate (MADR) in the closing phase of the glottis during VFs' vibration has been reported as a measure of the impact stress loading the VFs during collision,[4] thus, as a relevant parameter when considering biomechanical economy of phonation.[5]

The collision stress between the VFs can be estimated from basic physical principles[6]:

Assuming the mass of a tissue element at the medial surface of the VF edge to be

$$m = \rho\ \Delta x\ \Delta y\ \Delta z \qquad \text{(Eq. 1)}$$

where ρ is tissue density (1040 kg/m^3) and $\Delta x \Delta y \Delta z$ is a small volume, then, from Newton's second law, the average collision force over an impact interval Δt is

$$F = m\ \Delta v / \Delta t \qquad \text{(Eq. 2)}$$

where Δv is the change in velocity during impact. Jiang & Titze[7] estimated the impact interval to be of the order of

$$\Delta t = T_0\ /\ 10 \qquad \text{(Eq. 3)}$$

where T_0 is the fundamental period.

The velocity change in Equation (2) can be estimated by assuming a sinusoidal motion of amplitude A and radian frequency $\omega = 2\pi F_0$. The maximum velocity, which occurs near impact, is

$$v = \omega A = 2\pi\ F_0\ A \qquad \text{(Eq. 4)}$$

This velocity is reduced to zero during the collision interval, such that

$$\Delta\ v = v\ -\ 0 = 2\ \pi\ F_0\ A \qquad \text{(Eq. 5)}$$

Substituting (1), (3), and (5) in (2),

$$F = 20\ \pi\ A\ F_0^{\ 2}\ \rho\ \Delta x\ \Delta y\ \Delta z \qquad \text{(Eq. 6)}$$

when the impact starts from the phase when the VF velocity is at its maximum value.

If $\Delta y\ \Delta z$ is the impact surface and Δx the depth of the vibrating tissue, then the collision stress is

$$\sigma = F\ /\ \Delta y\ \Delta z = 20\ \pi\ A\ F_0^{\ 2}\ \rho\ \Delta x \qquad \text{(Eq. 7)}$$

To set these ideas on an example corresponding with modal male speech, for an amplitude of vibration of 10^{-3} m, a depth of vibration of 10^{-3} m, and a F_0 of 120 Hz, the stress is 9.4 hPa.

Another approach for estimating the impact stress between the colliding VFs is the Hertz impact theory according to which the maximum contact stress during the collision is approximately proportional to the velocity of glottal closing just before the complete closure (V_c), in line with the formula[8]:

$$\text{Impact stress} = \text{constant}\ .\ Vc^{2/5} \qquad \text{(Eq. 8)}$$

where the constant depends on the material properties and the curvature of impacting surfaces. Results of comparable magnitude were obtained with computer modeling. They are also in good agreement with measurements on artificial VFs as well as with results on excised larynges.[3,9]

In an *in vitro* study by Jiang and Titze,[7] the peak impact stress between the VF of a canine hemilarynx and the vertical plate during self-sustained oscillation was found to be in the range 5−50 hPa. Verdolini et al[10] have reported values in the range 0.3−5.3 kPa for impact stress in canine

Accepted for publication August 21, 2020.

From the *Federal Agency for Occupational Risks, Brussels, Belgium; and the †Institute of Neurosciences, University of Louvain, Brussels, Belgium.
Address correspondence and reprint requests to P.H. DeJonckere, Federal Agency for Occupational Risks, Brussels, Belgium. E-mail addresses: ph.dejonckere@outlook.com jean.lebacq@uclouvain.be

Journal of Voice, Vol. ■■, No. ■■, pp. ■■−■■
0892-1997

https://doi.org/10.1016/j.jvoice.2020.08.025

larynges: the VF impact stress was found to be strongly (and approximately linearly) related to the closed quotient (as estimated by electroglottogram [EGG]), an increase of 0.15 in the closed quotient (CQ) corresponding to about 1 kPa in impact stress.[10] Using an aeroelastic model, Horacek et al report, for conditions comparable to "normal" speech, maximum impact stress values of 2−3 kPa.[3]

Direct measurements on human subjects also yielded a range of 1−5 kPa[11,12]; Verdolini et al[13] found *in vivo* values of 1−4 kPa. Experimenting with excised human hemilarynges aerodynamically driven into self-sustained oscillation and using a dual-sensor intraglottal and/or subglottal pressure probe, Mehta et al.[14] found that, as subglottal pressure was increased, the peak amplitude of the collision pressure increased and typically reached a value below that of the average subglottal pressure, except for subglottal pressure values >60 cm H_2O. The ratio of mean peak collision pressure and/or mean subglottal pressure varied from 0.51 to 0.88 up to about 60 cm H_2O, but increased up to 1.09 at higher values of driving pressure.

Equation 7 shows that collision stress increases with F_0^2; however, both amplitude and depth of vibration are expected to decrease with F_0, making the exact stress uncertain. At habitual speaking frequency, when F_0 remains within a limited range, the *velocity change during impact* is the essential parameter (Equation 2). As velocity is reduced to zero during the collision interval, the peak velocity and the moment at which this peak occurs with respect to the impact are major determinants. Titze hypothesized that the maximum velocity occurs near impact.[4] This is indeed what can be expected in the case of a closed quotient of 0.5, ie, when the closed phase and the open phase have the same duration. It is interesting to know what happens when the closed quotient departs from 0.5.

Hence, determining the closing velocity is an interesting noninvasive method to estimate the impact stress.[8]

Very recently, Horacek et al[15] investigated the velocity of glottal width closing in a male subject during sustained modal phonation at comfortable speaking pitch (110−113 Hz) and loudness, using high speed video at 2000 frames/s with a resolution of 512 × 512 pixels. They observed that the maximum velocity of the glottal closing, given by the absolute value of the time derivative of the glottal width, was followed by *lower* values of glottal width decreasing velocity just before glottal closure. From cycle to cycle, the closing velocity decreased by 40% to 70%, and even to 90%. Their conclusion was that MADR is in fact not an adequate measure of impact stress between the colliding VFs, because the glottal closing speed is noticeably reduced just before the VFs' contact. However, they admit that these results were limited by the low sampling frequency: the closing phase of the glottis indeed comprised only seven area values, which is much too low for accurately computing the derivative.

Hence it is relevant to more precisely know the actual VFs velocity at the time they collide, and the relationship between impact velocity and voicing intensity. It would indeed be desirable to weigh increases in vocal loudness against increases in tissue stress to obtain a cost and/or benefit ratio for certain vocal productions like teaching, acting or speaking in public, and consequently to define the level absolutely requiring electrical amplification. The reasoning is similar to what is well known in occupational medicine for preventing tissue lesions caused by exposure to external vibrations.

MATERIAL AND METHODS

Experimental variables

Glottal area (light flow)

The glottal area was derived from a photometric record obtained by transilluminating the trachea, as described in previous work.[16] The light flux was detected by a nondirectional photovoltaic transducer positioned as dorsally as possible in the pharynx (photoglottography [PGG]).[17] The transducer, a BP104 silicon photodiode (Vishay Precision Group, Malvern, PA), was glued onto a small laryngoscopic mirror (nr. 3). The current produced by the photodiode was preamplified by a current-to-voltage converter with a linear response up to 2 kHz. During VF vibration, the photovoltaic transducer produces a current which is directly proportional to the light flux, hence to the glottal area. However, the precise position and orientation of the photodiode into the pharynx cannot be reproduced from record to record, which means that the amplitude of the absolute value of the photovoltaic current can vary from experiment to experiment, but not within a single voice utterance. For measurement of the glottal area and calibration of the photoglottographic signal, we first need to know the ventrodorsal length of the vibrating glottis, which may be assumed to be stable within the frequency range 100−125 Hz. This ventrodorsal length of the glottis during a vibration cycle is constant for a sustained modal phonation at controlled F_0, and can be measured (in mm) on a stroboscopic picture obtained in the same subject uttering a similar voice sound. In order to obtain this reference, a rigid 90° Wolf laryngeal telescope (4450.57; CE 0124) and an ATMOS Strobo 21 LED stroboscope (Atmos Medizin Technik, Lenzkirch, Germany) were used. The telescope has a magnifying facility, with narrow depth of field and critical sharpness adjustment; scaled paper was filmed at the same focal length, critical care being given to maximal sharpness. This way of proceeding is inspired by Fex et al[18], who used a microscope, and calculated the maximal error of measurement: 4.65 ± 3.10%. With our 90° telescope and the magnifying option, the range of sharpness was found to be at most 3−4 mm at a distance of 40−45 mm. As such, the ventrodorsal length of the glottis was estimated to be 13 mm, in line with the values found by Larsson & Hertegard,[19] who applied laser triangulation, a direct measurement using projection of a dual laser spot on the VF surface.[20] In principle, this method should be more exact, but actually, the beam size (around 1 mm for a distance of 4.5 mm) limits the accuracy. Contrary to the length, the maximal glottal width is

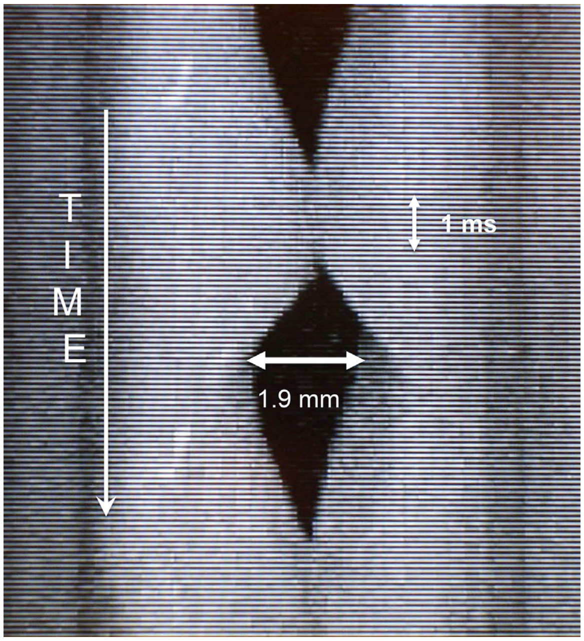

FIGURE 1. Videokymogram at a voice SPL of about 70 dB at 10 cm of the mouth. Closed quotient is about 0.25. Max. glottal width is 1.9 mm.

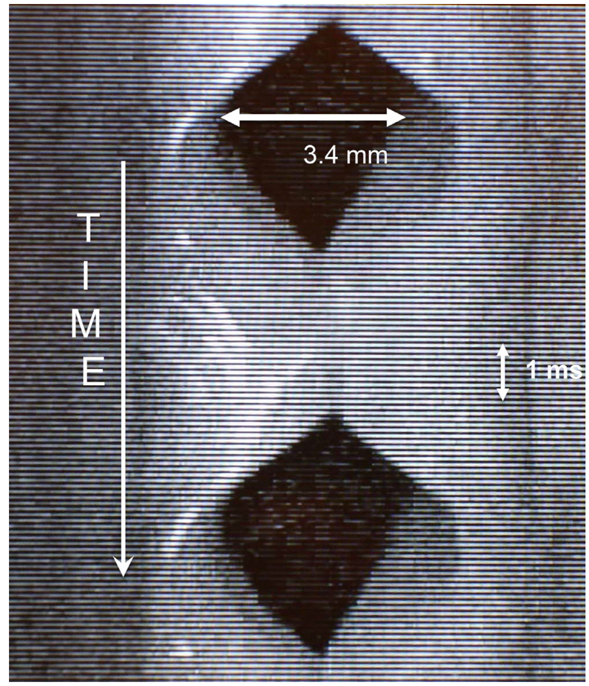

FIGURE 2. Videokymogram at a voice SPL of about 80 dB at 10 cm of the mouth. Closed quotient is nearly 0.50. Max. glottal width is 3.4 mm.

strongly variable and correlated (male subject; modal speaking frequency) with the intensity of voicing, as shown in previous work[21] (Figures 1 and 2). The results of a new calibration, based on the ratio width and/or length and made especially for the present study using videokymography (VKG, see below) with the same rigid telescope and in the same subject, are given in Figure 3, with a Spearman's correlation coefficient of 0.98. While making kymographic recordings, the magnitude of the whole glottal image can be regularly controlled by switching (pedal) the camera mode from VKG to normal (nonstroboscopic) video.

By inspecting still stroboscopic pictures at the time of maximal opening, we found that the contour of the glottal image could be well fitted with an ellipse, the major and the minor axes of which were the ventrodorsal length and the maximal width of the glottis picture respectively.[22] This is illustrated in Figure 4. In this case (Sound Pressure Level (SPL) 71 dB), the difference between the calculated area of the ellipse and the measured area of the glottis is less than 1%.

The sound signal, measured in Vrms, was first calibrated in dB by recording series of short (~ 4–5 s) voice utterances at stable SPL (controlled by visual feedback) at intervals of 5 dB, from ~ 55 dB on. This was achieved by simultaneously recording the sound signal with the microphone and with the Dawe Sound Level Meter. The relationship between the

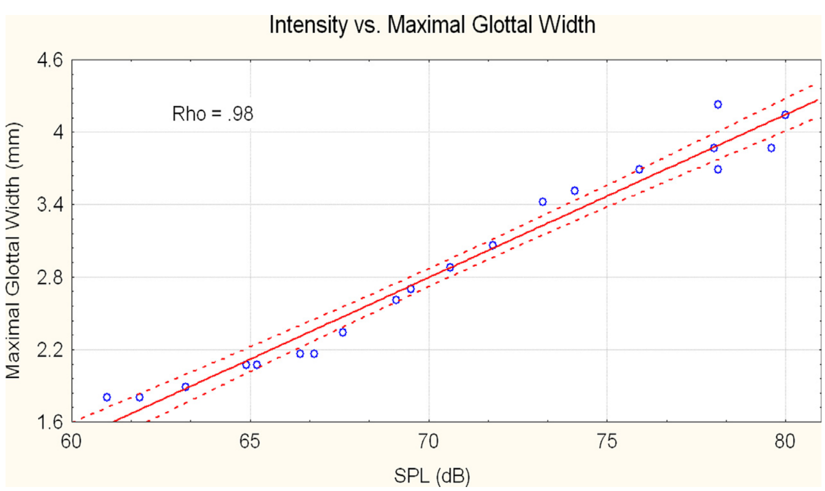

FIGURE 3. Correlation between the max. glottal width (mm) and intensity (dB), measured in 20 calibrated videokymograms. The positive correlation is very strong (Spearman's Rho = 0.98).

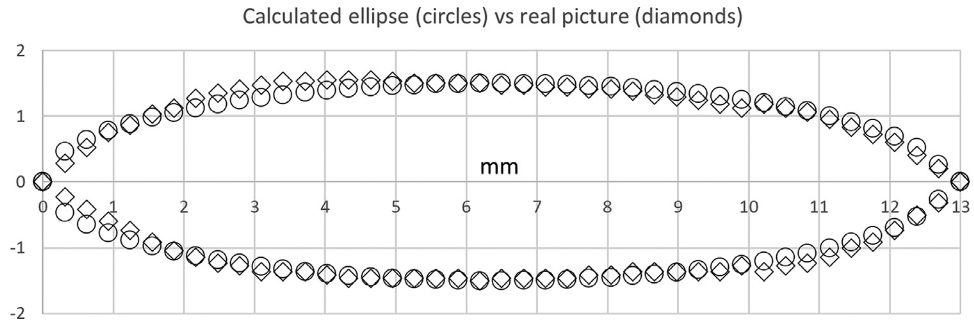

FIGURE 4. Match of the maximal glottal area during a vibration cycle with the calculated contour of an ellipse of similar major and minor axes. The glottal contour is traced out from a real videostroboscopic picture "frozen" at maximum amplitude.

two records was analyzed to deduce the empirical equation relating them. This equation was then used to transform the sound signal from Vrms to dB in all records. The value in dB thus obtained was introduced into the equation derived from Figure 3 to obtain the maximal glottal width (within one cycle) in each record. As the length of the glottis is constant, maximal width and glottal length were used to calculate the area (in mm^2) of the equivalent ellipse by applying the simple geometrical equation of the area of the ellipse, and this value is equivalent to the maximum value of the PGG signal during each cycle, or 100% of the glottal area. All values of the PGG signal, expressed in % of the maximum area, were then transformed into units of area; from area, the half width of the glottis was finally calculated by the equation of the ellipse. Similarly, all values of the derivative of the PGG signal were expressed in rate of change of area (in mm^2/s), from which the speed of the edge of each VF (in m/s) was obtained. This approach makes it easily possible to calculate the speed of each VF edge at its middle length, with the assumption that, in normal conditions, VFs are vibrating approximately symmetrically.

The photometric method also allows defining the closed quotient. The beginning of the opening phase and the end of the closing phase had to be clearly defined. As in previous experiments,[16,17] we considered, according to the method of Gerratt et al,[23] that these boundaries occur when the rising and falling trace intersects a horizontal line drawn at 90% down from the positive peak.

Another point to take into account is the time delay. As very critical time intervals related to the derivative of the PGG-signal are investigated in this study, it is necessary to precisely know the delays introduced by the electronic circuitry and to correct them in the calculations. A function generator (HP model 3311A, HPE, Palo Alto, CA) was used to produce calibrated sine waves added to a constant positive offset. This signal was fed to light up an LED placed in front of the detecting photodiode as used in the records. The aim of the constant base current was to keep the LED lit all the time, so as to avoid a possible threshold effect of the LED when applying low frequency sine waves. The light emitted by the LED was readily detected by the photodiode and the magnitude of the exciting currents was adjusted to produce photodiode signals of the same order of magnitude as those of signals from the glottis. By displaying both the LED current and the photodiode output on the oscilloscope screen, the exact time delay of the whole chain of photo signal recording was measured over a range of frequencies from 80 to 220 Hz. The result was a systematic mean delay of 0.102 ms with no significant change over the tested range of frequencies, with an effective time resolution of 5 μs. All PGG records were corrected accordingly.

High speed video provides a global view of the moving VFs and glottis, but photoglottography gives the most accurate measure of glottal area. The photoglottographic signals are more accurate than those provided by image processing from high speed video[24] or from VKG,[21,25,26] the high sampling frequency allowing adequate time resolution for computation of the derivative (for example Figures 9 and 10), much better than that of the imaging technique.

VKG

Single line scanning of VF vibrations (VKG) is an imaging method based on a special digital camera, fixed onto a rigid 90° endoscope. The system used in the present experiments comprises a Lambert Charge-Coupled Device (CCD) (Lambert Instruments, The Netherlands) Kymocam, a JVC HR J746 MS magnetoscope (The Victor Company of Japan, Yokohama), an Optomic Fibrolux 150 W (Optomic España SA, Madrid, Spain) light source and a monitor.

In the high-speed mode, the video camera delivers images from a single line selected in the whole image, at the rate of approximately 7875/7812.5 line-images/s and 720 × 1/768 pixels resolution, depending on the video format. The resulting high-speed image displays the vibratory pattern of the small selected part of the VF cycle by cycle.[27] Currently available high-speed video systems offer the possibility to extract and display the videokymograms (single line scans) of several selected lines.[28]

Videokymograms at midlength of the vibrating folds allow a reasonably precise measurement of the maximum glottal width (the minor axis of the ellipse), which is - as already mentioned - strongly correlated (male subject, modal speaking frequency) with the intensity of voicing, as shown in previous work dealing with the physiological background of the "messa di voce exercise"[21] (Figures 1

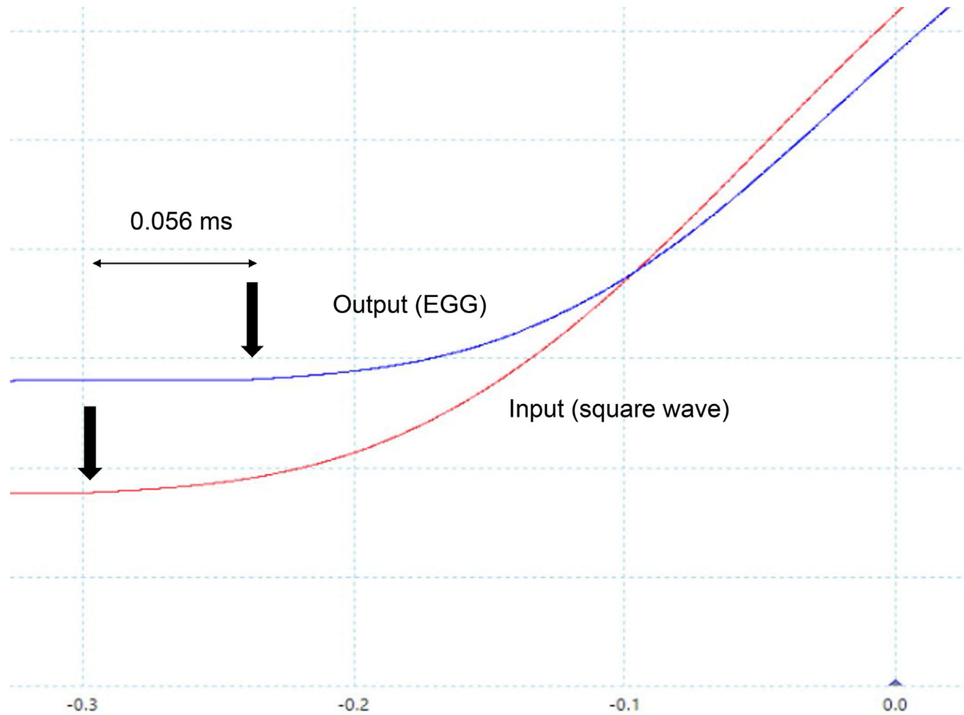

FIGURE 5. Instrumental delay of the EGG signal. Red trace: smoothed square signal applied to the EGG electrodes. Blue trace: output of the EGG device. Vertical scale: 40 mV/div. Horizontal scale: 0.1 ms/div. The response of the EGG signal is delayed by 0.056 ms. (For interpretation of the references to color in this figure legend, the reader is referred to the Web version of this article.)

and 2). In 20 videokymograms extracted from a continuous recording during a 'messa di voce' monitored by PRAAT for the SPL, the value of the maximum glottal width in mm at mid-length was obtained by relating this maximum width (within one cycle) to the (constant) ventro-dorsal dimension of the glottis, displayed when switching the camera mode from VKG to normal video. The correlation coefficient between the maximum glottal width and the SPL is 0.98.

EGG

Electroglottography measures the transglottic electrical impedance using an AC current at a frequency above 100 kHz and monitors the changes in contact surface of the VF. The method is patient-friendly and does not interfere with vocalization. It allows precise phonetic tasks, with acoustic control.[29,30]

The EGG-signal, used as a reference for monitoring the contact surface changes, was detected using a portable electroglottograph (Laryngograph Ltd, London, UK) Model EG90. As for the photoglottogram, the very high sampling frequency makes it possible to accurately compute the derivative. The positive peak of the EGG-derivative indicates the maximum rate of increase in VF contact.

However, as very critical time intervals with respect to this dEGG-peak are investigated in this study, it is necessary to precisely know and correct for the delay related to the electronic circuitry, in the same way as for PGG. This important piece of information cannot be found in the literature.

Values of the impedance of the larynx have been reported by Sarvayya et al.[30] According to these authors, the impedance is purely ohmic for exciting current frequencies in the range 100 kHz to 1 MHz, and its value is typically in the range 100−500 Ω. Our EGG device accepts values of source resistances in the range ~150 to ~500 Ω. Accordingly, a 500 Ω potentiometer was connected between the electrodes and it was adjusted so that the EGG device was in the middle of its characteristics, around 250 Ω. Both rectangular and sine wave signals were applied. Likely due to complex multipole filters inside the circuit, the EGG output showed multiple oscillations when a square input was applied to the electrodes. For this reason, very smoothed rectangular waves were used instead, as shown, on an expanded time-scale, by the red curve in Figure 5, which is a portion of a PicoScope screen. The blue curve shows the output signal of the EGG circuit, shifted to the right by 56 μs. Hundred Hertz sine-wave signals were also applied between the electrodes. Figure 6 shows the input and the output signals of such a sine wave on an expanded time-scale. To best evaluate the time shift between the input and the output, only the very expanded tips of the two sine waves are shown, so that the noise becomes very conspicuous. As in the case of a smoothed rectangular input, the output is shifted by 56 μs relative to the input.

Sound

The voice sound was recorded by a calibrated Dawe Type 1400 H Sound Level Meter (B.S. 3489; Dawe Instruments

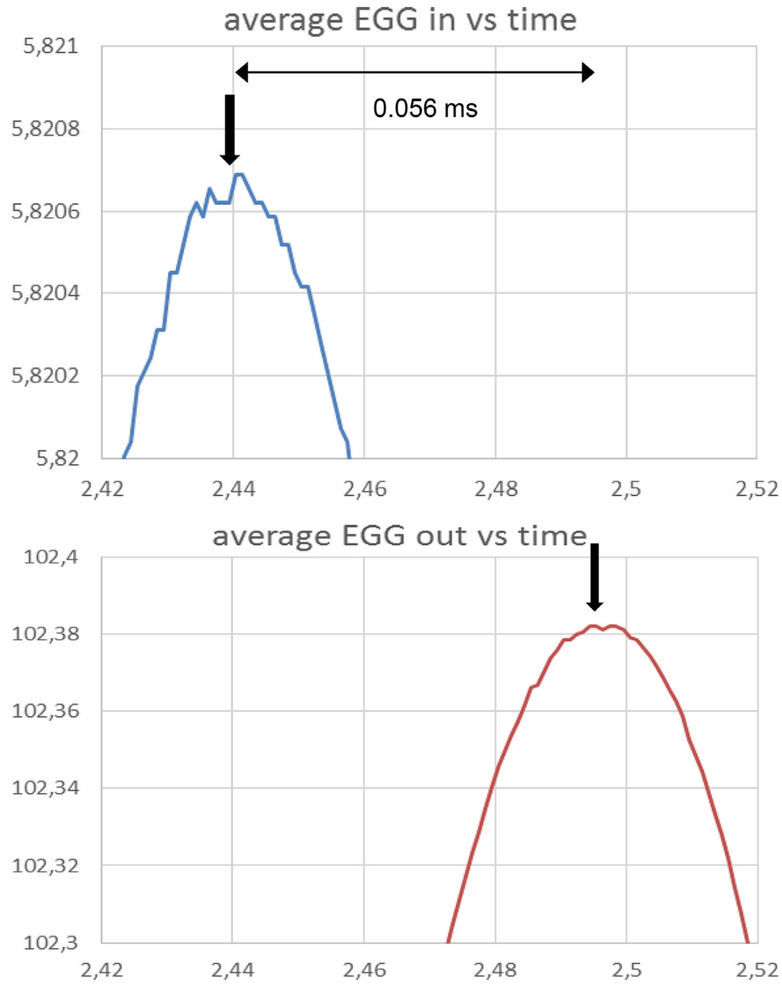

FIGURE 6. Instrumental delay of the EGG. 100 Hz sine wave input. Top panel: input signal applied to the electrodes. Vertical scale: 0.2 mV/div. Horizontal scale: 20 μs/div. Bottom panel: output of the EGG. Vertical scale: 20 μV/div. Horizontal scale: 20 μs/div. As in Figure 5, the response of the EGG signal is delayed by 0.056 ms.

Ltd, London) (frequency weighting A; time weighting S with 1 s time constant) and a Sennheiser MD 421 U microphone 10 cm from the lips, for analysis of SPL and F_0. All experiments took place in a quiet room (background noise 37 dB_A). For calibration of the rms sound oscillogram we used PRAAT 5.3.10, 2012, by P. Boersma and D. Weenink.[31] The calibration was achieved by recording series of short (~ 4−5 s) voice utterances at stable SPL (controlled by visual feedback) at intervals of 5 dB, from 55 dB on.

All signals were recorded by means of a 4-channels Pico Scope 3403D module (Pico Technology Ltd, St Neots, England, UK) driven by the Picoscope 6 programme, at a sampling frequency of 500 kHz on each channel, and stored in a computer.

Vocal material

First a corpus of about 140 recordings was created with short sustained vocal emissions on /ə/ with the photoglottograph *in situ*, and simultaneous EGG and sound monitoring, at spontaneous speaking pitch (F_0 between 95 and 125 Hz) in a large range of loudnesses, however avoiding extremes (intensities <60 and >90 dB at 10 cm of the lips). Out of this corpus, 32 records were selected. Criteria for selection were (1) full display of all traces in the central part of the recording; a full display of all traces means a portion of the record containing a series of consecutive cycles in which all traces were visible within the screen and stable over the whole screen. In practice, it meant selecting a section of about a third to a half of the utterance. (2) absence of drift of the light and impedance traces, (3) minimal noise in the derivative signals and (4) clearly defined peaks (positive for EGG, negative for PGG) in the derivative signals.

The subject was a healthy trained male vocalist, experienced in controlling voicing parameters.[16,17,21,22]

RESULTS

Figure 7 gives an example of an original raw tracing showing the three signals: the photoglottogram (PGG, glottal area), the EGG (translaryngeal electrical impedance) and the microphone signal. Intensity is moderate (70.24 dB), as is the closed quotient (0.35). Fundamental frequency is about 115 Hz. An estimate of the maximum glottal area

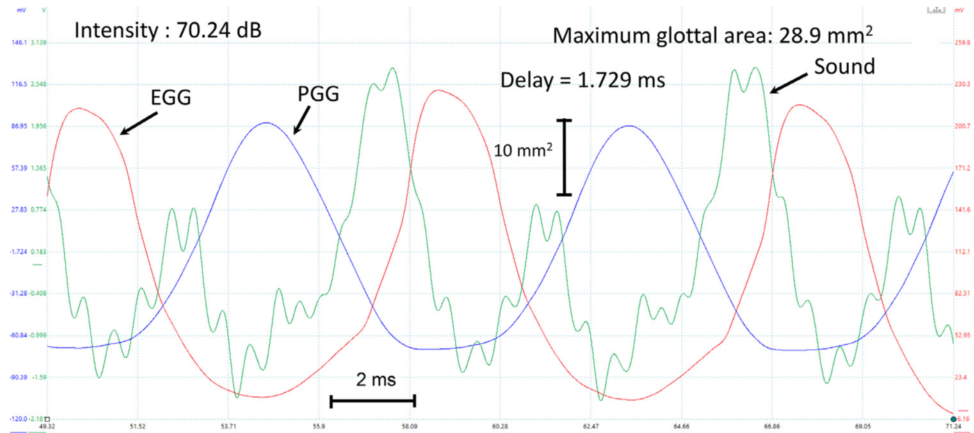

FIGURE 7. Example of an original raw tracing showing the three signals: the photoglottogram (PGG, glottal area), the electroglottogram (EGG, translaryngeal electrical impedance) and the microphone signal. Intensity is moderate (70.24 dB), as is the closed quotient (0.35). The horizontal axis is time (ref. = 2 ms). The y axis represents the calibrated glottal area (increasing upwards, ref. = 10 mm^2) for the PGG, the translaryngeal electrical impedance (decreasing upwards) for the EGG and the acoustic pressure (microphone). Fundamental frequency is about 115 Hz. An estimate of the maximum glottal area during one cycle of a sustained phonation (28.9 mm^2) is obtained by using video-kymography and videostroboscopy in similar voicing conditions and in the same subject. The delay (1.729 ms) is computed by using the derivatives of the EGG and PGG signals: it is the time distance between the maximum glottal closing velocity and the collision peak.

during a sustained phonation (28.9 mm^2) is obtained by using VKG and videostroboscopy in similar voicing conditions and in the same subject (see above).

Another example, at a higher intensity (82.90 dB), is shown in Figure 8. In this case, the closed quotient is 0.58. Fundamental frequency is about 117 Hz. The estimate of the maximum glottal area during a sustained phonation is 46.3 mm^2.

Examples of the first derivatives of the PGG (dPGG/dt) and EGG (dEGG/dt) signals, calculated from two original tracings corrected for the respective time delays of 0.102 ms and 0.056 ms can be compared in Figures 9 and 10. The vertical arrows indicate the positive peak for EGG (ie, the max. rate of increase in VF contact when VF collide, considered as the collision peak) and the negative peak for PGG (ie, the max. glottal closing velocity). At the lower intensity (70.95 dB) (Figure 9), the delay between the two peaks is quite large (1.419 ms). At the higher intensity (86,89 dB) (Figure 10), the delay between the two peaks becomes much shorter (0.188 ms). When a vertical straight line is drawn through the peak of the EGG-derivative, it can be seen that the time of this peak (collision peak) corresponds in Figure 9 (horizontal arrow) to a value of 6.9% of the maximum negative amplitude of the PGG-derivative (= max. glottal closing velocity), giving a calculated VF velocity of 0.04 m/s. Similarly, in Figure 10, the peak of the EGG-derivative (collision) corresponds (horizontal arrow) to 44.2% of the maximum negative amplitude of the PGG-derivative and to a calculated VF velocity of 1.02 m/s.

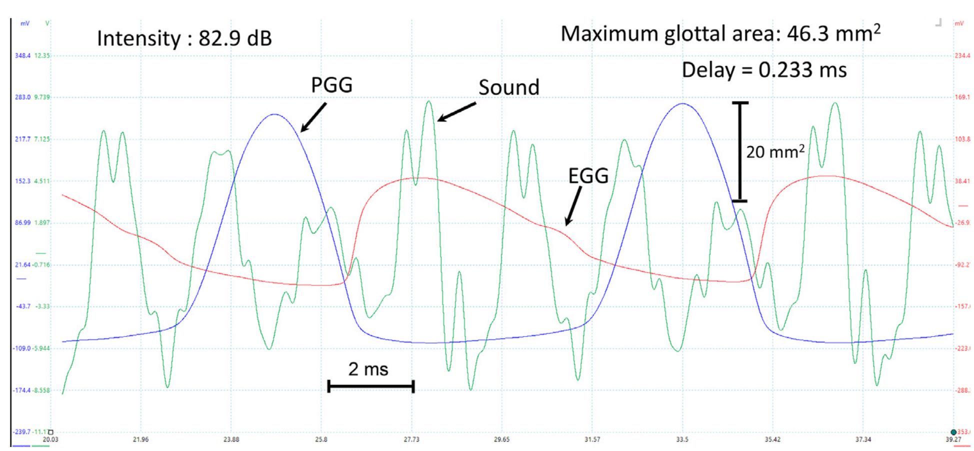

FIGURE 8. As Figure 7 but at a higher intensity (82.90 dB) with a larger closed quotient (0.58). Fundamental frequency is about 117 Hz. The estimate of the maximum glottal area during one cycle of a sustained phonation is here 46.3 mm^2. The delay (0.233 ms) is computed by using the derivatives of the EGG and PGG signals: it is the time distance between the maximum glottal closing velocity and the collision peak.

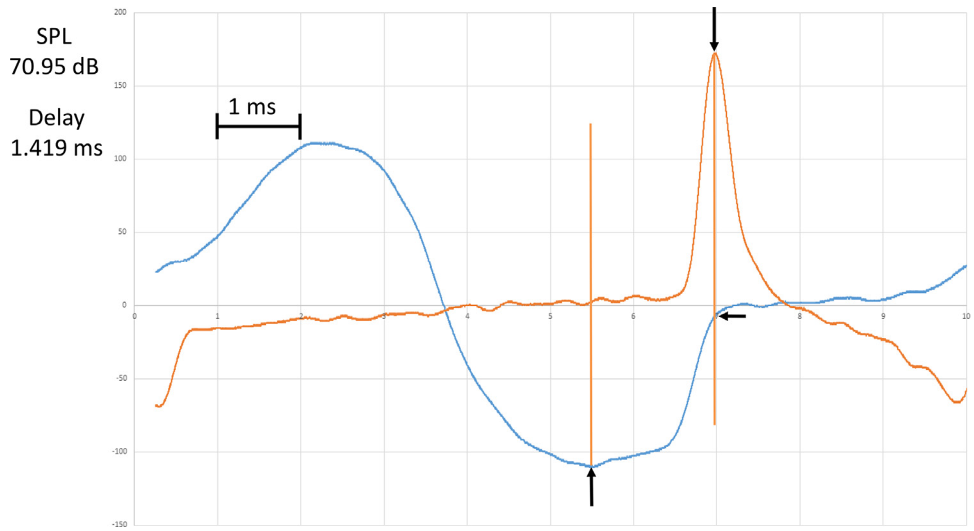

FIGURE 9. Example of the first derivatives of the PGG (dPGG/dt) and EGG (dEGG/dt) signals, calculated from an original raw tracing. Corrections of instrumental delays have been applied to both signals (0.102 ms and 0.056 ms). The arrows indicate the positive peak for EGG (ie, the max. rate of increase in VF contact once the glottis is closed) and the negative peak for PGG (ie, the max. glottal closing velocity). Fundamental frequency is about 104 Hz. At this rather low SPL (70.95 dB), the delay between the two peaks is large (1.419 ms). The time of the positive peak of the EGG-derivative (collision peak) corresponds (horizontal arrow) to a value of 6.9 % of the maximum negative amplitude of the PGG-derivative.

The data corresponding to the 32 records are given in Table 1.

The delay between the max. glottal closing velocity and the collision is plotted as a function of the intensity of voicing in Figure 11. The correlation is strongly negative (Rho = −0.80; $P < 0.0001$): the louder the voice, the shorter the delay.

Furthermore, the maximum VF velocity, indicated by the maximum negative amplitude of the PGG-derivative, increases when the voice gets more power (Figure 12; Rho = 0.89; $P < 0.0001$). This approximately fivefold (0.5 −2.5 m/s) increase of maximum velocity from soft to loud voicing actually results from a combined effect: (1) a larger maximum glottal width (Figures 2−4) and (2) a shorter open time, which affects both the opening and the closing phases. The closed quotient (duration of closed phase / period) is well known to reflect the glottal resistance to the lung pressure, and increases with intensity of phonation in

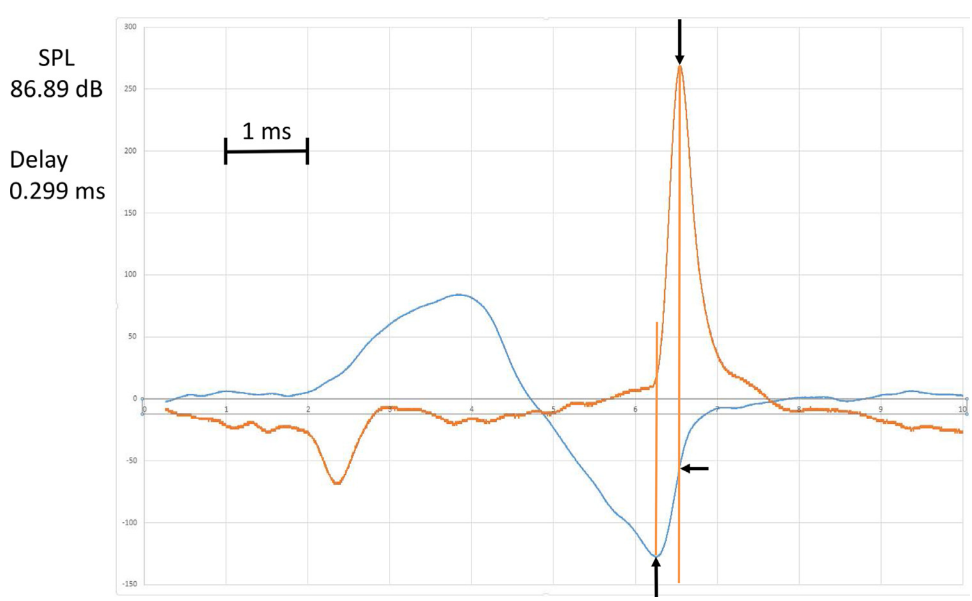

FIGURE 10. As Figure 9 but at a higher intensity (86,89 dB). Fundamental frequency is about 120 Hz. The delay between the two peaks is shorter (0.188 ms). The positive peak of the EGG-derivative (collision) corresponds (horizontal arrow) to 44,2 % of the maximum negative amplitude of the PGG-derivative.

TABLE 1.
Data Corresponding to the 32 Records

SPL (dB$_A$)	Closed Quotient (PGG)	Closed Quotient (EGG)	Max. VF Velocity (m/s)	Delay [Max. VF Velocity - Collision Peak] (ms)	% of Max. VF Velocity at Collision Peak	Actual VF Velocity at Collision Peak (m/s)
81.61	0.46	0.52	1.37	0.40	36.11	0.49
82.19	0.46	0.53	1.46	0.56	34.89	0.51
83.18	0.47	0.56	1.55	0.17	45.84	0.71
83.49	0.46	0.49	1.57	0.21	44.43	0.70
67.65	0.35	0.50	0.53	1.05	12.95	0.07
72.10	0.39	0.51	0.78	0.75	19.32	0.15
71.45	0.40	0.48	0.69	0.79	24.50	0.17
69.35	0.39	0.53	0.61	0.98	21.75	0.13
77.74	0.43	0.49	1.25	0.69	26.44	0.33
79.42	0.43	0.48	1.39	0.68	27.27	0.38
79.66	0.43	0.47	1.22	0.72	29.52	0.36
77.52	0.42	0.47	1.07	0.73	33.54	0.36
74.04	0.37	0.44	0.96	0.93	23.47	0.23
75.71	0.39	0.48	1.00	0.88	25.03	0.25
80.07	0.38	0.45	1.25	0.81	28.01	0.35
80.69	0.38	0.46	1.30	0.86	28.35	0.37
80.14	0.40	0.46	1.13	0.77	34.37	0.39
67.32	0.32	0.46	0.68	1.81	1.84	0.01
70.29	0.34	0.46	0.51	1.67	2.57	0.01
74.09	0.31	0.42	0.67	1.95	11.53	0.08
86.89	0.54	0.63	2.31	0.30	44.17	1.02
86.27	0.63	0.67	1.98	0.24	41.87	0.83
86.04	0.56	0.58	1.95	0.22	44.23	0.86
70.64	0.49	0.54	0.70	0.84	25.18	0.18
70.95	0.39	0.39	0.61	1.42	6.87	0.04
77.73	0.59	0.49	1.22	0.53	31.15	0.38
65.46	0.33	0.47	0.37	0.68	50.81	0.19
79.34	0.45	0.52	0.11	0.25	53.31	0.06
81.57	0.49	0.53	1.15	0.18	52.09	0.60
82.90	0.58	0.58	0.95	0.18	68.84	0.66
83.95	0.54	0.62	2.20	0.17	49.07	1.08
80.89	0.41	0.55	1.43	0.42	56.11	0.80

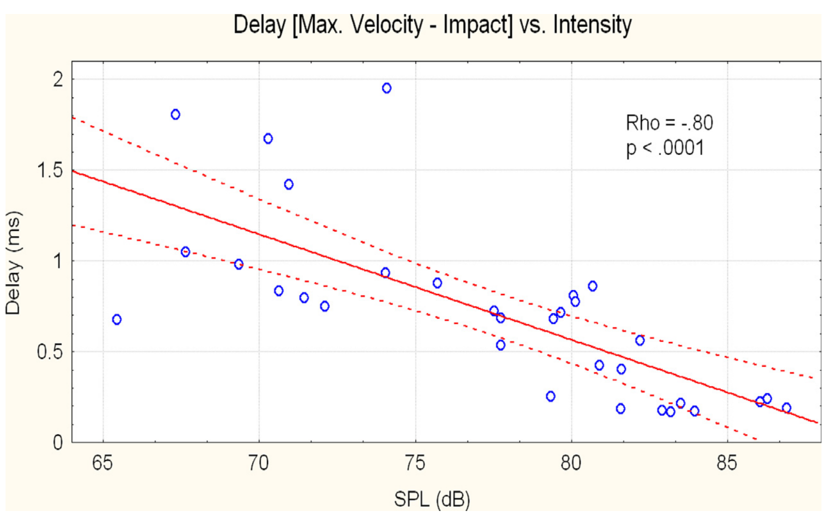

FIGURE 11. Correlation between the delay [max. velocity − impact] (ms) and the SPL of voicing (dB). The dashed lines represent the confidence bands (0.95) of the regression line. The louder the voice, the shorter the delay between max. VF velocity and impact collision.

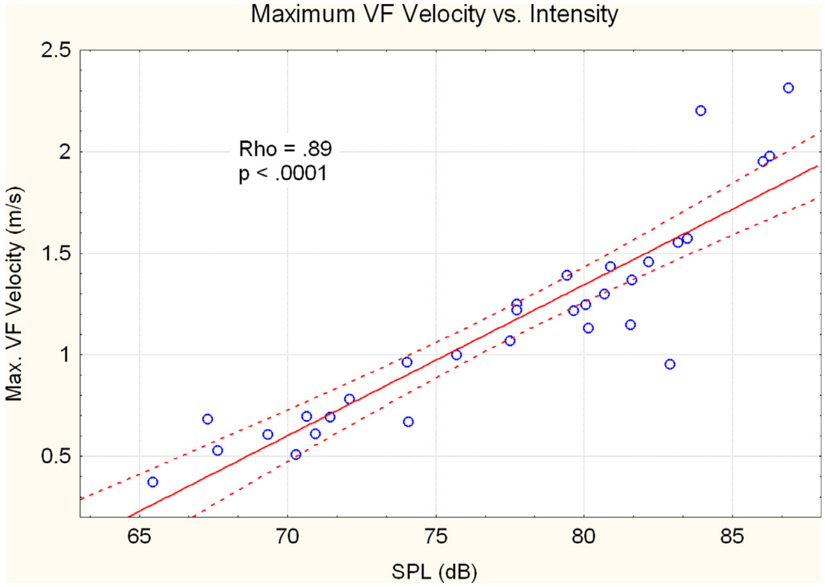

FIGURE 12. Correlation between the maximum VF closing velocity and the SPL of voicing (dB). The dashed lines represent the confidence bands (0.95) of the regression line. The maximum VF velocity, indicated by the maximum negative amplitude of the PGG-derivative, increases when the voice gets more power.

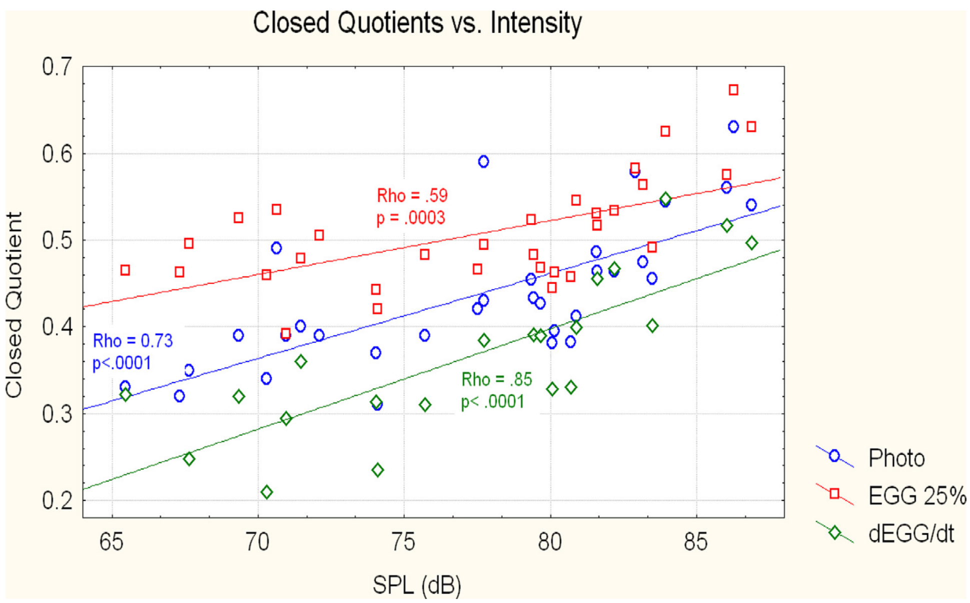

FIGURE 13. Correlation between closed quotient and SPL. Whatever the signal (EGG/PGG) and the way of computing it, the closed quotient significantly increases with SPL. Three different ways of computing the closed quotient are plotted: (1) the photometric method, (2) the width of the EGG signal at 25% of the signal magnitude and (3) the time distance between the closing and opening peaks of the EGG derivative. The strongest correlation is observed with the method based on the peaks of the EGG derivative (Rho = 0.85; $P < 0.0001$).

normal subjects.[21,32–34] Figure 13 illustrates, with the data of the current experiments, this positive correlation of closed quotient with intensity. Three different ways of computing the closed quotient are plotted: (1) our photometric method as explained above, (2) the width of the EGG signal at 25% of the signal magnitude[37] and (3) the time interval between the closing and opening peaks of the EGG derivative.[35] The strongest correlation is observed with the method based on the peaks of the EGG derivative (Rho = 0.85; $P < 0.0001$), although a few data points are missing, because the peak of the dEGG opening could not be accurately defined. Figure 14 shows that the maximum VF closing velocity positively correlates with the closed quotient (photometric method).

However, the most relevant information is the actual velocity of one single VF (m/s) at the time of collision (cf. Equation 2). This velocity depends on (1) the maximum velocity, given by the magnitude of the negative PGG peak, and (2) the extent of the reduction (in %) of this maximum velocity during the interval between the time this maximum

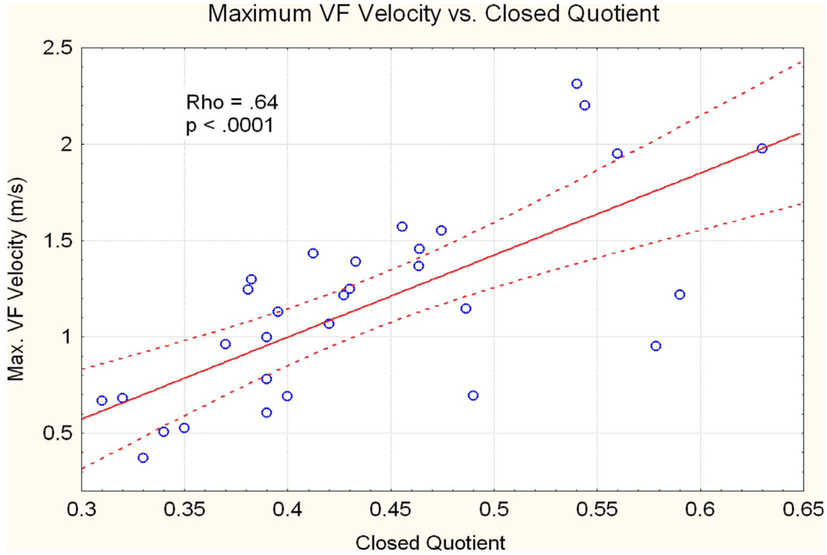

FIGURE 14. Max. VF velocity (m/s) during the closing phase of the glottis as a function of closed quotient. The max. VF velocity clearly increases with the open time shortens.

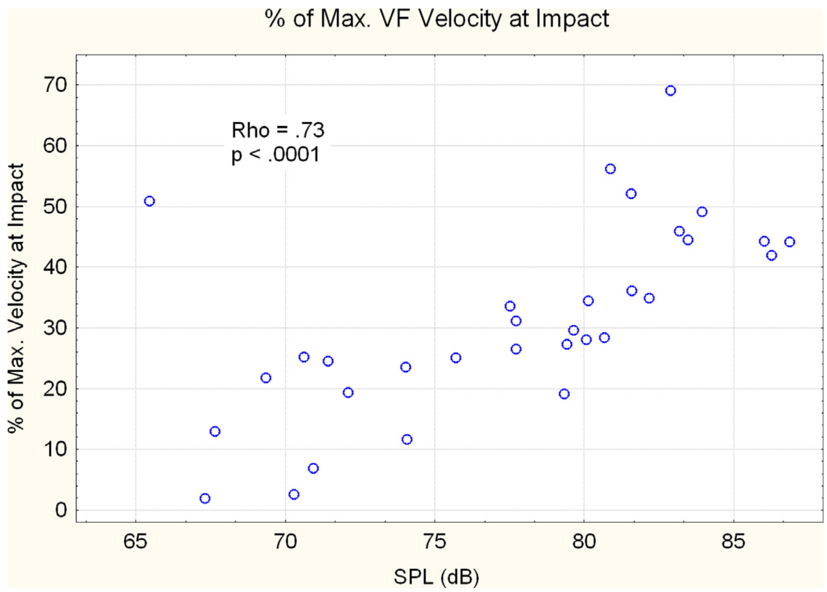

FIGURE 15. Percentage of maximum velocity at time of impact as a function of intensity of voicing. The velocity at impact is always a fraction of the maximum velocity, ranging from less than 5%–70%. The percentage increases with intensity.

closing velocity is reached and the VF collision peak. The extent of the deceleration depends itself on this delay, but also on the shape of the closing phase, which is not linear and involves, at its terminal phase, aspects like tissue compression and deformation, ie, what occurs between the first contact of the VF and the maximum rate of increase in VF contact when VF collide, which has been considered as the collision peak. Figure 15 shows that, at low voicing intensities, the collision velocity is not more than 5%–25 % of the maximum velocity, while at higher intensities the percentage promptly becomes 25%–70%. To oversimplify, if we compare the oscillation pattern of the VF edge with a sinusoidal motion clipped at half-height (one-mass model), clipping begins at the time of maximum velocity. This is the equivalent of a closed quotient of 0.5. In a situation where VFs freely oscillate without making contact, there is no glottal closure, the closed quotient is 0 and the velocity is also 0 at the time the VF edges are closest to each other. Between these extremes, every percentage (0%–100%) of the maximum velocity at the moment of clipping (contact) is possible.

The actual velocity of one single VF at the collision peak is plotted as a function of intensity (dB) in Figure 16. Globally, the velocity at the time of impact clearly increases with intensity. Yet the relation is not linear, and two different patterns can be identified: from about 78 dB on, the regression slope becomes substantially steeper, even if the correlation is strong and highly significant both in the range 65–78 dB

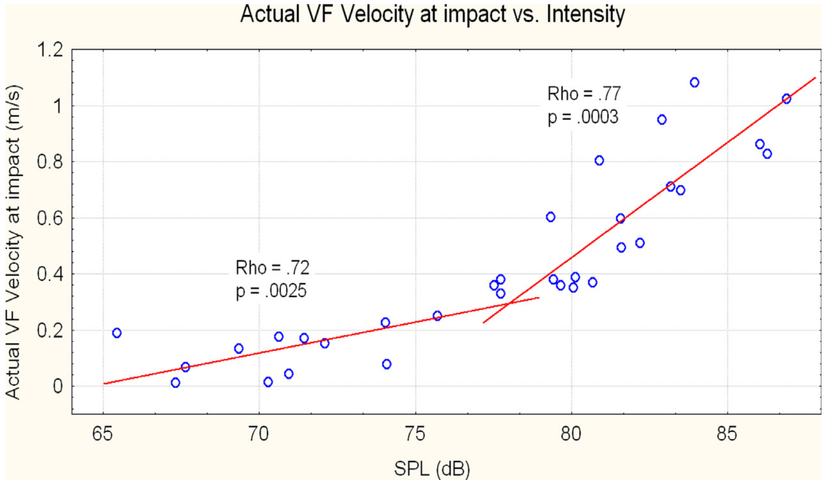

FIGURE 16. Actual velocity of one VF (m/s) at the time of impact as a function of intensity (dB). Globally, the velocity at impact clearly increases with intensity. However, two different patterns can be identified: from about 79 dB on, the regression slope becomes significantly steeper. Both in the range 65−78 dB and in the range 78−87 dB, the correlation is strong and highly significant.

and in the range 78-87dB. Covariance analysis demonstrates that the difference in slope between the two regression lines is highly significant (t = −3.3029; P = 0.0026).

DISCUSSION

The results of the present work are basically based on accurate measurement of three values: the maximum glottal area, the maximum glottal closing velocity and the time of maximum collision force (collision peak). The maximum glottal area is directly computed from the photometric signal, after calibration based on imaging, with a precision of 5%−10%. Determination of the maximum closing velocity uses the first derivative of the glottal area, which requires a high quality, noise-free signal and a high sampling frequency. In this scope, our photometric method far outperforms the imaging techniques. High-speed video images are limited by the number of pixels (resolution), but merely by the frequency of the measurement moments, as demonstrated by the experiments of Horacek et al,[15] in which, eg, at 2.000 images/s, for a F_0 of 100 Hz and a closed quotient of 0.5, only 5 points are measured during the closing phase.

As to the instant of maximum collision force in the conditions of the present experiments, and considering that contact is initiated along the lower margins of the VFs, then propagating to the upper margins,[12] it seems reasonable to assimilate the instant of maximum collision force with the maximum rate of increase in VF contact when VF collide. This maximum rate of increase in VF contact can be objectively and precisely inferred from the peak of the derivative of the EGG signal, provided that the glottal impedance signal is of high quality, noise-free and sampled at a sufficiently high frequency.

Verdolini et al[36] already found that closed quotient (and thus intensity) and impact stress were strongly related: within the region of linear relationship, an increase of 0.15 in the closed quotient corresponded to an increase in impact stress of about 1 kPa (excised canine larynges).

It seemed important to us to compare the *actual* VF collision velocity to the *maximum* closing velocity, and to show that, as a rule, this actual VF collision velocity is lower than the maximum closing velocity, that it is considerably lower (about 10%−30% of the maximum closing velocity) at intensities up to about 78 dB, but that it tends to strongly increase (up to about 70% of the maximum closing velocity) at higher intensities and higher values of closed quotient. This states more precisely the actual significance of the MADR.

The biphasic shape of the relationship between velocity at impact and intensity (Figure 16) is a new finding, but actually, it is not surprising. On the basis of the open quotient, skewing quotient and maximum flow declination rate values, Titze has hypothesized (using computer modelling) that there is an open quotient value at which an optimal voice power is obtained: it should occur around a closed quotient of 0.5, ie, when the open and closed phases have approximately equal durations.[38,17]

Furthermore, in our own study [17] of the dynamics of the driving force during the normal VF vibration cycle, we computed the ratio of the intraglottal pressure during the opening phase of the glottis to the intraglottal pressure during the closing phase, at different intensities of male modal chest voice. This ratio accounts for the extent to which the driving force is imparted to the VFs. A biphasic shape was observed: there is a zone of intensity values (around 76.6 dB) where the ratio of pressures is maximal, indicating that it is the zone of optimal mechanical efficiency at the purely glottal level.

What could explain the biphasic shape of the relation between collision velocity and intensity? It is known that the photoglottographic speed quotient (time quotient opening and/or closing) does not vary significantly with intensity in normal male subjects.[39] Moreover, the maximum velocity expressed as a function of intensity does not show a clearly biphasic shape (Figure 12), but it is less clear for the percentage of maximum velocity at the time of impact the increase of which seems enhanced at higher intensities (Figure 15). The increased reduction of the deceleration ("braking") effect above 78 dB could be due to subtle changes in the shape of the closing motion of the VFs, admitting that the assimilation to a sinusoidal movement is not more than an approximation.

It may be assumed that, as regards this relation between intensity of voicing and collision stress, there exists an interindividual variation. This makes its nonintrusive assessment relevant for the voice hygiene of the individual professional voice user. It may be expected that technical improvements will soon make high speed and/or high resolution glottal imaging via a noninvasive transnasal fiberscope possible, so that the dynamics of the glottal area can be displayed and quantified with sufficient accuracy to derive the parameters investigated in the present study.

CONCLUSION

Accurate *in vivo* measurements of the glottal area, the maximum glottal closing velocity and the instant of maximum collision force (collision peak) during sustained phonation allow computing the actual VF collision velocity and its relation to the maximum glottal closing velocity. At modal speaking pitch, this actual VF collision velocity is significantly lower than the maximum closing velocity, and the extent of the deceleration effect strongly depends on the intensity of voicing. Moreover, the relationship between collision velocity and intensity shows a biphasic shape: the deceleration ("braking") effect is increasingly reduced at loud voicing, from about 78 dB on. Hence the MADR in the closing phase of the glottis during VFs' vibration may not be considered as a measure of the impact stress loading the VFs during collision.

Mechanical stress has been considered as the key to the etiology of VF nodules.[7] In depth understanding of physiological variables that influence VF collision forces provides relevant insight into the pathophysiology and the prevention of voice disorders associated with phonotraumatic vocal hyperfunction.

REFERENCES

1. Li Z, Bakhshaee H, Helou L, et al. Evaluation of contact pressure in human vocal folds during phonation using high-speed videoendoscopic, electroglottography, and magnetic resonance imaging. *Proc Meet Acoust (Acoustical Society of America)*. 2013;19:1–8. https://doi.org/10.1121/1.4800732. 060306.
2. DeJonckere P H, Kob M. Pathogenesis of vocal fold nodules: new insights from a modelling approach. *Folia Phoniatr Logop*. 2009;61:171–179.
3. Horáček J, Laukkanen AM, Šidlof P, et al. Comparison of acceleration and impact stress as possible loading factors in phonation: a computer modeling study. *Folia Phoniatr Logop*. 2009;61:137–145.
4. Titze IR. Theoretical analysis of maximum flow declination rate versus maximum area declination rate in phonation. *J Speech Lang Hear Res*. 2006;49:439–447.
5. Titze IR, Laukkanen AM. (2007) Can vocal economy in phonation be increased with an artificially lengthened vocal tract? A computer modeling study. *Logoped Phoniatr Vocol*. 2007;32:147–156.
6. Titze IR. Mechanical stress in phonation. *J Voice*. 1994;8:99–105.
7. Jiang JJ, Titze IR. Measurement of vocal fold intraglottal pressure and impact stress. *J Voice*. 1994;8:132–144.
8. Horáček J, Radolf V, Bula V, et al.. Experimental and computer modelling study of glottal closing velocity during phonation. 25th International Conference Engineering Mechanics 2019. May 13 − 16, 2019, Svratka, Czech Republic. Editors: Igor Zolotarev and Vojtěch Radolf. Copyright © 2019 Institute of Thermomechanics 2019; Vol. 25: 153−156.
9. Horáček J, Bula V, Radolf V, et al. Impact stress in a self-oscillating model of human vocal folds. *J Vib Eng Technol*. 2016;4:183–190.
10. Verdolini-Abbott K, Chan RW, Titze IR, et al. Correspondence of electroglottographic closed quotient to vocal fold impact stress in excised canine larynges. *J Voice*. 1998;12:415–423.
11. Gunter HE, Howe RD, Zeitels SM, et al. Measurement of vocal fold collision forces during phonation: Methods and preliminary data. *J Speech Lang Hear Res*. 2005;48:567–576.
12. Hess M, Verdolini K, Bierhals W, et al. Endolaryngeal contact pressures. *J Voice*. 1998;12:50–67.
13. Verdolini K, Hess M, Titze IR, et al. Investigation of vocal fold impact stress in human subjects. *J Voice*. 1999;13:184–202.
14. Mehta DM, Kobler JB, Zeitels SM, et al. Toward development of a vocal fold contact pressure probe: bench-top validation of a dual-Sensor probe using excised human larynx models. *Appl. Sci*. 2019;9:4360–4377.
15. Horáček J, Radolf V, Bula V, et al.. Experimental modelling of glottal area declination rate in vowel and resonance tube phonation. In Models and Analysis of Vocal Emissions for Biomedical Applications: 11th International Workshop, December, 17-19, 2019, Claudia Manfredi (ed.), © 2019 FUP, CC BY 4.0 International, published by Firenze University Press (www.fupress.com), ISSN 2704-5846 (online), ISBN (online PDF) 978-88-6453-961-4. 205−207.
16. DeJonckere PH, Lebacq J. In Vivo quantification of the intraglottal pressure: modal phonation and voice onset. *2019 J Voice*. 2020;34:645.e19–645.e39. https://doi.org/10.1016/j.jvoice.2019.01.001. Epub 2019 Jan 16.
17. DeJonckere PH, Lebacq J, Titze IR. Dynamics of the driving force during the normal vocal fold vibration cycle. *J Voice*. 2017;31:649–661.
18. Fex S, Fex B, Hirano M. A clinical procedure for linear measurement at the vocal fold level. *J Voice*. 1991;5:328–331.
19. Larsson H, Hertegard S. Vocal fold dimensions in professional opera singers as measured by means of laser triangulation. *J Voice*. 2008;22:734–739.
20. Lee Y-O, Kim B-C, Lee J-H, et al. Development of laser ruler in rigid laryngoscope. *Clin Exp Otorhinolaryngol*. 2011;4:199–203.
21. DeJonckere PH, Lebacq J, Bocchi L, et al. Automated tracking of quantitative parameters from single line scanning of vocal folds: a case study of the 'messa di voce' exercise. *Logop. Phoniatr. Vocol*. 2015;40:44–54.
22. DeJonckere PH, Lebacq J. Intraglottal aerodynamics at vocal fold vibration onset. *J Voice*. 2019. https://doi.org/10.1016/j.jvoice.2019.08.002. S0892-1997(19)30237-1 Online ahead of print.
23. Gerratt BR, Hanson DG, Berke GS, et al. Photoglottography: a clinical synopsis. *J Voice*. 1991;5:98–105.
24. Köster O, Marx B, Gemmar P, et al. Qualitative and quntitative analysis of voice onset by means of a multidimensional voice analysis system (MVAS) using high-speed imaging. *J Voice*. 1999;13:355–374.

25. Piazza C, Mangili S, Del Bon F, et al. Quantitative analysis of videokymography in normal and pathological vocal folds: a preliminary study. *Eur Arch Oto-Rhino-Laryngol*. 2012;269:207–212.
26. Manfredi C, Bocchi L, Cantarella G, et al. Videokymographic image processing: objective parameters and user-friendly interface. *Biomed Signal Process Control*. 2012;7:192–201.
27. Svec J, Schutte HK. Videokymography: high-speed line scanning of vocal fold vibration. *J Voice*. 1996;10:201–205.
28. DeJonckere PH, Versnel H. High-speed imaging of vocal fold vibration: analysis by four synchronous single-line scans of onset, offset and register break. In: Passali D, ed. *Proc. XVIIIth IFOS World Congress*. 20051–8. Rome.
29. Fourcin A, Abberton E. First applications of a new laryngograph. *Volta Rev*. 1972;69:507–508.
30. Sarvaiya JN, Pandey PC, Pandey VK. An impedance detector for glottography. *IETE J Res*. 2011;55:100–105.
31. Boersma P, Weenink D. Available at: www.praat.org. Accessed February 10, 2020.
32. Sodersten M, Lindestadt P. Glottal closure and perceived breathiness during phonation in normally speaking subjects. *J Speech Hear Res*. 1990;33:601–611.
33. Dejonckere P. Control of fundamental frequency and glottal impedance with increasing sound pressure in normal and pathological voices. *Voice*. 1994;3:10–16.
34. Henrich N, d'Alessandro C, Doval B, et al. Glottal open quotient in singing: measurements and correlation with laryngeal mechanisms, vocal intensity and fundamental frequency. *J Acoust Soc Am*. 2005; 117:1417–1430.
35. Henrich N, d'Alessandro C, Castellengo M, et al. On the use of the derivative of electroglottographic signals for characterization of nonpathological phonation. *J Acoust Soc Am*. 2004;115:1321–1332.
36. Verdolini K, Chan R, Titze IR, et al. Correspondence of electroglottographic closed quotient to vocal fold impact stress in excised canine larynges. *J Voice*. 1998;12:415–423.
37. Scherer RC, Vail VJ, Rockwell B. Examination of the laryngeal adduction measure EGGW. In I. Titze (Ed.), NCVS Status and Progress Report1993 (Vol. 5): 73-82.
38. Titze IR. *Principles of Voice Production. 2nd Printing*. Iowa City IA: National center for voice and speech; 2000.
39. Hanson D, Gerratt BR, Berke G. Frequency, intensity and target matching effects on photoglottographic measures of open quotient and speed quotient. *J Speech Hear Res*. 1990;33:45–50.