Physiology and Acoustics of Inspiratory Phonation

*Françoise Vanhecke, †Jean Lebacq, ‡Mieke Moerman, §Claudia Manfredi, ¶Godfried-Willem Raes, and *,**Philippe H. DeJonckere, *‡¶Ghent, Belgium; †**Brussels, Belgium; and §Firenze, Italy

Summary: Introduction. Inspiratory phonation (IP) means phonating with inspiratory airflow. Some vocalists remarkably master this technique, to such an extent that it offers new dramatic, aesthetic, and functional possibilities in singing specific contemporary music. The present study aims to a better understanding of the physiological backgrounds of IP.

Material and methods. A total of 51 inhaling utterances were compared with 61 exhaling utterances in a professional soprano highly skilled in inhaling singing, by means of high-speed single-line scanning and advanced acoustic analysis. Ranges of intensity and Fo were kept similar.

Results. The main differences are: (1) an inversion of the mucosal wave, (2) a smaller closed quotient in IP, (3) a larger opening/closing quotient in IP with the additional difference that in IP, the quotient is larger than 1 (opening slower than closing), whereas it is less than 1 in expiratory mode (opening faster than closing), (4) a larger vocal-fold excursion in IP, (5) higher values of adaptive normalized noise energy in IP, and (6) a steeper slope of harmonic peaks in IP. However, jitter values are similar (within normal range), as well as damping ratios and central formant frequencies. The two voicing modes cannot be differentiated by blind listening.

Conclusion. The basic physiological mechanisms are comparable in both voicing modes, although with specific differences. IP is actually to be considered as an "extended vocal technique," a term applied to vocalization in art music, which falls outside of traditional classical singing styles, but with remarkable possibilities in skilled vocalists.

Key Words: Inspiratory voice–Videokymography–Jitter–NNE–Closed quotient.

INTRODUCTION

Inspiratory phonation means voice production with inspiratory airflow. The first report of this phenomenon seems to be from Franz Merkurius Van Helmont (John Baptist's son) in 1657.[1] Physiologically, inspiratory or reverse phonation is well known as a transitional phenomenon occurring during normal speech: it occurs naturally in different situations, including laughing, crying, and sighing.[2] Birth cry is of course inspiratory, and children commonly produce inspiratory cries.[3] According to Eklund[4], transitional ingressive (or inspiratory) phonation has been used for hundreds of years as a deliberate means of speech or sound production to achieve specific effects, and it is still being used for the same purposes, eg by shamans and ventriloquists. In normal spoken and spontaneous conversation, ingressive speech is found in all continents in genetically unrelated languages. Where ingressive speech occurs, it serves more or less the same paralinguistic functions, such as a feedback marker in a dialog. Further, pulmonic inspiratory phonation is not exclusively used by humans but is encountered in the phonation of many animals, as already noticed eg by Charles Darwin[5] in 1872: frogs, dogs, foxes, cats, horses, donkeys, several monkeys, and even birds, for example, make use of ingressive phonation.

Behlau and Pontes[6] consider inspiratory phonation as an adequate exercise to loosen supraglottic constriction. According to Boone and MacFarlane,[7] it should "tend to relax and open up the pharynx and laryngeal aditus." Indeed, inspiratory phonation has been reported by radiologists[8] as a possible maneuver for improving contrast visualization of the endolarynx. Also Kollar[9] has recommended reverse phonation for indirect laryngoscopy when an endolarynx is difficult to visualize.

Habitual reverse phonation has also been reported as a pathological entity in a particular psycho-dysfunctional context.[10]

Inspiratory speech has implications in voice and speech therapy: it has been suggested for treating psychogenic aphonia, mutational dysphonia, ventricular dysphonia and spasmodic dysphonia, but also unilateral vocal-fold paralysis and even severe stuttering[10–12]

Inspiratory phonation also has artistic implications in contemporary vocal music, eg Helmut Lachenmann's temA (1968) for flute, voice (mezzosoprano), and cello, and Nicholas DeMaison's Ursularia (2006), a chamber opera in one act.[1] DeBoer[13] made an analysis of these and other works with regard to vocal technique, notation, aesthetic orientation, and dramatic implications in her Doctor of Musical Arts dissertation. According to this author, inspiratory phonation offers a multitude of dramatic, aesthetic, and functional possibilities for composers and performers.

Reports about physiological and acoustical characteristics of inspiratory phonation are scarce: Orlikoff et al[11] collected the electroglottographic (EGG), flow, and microphonic signals in eight normal men and eight normal women, and performed videostroboscopy in four of them. The subjects were asked to alternate between inspiratory and expiratory voice. The authors found for the inspiratory utterances a caudal displacement of the larynx, a larger airflow, a higher Fo, a larger jitter %, a prolonged interval of increasing contact (EGG), and an approximation

of the upper margins of the vocal folds that preceded that of the lower margins.

Kelly and Fisher[14] also performed stroboscopic and acoustic measures in 10 normal females, shortly trained to produce an inspiratory phonation immediately followed by expiratory phonations, at comfortable pitch and loudness. A decreased membranous vocal-fold contact was noticed during inspiratory phonation relative to expiratory phonation, as well as an Fo increase.

DeBoer,[13] a trained female singer specializing in contemporary music, made EGG, flow, and intraoral measurements (/pæ pæ pæ/ sequences) on herself while performing excerpts from the two above-mentioned musical sores (Lachenmann's temA and DeMaison's Ursularia). She observed (although without statistical analysis) less intensity in higher formants for inspiratory phonation, as well as less vocal-fold adduction, but similar ranges for the aerodynamic parameters.

The present study deals for the first time with an exhaustive and documented physiological and acoustical investigation of inspiratory phonation using high-speed single-line scan, in which the basic parameters of intensity and Fo are controlled over a wide and continuous range, however, without exploring extreme possibilities. In this way, a detailed comparison can be made between the expiratory and inspiratory modes for quite a wide range of comparable vocal emissions in the same outstanding vocalist. All recordings were achieved from a highly experienced professional singer, classically educated, who has particularly developed the technique of inhaling singing for years.

MATERIAL AND METHODS

Subject
The subject for the present study was Françoise Vanhecke (FV)[15] (aged 56), a professional soprano who has particularly developed the inhaling singing technique.

All recordings were made during one single session. Simultaneous records of image and sound were taken. The subject was asked to utter sustained tones on an open vowel at various pitches and loudness, but always in a comfortable way, with a 90° larynx telescope in mouth and oropharynx. The presence of the scope of course hinders articulation as well as production of extremes in pitch and loudness. After some habituation trials, several series of recordings were made of the inspiratory voice, followed by several series of recordings of normal, expiratory phonation. Extreme possibilities of voice were not explored. The whole session lasted about 3 hours, with some resting intervals.

Kymography system
Single-line scanning of vocal-fold vibrations (kymography or videokymography [VKG])[16] is an imaging method based on a special digital camera, which can operate in two different modes: standard and high-speed. The camera is fixed on a rigid endoscope. In the standard mode, the camera provides standard images displaying the whole vocal folds at standard video frame rate (25 frames per second, with 768×576 pixels resolution). In the high-speed mode, the video camera delivers images from a single

line selected from the whole image, at the rate of 7812.5 line images per second (ie 0.128 millisecond per line) with a 768×1 pixels resolution. The selected line is chosen at the level of the midportion of the vibrating folds. The resulting high-speed image, called a "kymogram," displays the vibratory pattern of the small selected part of the vocal folds cycle by cycle. A kymographic recording is divided into video frames, and one screen corresponds to 18.4-millisecond duration (time is on the vertical axis, from top to bottom of the screen). Time calibration was verified by making a VKG of a tuning fork vibrating at 125 Hz.

When correctly applied, the technique allows a clear visualization of some essential physiological parameters of the vocal-fold vibration: period duration; duration of opening, closing, and closed phases; maximal amplitude of the vibration and right-left symmetry (Figure 1). Kymography has been applied successfully to voice pathology,[17–20] particularly for situations where traditional videostroboscopy is failing, as in the case of very irregular vibrations, or when the left and right vocal folds do not vibrate at the same frequency, or in cases of short "accidents," eg register breaks.

The single-line scanning system used in these experiments comprises a Lambert CCD (Lambert Instruments, Groningen, The Netherlands) Kymocam with technical characteristics corresponding to the above-reported description, a rigid 90° 4450.57 R. Wolf laryngeal telescope (Richard Wolf GmbH, Knittlingen, Germany), a JVC HR J746 MS magnetoscope (The Victor Company of Japan, Yokohama, Japan), and a monitor. The telescope has a magnifying facility, with a narrow depth of field and critical sharpness adjustment.
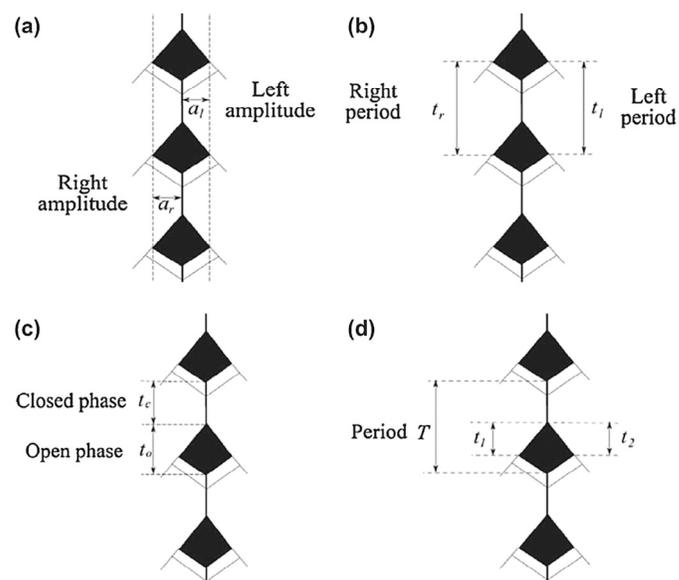

FIGURE 1. Schematic view of the physiological parameters of vocal-fold vibration, as shown by a videokymogram (from Manfredi et al 2012,[20] with permission). **A.** The maximal amplitude (right/left) of vocal-fold oscillation. **B.** The separate value for the period of the right and the left vocal fold (left = right in a normal voice). **C.** The timing of open an closed phases. **D.** The duration of the period and also the duration of the opening phase (t1 and t2 are equal in normal cases). The closing phase is the difference between the open phase and the opening phase.

Sound was also recorded using a Sennheiser MD421-U microphone (Sennheiser, Wedemark, Germany) held at 15 cm from the lips, for analysis of sound pressure level (SPL) and fundamental frequency (Fo) (using *PRAAT* 5.3.10, 2012, by P. Boersma and D. Weenink; www.praat.org).

Afterward, the videotape (including the voice sound) was digitized using a commercial Roxio Easy VHS to DVD 3 Plus converter (Corel Corporation, Ottawa, Canada). As to the audio recording, the S/N ratio was 45 dB, the frequency range was 20–20.000 Hz, the total wow and flutter was less than 0.05%, and the sampling frequency was 44.100 Hz. SPL calibration was achieved with different voice emissions within the range of 55–90 dB, in the same recording setting, by matching the direct sonometric values with the intensity measurements provided by *PRAAT*.

Image analysis

After a complete review, 112 short segments were selected for analysis, on the basis of a good quality of simultaneous kymographic images as well as sustained voice sound. Selected segments typically have a 300- to 700-millisecond duration. A total of 51 segments relate to the inspiratory condition, and 61 to the normal condition.

The images were enlarged and printed for manual measurements of the following parameters, as schematically shown in Figure 1: period duration; duration of opening, closing, and closed phases; and maximal amplitude of the vibration. All measurements were made manually, independently by two observers, with a high overall correlation (r = 0.96). In the cases of a slight difference, which never exceeded 7%, the two values were averaged. In a similar setting, such manual measurements have demonstrated a strong correlation with those achieved by an automatic analysis program using a digital image-processing algorithm developed and optimized for VKG by one of the authors.[20,21] The closed quotient is the quotient "duration of closed phase/ duration of cycle." The "opening/closing quotient" (sometimes called "speed quotient") is the ratio of the duration of the opening phase to that of the closing phase.

Acoustic analysis

Two computer programs, complementing each other, were used for acoustic analysis. This choice was based on previous studies, mainly including synthesized deviant voices.[22–29]

(1) *PRAAT* 5.3.10, 2012, by P. Boersma and D. Weenink (www.praat.org) for intensity and power spectrum.
(2) *BioVoice2* for fundamental frequency, central frequency of formants 1, 2, and 3, jitter and adaptive normalized noise energy (ANNE).

BioVoice is a software tool developed under Matlab® package.[22] Extended references can be found in Manfredi et al.[23]

Jitter

To calculate jitter, the following (traditional) formula is used: this parameter (×100 in percentage) measures the very short-term (cycle to cycle) irregularity of the pitch period of the voice:

$$relative\ Jitter = \frac{\frac{1}{N-1}\sum_{i=1}^{N-1}|T_i - T_{i+1}|}{\frac{1}{N}\cdot\sum_{i=1}^{N}T_i}$$

where T is the fundamental period for each frame (glottal cycle length) and N is the number of cycles.

ANNE

To estimate the noise component in the voiced signal, BioVoice uses an adaptive version of the normalized noise energy, named ANNE.[24] ANNE is the ratio of the spectral energy between the harmonics and the whole signal energy. A signal x(n) is considered to be composed of a periodic component s(n) and additive noise w(n): x(n) = s(n) + w(n), and their Discrete Fourier Transformations (DFT) are X(k), S(k), and W(k), respectively. The ANNE is calculated as:

$$ANNE(k) = 10log\left(\frac{\left|\sum_{m=N_L}^{N_H}\hat{W}_m(k)\right|^2}{\left|\sum_{m=N_L}^{N_H}X_m(k)\right|^2}\right)$$

with $k = N_L, \ldots N_H$, $N_L = Nf_LT$, $N_H = Nf_HT$, N is the number of DFT points, L is the number of frames in the analysis interval, f_L and f_H are the lowest and the highest frequencies respectively in the frequency band of interest and T is the sampling period. $\left|\hat{W}_m(k)\right|^2$ is an estimate of the unknown noise energy and $|X_m(k)|^2$ is the signal energy. In the harmonic dip intervals D_i, $\left|\hat{W}_m(k)\right|^2$ is given by $|X_m(k)|^2$, whereas in the harmonic peak intervals P_i, it may be estimated by interpolating between the values in the dip intervals left and right of the peak intervals P_i. The values of ANNE are negative: the higher (ie closer to 0) the ANNE index, the higher the noise in the signal.

The validity of these measurements has been tested using synthetic voices with variable jitter and noise inputs.[25–29]

BioVoice also computes the central frequency of formants 1, 2, and 3.

RESULTS

At the first inspection of the kymograms, an obvious systematic difference between the inspiratory and the expiratory voices is clear: the inverse locations of the grayish gradient on the glottal margin corresponding to the phase shift between the upper and lower margins of the vocal fold. The medial surface of the vocal fold is then temporarily visible. In normal expiratory voice, the gradient appears during the closing phase, which reflects the progression of the wave from caudal to cranial, whereas in inspiratory voice, the gradient is visible during the opening phase, revealing an inverted progression of the mucosal wave (Figures 2 and 3).

In contrast, it is impossible for the authors to discriminate the two types of voice segments based on sole listening.

Distribution

The values of all analyzed parameters can be considered as normally distributed in each category of utterances (inspiratory or expiratory), as demonstrated by the Kolmogorov-Smirnov test).
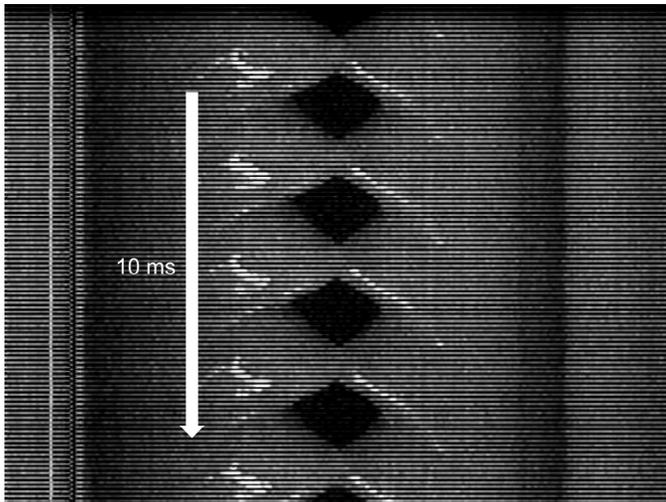
FIGURE 2. Single-line scan (videokymogram) in case of normal (expiratory) phonation. Time is progressing from top to bottom. Mucosal wave is normal.

Comparability of the two categories of vocal utterances

Figures 4 and 5 show the mean and confidence interval (standard deviations [SDs] = ±2) values for Fo and intensity in the inspiratory and the expiratory samples respectively: no statistically significant difference in means or in variances is found between the two groups. The Fo range is 192–360 Hz (mean = 271, SD = 54, standard error of the mean [SEM] = 8) for inspiratory voicing and 156–358 Hz (mean = 290, SD = 60, SEM = 8) for expiratory voicing. Intensities range from 73 to 90 dB (mean = 81, SD = 4.5, SEM = 0.6) for inspiratory voicing and from 68 to 91 dB (mean = 80, SD = 6, SEM = 0.74) for expiratory voicing.
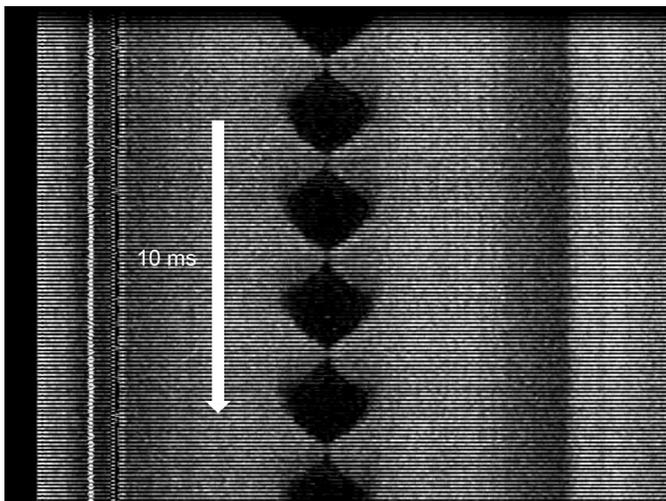

FIGURE 3. Single-line scan (videokymogram) in case of inspiratory phonation. Time is progressing from top to bottom. Mucosal wave is inverted.
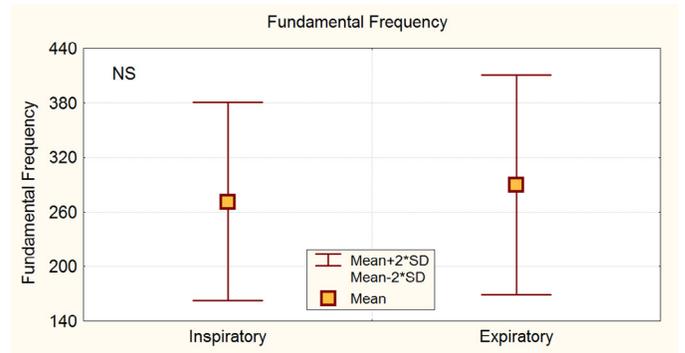

FIGURE 4. Fundamental frequency (Hz): mean values and standard errors of the 51 inspiratory and the 61 expiratory utterances.
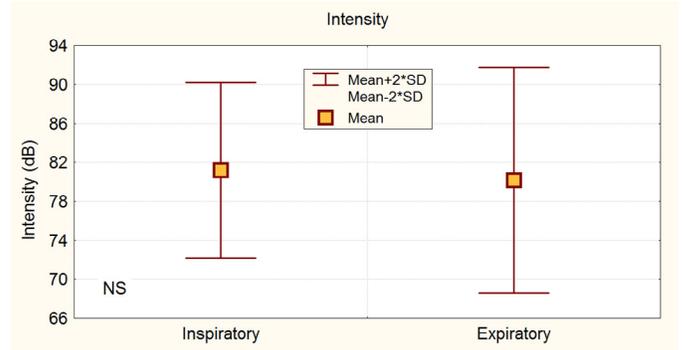

FIGURE 5. Intensity (dB): mean values and standard errors of the 51 inspiratory and the 61 expiratory utterances.

Physiological parameters
Closed quotient

Figure 6 shows that the closed quotient is significantly larger in expiratory than in inspiratory voicing (*t* test for independent samples: mean closed quotient for expiratory = 0.292 (SD = 0.11, SEM = 0.02) *vs* mean closed quotient for inspiratory = 0.216 (SD = 0.10, SEM = 0.01); *P* = 0.0003; df = 110). This points out a globally longer closed phase in expiratory phonation for comparable Fo and intensity.
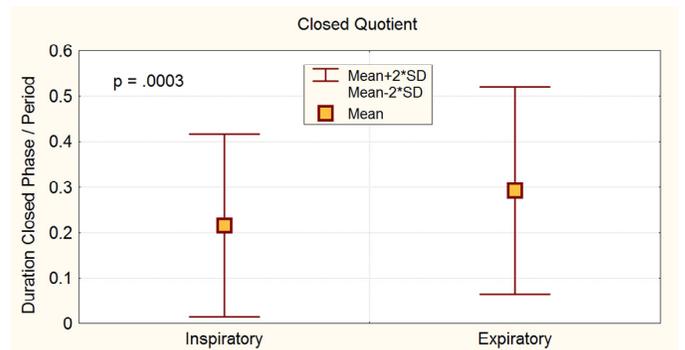

FIGURE 6. Closed quotient (duration closed phase/period): mean values and standard errors of the 51 inspiratory and the 61 expiratory utterances.
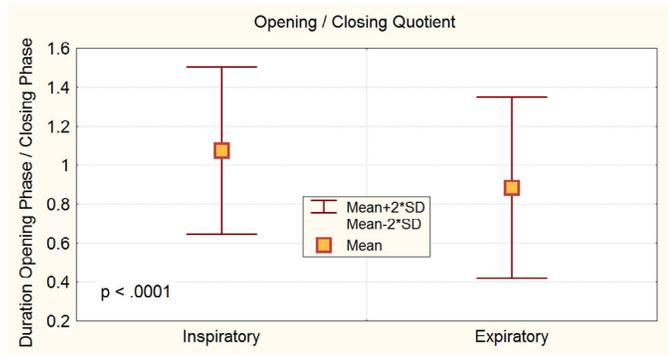
FIGURE 7. Opening/closing quotient (speed quotient): mean values and standard errors of the 51 inspiratory and the 61 expiratory utterances.

Opening/closing quotient

Figure 7 compares the opening/closing (or speed) quotients. They differ significantly between the two categories (*t* test for independent samples; $P = 0.0001$; df = 110). Moreover, an important difference is observed: in inspiratory voicing, the quotient is larger than 1 (mean = $1.075 \pm$ SD = 0.214; SEM = 0.03) (opening slower than closing), while it is less than 1 in expiratory voicing (mean = $0.884 \pm$ SD = 0.232; SEM = 0.03) (opening faster than closing).

Maximal amplitude

Figure 8 points out that the maximal amplitude is larger in inspiratory than in expiratory phonation (mean = $39.25 \pm$ SD = 13.31; SEM = 1.87) *vs* mean = $31.51 \pm$ SD = 6.98; SEM = 0.89) and has a larger variability (*t* test for independent samples; $P = 0.0001$; p $_{variances} < 0.0001$; df = 110).

Damping at end of emission

In a few recordings, it was possible to visualize (a part of) the damping of the oscillation, ie the progressively declining amplitude of the oscillation of each vocal fold between the moment of last contact on the midline until the complete immobilization during the abducting phase. This is an important parameter, accounting for the mechanical properties of the vocal folds as oscillators.[30,31] Even if in the current study data are limited, it is interesting to notice, as Figure 9 shows, that the amplitude
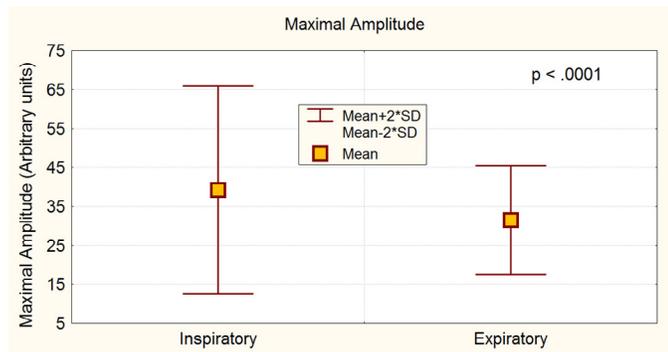

FIGURE 8. Maximal amplitude of vibration (maximal distance between the vocal-fold edges): mean values and standard errors of the 51 inspiratory and the 61 expiratory utterances.
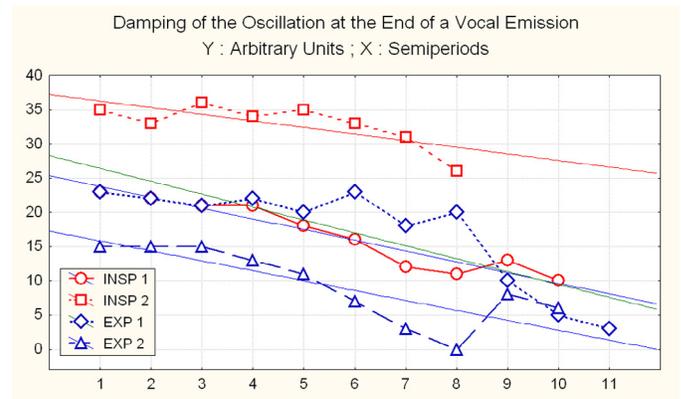

FIGURE 9. Distance between the most lateral and the most medial position of the vocal-fold edge (two measures per cycle; average of both folds) after the last contact phase (last cycle with glottal closure): four to five free oscillations. The two inspiratory and two expiratory utterances are compared.

decrement is very comparable in both types of emission, with an order of magnitude of 9–18% per cycle.

Vocal accidents

A few very short episodes of irregular vibrations were observed in inspiratory as well as in expiratory singing (see Figure 10 for an example). Aside from the inverted mucosal wave, they seem to have similar patterns and correspond to what can be episodically observed in normal subjects.

Acoustical parameters

Jitter

Jitter percentage values do not significantly differ between inspiratory (mean = 0.79, SD = 0.75, SEM = 0.10) and expiratory (mean = 0.61, SD = 0.66, SEM = 0.09) voicing and are within the normal range, ie less than 1%[32] (Figure 11).
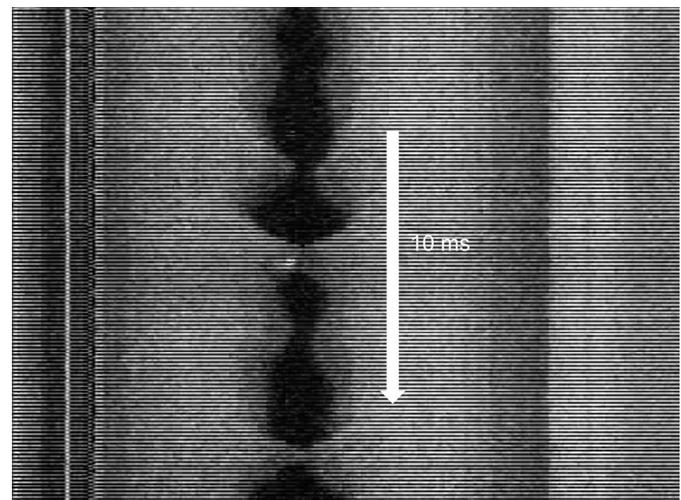

FIGURE 10. Videokymogram of a short "creak" in the voice at a pitch change (inspiratory voicing).
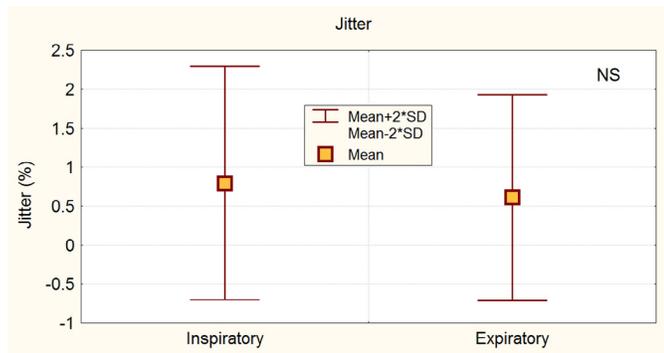
FIGURE 11. Jitter %: mean values and standard errors of the 51 inspiratory and the 61 expiratory utterances.

Adaptive normalized noise energy
ANNE is significantly larger (less negative, thus more noise, although unaudible) in inspiratory phonation (mean = $-27.80 \pm$ SD = 4.90 dB, SEM = 0.69 vs mean = $-30.10 \pm$ SD = 3.11 dB, SEM = 0.40 in expiratory phonation) (t test for independent samples; $P = 0.003$; df = 110) (Figure 12).

Harmonic decay
In each power spectrum (as generated by *PRAAT*) (a computer program made by P. Boersma and D. Weenink, Phonetic Sciences, University of Amsterdam, Amsterdam, The Netherlands), the magnitudes of harmonic peaks 1–12 (Fo is not taken in account) were measured and plotted comparatively in Figure 13. A variance analysis demonstrates a globally significant difference between the two conditions ($P = 0.03$). As can be seen in Figure 13, this is mainly determined by harmonics 6–9, that are significantly stronger ($P < 0.01$) in expiratory voicing.

Formants
Table 1 reports the mean central frequencies of the first three formants in inspiratory and expiratory utterances. No significant difference is observed.

Correlations
Tables 2 and 3 provide the correlation matrix of the most relevant parameters for inspiratory and expiratory singing, respectively.
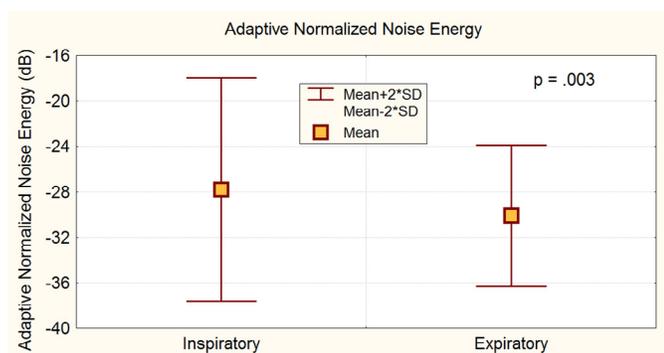

FIGURE 12. Adaptive normalized noise energy (ANNE): mean values and standard errors of the 51 inspiratory and the 61 expiratory utterances.
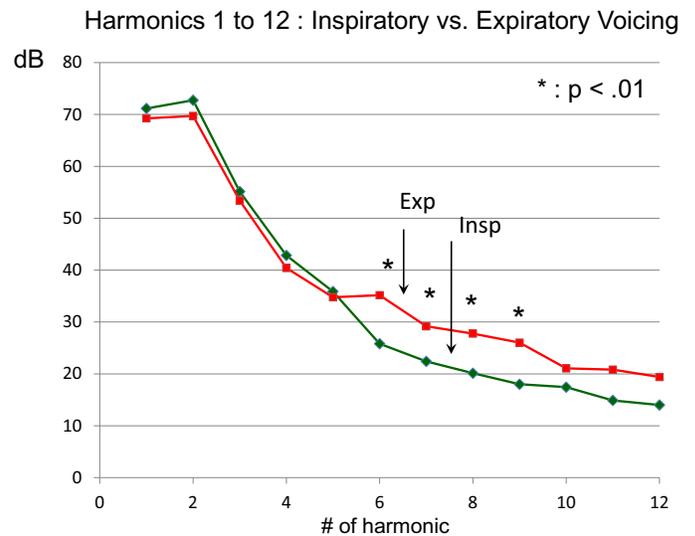

FIGURE 13. Magnitude (dB) of the harmonics 1–12: mean values of the 51 inspiratory and the 61 expiratory utterances. Significantly lower values are found for inspiratory phonation in harmonics 6–9.

Both in inspiratory and in expiratory phonations, the intensity is positively correlated with the closed quotient ($r = 0.41$; $P < 0.01$ and $r = 0.30$; $P < 0.05$) (Figures 14 and 15). However, only in the expiratory samples is the intensity also weakly correlated with the maximal oscillation amplitude ($r = 0.264$; $P < 0.05$) (Figure 16).

In expiratory voicing, there is a weak positive correlation between maximal amplitude and closed quotient ($r = 0.29$; $P < 0.05$), but in inspiratory voicing, it is the opposite ($r = -0.51$; $P < 0.01$).

In both categories, relative noise decreases with intensity ($r = -0.27$; $P < 0.05$ and $r = -0.49$; $P < 0.001$).

DISCUSSION
Comparability of inspiratory and expiratory conditions
A crucial condition for a valid comparison is the similarity of the basic parameters: intensity, Fo, and recording conditions. Therefore, all recordings were made in a single subject, achieved in one single session in exactly the same conditions and using the same equipment. Intensity and Fo ranges of inspiratory voicing *quasi* perfectly match with those of exhaling sound production (Figures 4 and 5).

Mucosal wave
The inverted progression of the mucosal wave was already noticed by Orlikoff et al.[11] It is absolutely typical and allows to differentiate immediately the inspiratory voicing by simple inspection (Figures 2 and 3). This can also account for the inversion of the opening/closing quotient: in inspiratory voicing, the quotient is larger than 1 (opening slower than closing), whereas it is less than 1 in expiratory voicing (opening faster than closing) (Figure 7). Warhurst et al[33] using high-speed video find a quotient of 0.97 in normal subjects (expiratory voicing). The differences between inspiratory and expiratory

TABLE 1.
Mean Central Frequencies of Formants 1, 2, and 3, With SD, for Inspiratory and Expiratory Voicing, Respectively

Formant	Mean Central Frequency IN (Hz)	SD	Mean Central Frequency EX (Hz)	SD	P	p variance
1	635	172	594	137	0.17	0.09
2	3544	892	3509	884	0.83	0.83
3	5866	741	5837	873	0.85	0.85

Abbreviation: IN, inspiratory; EX, expiratory.

TABLE 2.
Correlation Matrix for Inspiratory Voicing

	Intensity	Closed Quotient	Maximum Amplitude	Fo	Jitter	ANNE
Intensity	1					
Closed quotient	0.41**	1				
Maximum amplitude	−0.11	−0.51**	1			
Fo	−0.05	−0.04	0.16	1		
Jitter	−0.01	−0.03	0.23	−0.02	1	
ANNE	−0.27*	−0.07	−0.07	−0.39**	0.27*	1

* $P < 0.05$; ** $P < 0.01$.

TABLE 3.
Correlation Matrix for Expiratory Voicing

	Intensity	Closed Quotient	Maximum Amplitude	Fo	Jitter	ANNE
Intensity	1					
Closed quotient	0.33**	1				
Maximum amplitude	0.28*	0.29*	1			
Fo	−0.15	−0.28*	0.28*	1		
Jitter	−0.01	0.15	0.16	−0.02	1	
ANNE	−0.49**	−0.26*	−0.13	−0.01	0.09	1

* $P < 0.05$; ** $P < 0.01$.

vibration patterns may probably be related to anatomical properties of the vocal fold: the triangular shape in the frontal section is asymmetric, with a more horizontal cranial than caudal side.

Glottal mechanics
In normal voicing conditions and in modal register, the physiological mechanism of intensity regulation seems to consist basically in an increase/decrease of glottal impedance in
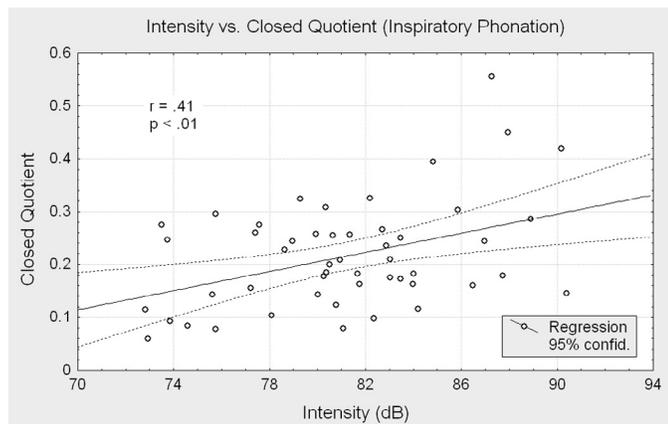

FIGURE 14. Closed quotient as a function of intensity in inspiratory phonation (51 utterances).
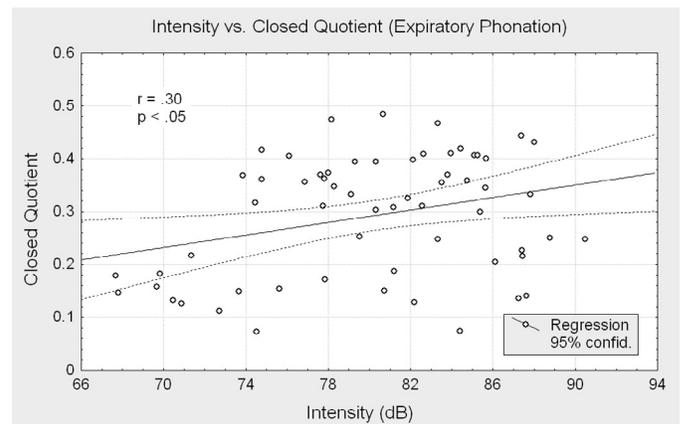

FIGURE 15. Closed quotient as a function of intensity in normal (expiratory) phonation (61 utterances).
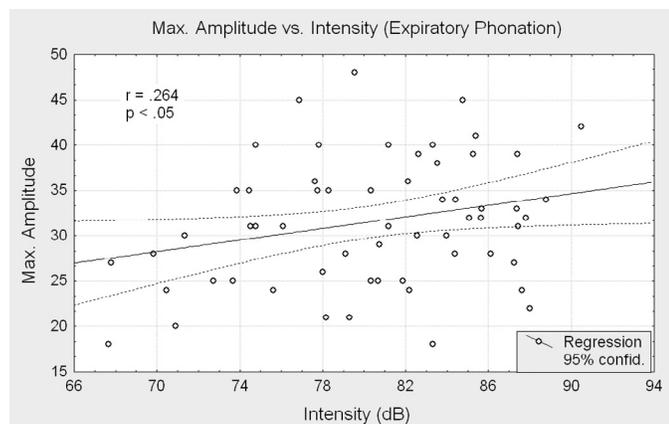
FIGURE 16. Maximal amplitude of vibration (maximal excursion of the vocal-fold edges) as a function of intensity in normal (expiratory) phonation (61 utterances).

parallel with the increase/decrease of the SPL, the SPL itself being closely related to the lung pressure.[21,34–36] Videostroboscopy of the vocal-fold vibration has demonstrated that the degree of glottal closure increases with the intensity of phonation in normal subjects.[37] Furthermore, also in normal subjects, the glottal impedance is well reflected by the closed quotient, and the quotient clearly increases with voice SPL; this increase in vocal-fold closure is plausibly related to an active thickening of the vocal-fold edge (in modal register) and to a longer closed phase of the vibratory cycle.[34] This may account for the limited concomitant increase in transglottal flow.[38] In the same way, Henrich et al[39] found, in singers achieving crescendos and decrescendos, a negative correlation between open quotient and vocal intensity.

All these basic mechanisms are confirmed as well in expiratory as in inspiratory phonation (Figure 14; Tables 1 and 2).

A reduced vocal-fold contact (Figure 6) in inspiratory mode is consistent with the observations of Kelly and Fisher[14] and DeBoer.[13] Less vocal-fold contact area would indeed generally correlate with more airflow[40] and an increased transglottal airflow in general is associated with enhanced turbulence,[41] hence with an increase of ANNE (Figure 12).

In both voicing modes, increasing SPL, which improves glottal closure, reduces the noise ratio (Tables 1 and 2), as was already shown in patients with functional dysphonia and normal vocal folds.[42]

The expected positive correlation between SPL and lateral excursion of the vibrating folds[35] is observed only in expiratory mode. Actually, the SPL increase is assumed to be related to a combination of a larger lateral vocal-fold excursion as a longer closed phase, the latter reflecting to some extent a virtual part of the amplitude. In inspiratory phonation, the stronger correlation of intensity with closed quotient (0.41 *vs* 0.033) could suggest that in this mode, glottal closure is more involved than vibration amplitude.

Another interesting difference between inspiratory and expiratory phonations at similar intensities and Fo values is the difference in magnitude of higher harmonics. As also noticed by DeBoer,[13] the spectral slope is steeper in the inspiratory con-

dition (Figure 13), and this fits with the lower closed quotient values (Figure 6). A positive association (in normal subjects) between closed quotient (measured with EGG) and shallowness of the spectral slope was also found by Barlow and Lovetri.[43]

Voice quality

It is noticeable that jitter values do not differ between inhaling and exhaling utterances and are within normal values. Orlikoff et al[11] found clearly increased jitter in inspiratory phonation, but his subjects were untrained and "naïve." ANNE values are also consistent with what is observed in normal subjects and singers (Figures 11 and 12), even if there is more (although unaudible) measured noise in the inspiratory utterances. Actually, it is practically not possible for the authors of this article to differentiate inhaling from exhaling utterances by blind listening to the recordings.

Formants

No significant difference could be found in the location of the three first formants. It could be expected that inspiratory phonation, due to the negative intrathoracal pressure, has an effect of pulling the larynx caudally, and this was also noticed by Orlikoff et al.[11] Current magnetic resonance imaging investigation does not support this assumption.[44] We observe a nonsignificant lowering of the central frequency of first formant F1 from 635 to 594 Hz. When applying the equation $L = c / 4F1$ (L is the length of the vocal tract, c the sound speed) for the lowest resonant frequency, it appears that the corresponding values for L are 13.38 and 14.30 cm, respectively. To reach the significance level, the difference in vocal tract length should be 1.40 cm, which seems improbable in a small person (1.60 m).

Damping

As vocal folds may be considered as an oscillator, it is relevant to observe its behavior when the driving force (the air pressure) vanishes, ie at the end of a vocal emission if the oscillation is not interrupted by a laryngeal closure. When the fold edges can continue to oscillate freely while abducting, a progressive damping occurs due to internal and external frictional forces. This damping is typically related to the mechanical properties of the vocal fold and gives an indication on the mechanical efficiency of the voice production.[29,30] Only a few single-line scans provide a visualization of this short critical phase, but the amplitude decrement seems similar in normal and inhaling conditions (Figure 9).

The artist's point of view

The current study limits to physiological and acoustical aspects, and only "standard" sustained vocal emissions, with a stable pitch and loudness, were considered to make a valid comparison with equivalents in normal exhaling singing. Inspiratory phonation is to be considered as an "Extended vocal technique," a term applied to vocalization in art music, which falls outside of traditional classical singing styles.

DeBoer[13] has identified several "ingressive phonation techniques," like alternation (ingressive/egressive), breath sounds, ingressive glissando, ingressive multiphonics, ingressive vocal

fry, ingressive whistle register, pitched ingressive phonation, all with specific notations on the musical score.[13] This is what is meant with "offering a multitude of dramatic, aesthetic and functional possibilities for composers and performers." This falls of course out of the scope of this study.

CONCLUSION

Inspiratory phonation means phonating with inspiratory airflow. Some singers remarkably master this technique to such an extent that it can offer new dramatic, aesthetic, and functional possibilities in contemporary music.

The current work pertains to the physiological and acoustical characteristics of inspiratory phonation, compared with expiratory phonation, as investigated in the same subject, a professional soprano highly skilled in inhaling singing.

A total of 51 inspiratory utterances were compared with 61 expiratory utterances by means of high-speed single-line scanning and advanced acoustic analysis. The ranges of intensity and Fo were similar.

The main differences are the following: (1) an inversion of the mucosal wave in inspiratory phonation, (2) a lower closed quotient in inspiratory phonation, (3) a higher opening/closing quotient in inspiratory phonation, with the additional difference that in inspiratory mode, the quotient is larger than 1 (opening slower than closing), whereas it is less than 1 in expiratory mode (opening faster than closing), (4) a larger vocal-fold excursion in inspiratory phonation, (5) higher values of ANNE in inspiratory phonation, and (6) a steeper slope of harmonic peaks in inhaling voice.

However, jitter values are similar (within the normal range), as well as damping ratios and central formant frequencies.

The current experiment is limited to voicing within a comfortable range of frequencies and intensities, for comparing as well as possible the two modes of voicing. Exploring the extremes of voice possibilities as well as considering artistic aspects of singing is out of the scope of this study.

REFERENCES

1. Van Helmont FM. Alphabeti vere naturalis Hebraici brevissima elineation. Sulzbach, 1657. German edition, 1667; Dutch edition, 1697.
2. Timcke R, von Leden H, Moore P. Laryngeal vibrations: measurements of the glottic wave part II : physiologic variations. *AMA Arch Otolaryngol*. 1959;69:438–444.
3. Grau SM, Robb MP, Cacace AT. Acoustic correlates of inspiratory phonation during infant cry. *J Speech Hear Res*. 1995;38:373–381.
4. Eklund R. Pulmonic ingressive phonation: diachronic and synchronic characteristics, distribution and function in animal and human sound production and in human speech. *J Int Phon Assoc*. 2008;38:235–324.
5. Darwin C. *The Expression of the Emotions in Man and Animals*. London, J Murray; 1872.
6. Behlau M, Pontes P. *Avaliaçao e tratamento das disfonias*. Sao Paulo, Brazil: Lovise; 1995.
7. Boone DR, MacFarlane SC. *The Voice and Voice Therapy*. 5th ed. Englewood Cliffs, NJ: Prentice Hall; 1994.
8. Powers WE, Holtz S, Ogura J. Contrast examination of the larynx and pharynx. Inspiratory phonation. *Am J Roentgenol Radium Ther Nucl Med*. 1964;92:40–42.
9. Kollar A. Indirect laryngoscopy. *Cesk Otolaryngol*. 1989;38:114–115.
10. Wattremez A, Delpech C, De Brugiere C, et al. Reverse phonation: pathological and therapeutic aspects. Study of a clinical case. *Rev Laryngol Otol Rhinol (Bord)*. 2011;132:51–55.
11. Orlikoff RF, Baken RJ, Kraus DH. Acoustic and physiologic characteristics of inspiratory phonation. *J Acoust Soc Am*. 1997;102:1838–1845.
12. Kenjo M. Treatment featuring direct speech therapy for a school age child with severe stuttering : a case report. *Jpn J Logop Phoniatr*. 2005;46:21–28.
13. DeBoer A. Ingressive phonation in contemporary vocal music [DMA dissertation]. Bowling Green State University. 2012.
14. Kelly CL, Fisher KV. Stroboscopic and acoustic measures of inspiratory phonation. *J Voice*. 1999;13:389–402.
15. www.francoisevanhecke.be/; www.inhalingsinging.be; www.isfv.be.
16. Svec J, Schutte HK. Videokymography: high-speed line scanning of vocal fold vibration. *J Voice*. 1996;10:201–205.
17. Manfredi C, Bocchi L, Bianchi S, et al. Objective vocal fold vibration assessment from videokymographic images. *Biomed Signal Process Control* 2006;1:129–136.
18. Svec J, Sam F, Schutte HK. Videokymography in voice disorders: what to look for? *Ann Otol Rhinol Laryngol*. 2007;116:172–180.
19. Piazza C, Mangili S, Del Bon F, et al. Quantitative analysis of videokymography in normal and pathological vocal folds: a preliminary study. *Eur Arch Otorhinolaryngol*. 2012;269:207–212.
20. 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.
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. *Logoped Phoniatr Vocol*. 2015;40:44–54.
22. Manfredi C, Bocchi L, Cantarella G. A multipurpose user-friendly tool for voice analysis : application to pathological adult voices. *Biomed Signal Process. Control* 2009;4:212–220.
23. Manfredi C, Barbagallo D, Baracca G, et al. Automatic assessment of acoustic parameters of the singing voice: application to professional western operatic and jazz singers. *J Voice*. 2015;29:517.e1–517.e9. doi:10.1016/j.jvoice.2014.09.014. Epub 2015 Mar 17.
24. Manfredi C. Adaptive noise energy estimation in pathological speech signals. *IEEE Trans Biomed Eng*. 2000;47:1538–1543.
25. Dejonckere PH, Schoentgen J, Giordano A, et al. Validity of jitter measures in non-quasi-periodic voices. Part I. Perceptual and computer performances in cycle pattern recognition. *Logoped Phoniatr Vocol*. 2011;36:70–77.
26. Manfredi C, Giordano A, Schoentgen J, et al. Dejonckere PH. Validity of jitter measures in non-quasi-periodic voices. Part II. The effect of noise. *Logoped Phoniatr Vocol*. 2011;36:78–89.
27. Dejonckere PH, Giordano A, Schoentgen J, et al. To what degree of voice perturbation are jitter measurements valid ? A novel approach with synthesized vowels and visuo-perceptual pattern recognition. *Biomed Signal Process Control* 2012;7:37–42.
28. Manfredi C, Giordano A, Schoentgen J, et al. Perturbation measurements in highly irregular voice signals: performances/validity of analysis software tools. *Biomed Signal Process Control* 2012;7:409–416.
29. Rruqja N, Dejonckere PH, Cantarella G, et al. Testing software tools with synthesized deviant voices formedicolegal assessment of occupational dysphonia. *Biomed Signal Process Control* 2014;13:71–78.
30. Dejonckere PH. Damping biomechanics of vocal fold oscillation. In: Gauffin J, Hammarberg B, eds. *Vocal Fold Physiology: Acoustic, Perceptual and Physiological Aspects of Voice Mechanisms*. San Diego, CA: Singular Publishing Group, Inc.; 1991:105–111.
31. Dejonckere PH, Versnel H. High-speed imaging of vocal fold vibration: analysis by 4 synchronous single-line scans of onset, offset and register break. Proceedings of the XVIII I.F.O.S. (International Federation of Oto-rhino-laryngological Societies) World Congress 25–30 June, D. Passali, Ed. Rome 2005.
32. Sigmund M. *Voice recognition by computer*. Marburg, Germany: Tectum Verlag Marburg; 2003.
33. Warhurst S, McCabe P, Heard R, et al. Quantitative measurement of vocal fold vibration in male radio performers and healthy controls using high-speed videoendoscopy. *PlosOne*. 2014;9:e101128.

34. Colton RH. Physiology of Phonation. In: Benninger MS, Jacobson BH, Johnson AF, ed. *Vocal Arts Medicine*. New York, NY: Thieme; 1993: 30–60.

35. Dejonckere PH. Control of fundamental frequency and glottal impedance with increasing sound pressure in normal and pathological voices. *Voice*. 1994;3:10–16.

36. Kent RD. *The MIT Encyclopaedia of Communication Disorders*. Cambridge, Massachussetts, US: Massachusetts Institute of Technology; 2004.

37. Sodersten M, Lindestadt PA. Glottal closure and perceived breathiness during phonation in normally speaking subjects. *J Speech Hear Res*. 1990;33:601–611.

38. Smitheran JR, Hixon TJ. A clinical method for estimating 10. laryngeal airway resistance during vowel production. *J Speech Hear Disord*. 1981;46:38–46.

39. Henrich N, d'Alessandro C, Doval B. Glottal open quotient in singing: measurements and correlation with laryngeal mechanisms, vocal intensity and fundamental frequency. *J Acoust Soc Am*. 2005;117:1417–1430.

40. Rothenberg M. Some relations between glottal air flow and vocal fold contact area. Proceedings of the Conference on the Assessment of Vocal Pathology, ASHA Reports No, 11, pp. 88–96, 1979.

41. Baken RJ, Orlikoff RF. *Clinical Measurement of Speech and Voice*. 2nd ed. San Diego, CA: Singular Publishing Group; 2000.

42. Dejonckere PH. Effects of louder voicing on acoustical measurements in dysphonic patients. *Log Phon Vocol* 1998;23:79–84.

43. Barlow C, Lovetri J. Closed quotient and spectral measures of female adolescent singers in different singing styles. *J Voice*. 2010;24:314–318.

44. Moerman M, Vanhecke F, Van Assche L, et al. Vocal tract morphology in inhaling singing: a MRI-based study. *J Voice*. 2015. In press.