



**HAL**  
open science

## Ventricular-Fold Dynamics in Human Phonation

Lucie Bailly, Nathalie Henrich Bernardoni, Frank Müller, Anna-Katharina Rohlfs, Markus Hess

► **To cite this version:**

Lucie Bailly, Nathalie Henrich Bernardoni, Frank Müller, Anna-Katharina Rohlfs, Markus Hess. Ventricular-Fold Dynamics in Human Phonation. *Journal of Speech, Language, and Hearing Research*, 2014, 57 (4), pp.1219-1242. 10.1044/2014\_JSLHR-S-12-0418 . hal-00998464

**HAL Id: hal-00998464**

**<https://hal.science/hal-00998464>**

Submitted on 8 Oct 2014

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

## Research Article

## Ventricular-Fold Dynamics in Human Phonation

Lucie Bailly,<sup>a</sup> Nathalie Henrich Bernardoni,<sup>b</sup> Frank Müller,<sup>c</sup>  
Anna-Katharina Rohlf,<sup>c</sup> and Markus Hess<sup>c</sup>

**Purpose:** In this study, the authors aimed (a) to provide a classification of the ventricular-fold dynamics during voicing, (b) to study the aerodynamic impact of these motions on vocal-fold vibrations, and (c) to assess whether ventricular-fold oscillations could be sustained by aerodynamic coupling with the vocal folds.

**Method:** A 72-sample database of vocal gestures accompanying different acoustical events comprised high-speed cinematographic, audio, and electroglottographic recordings of 5 subjects. Combining the physiological correlates with a theoretical model of phonation, the vocal–ventricular aerodynamic interactions were investigated.

**Results:** A ventricular-fold motion is found during (de)crescendos, shout, throat singing, yodel, growls, and glides with transitions between registers. Three main types of dynamics are identified: slow nonoscillatory motion and

fast oscillatory motion with aperiodical or periodical vibrations. These patterns accompany a change in voice quality, pitch, and/or intensity. Alterations of glottal-oscillatory amplitude, frequency, and contact were predicted. It is shown that a ventricular oscillation can be initiated and sustained by aerodynamic coupling with the vocal folds.

**Conclusions:** Vocal–ventricular aerodynamic interactions can alter, enhance, or suppress vocal-fold vibrations or leave them unchanged, depending on the ventricular-fold dynamics involved. Depending on its variation in time, a similar level of ventricular-fold adduction impacts the glottal vibratory magnitude and contact much differently.

**Key Words:** voice, ventricular folds, aerodynamics, high-speed cinematography, phonation modeling

The ventricular folds, also called *false vocal folds* or *vestibular folds*, are two laryngeal structures located above and in the vicinity of the vocal folds. As illustrated in Figure 1a, they are separated from the glottal plane by a laryngeal cavity, the Morgagni ventricle. Typical anatomical dimensions extracted from previous investigation of phonation using X-ray imaging techniques are summarized in Table 1 (Agarwal, 2004; Agarwal, Scherer, & Hollien, 2003; Hirano, Kurita, & Nakashima, 1983; Hollien & Colton, 1969; Wilson, 1976). Ventricular folds have a less differentiated layered structure than the vocal folds. They are composed of a mucous membrane, abundant adipose tissues, and seromucous glands (Kutta, Steven, Varoga, & Paulsen, 2004). Collagenous and elastic fibers are found irregularly arranged in the ventricular fold (Guida & Zorzetto,

2007), thus differing from the specific vocal laminar microstructure. Muscle and ligament fibers also exist, however sparse and in the margins (Guida & Zorzetto, 2007; Kimura et al., 2002; Kotby, Kirchner, Kahane, Basiouny, & El-Samaa, 1991; Reidenbach, 1998). The ventricular folds present a greater viscosity and lower stiffness than vocal folds (Chan, Fu, & Tirunagari, 2006; Haji, Mori, Omori, & Isshiki, 1992). These differences observed at both micro- and macroscale tend to make the ventricular folds a priori poor biomechanical oscillators (Agarwal, 2004).

Comprising a large number of subepithelial glands (128 glands/cm<sup>3</sup>) compared to the vocal folds (13 glands/cm<sup>3</sup>), the ventricular folds are often described as playing a key role in lubrication of laryngeal tissues and air moistening (Bak-Pedersen & Nielsen, 1986; Guida & Zorzetto, 2007; Kutta, Steven, Kohla, Tillmann, & Paulsen, 2002; Pepinsky, 1942). Their function in antigen defense has been evidenced by Kutta et al. (2002), who reported their role in preventing invasion by pathogenic agents thanks to constitutive antimicrobial peptides. The ventricular folds are also known to ensure a sphincter-like action preserving the lower respiratory tract during swallowing (Kawasaki, Fukuda, Shiotani, & Kanzaki, 2001) and cough (Pepinsky, 1942). They are involved in phonation, yet their effective contribution to the laryngeal sound source is still a topic for current research

<sup>a</sup>Aix Marseille Université, CNRS, Centrale Marseille, IRPHE UMR 7342, 13384, Marseille, France

<sup>b</sup>GIPSA-lab (CNRS, Grenoble INP, University of Grenoble), Grenoble, France

<sup>c</sup>University Medical Center Hamburg-Eppendorf, Hamburg, Germany

Correspondence to Lucie Bailly: lucie.bailly@irphe.univ-mrs.fr

Editor: Jody Kreiman

Associate Editor: Scott Thomson

Received January 7, 2013

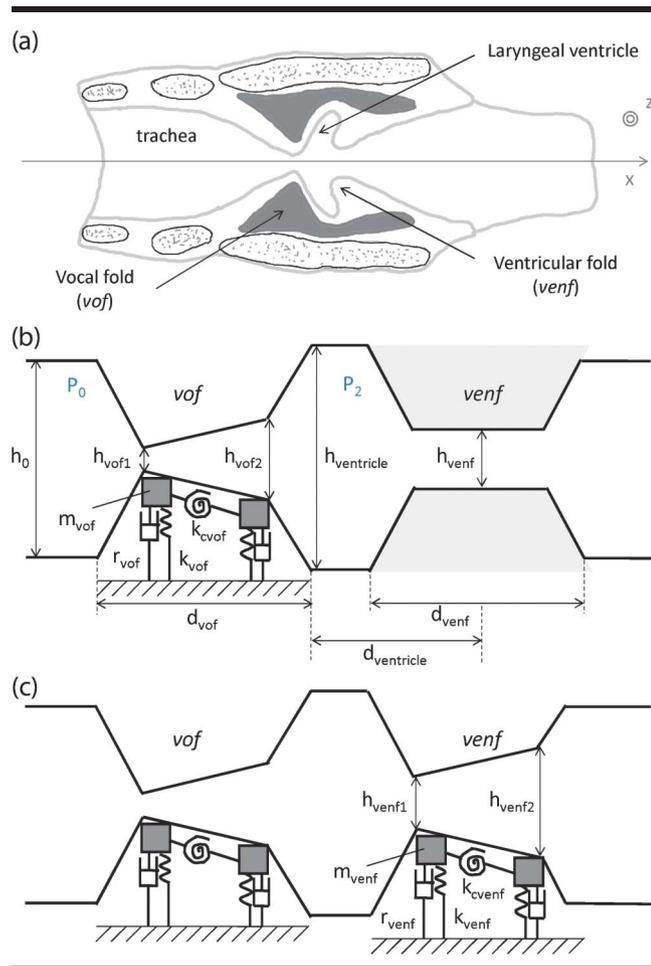
Revision received June 16, 2013

Accepted December 28, 2013

DOI: 10.1044/2014\_JSLHR-S-12-0418

**Disclosure:** The authors have declared that no competing interests existed at the time of publication.

**Figure 1.** (a) Rear view coronal section of the larynx. (b) Sketch of the two-mass model of the vocal folds ( $M2M_{vof}$ ) combined with a ventricular-fold constriction of controlled aperture,  $h_{venf}$ . Model used for processing of Cases 1 to 4. (c) Sketch of the  $M2 \times M2$  model used in this study, constituted by a two-mass model of vocal folds  $M2M_{vof}$ , and a two-mass model of ventricular folds  $M2M_{venf}$ . Model used for simulation of vocal–ventricular vibrations (not to scale).



**Table 1.** Anatomical characteristics of an adult larynx in phonation.

Characteristic	Range of values	Mean value
Glottal length, $\bar{d}_{vof}$ (mm)	3–9	4
Ventricular-fold length, $d_{venf}$ (mm)	1.1–10.6	5.5
Ventricle length, $d_{ventricle}$ (mm)	2.2–7.5	4.7
Glottal aperture, $h_{vof}$ (mm)	0–3	1
Ventricular-fold aperture, $h_{venf}$ (mm)	0–9	5.5
Ventricle height, $h_{ventricle}$ (mm)	9–23	15.7
Aspect ratio, $h_{venf}/h_{ventricle}$	0–1	0.3
Aspect ratio, $h_{venf}/\bar{d}_{ventricle}$	0–4	1.2
Aspect ratio, $h_{venf}/\max(h_{vof})$	0–3	1.8

*Note.* These are typical combined data for male and female adults (obtained from Agarwal, 2004; Agarwal et al., 2003; Hirano et al., 1983; Hollien & Colton, 1969; Wilson, 1976). The mean value comes from non-singer speakers.

(see, e.g., Agarwal, 2004; Bailly, 2009; Moisik, 2013; Stager, 2011). In this context, the main purpose of the present study is to characterize the ventricular-fold dynamics involved during human voicing, and to study the induced vocal–ventricular fold interactions from a physical point of view.

### Involvement of the Ventricular Folds in Voice Production

The medial adduction of the ventricular folds during human phonation can be observed in endoscopic examinations of individuals with no voice complaints, normal laryngeal structure, and normal voice quality (Behrman, Dahl, Abramson, & Schutte, 2003; Stager, 2011; Stager, Bielamowicz, Regnell, Gupta, & Barkmeier, 2000; Stager et al., 2001, Stager, Neubert, Miller, Regnell, & Bielamowicz, 2003). Although supraglottic activity has been largely considered as a compensatory behavior occurring in patients with glottis insufficiency or voice disorders (Lindstad, Blixt, Pahlbert-Olson, & Hammarberg, 2004; Nasri et al., 1996; Pinho, Pontes, Gadelha, & Biasi, 1999; Tsai, Celmer, Berke, & Chhetri, 2007; Von Doersten, Izdebski, Ross, & Cruz, 1992), Behrman et al. (2003) and Stager et al. (2000, 2001) demonstrated that medial compression of the ventricular folds can be a normal articulatory posture rather than one of dysfunction. In addition, several studies showed that these supraglottic structures can move closer to each other, come into contact, and even vibrate for specific sound productions, such as glottal stops (Esling, Zeroual, & Crevier-Buchman, 2007; Stager et al., 2000), whisper phonation (Crevier-Buchman, Vincent, & Hans, 2008; Rubin, Praneetvatakul, Gherson, Moyer, & Sataloff, 2006; Zeroual, Esling, & Crevier-Buchman, 2005), pressed or breathy voice qualities (Granqvist & Lindstad, 2001; Lindstad et al., 2004), or even during some musical performances such as lyric opera, growls, Japanese folk music, period-doubling, and throat singing (Bailly, Henrich, & Pelorson, 2010; Chevillier, Guilbault, Renard, & Tran Ba Huy, 2008; Fuks, Hammarberg, & Sundberg, 1998; Lindstad, Sodersten, Merker, & Granqvist, 2001; Sakakibara, Imagawa, Niimi, & Tayama, 2004; Zangger Borch, Sundberg, Lindstad, & Thalen, 2004).

The *ventricular-fold motion*, defined by any displacement of the ventricular-fold free edges in phonation, can come along with (a) an antero-posterior supraglottic compression deriving from a displacement of the petiole of the epiglottis and the arytenoids cartilages; or (b) a medial supra-glottic compression deriving from the contraction of the thyro-arythenoid muscle. The primary classification of the ventricular-fold motion in the laryngeal articulation relies on the ventricular folds’ degree of adduction toward mid-line, which can range from partial covering of the vocal folds to full covering (Stager, 2011). More specifically, two types of ventricular-fold compression in the medial-lateral direction were identified during a speech task, respectively referred to as “static compression” or “dynamic compression” (Stager et al., 2000, 2001). In *static compression* occurrences, the ventricular folds move into place at the initiation of the task, maintain the position throughout the whole task,

and then move apart back to the rest position at the end of the task. In *dynamic compression*, ventricular-fold adductory movements are rapid and they can occur several times within a speech task, depending on the phonetic context.

### ***Vocal–Ventricular Fold Interaction: A Biomechanical, Aerodynamic, and Acoustical Coupling***

In light of the previous laryngeal examinations during voicing, the thesis of a strong physical vocal–ventricular fold interaction is largely supported (e.g., Bailly et al., 2010; Fuks et al., 1998; Lindestad et al., 2001; Sakakibara et al., 2004). Therefore, during recent decades, numerous computational and experimental studies have investigated the biomechanical, aerodynamic, and acoustical effects underlying this interaction, so as to better understand the ventricular-fold effective contribution to the laryngeal sound source.

Many physical models of phonation were proposed to predict the impact of both the presence and the degree of adduction of the ventricular folds on the glottal-jet dynamics. Thereby, the translaryngeal airflow was extensively described under the assumption of rigid ventricular walls, combined with either rigid glottal boundaries (Iijima, Miki, & Nagai, 1992; Pelorson, Hirschberg, Wijnands, & Baillet, 1995; Nomura & Funada, 2007a), rigid forced-vibrating glottal walls (Zhang, Zhao, Frankel, & Mongeau, 2002), continuum-based finite-element modeling of deformable glottal tissues (Oliveira Rosa & Pereira, 2008; Oliveira Rosa, Pereira, Grellet, & Alwan, 2003; Zheng, Bielamowicz, Luo, & Mittal, 2009; Zheng, Mittal, Xue, & Bielamowicz, 2011), or low-dimensional distributed models of glottal self-sustained oscillations (Ikeda, Matsuzaki, & Aomatsu, 2001; Nomura & Funada, 2007b). In parallel, some experimental works were conducted to mimic physical phenomena likely to occur in the human larynx, but also to assess and validate the relevance of the theoretical outcomes. To this end, many authors conceived simplified mock human phonatory systems and performed *in vitro* experiments, thus providing an unmatched control in terms of reproducibility, repeatability, measurement accuracy, or initial experimental conditions (Agarwal, 2004; Brücker, Triep, & Kob, 2004; Gauffin, Binh, Ananthapadmanabha, & Fant, 1983; Kucinski, Scherer, DeWitt, & Ng, 2006; Li, Wan, & Wang, 2008; Pelorson et al., 1995; Shadle, Barney, & Thomas, 1991; Triep, Brücker, & Schröder, 2005). Many of these studies were carried out using flow visualization and pressure measurements in (quasi-)steady flow conditions through rigid motionless replicas of both vocal and ventricular folds (Agarwal, 2004; Gauffin et al., 1983; Kucinski et al., 2006; Pelorson et al., 1995). A few studies provided noticeable advances employing a static ventricular-fold constriction downstream to a forced-vibrating replica of the vocal folds driven in simple harmonic motion (Shadle et al., 1991), or mimicking the cyclic variations in the 3-D contour of the glottal gap and performing time-resolved velocity measurements (Brücker et al., 2004; Triep & Brücker, 2010; Triep et al., 2005). Spare studies were dedicated to the changes implied by the geometrical configuration of a static ventricular-fold replica on the vibratory properties

of mock vocal folds able to sustain self-oscillations (Bailly, Pelorson, Henrich, & Rutu, 2008; Drechsel & Thompson, 2008). Finally, the phonatory effects of the ventricular-fold presence and geometrical position were also evidenced on vibrating excised canine larynges by the recordings of the subglottal pressure and the electroglottographic (EGG), flow rate, audio, and sound pressure signals (Alipour & Finnegan, 2013; Alipour, Jaiswal, & Finnegan, 2007; Finnegan & Alipour, 2009), and air pressures within the laryngeal ventricle (Alipour & Scherer, 2012).

When the ventricular folds are located in the vocal tract, these previous numerical, *in vitro* and *ex vivo* experimental investigations have shown the following:

- Additional high-frequency components in the acoustical power may appear in the range 2–2.5 kHz, which is around the acoustic resonance frequency of the ventricle (Imagawa, Sakakibara, Tayama, & Niimi, 2003; Sakakibara et al., 2001; Zhang et al., 2002).
- Below 2 kHz, the results are still controversial: Little effect on the propagated pressure wave from the voice source was found in Zhang et al. (2002) and McGowan and Howe (2010), whereas the sound intensity level increased with the ventricular-fold medial compression in Alipour et al. (2007), Finnegan and Alipour (2009), and Zheng et al. (2009); an enhancement of the second partial of the acoustical signal was also measured in Finnegan and Alipour (2009), together with a decrease of low-frequency noise below 300 Hz.
- Under steady flow conditions, the translaryngeal airflow resistance through stationary laryngeal geometries increases or decreases by about 10%–25%, depending on the ratio of the ventricular gap over the glottal diameter,  $\gamma$  (Agarwal, 2004; Bailly et al., 2008; Li et al., 2008); under unsteady flow conditions also, an increase flow rate by about 20%, hence a reduction of the dynamic translaryngeal flow resistance can be generated through a moving glottis (Triep & Brücker, 2010; Zhang et al., 2002; Zheng et al., 2009) or through excised larynges vibrating with flow rates lower than 600 mL/s (Alipour et al., 2007).
- Antero-posterior compression of the ventricular folds has minor effects on glottal flow resistance compared to medial compression, although both contributed positively to the sound intensity (Alipour et al., 2007).
- In the convergent opening phase of the phonatory cycle, the glottal jet emerges and tends to attach randomly to one of the ventricular-fold side walls or the other from cycle to cycle (Brücker et al., 2004; Kucinski et al., 2006; Triep & Brücker, 2010).
- In the divergent closing phase of the cycle, the jet is straightened out into the axial direction and stabilized (Drechsel & Thompson, 2008; Kucinski et al., 2006; Li et al., 2008; Pelorson et al., 1995; Shadle et al., 1991; Triep & Brücker, 2010; Zheng et al., 2009).
- A downstream shift of the glottal separation point takes place, inducing a conservation of the flow

laminar properties over a longer distance (Kucinski et al., 2006; Li et al., 2008).

- An aid in the ability to phonate may be expected. Indeed, the inclusion of the ventricular folds in the vocal tract can lead to larger glottal oscillatory amplitude by about 33% for a given translaryngeal pressure drop (Zheng et al., 2009), and to a decrease of phonation threshold pressures for  $\gamma$  ranging between 1 and 7 (Bailly et al., 2008).
- Fundamental frequency ( $f_0$ ) of glottal oscillations is also affected by the ventricular-fold degree of adduction, yet in a controversial manner. A progressive  $f_0$  drop-off is obtained with an increasing ventricular level of adduction for  $\gamma < 1$  in Bailly et al. (2008), whereas it comes along with a rise in oscillation frequencies in Alipour et al. (2007) and Finnegan and Alipour (2009).

### Issues, Focus, and Objectives

In summary, past studies have shown the relevance of the ventricular-fold presence and degree of adduction on both translaryngeal airflow and resulting sound. However, several limitations make the issue of their physical contribution to phonation not fully settled:

1. As pointed out by Zheng et al. (2009), *the impact of the ventricular folds on the glottal self-sustained oscillations* has been addressed in very few studies (Alipour et al., 2007; Finnegan & Alipour, 2009; Zheng et al., 2009);
2. Similarly, the majority of *previous numerical and in vitro studies tackled with the static constriction of the ventricular folds*, that is, considering a constant level of adduction throughout phonation. Only few authors have provided a first theoretical description of the ventricular-fold dynamics, that is, considering a varying level of adduction throughout phonation. Either a one-mass (Miller, Pereira, & Thomas, 1988) or a two-mass model (Imagawa et al., 2003; Sakakibara, Imagawa, Niimi, & Osaka, 2002) has been used to simulate ventricular-fold vibrations and explore their aeroacoustic effects;
3. Last, to the best of our knowledge, *a quantitative connection between the current numerical models of vocal–ventricular fold interaction and the vocal–ventricular-fold dynamics as really observed during human phonation is still lacking*.

Thus, in the current study, focus is given to the ventricular-fold dynamics and the aerodynamic aspect of the vocal–ventricular fold interaction in human phonation. The objective of this article is threefold: (a) to provide a detailed classification of the ventricular folds' motion during human voice production by means of high-speed cinematography, (b) to study the impact of ventricular-fold motion on the glottal self-sustained oscillations for each identified category, and (c) to assess whether ventricular-fold

self-sustained oscillations could be simulated using a common and simplified aerodynamical modeling approach.

An original database was recorded to explore the phonatory gestures of five subjects, using high-speed cinematographic, audio, and EGG recordings. A quantitative characterization of ventricular-fold dynamics was derived from the detection of ventricular areas, the kymographic processing of the laryngeal images, and the analysis of EGG signals. Combining these physiological correlates with a simplified physical modeling of phonation, the aerodynamic interactions between vocal and ventricular folds were simulated.

## Method

### *In Vivo Recordings*

The recording sessions took place at the Department of Voice, Speech and Hearing Disorders of the University Medical Center Hamburg-Eppendorf (UKE, Hamburg, Germany).

*Subjects.* Five subjects (four male, one female) participated in the study. One subject (S1) was a professional singer. No subject had known voice problems nor any laryngeal anomaly. Table 2 lists the subjects' characteristics.

*Voice tasks and database.* Each subject was asked to produce several voice tasks, including to sustain a vowel in normal voice (modal phonation using laryngeal mechanism M1) and in different vocal registers (vocal fry using laryngeal mechanism M0; falsetto or head register using M2, whistle register using M3 when possible),<sup>1</sup> to perform glides with transitions between registers, and to yodel when possible, to increase or decrease vocal effort (crescendo, shout, decrescendo), to produce growls,<sup>2</sup> to do throat singing when possible, and finally to produce any sound with which the subject had the feeling of using ventricular folds.

When possible, the voice tasks were recorded for several vowels and pitches. In the end, the database comprised 72 recorded sequences. The number of recorded sequences and the voice tasks managed by each subject are detailed in Table 2. Only Subject S1 was able to produce throat singing. Only two subjects could produce yodel sequences: S1 (with a professional mastery), and S4 (as a vocal exercise). Subject S1 was the only one who could produce a sound during which he felt he was using his ventricular folds (Louis-Armstrong-type growl).

### *Measurements*

The laryngeal exploration was conducted using high-speed cinematography synchronized with audio and EGG recordings.

<sup>1</sup>See Henrich (2006) and Roubeau, Henrich, and Castellengo (2009) for more details about laryngeal mechanisms and vocal registers.

<sup>2</sup>*Growl* is a vocal effect that can be used in jazz, blues, or pop music and that perceptually sounds hoarse and low-pitched (Loscos & Bonada, 2004; Sakakibara et al., 2004). By contrast, *shout* is produced with a louder voice, high pitch, and perceptible vocal effort.

**Table 2.** Descriptions of the subjects and of their high-speed-recorded voice productions.

Subject	Gender	Age	N	Voice productions (number of sequences)
S1	M	41	8	M0 (1), throat singing (1), yodel (3), growl (2), shout (1)
S2	M	33	10	normal voice (2), glide using M1 (1), M0 (1), growl (4), shout (2)
S3	M	37	24	normal voice (1), M0 (1), growl (9), shout (1), glide (7) [M1 (4), M1–M2 (1), M1–M0 (2)], (de)crescendo (5)
S4	M	49	16	normal voice (1), M0 (3), growl (3), shout (2), yodel (1), glide (6) [M1–M2 (2), M3–M2 (3), M0–M1 (1)]
S5	F	25	15	normal voice (1), glide using M1 (2), shout (11), crescendo (1)

Note. Each production sequence allowed the assessment of ventricular motion, apart from a Louis-Armstrong-type growl produced by Subject S1, for whom aryepiglottic constriction prevented observing the vocal and ventricular spaces. M = male; F = female; N = number of completed tasks per subject.

*High-speed cinematography recording.* A digital 8-bit grayscale high-speed camera (Wolf HS-Endocam 5560) attached to a rigid endoscope (Wolf 90°E 60491) with a continuous xenon fiber-optic light source (Wolf 5131) was used. The high-speed images were recorded at a rate of 2,000 frames/s with a resolution of  $256 \times 256$  pixels and a total recording time of 4 s.

*Audio and EGG recording.* The audio signal was recorded on the high-speed-camera internal microphone (Wolf 5052.801) at a sampling frequency of 44170 Hz. A dual-channel electroglottograph rate (EG2, Glottal Enterprises, Syracuse, NY; Rothenberg, 1992) was used to record the EGG signal simultaneous to the audio signal and at the same sampling. The quality of the EGG signal was monitored during experiments on an analog–digital oscilloscope.

### Data Processing

The data were processed using MATLAB. The nomenclature used in the following is detailed in Appendix A. In particular, each geometrical variable  $X(t)$  introduced in the theoretical description as a function of time,  $t$ , refers to a real quantity measurable in vivo, noted  $\tilde{X}$ .

*Audio and EGG signal processing.* The audio and EGG signals were synchronized with the images, with a precision of less than one image time-lag. The EGG signals were differentiated (*DEGG signal*) to estimate the speed of glottal contact and detect glottal closing and opening instants, using a threshold-based peak detection method (Henrich, d’Alessandro, Castellengo, & Doval, 2004). The acoustical fundamental frequency  $\tilde{f}_0$  was extracted from the audio signal using the YIN-autocorrelation method developed by de Cheveigné and Kawahara (2002).

*Image processing for ventricular and glottal apertures’ estimation.* The methods for high-speed images’ processing and in vivo–data estimation have been previously described in Bailly et al. (2010). Kymographic plots were derived from the image lines perpendicular to the detected glottal median axis. Ventricular edges were manually detected on each frame by means of Bezier-curves. Ventricular space area ( $\tilde{A}_{\text{venf}}$ ) was derived by fitting generic surface triangular meshes to the detected contours (Serrurier & Badin, 2008), as illustrated in Figure 2a. The area signal was normalized to its maximal value on the selected time sequence (see Figure 2b), and up-sampled to the EGG sampling

frequency (44170 Hz) by cubic interpolation. Ventricular aperture  $\langle \tilde{h}_{\text{venf}} \rangle$  was derived to be used in theoretical modeling (see Figure 2c). When glottal plane was visible all along the sequence, glottal aperture  $\langle \tilde{h}_{\text{vof}} \rangle$  was estimated in the same way.

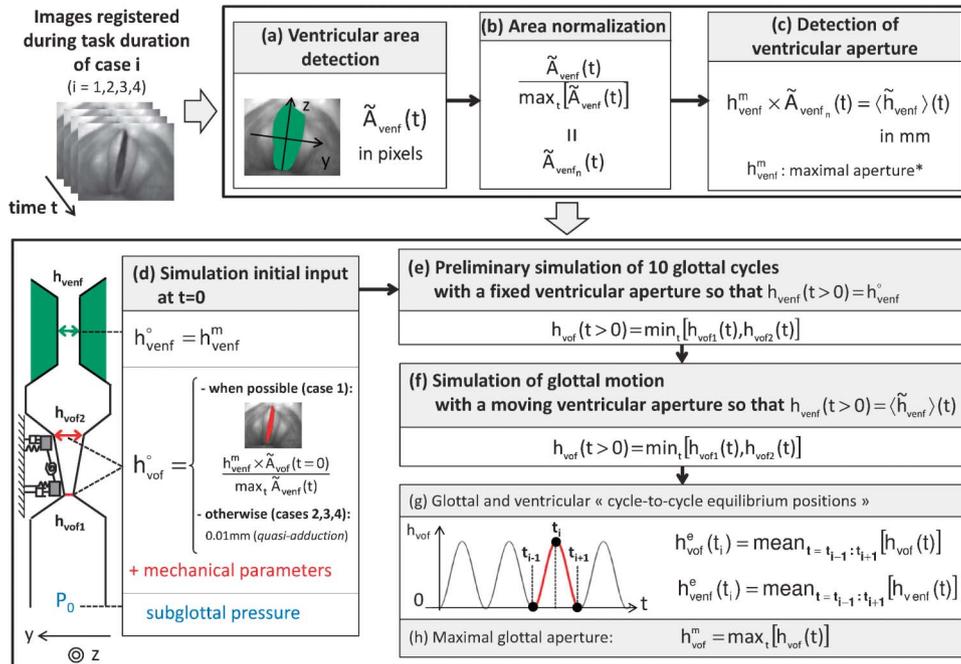
### Theoretical Modeling

A simplified two-mass model ( $M2M_{\text{vof}}$ ) was used to simulate vocal-fold self-sustained oscillations (Bailly et al., 2008; Lous, Hofmans, Veldhuis, & Hirschberg, 1998). The ventricular motion was either an input signal derived from in vivo data (see Figure 1b), or an output signal simulated by an additional two-mass model of ventricular folds ( $M2M_{\text{venf}}$ ) attached downstream to the  $M2M_{\text{vof}}$  model (see Figure 1c). Using ventricular motion as an input signal allowed the simulation of ventricular impact on vocal-fold vibration. Deriving it as an output signal enabled us to simulate vocal–ventricular vibrations and investigate possible physical reasons for ventricular motion.

*Simulation of ventricular impact on vocal-fold vibration.* A representative sketch of the phonation model considered in this section is given in Figure 1b, along with relevant mechanical, geometrical and aerodynamic parameters. The  $M2M_{\text{vof}}$  model yields the simulation of glottal motion in time, through the predictions of apertures  $h_{\text{vof1}}$  and  $h_{\text{vof2}}$  (Lous et al., 1998; Ruty, Pelorson, Van Hirtum, Lopez Arteaga, & Hirschberg, 2007). Simulations were conducted for four cases of ventricular-fold aperture’s time-variations, numbered from 1 to 4, representative of the recorded database as detailed in the next section. Figure 2 summarizes the simulation steps used for the four cases.

*Mechanical input parameters.* The reduced mechanical model is a function of mass  $m_{\text{vof}}$ , spring stiffness  $k_{\text{vof}}$ , coupling constant  $k_{\text{cvof}}$ , and damping  $r_{\text{vof}}$ . The value of  $m_{\text{vof}}$  was chosen in agreement with previous studies on voice physical modeling (Ishizaka & Flanagan, 1972; Miller et al., 1988). For each case, stiffness  $k_{\text{vof}}$  and damping  $r_{\text{vof}}$  input parameters were fitted so that the acoustical fundamental frequency was identical to the measured one from the subjects ( $\tilde{f}_0$ ) with an average discrepancy below 5%. Their values are given in Table 3.

**Figure 2.** Flow chart from the post-processing of the in vivo high-speed images to the simulation of the ventricular forced-motion's impact on glottal vibrations. Procedure applied for use of the model displayed in Figure 1b, and for Cases 1 to 4 (see Table 4).



The coupling constant  $k_{cvof}$  was set equal to  $0.5 k_{vof}$  (Ruty et al., 2007; Vilain, 2002).

**Geometrical input parameters.** A part of the M2M<sub>vof</sub> model geometrical input parameters was kept constant whatever the simulated case (1 to 4). Those parameters were chosen in line with anatomical data available in the literature and summarized in Table 4. Two other key parameters were varied accordingly to the case under study: the initial glottal aperture (at  $t = 0$ ) and the ventricular-fold aperture's time variations during the task. The initial glottal aperture (see  $h_{vof}^o$  in Figure 2d and Table 3) was either derived from in vivo data when possible (for Case 1 only), or

set at a fixed predefined value (for Cases 2 to 4). A preliminary simulation of 10 glottal cycles with a fixed ventricular aperture was performed to ensure established glottal oscillations at the adjusted mechanical parameters (see Figure 2e). Then, time variations of the ventricular-fold aperture as derived from in vivo images were imposed as a forced input motion (see Figure 2f). In the following, this input ventricular motion is characterized by the maximal ventricular-fold aperture achieved during the motion,  $h_{vent}^m$  (see Figure 2c).

**Laryngeal airflow model.** General thermodynamic conditions and aerodynamic parameters are reported in

**Table 3.** Geometrical, aerodynamic, and mechanical parameters of the M2M<sub>vof</sub> simulations corresponding to the ventricular-fold static case (Case 1) and to the three dynamic cases under study.

Ventricular static or dynamic case	Case 1	Case 2	Case 3	Case 4
	No motion	Slow nonoscillatory motion	Fast motion aperiodical vibrations	Fast motion periodical vibrations
Acoustical fundamental frequency (mean value), $f_0$ (Hz)	119 <sup>a</sup>	320 <sup>a,b</sup>	312 <sup>a,b</sup>	75 <sup>a</sup>
Initial glottal aperture, $h_{vof}^o$ (mm)	0.6 <sup>a</sup>	0.01	0.01	0.01
Initial ventricular aperture, $h_{vent}^o$ (mm)	5.5	1.5	1.5	1.5
Maximal glottal aperture, $h_{vof}^m$ (mm)	2.7	1.7	1.8	1.9
Maximal ventricular aperture, $h_{vent}^m$ (mm)	5.5	1.5	1.5	1.5
Subglottal pressure, $P_0$ (Pa)	900	4500	4500	900
Spring mass, $m_{vof}$ (g)	0.05	0.05	0.05	0.05
Spring stiffness, $k_{vof}$ (N/m <sup>-1</sup> )	19.7	164.9	207.2	44.4
Damping, $r_{vof}$ ( $\times 10^{-3}$ N/s/m <sup>-1</sup> )	0.8	2.3	2.3	1.0

<sup>a</sup>Derived from in vivo measured data. <sup>b</sup>Evaluated on the first five glottal cycles detected on DEGG<sub>n</sub> signal.

**Table 4.** Geometrical and aerodynamic constant input parameters in the M2M<sub>vof</sub> and M2 × 2M modeling.

Parameter	Value
Geometrical parameter	
Trachea height, $h_0$ (mm)	15.7
Ventricle height, $h_{ventricle}$ (mm)	15.7
Glottal length, $d_{vof}$ (mm)	4.0
Ventricular length, $d_{venf}$ (mm)	5.5
Ventricle length, $d_{ventricle}$ (mm)	4.7
Aerodynamic parameter	
Temperature (°)	37.2
Atmospheric pressure (kPa)	101.1
Air density, $\rho$ (kg.m <sup>-3</sup> )	1.13
Air dynamic viscosity, $\mu$ (kg.m <sup>-1</sup> .s <sup>-1</sup> )	$1.85 \times 10^{-5}$
Speed of sound, $c_0$ (m.s <sup>-1</sup> )	353.3

Table 3. The pressure  $P_2$  in the laryngeal ventricle was simulated as previously detailed in Bailly et al. (2008, 2010), assuming a semi-empirical Liljencrant's flow separation model at vocal and ventricular folds, a jet-flow linear geometrical expansion in the ventricle, dissipation in the jet being neglected, a zero-pressure at the outlet of the ventricular folds, and a quasi-steady Bernoulli flow dynamics description. The pressure  $P_0$  upstream of the two-mass model was parametrically controlled as input data (see Table 3). For low-pressure simulations, it was set to 900 Pa, within the range of subglottal pressures commonly measured in the trachea (Holmberg, Hillman, & Perkell, 1988). For high-pressure simulations (growl, shout), it was set to 4500 Pa, that is, five times higher than the common value but similar to what was measured by Kitajima and Fujita (1990).

*Summary of the model output.* Finally, the output variables of the M2M<sub>vof</sub> model are as follows:

- the glottal aperture as a function of time,  $h_{vof}(t)$  (see Figure 2f);
- the “glottal cycle-to-cycle equilibrium position,”  $h_{vof}^e(t)$ , calculated as the mean value of  $h_{vof}(t)$  during each simulated glottal cycle (see Figure 2g);
- the “ventricular cycle-to-cycle equilibrium position,”  $h_{venf}^e(t)$ , calculated as the mean value of  $h_{venf}(t)$  during each simulated glottal cycle (see Figure 2g);
- the maximal glottal aperture achieved during glottal motion,  $h_{vof}^m$  (see Figure 2h).

*Simulation of vocal–ventricular vibrations.* The proposed M2 × 2M model of phonation combines the M2M<sub>vof</sub> model presented above with the M2M<sub>venf</sub> model of ventricular motion as illustrated in Figure 1c. Similar to the M2M<sub>vof</sub> model, the M2M<sub>venf</sub> model is a function of a set of mechanical parameters comprising mass  $m_{venf}$ , spring stiffness  $k_{venf}$  and damping  $r_{venf}$ . These mechanical input parameters were chosen to give agreement with the frequency response functions of the ventricular folds identified by Svec, Horacek, Sram, and Vesely (2000) from a 50–200-Hz frequency sweep of sinusoidal excitation combined with laryngoscopy. Resonance frequency

of the ventricular folds was measured around 70 Hz ( $\tilde{f}_r$ ), with a typical half-power (3 dB) bandwidth of 24 Hz, yielding to a quality factor  $\tilde{Q}_r$  of about 3. Assuming comparable masses of vocal and ventricular folds ( $2m_{vof} = 2m_{venf} = 0.10$  g), the M2M<sub>venf</sub> stiffness and damping parameters could then be calculated (Ruty et al., 2007):

$$k_{venf} = m_{venf} \cdot \left(2\pi\tilde{f}_r\right)^2 = 9.7\text{N} \cdot \text{m}^{-1}$$

$$r_{venf} = m_{venf} \cdot \frac{2\pi\tilde{f}_r}{\tilde{Q}_r} = 7.5 \times 10^{-3}\text{N} \cdot \text{s} \cdot \text{m}^{-1}. \quad (1)$$

The driving pressure upstream of the M2M<sub>venf</sub> model equals the supraglottal pressure  $P_2(t)$  as assessed in the ventricle by the laryngeal airflow model at each M2M<sub>vof</sub> iteration. Geometrical, aerodynamical and mechanical input parameters are summarized in Table 5. Finally, the M2 × 2M model yields the simulation of both glottal and ventricular motion in time, through the predictions of the apertures  $h_{vof1}(t)$ ,  $h_{vof2}(t)$ , and  $h_{venf1}(t)$ ,  $h_{venf2}(t)$ . As in the M2M<sub>vof</sub> model, the predicted ventricular aperture is defined by  $h_{venf}(t) = \min_t[h_{venf1}(t), h_{venf2}(t)]$  and the derived ventricular cycle-to-cycle equilibrium position is referred as  $h_{venf}^e(t)$ .

## Results

In the following, all online supplemental materials are available at <http://hal.archives-ouvertes.fr/hal-00998464>.

### Time-Based Classification of Ventricular-Fold Dynamics

The high-speed sequences recorded in the database demonstrated a large variety of ventricular-fold dynamics, characterized by different velocities and amplitudes of narrowing or widening, possible occurrences of ventricular-fold contact, vibratory patterns, and stabilized periodicity. We do not take into account here the motions co-occurring with vocal-fold adduction and abduction at phonation starts and stops, but the ones detected throughout the voice tasks. Note that a part of the sequences showed a level of ventricular-fold adduction that does not vary in time. Most

**Table 5.** Geometrical, aerodynamic, and mechanical input parameters of the M2 × 2M simulations.

Parameter	Value
Initial glottal aperture, $h_{vof}^o$ (mm)	0.1
Maximal ventricular aperture, $h_{venf}^m$ (mm)	5.5
Subglottal pressure, $P_0$ (Pa)	4500
Spring mass, $m_{vof}$ (g)	0.05
Spring stiffness, $k_{vof}$ (N.m <sup>-1</sup> )	207.2
Damping, $r_{vof}$ ( $\times 10^{-3}$ N.s.m <sup>-1</sup> )	2.3
Spring mass, $m_{venf}$ (g)	0.05
Spring stiffness, $k_{venf}$ (N.m <sup>-1</sup> )	9.7
Damping, $r_{venf}$ ( $\times 10^{-3}$ N.s.m <sup>-1</sup> )	7.5

often, the motionless ventricular constriction is placed so that the glottal plane is visible all along the sequence. This case is comparable to the static component of ventricular-fold compression described by Stager et al. (2000, 2001), and is referred to as the “no motion” case hereafter.

Based on the temporal variation of ventricular adduction, two elementary categories of ventricular-fold dynamics were identified in the whole database:

1. *Slow nonoscillatory motion* (see Appendix B, Figure B1). The level of ventricular-fold adduction moves gradually from wide to narrow or from narrow to wide. Time duration spent by the ventricular folds to move from a minimal to a maximal level of adduction (or vice versa) is typically beyond the range of 10 glottal periods. Thus, this motion can be considered as a slow motion relative to the glottal period.
2. *Fast oscillatory motion* (see Appendix B, Figure B2). The level of ventricular-fold adduction oscillates around an equilibrium position. Time duration spent by the ventricular folds to move from a minimal to a maximal level of adduction (or vice versa) during an oscillation is typically below the range of 10 glottal periods. Thus, this motion can be considered as a fast motion because it is on the same order as the glottal period.

Similar to electric current or airflow—physical entities that can be characterized by the combination of both a continuous (slow time-varying) component and an alternative (fast time-varying) one—these categories of ventricular-fold motion may be combined within a sequence. Most often, a fast oscillatory motion is preceded by a slow nonoscillatory narrowing. Oscillations may start during the narrowing movement.

Table 6 lists the recorded voice tasks and relates them to the two elementary categories of ventricular-fold dynamics, as well as to the “no-motion” case. Thirteen sequences out of 72 showed a constant ventricular-fold level of adduction, associated with four normal-voice tasks, seven glides, one M0, and one growl sequence. In the remaining sequences (82% of the database), a slow nonoscillatory motion, a fast oscillatory one, or a combination of both movements were observed.

### *Slow Nonoscillatory Motion*

As shown in Table 6, a ventricular narrowing or widening with no superimposed oscillatory motion was found in

all the yodel sequences (4/4), in most M0 sequences (4/6), in half of the glides (8/16), in few crescendos and decrescendos (2/6) and in one normal-voice task (related to a glottal attack). The ventricular narrowing motion was characterized either by an anterior narrowing of the ventricular folds (“V” shape; e.g., see Figure 4, High-Speed Image [c]) or a median one (“U” shape; e.g., see Figure 5, High-Speed Image [d]), most often symmetric with respect to the glottal midline. In M0, for instance, the narrowing occurred systematically in the antero-median part of the ventricular fold. It went with an epiglottis tilt toward the glottal plane. In most M0 sequences characterized by a ventricular-fold adductory motion, the ventricular movement ended with a partial ventricular contact (over part of the ventricular-fold length), or a total contact (complete ventricular closure), in which case phonation stops (vanishing periodicity on DEGG signals).

A ventricular motion with no additional oscillation was observed neither in growls nor in throat singing. It was found in shout for only two sequences (2/17), the other shout sequences showing a slow motion combined with fast ventricular-edge oscillations. In most sequences of growls, throat singing, and shout, the narrowing initiated a ventricular vibration.

The following figures illustrate several slow nonoscillatory motions with an increasing degree of ventricular narrowing. Ventricular-fold dynamics vary from a slight displacement of ventricular edges throughout yodel to a narrowing-to-contact motion throughout a descending glide from laryngeal mechanisms M2 (falsetto) to M0 (vocal fry).

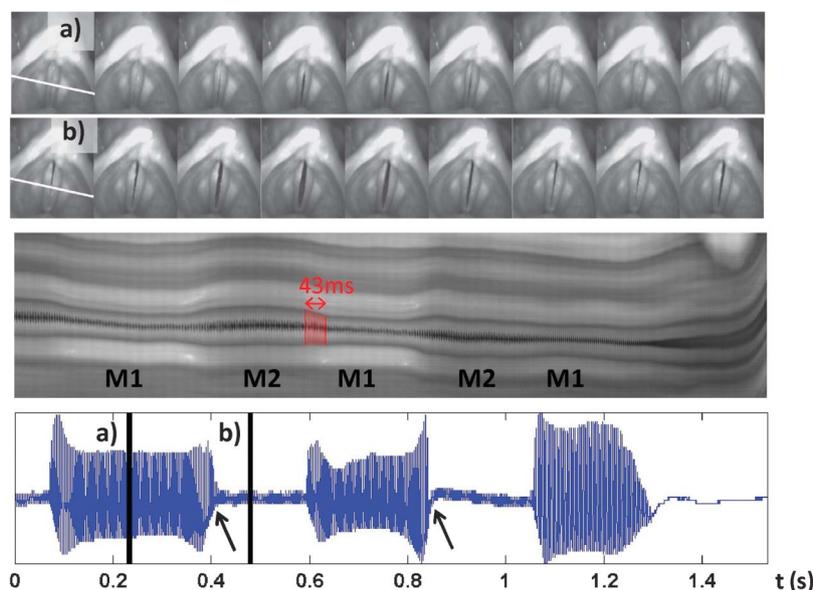
A typical ventricular narrowing during yodel task performed by Subject S1 is shown in Figure 3. The singer played on the transition between laryngeal mechanisms M1 and M2 to produce noticeable pitch jumps, representative of yodel-singing style. The EGG-signal envelope (Figure 2, bottom panel) reflects the sudden decrease in vocal-fold contact area when switching from M1 to M2 (see arrows plotted on EGG signal). The production in M2 is characterized by a lack of full contact, as illustrated in the image sequence (top, Panel [b]) and reflected on EGG signal (no contact detection; Roubeau et al., 2009). Ventricular motion is shown in the middle panel, by means of a kymographic analysis along the glottal line plotted in High-Speed Images (a) and (b). Time duration spent by the ventricular folds to move from a minimal to a maximal level of

**Table 6.** For each voice task, percentage of sequences that did not demonstrate ventricular motion and that did demonstrate ventricular motion of each type.

Ventricular motion	Normal voice ( <i>n</i> = 5)	Yodel ( <i>n</i> = 4)	Glides ( <i>n</i> = 16)	M0 ( <i>n</i> = 6)	Crescendos/ decrescendos ( <i>n</i> = 6)	Shout ( <i>n</i> = 17)	Growls ( <i>n</i> = 17)	Throat singing ( <i>n</i> = 1)
No motion	80	0	44	16.5	0	0	6	0
Slow nonoscillatory motion	20	100	50	67	33	12	0	0
Fast oscillatory motion	0	0	6	16.5	67	88	94	100

Note. *n* = number of sequences recorded for each voice task.

**Figure 3.** Illustration of slight ventricular adductory and abductory motions on a yodel produced by Subject S1 (slow nonoscillatory motion). Top panels show the laryngeal configuration during a glottal cycle in mechanisms M1 (a) and M2 (b) (see higher-resolution images in online supplemental materials Figure 1). The kymographic analysis is shown in middle panel. Synchronized EGGn variations are given in bottom panel as a function of time (in seconds).



adduction lasted about 13 glottal cycles (43 ms), as illustrated in the kymogram. Within this delay, a slight ventricular narrowing occurred when going from M2 to M1, and a slight widening from M1 to M2.

A greater narrowing movement is illustrated in Figure 4 during a decrescendo task produced by Subject S3. The kymographic analysis corresponding to a median glottal line (top panel) shows the gradual ventricular narrowing occurring from shot instants t(a) to t(d) (duration = 1500 ms). The degree of constriction remained stable at this minimal opening during nearly 250 ms, followed by the release of the ventricular constriction, in about half a second. At the end of the ventricular adductory motion at t(d), the ventricular maximal aperture along the ventricular-fold length got smaller than the vocal one, as shown by the high-speed images displayed in the middle panel at different shot instants. The acoustic analysis of the sequence is also presented in the bottom panels, showing a spectrogram of the sequence (left) and the variations of vocal intensity (sound-pressure level) as a function of time. A 6-dB drop-off was measured during the abductory motion of the ventricular folds, from shot instants t(d) to t(g). The ventricular movement from t(a) to t(f) is reflected on the spectral-energy repartition in the 3–4-kHz frequency band. As illustrated in the spectrogram, the spectral energy boost induced by the fourth formant  $F_4$  decreased from about 3.5 kHz, t(b), to 3 kHz, t(d).

Finally, Figure 5 illustrates a narrowing-to-contact ventricular motion observed during a 2-s descending glide from M2 to M1 to M0 produced by Subject S3. Instants of mechanisms' changes are derived from frequency jumps

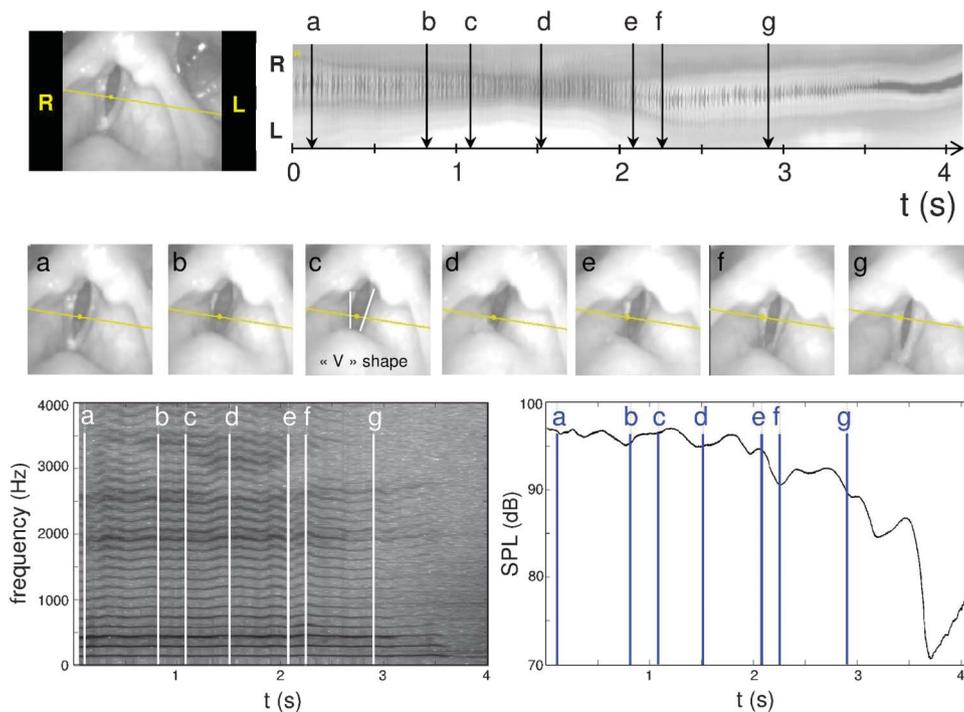
measurable in the sequence's spectrogram (bottom panel), together with the changes of the glottal vibratory patterns detected in the synchronized kymogram (middle panel) corresponding to a given median glottal line (top panel). A progressive ventricular adductory motion is clearly evidenced on the kymogram, starting from a wide and constant aperture during M2 (from E4-330 Hz to C4-260 Hz), a narrowing of the ventricular edges emphasized during M1 (from F3-170 Hz to G2-98 Hz), and continued in M0 (below G2-98 Hz) to achieve a complete ventricular closure, beyond which phonation vanished. This gradual ventricular narrowing took place during about 1680 ms.

### *Fast Oscillatory Motion*

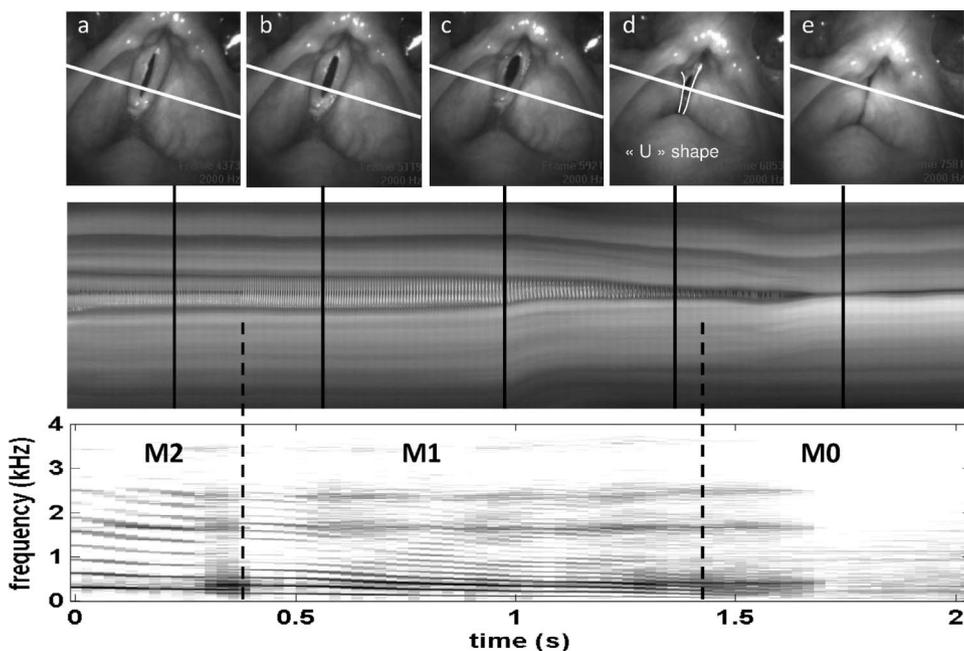
A ventricular-fold fast oscillatory motion was observed in 53% of the recorded sequences, as shown in Table 6 and Appendix B, Figure B2. It was systematically observed in throat singing (one sequence), and very commonly found in growls (16/17 growl sequences) and shout tasks (15/17 shout sequences). It was observed in crescendos-decrescendos (4/6 crescendos-decrescendos sequences) and was occasionally found in vocal fry with laryngeal mechanism M0 (1/6 sequence) and during one glide.

This oscillatory motion consisted of different schemes of ventricular-fold/vocal-fold co-vibrations. Representative examples of aperiodical or periodical schemes of ventricular-fold vibration are illustrated on kymograms in Figure 6 during several growl and throat-singing sequences. The oscillatory motion is also shown for the whole database in Appendix B, Figure B2. A ventricular contact was not

**Figure 4.** Illustration of strong ventricular adductory motion on a decrescendo produced by Subject S3 (slow nonoscillatory motion). A kymographic analysis of the sequence is presented in top right panel, which is a time representation of the glottal line shown in top left panel. Laryngeal configuration is illustrated at different instants in middle panels (see higher-resolution images in online supplemental materials Figure 2). Acoustic analysis is presented in bottom panels: time-frequency analysis of the sequence (left) and vocal intensity (SPL) as a function of time (right).



**Figure 5.** Illustration of ventricular narrowing-to-contact motion on a descending glide from M2 to M0, produced by subject S3 (slow nonoscillatory motion). Laryngeal configuration is given for several instants in top panels (see higher-resolution images in online supplemental materials Figure 3). The kymographic analysis is shown in bottom panel.



mandatory to this vibratory movement (e.g., see Figure 6a). Yet, the ventricular vibration went together with a ventricular contact in 70% of cases. Interestingly, the subjects did not have a proprioceptive feeling of the ventricular contact or vibration.

*Aperiodical vibrations.* Aperiodical ventricular-fold vibrations were commonly observed during shouts and growls. Typical kymograms obtained in shouts and growls are illustrated in Figures 6 and 7 and Appendix B, Figure B2. The aperiodical ventricular vibrations were observed independent of the vowel ([e], [a], [ae], [o], [ɛ]) and for several pitches. They were perceived as a specific hoarse timbre. Ventricular-fold aperiodical vibratory movement followed a noticeable ventricular adductory motion, which was most often done with contact. The vibration was characterized by an irregularity of the oscillatory amplitude. Although some ventricular vibratory “cycles” could sometimes be observed, the frequency of these oscillations was often unstable throughout the task (e.g., see S5<sub>06</sub>, S4<sub>04</sub>, S4<sub>016</sub> in Appendix B, Figure B2). In the case of vibration with ventricular contact, the vibration irregularity could also be related to the contact occurrence and duration, and to the ventricular length in contact. In some cases, additional supraglottal structures also vibrated (aryepiglottic folds, laryngeal mucous membrane).

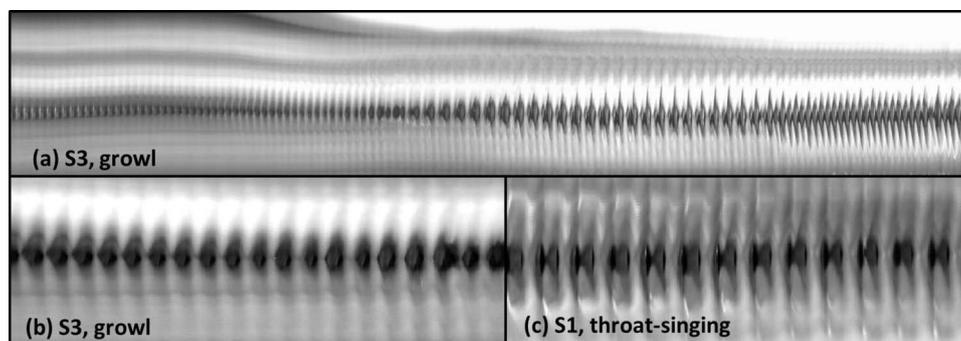
Figure 7 illustrates several features of a typical ventricular aperiodical vibratory movement registered in the case of a shout (Subject S4). Two ventricular vibrations could be observed during the 12 glottal cycles shown on the EGG signal and its derivative. During the ventricular contact phase, the contact propagated along the whole length of the ventricular folds, as evidenced by the five kymograms matched to the lines perpendicular to the glottal median axis and plotted on the top left laryngeal image. Vibratory amplitude and contact duration were modified from one ventricular-fold oscillation to the other (e.g., see differences of ventricular-fold contours between both vibrations on Kymogram 2). However, the time delay between two successive glottal closing instants, which were detected by the DEGG signal maxima (e.g., at shot instants t[a] to t[d]),

remained unchanged in spite of the downstream ventricular-fold vibrations. Therefore, the periodical glottal vibration was maintained at a fundamental frequency around 311 Hz, and it did not stop during a ventricular contact phase.

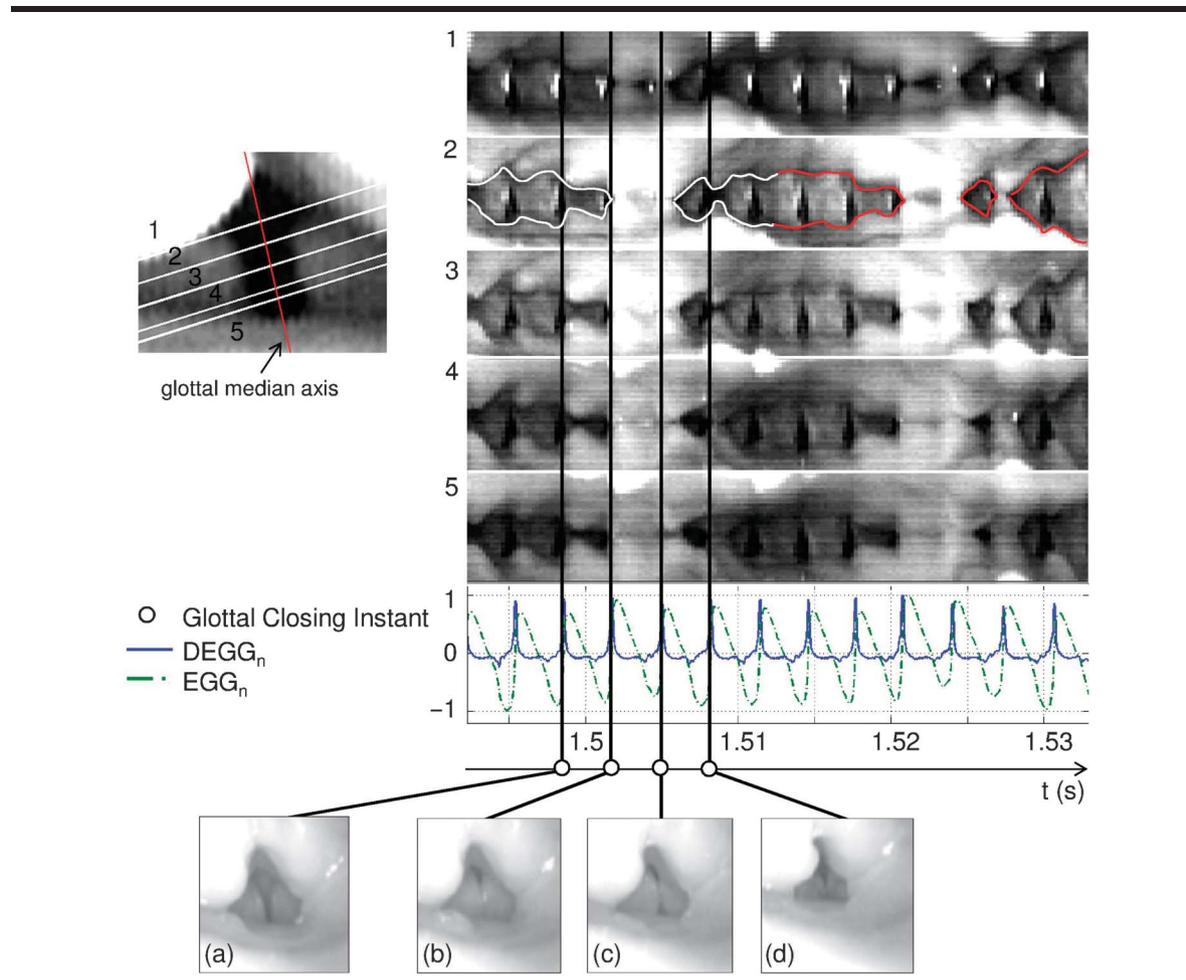
*Periodical vibrations.* Periodical ventricular-fold vibrations were observed during throat singing. Occasionally they were observed during laryngeal mechanism M0 and shout. Two types of ventricular-fold vibrations can be distinguished according to their fundamental cycle duration, relative to the glottal one.

The ventricular periodicity can be equal to the glottal one. This was typically observed during crescendos in the starting phase of shouts and growls. Figure 8 illustrates a representative ventricular periodical movement during a crescendo sequence prior to a growly phonation (Subject S5). As shown by the kymogram in the top panel, the ventricular vibration was initiated without ventricular contact. The amplitude of the ventricular vibrations was first negligible as compared to that of the vocal-fold vibrations. Throughout the task, the ventricular folds moved closer to each other so that the vibration gradually tended to a stable oscillation with amplitudes comparable to vocal-fold motion and a periodical ventricular contact. The zoomed kymogram illustrates the similarity in cycle duration between the vocal folds and the ventricular folds, and the phase difference in opening and closing. Ventricular opening commonly preceded glottal opening, and ventricular closing occurred during glottal open phase. As shown in the spectrographic sound analysis (“transition” region), despite the ventricular vibrations, the glottal fundamental frequency ( $\tilde{f}_0 = 301$  Hz in average) and higher harmonics were preserved, and no additional frequencies were observed. Therefore, the periodical ventricular covibration did not seem to noticeably affect the glottal-source properties. This result is in line with previous observations by Lindestad, Hammarberg, Larsson, and Granqvist (1998) and Lindestad et al. (2004), who showed that the ventricular folds may vibrate in the same phase as the true vocal folds without interfering with sound production and necessarily adding rough quality to the voice.

**Figure 6.** Examples of ventricular-fold oscillatory motions. (a) Periodical and aperiodical ventricular vibrations, with and without ventricular contact. (b) Periodical ventricular vibration, in phase with glottal vibration and at same periodicity. (c) Periodical ventricular vibration with ventricular contact every two glottal cycles.



**Figure 7.** Illustration of ventricular aperiodical vibration during a shout. Subject S4,  $\tilde{f}_0 \sim 311\text{Hz}$  (D#4). Top left panel shows a laryngeal image and the five selected kymographic lines. Top right panels show the kymogram analysis for each line. The middle panel gives the corresponding EGG<sub>n</sub> signal (dashed line) and its derivative (solid line). The laryngeal configuration is illustrated for four instants (a) to (d); see higher-resolution images in online supplemental materials Figure 4. Contours of ventricular-folds' free edges are highlighted for the two ventricular vibrations in Kymogram 2 (in white and red).



Ventricular folds can also periodically vibrate at a fundamental frequency different from that of the vocal folds. This ventricular vibration was observed solely during the throat-singing task. In this case, the ventricular folds vibrated with twice the glottal fundamental period, yielding a period-doubling phenomenon (see Panel [c] in Figure 6). The physiological and aerodynamical correlates of this type of ventricular vibration were studied in a previous article, and we refer the reader to Bailly et al. (2010) for all the details.

### ***Impact of Ventricular-Fold Dynamics on Vocal-Fold Vibrations: Theoretical Predictions From In Vivo Ventricular Area Signals***

In this section, we present the aerodynamic impact of the different types of ventricular motions described in the previous sections, on both the transglottal pressure

distribution and the vocal-fold self-oscillations. The aerodynamic impact in the “no-motion” case is also discussed, for a total four cases that were distinguished and simulated: no ventricular motion (Case 1), slow nonoscillatory motion (Case 2), and fast oscillatory motion with aperiodical vibrations (Case 3) or periodical vibrations (Case 4). Each case is represented by a typical ventricular-fold pattern selected from the database. For each case, two simulations of glottal oscillations using the M2M<sub>vof</sub> model are presented and compared:

- The first simulation, referred to as the *reference* configuration, was processed without any ventricular constriction downstream to the glottis.
- The second simulation, referred to as the *ventricular* configuration, included the ventricular dynamics as an input signal derived from in vivo ventricular area measurements, as described in Figure 2.

**Figure 8.** Illustration of ventricular periodical vibration during a crescendo task preceding a growl. Subject S5,  $\tilde{f}_0 \sim 301$  Hz (D4). Panel (a) shows a laryngeal image and kymographic line; Panel (b) displays the corresponding kymographic analysis, with a zoom on a few cycles. Panel (c) shows the spectrographic analysis of the whole sequence; Panel (d) plots glottal fundamental frequency,  $\tilde{f}_0$ , as a function of time. Duration of kymogram (b) is illustrated by the arrow on Panels (c) and (d).

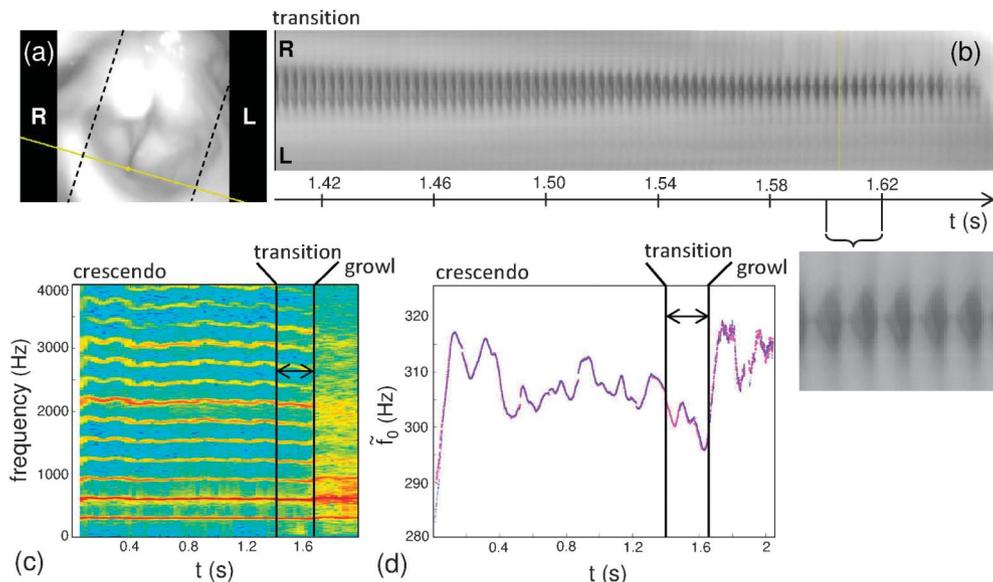


Figure 9 (respectively [resp.] 10) displays the glottal vibrations and pressure within the laryngeal ventricle simulated for Cases 1 and 2 (resp. 3 and 4). In the first row, the normalized glottal aperture ( $h_{\text{vof}}(t)/h_{\text{vof}}^m$ ) predicted in the *reference* configuration is presented (no ventricular constriction). In the second row, the normalized glottal aperture ( $h_{\text{vof}}(t)/h_{\text{vof}}^m$  in red, solid line) and ventricular aperture ( $h_{\text{venf}}(t)/h_{\text{venf}}^m$  in green, dashed line) simulated in the *ventricular* configuration are compared. In the third row, the pressure ratio  $P_2(t)/P_0$  simulated in the *ventricular* configuration is displayed as function of time. Finally, the corresponding EGGn and DEGGn time variations recorded during the selected sequence are reported. Different features can be highlighted regarding each simulated case.

*No ventricular motion (Case 1).* This case is typically represented by a constant and wide ventricular-fold aperture downstream to vocal-fold oscillations. It is illustrated in the left column of Figure 9 during a 6-glottal-cycle sequence of normal voice performed by Subject S4 (see green dashed line in second row).

No pressure recovery is predicted in the laryngeal ventricle ( $P_2(t) = 0$ ) in both *reference* and *ventricular* configurations. The  $M2M_{\text{vof}}$  model oscillates around the same constant equilibrium position ( $h_{\text{vof}}^e(t)/h_{\text{vof}}^m = 0.49 \pm 0.009$ ), with a periodical glottal contact occurring at a quasi-constant fundamental frequency  $f_0 = 120 \pm 3$  Hz, as shown in the first two rows. This prediction is consistent with DEGG and EGG variations displayed in the last two rows (average discrepancy between simulated  $f_0$  and measured  $\tilde{f}_0$  being 1.20%). Therefore, in Case 1, the presence of the ventricular folds downstream to the glottis does not

alter the translottal pressure distribution nor the  $M2M_{\text{vof}}$  vibratory pattern.

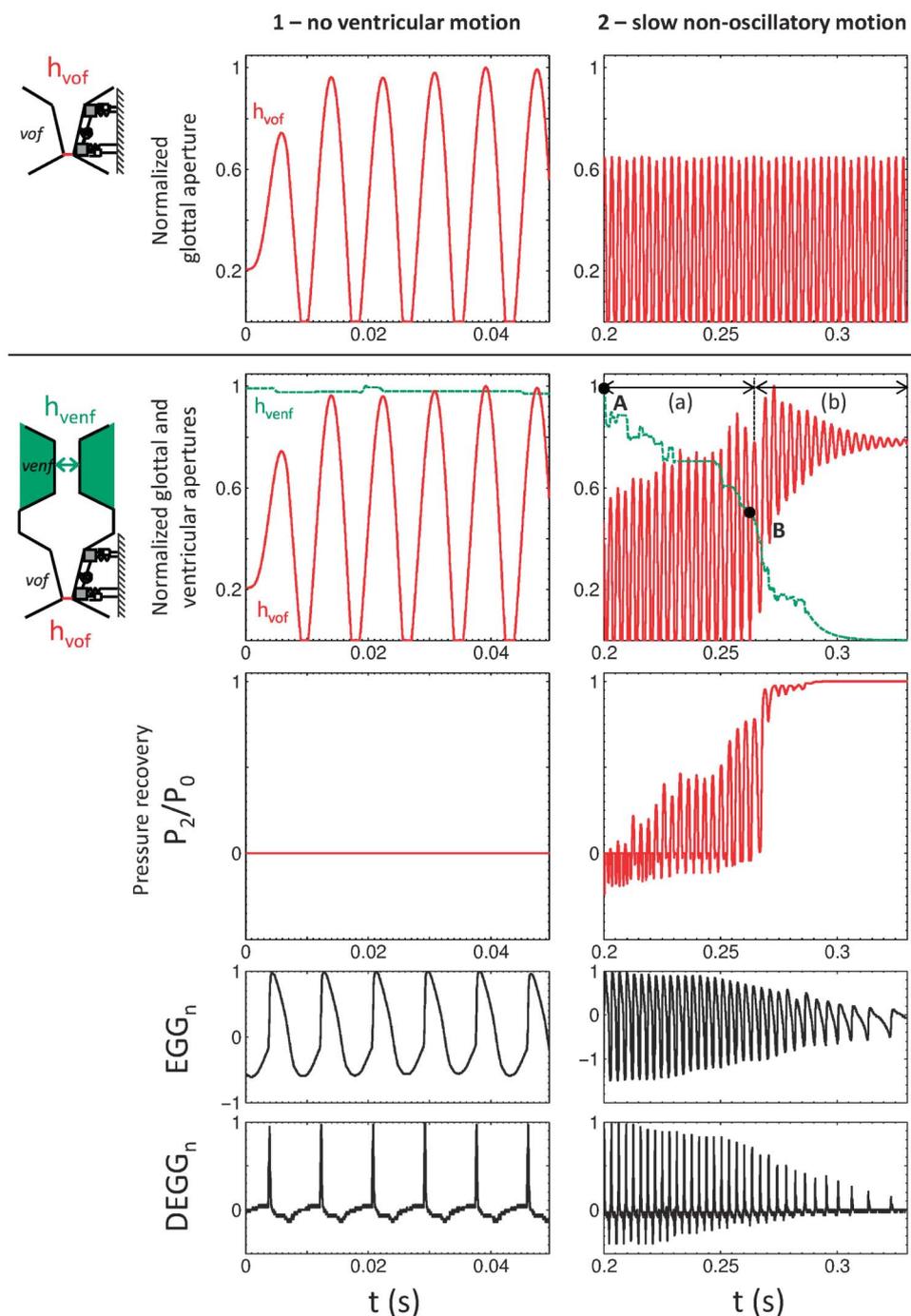
*Slow nonoscillatory motion (Case 2).* This case of ventricular-fold dynamics is typically represented by a slow and progressive adduction of the ventricular-fold constriction from partial to complete closure. It is illustrated in the right column of Figure 9 during a 32-glottal-cycle sequence of shout (Subject S1; see green dashed line in second row, and boxed kymogram S1<sub>08</sub> in Appendix B, Figure B1).

In the *reference* configuration (see first row), the  $M2M_{\text{vof}}$  oscillations are characterized by a constant equilibrium position ( $h_{\text{vof}}^e(t)/h_{\text{vof}}^m = 0.28 \pm 0.003$ ) and a periodical glottal contact occurring at a fundamental frequency  $f_0 = 312 \pm 3$  Hz. These *reference* glottal vibrations would be those predicted when neglecting the aerodynamic impact induced by the ventricular-fold presence and motion.

In the *ventricular* configuration (see second row), the progressive decrease of the ventricular-fold aperture from partial to complete closure generates different alterations of both glottal oscillatory magnitude and contact in time. Two trends can be distinguished, depending on the geometrical ratio  $h_{\text{venf}}^e(t)/h_{\text{vof}}^e(t)$ :

1. For  $h_{\text{venf}}^e(t)/h_{\text{vof}}^e(t)$  ranging from 3.7 to 1 (Point A to Point B in Figure 9), a complete  $M2M_{\text{vof}}$  closure is predicted from one cycle to another (see Phase [a] in the second row). The glottal oscillatory magnitude is enhanced by the ventricular-fold narrowing, as shown by a 68% rise of the vocal-fold equilibrium position with respect to the *reference* configuration. The ventricular-fold motion induces an increase in

**Figure 9.** Theoretical predictions and corresponding in vivo measurements during the no-motion case (Case 1) and one case of ventricular-fold dynamics (Case 2). Top to bottom: Two-mass simulation of the normalized glottal aperture as a function of time in the *reference* configuration (no ventricular constriction); similar simulation in the *ventricular* configuration (dashed line = ventricular input motion; solid line = glottal output motion); predictions of pressure ratio  $P_2(t)/P_0$ ; normalized EGG (EGG<sub>n</sub>) and DEGG (DEGG<sub>n</sub>) signals measured in each case.



the pressure drop at the ventricular folds, reaching up to 76% of the driving tracheal pressure (see Phase [a] in the third row). The aerodynamic impact of the ventricular-fold narrowing on the glottal oscillatory magnitude comes along with a decrease of the glottal

fundamental frequency  $f_0$ , as shown in previous work (Bailey et al., 2008). Indeed, a frequency decrease of 60 Hz is predicted in comparison with the *reference* configuration. Comparatively, DEGG variations show a 20-Hz decrease of the measured fundamental

frequency, as well as a progressive decrease of the speed of vocal fold contact.

2. For  $h_{\text{venf}}^e(t)/h_{\text{vof}}^e(t)$  less than 1 (after Point B in Figure 9), the  $M2M_{\text{vof}}$  oscillates without complete glottal closure (see Phase [b] in the second row). A rise of pressure  $P_2$  is predicted in the ventricle up to cancellation of the transglottal pressure drop (see Phase [b] in the third row). In such a case, the aerodynamic coupling vanishes in the model description, and the  $M2M_{\text{vof}}$  reduces to a damped harmonic oscillator with steady externally applied aerodynamic forces. The model vibrates with a constant frequency below that predicted in the *reference* configuration ( $f_0 = 287$  Hz), and with an amplitude decreasing with time. Finally, the vibratory system exponentially decays to an equilibrium position around three times higher than the simulated *reference* one ( $h_{\text{vof}}^e/h_{\text{vof}}^m = 0.78$ ). Thereby, in this geometrical range, the ventricular-fold narrowing tends to suppress the vocal-fold vibrations, as observed *in vivo*. Note that a glottal contact is still recorded on DEGG variations during the ventricular-fold complete closure, coming along with a continuous frequency drop-off, before achieving oscillations extinction ( $\tilde{f}_0$  varying from 286 to 102 Hz).

*Fast oscillatory motion: Aperiodical vibrations (Case 3).*

This case of ventricular-fold dynamics is typically represented by ventricular-fold aperiodical vibrations, consisting of irregular oscillatory amplitude, frequency, ventricular contact occurrences, and duration. It is illustrated in the left column of Figure 10 during a 21-glottal-cycle sequence of vocal effort (Subject S4; see green dashed line in second row, and boxed Kymogram S4<sub>04</sub> in Appendix B, Figure B2).

In the *reference* configuration (see first row), the  $M2M_{\text{vof}}$  periodically oscillates around a constant equilibrium position such as  $h_{\text{vof}}^e(t)/h_{\text{vof}}^m = 0.20 \pm 0.001$ . It is characterized by a complete glottal closure at a constant fundamental frequency of 350 Hz.

In the *ventricular* configuration, the ratio of ventricular to glottal apertures ( $h_{\text{venf}}^e(t)/h_{\text{vof}}^e(t)$ ) ranges from 0.2 to 4 (as derived from the second row). A modulation of pressure  $P_2$  within the laryngeal ventricle is predicted, depending on the ventricular-fold dynamics (see the third row). A large pressure recovery is simulated in the ventricle. It even reaches the driving pressure  $P_0$  each time the ventricular folds are in contact, thereby cancelling the transglottal pressure drop. In comparison with the *reference* configuration, this pressure recovery implies an alteration of the vocal-fold oscillatory magnitude in time shown in the second row, with a ratio  $h_{\text{vof}}^e(t)/h_{\text{vof}}^m$  varying from 0.18 to 0.46. Therefore, the ventricular-fold motion implies either a decrease or an increase of the glottal oscillatory magnitude depending on the considered shot instant. The glottal cycle fundamental frequency is also altered, fluctuating between 263 and 402 Hz (average value = 318 Hz). Instances of ventricular contact induce a noncomplete glottal closure in that case.

*Fast oscillatory motion: Periodical vibrations (Case 4).*

This case of ventricular-fold dynamics is typically represented

by vocal–ventricular periodical vibrations. It is illustrated in the right column of Figure 10 during a 9-glottal-cycle sequence of throat singing (Subject S1; see green dashed line in second row, and boxed Kymogram S1<sub>03</sub> in Appendix B, Figure B2).

In the *reference* configuration (see the first row), the  $M2M_{\text{vof}}$  periodically oscillates around a constant equilibrium position such as  $h_{\text{vof}}^e(t)/h_{\text{vof}}^m = 0.20 \pm 0.001$ . It is characterized by a complete glottal closure at a constant fundamental frequency of 163 Hz.

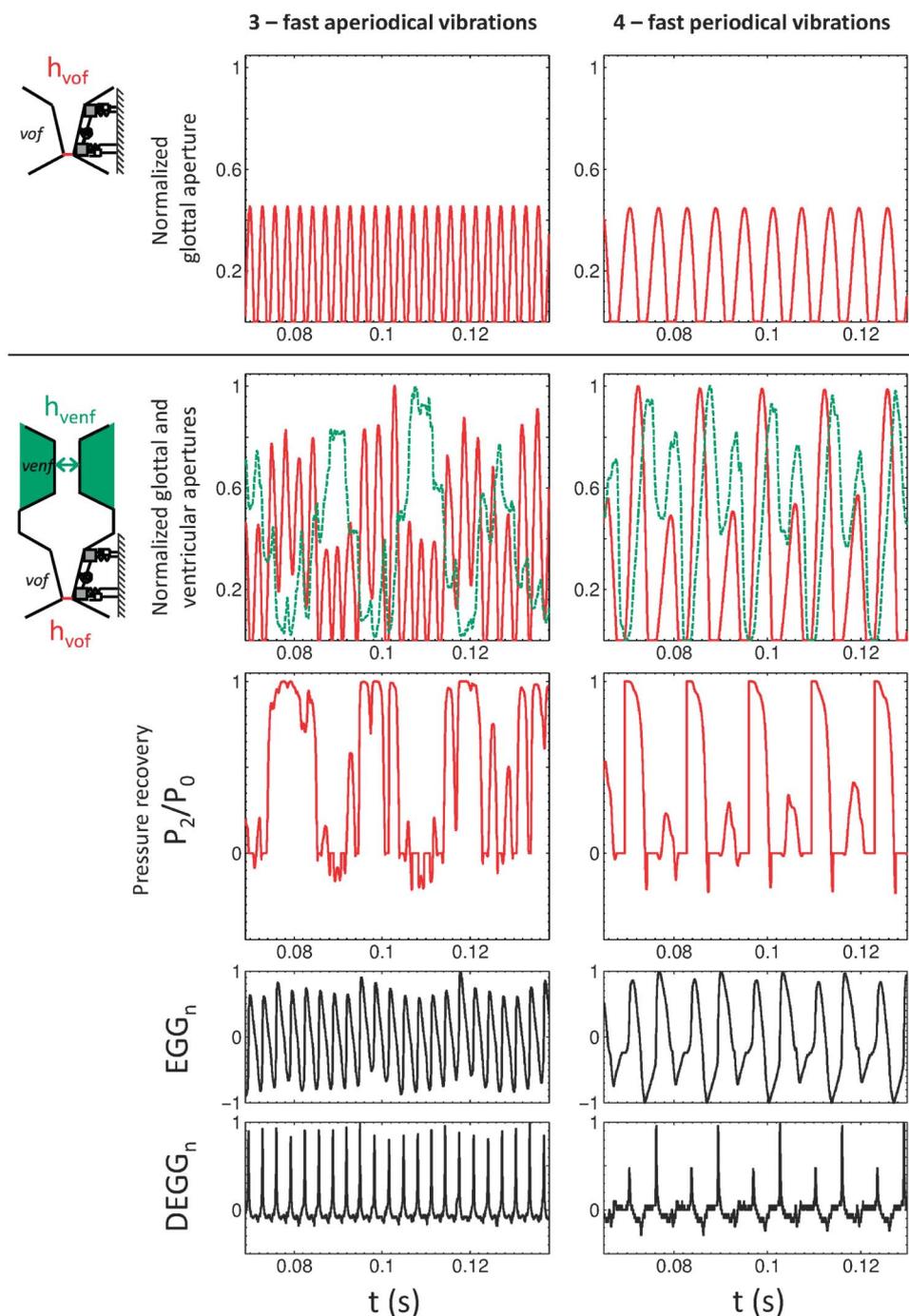
In the *ventricular* configuration, the ratio of ventricular to glottal apertures ( $h_{\text{venf}}^e(t)/h_{\text{vof}}^e(t)$ ) ranges from 0.9 to 2 (as derived from the second row). The large pressure recovery simulated in the ventricle induces an alteration of the vocal-fold oscillatory magnitude in comparison with the *reference* configuration, so that the vocal-fold equilibrium position increases by a factor of two, every two glottal vibrations. More specifically, a period-doubling phenomenon is simulated at the glottis: A second frequency appears in the vocal fold vibratory pattern, which equals to the fundamental frequency of the ventricular fold oscillation (average value = 75 Hz). A duration difference between two consecutive glottal events is also predicted (Bailly et al., 2010).

### ***Aerodynamical Modeling of Self-Sustained Ventricular Oscillations***

This section explores whether the ventricular-fold self-sustained vibration could be simulated using a common aerodynamical modeling approach based on the  $M2 \times 2M$  model of phonation (see Figure 1c). In this case, the ventricular motion is an output signal simulated by an additional two-mass model of ventricular folds ( $M2M_{\text{venf}}$ ) attached downstream to the  $M2M_{\text{vof}}$  model. The  $M2M_{\text{vof}}$  initialization was chosen identical to that identified in the previous section in the case of ventricular-fold oscillatory motion (Case 3). Regarding the  $M2M_{\text{venf}}$  initialization, the mechanical parameters were fixed in line with *in vivo* data of ventricular tissue biomechanics as detailed previously (see Table 5). The initial geometrical laryngeal configuration was parametrically varied through the geometrical aspect ratio  $\xi = h_{\text{venf}}^o/h_{\text{venf}}^m$ . Case  $\xi = 0$  corresponds to a complete initial ventricular adduction, whereas case  $\xi = 1$  refers to the maximal initial ventricular aperture. In the following, two illustrative configurations are discussed ( $\xi = 1$  and  $\xi = 0.21$ ). The normalized glottal aperture, ventricular aperture, and pressure in the ventricle as predicted under both configurations are depicted in Figure 11 in the top, middle, and bottom panels, respectively.

For  $\xi = 1$  (corresponding to  $h_{\text{venf}}^e(t)/h_{\text{vof}}^e(t) = 11 \pm 0.40$ ), the  $M2M_{\text{vof}}$  model predicts a self-sustained quasi-sinusoidal oscillation of the glottal aperture, unaffected by the large downstream ventricular constriction that causes a zero-pressure in the ventricle. The  $M2M_{\text{venf}}$  model does not predict any ventricular aerodynamic-induced motion in that case (see middle panel).

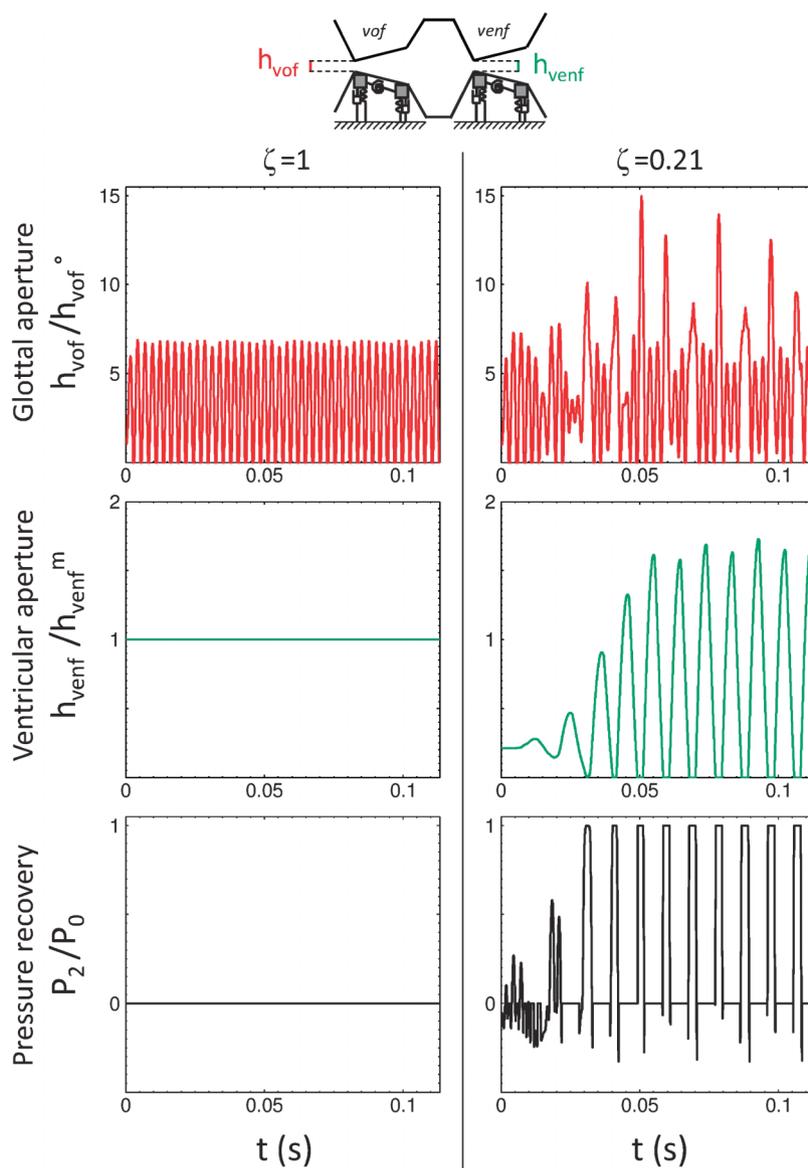
**Figure 10.** Theoretical predictions and corresponding in vivo measurements during Cases 3 and 4 of ventricular-fold dynamics. Top to bottom: Two-mass simulation of the normalized glottal aperture as a function of time in the reference configuration (no ventricular constriction); similar simulation in the ventricular configuration (dashed line = ventricular input motion; solid line = glottal output motion); predictions of pressure ratio  $P_2(t)/P_0$ ; normalized EGG (EGG<sub>n</sub>) and DEGG (DEGG<sub>n</sub>) signals measured in each case.



For  $\xi = 0.21$ , the aerodynamic coupling between  $M2M_{vof}$  and  $M2M_{venf}$  models enables the onset of periodical sustained ventricular vibrations, as shown in the second panel. The geometrical ratio  $h_{venf}^e(t)/h_{vof}^e(t)$  varies from 0.3 to 16 in that case. The vocal-fold oscillatory frequency is

altered, with an average frequency of 53 Hz occurring in the vibratory pattern, whereas the ventricular folds vibrate at around  $104 \pm 3$  Hz. The transitional alteration of the glottal vibration in frequency and amplitude is illustrated. The simulated outcomes support the thesis that

**Figure 11.** Theoretical predictions of the M2 × 2M model of phonation corresponding to two initial ventricular configurations, such as  $\xi = 1$  (left) and  $\xi = 0.21$  (right). Top to bottom: Simulation of the glottal aperture ratio  $h_{\text{vof}}(t)/h_{\text{vof}}^\circ$ ; simulation of the ventricular aperture ratio  $h_{\text{vent}}/h_{\text{vent}}^m$ ; simulation of pressure ratio  $P_2(t)/P_0$  as a function of time.



the oscillation of the ventricular folds can settle in the larynx and be sustained by aerodynamic coupling with the vocal folds, providing an adequate initial ventricular-fold narrowing.

## Discussion and Conclusion

### *Ventricular-Fold Motions and Vocal Output*

The laryngoscopic examinations by high-speed imaging have shown a large variety of ventricular-fold motions under different phonatory tasks: normal voicing, glides, M0 (vocal fry), throat singing, yodel, growls, shout, and

crescendos and descendos. Aside from cases of no ventricular motion, three main types of ventricular-fold dynamics were identified from the database with regard to the temporal variation of ventricular adduction level: (a) *slow nonoscillatory motion*, *fast oscillatory motion* composed of either (b) *aperiodical vibrations* or (c) *periodical vibrations*. In most cases, these different ventricular-fold dynamics were found to go along with a perceived change in voice quality, pitch, and/or vocal intensity.

Addition of roughness in voice quality was commonly found in growls and shouts, together with oscillatory motions of the ventricular folds. The latter was predominantly composed of ventricular-fold aperiodical vibrations,

in agreement with previous studies of vocal growls (e.g., Chevaillier et al., 2008; Fuks et al., 1998; Lindestad et al., 2001; Sakakibara et al., 2004; Zangger Borch et al., 2004). The hoarse timbre of the resulting sound vanished during occurrences of ventricular-fold periodical vibrations at the same frequency as the glottal oscillations, which corroborates the findings by Lindestad et al. (2004).

An interesting finding is that the alternance of laryngeal mechanisms M1 and M2 in yodel induces slow ventricular motions at the laryngeal transitions. A systematic slight ventricular narrowing (resp. widening) was observed when going from M2 to M1 (resp. from M1 to M2) during yodel tasks. This trend was not observed on glides with laryngeal transitions. It does not go along with the observations of Agarwal and colleagues (Agarwal et al., 2003; Agarwal, 2004) on laminagrams from experiments by Hollien and Colton (1969) and Wilson (1976), who observed larger ventricular aperture in modal register (M1) than in falsetto (M2), yet with no statistical significance. The possibility that ventricular folds may play a role in the specificity of pitch jumps in yodel singing calls for further exploration.

A decrease in fundamental frequency was found when simulating the aerodynamical impact of slow and nonoscillatory ventricular narrowing on glottal oscillations, in line with *in vivo* recordings of EGG variations. The drop-off of the acoustical fundamental frequency  $f_0$  with an increasing level of ventricular-fold adduction was demonstrated in an earlier experimental study (Bailly et al., 2008) on mock vocal folds able to sustain flow-induced self-oscillations. Yet, previous measurements by Alipour et al. (2007) and Finnegan and Alipour (2009) on excised canine larynges yielded to an  $f_0$ -rise obtained with an increased level of ventricular-fold adduction. The present *in vivo* measurements strengthen our previous findings on the controversial matter of the impact of ventricular-fold level of adduction on phonation frequency.

Finally, changes in vocal intensity mainly occurred in tasks performed with increasing vocal effort, that is, crescendos and shout. The latter productions were characterized by a systematic ventricular-fold progressive narrowing, followed by aperiodical vibrations in case of shout, which was not reported in past studies. This study showed that the ventricular motion could also be reflected on the spectral-energy repartition in the 3–4-kHz-frequency band. These measurements performed during *in vivo* human voicing provide further demonstration that the ventricular folds can be induced in the phonatory gesture to improve vocal efficiency, which was essentially known from computational predictions (Imagawa et al., 2003; Nomura & Funada, 2007b; Sakakibara et al., 2001; Zhang et al., 2002; Zheng et al., 2009), or *ex vivo* measurements on canine larynges so far (Alipour & Finnegan, 2013; Alipour et al., 2007; Finnegan & Alipour, 2009).

### ***Nonoscillatory and Oscillatory Ventricular Motions: A Different Impact on Glottal Vibrations***

For each pattern of ventricular-fold motion characterized from the *in vivo* study (with the exception of the

no-motion case), the theoretical outcomes showed an impact on the transglottal pressure drop, in agreement with previous studies (Agarwal, 2004; Bailly et al., 2008, 2010; Kucinski et al., 2006; Li et al., 2008; Nomura & Funada, 2007a; Shadle et al., 1991; Triep & Brücker, 2010; Triep et al., 2005; Zhang et al., 2002; Zheng et al., 2009). Due to the time-varying pressure recovery generated in the laryngeal ventricle, an alteration of the vocal-fold oscillatory amplitude, fundamental frequency, contact duration, contact speed, and occurrences was demonstrated. It was found that the vocal-ventricular aerodynamic interactions can alter, enhance, or suppress the vocal-fold vibrations, or leave them unchanged, depending on the ventricular-fold dynamics involved.

The effect induced by a *slow nonoscillatory motion* of the ventricular folds on glottal self-sustained vibrations is comparable to the one induced by the presence of a *static* ventricular-fold constriction in the larynx, which is the major case discussed in literature so far (Alipour et al., 2007; Bailly et al., 2008; Zheng et al., 2009). In fact, the ratio of the ventricular-fold equilibrium aperture over the vocal-fold equilibrium aperture is confirmed to be a critical parameter. Our findings showed that the ventricular-fold adductory motion induces a gradual enhancement of the glottal oscillatory magnitude while this geometrical ratio remains higher than 1. In the opposite, the ventricular-fold motion comes along with a progressive decrease of the glottal oscillatory magnitude for a ratio lower than 1. In particular, the complete *quasi-static* closure of the ventricular-fold aperture yields to the cancellation of glottal vibration. These results are in line with prior results of Bailly et al. (2008) and Zheng et al. (2009), confirming that a specific geometrical positioning of the ventricular folds can either facilitate or impede the phonation.

By contrast, however, the effect induced by a *fast oscillatory motion* of the ventricular folds on glottal self-sustained vibrations cannot be further described by this single spatial parameter. In particular, a short complete closure of the ventricular-fold aperture does not yield to the decrease of glottal oscillatory magnitude and to phonation stop, but to the enhancement of the glottal oscillatory magnitude. Depending on its variation in time, a similar level of ventricular-fold adduction will impact glottal contact much differently. Therefore, this study shows that the classical criterion used in (quasi)-static ventricular-fold conditions to predict the alterations induced on glottal vibrations is no more valid in the case of fast time-varying ventricular-fold conditions, as the ones found in shout, for instance.

### ***A Second Laryngeal Vibrator***

An important finding of this study relates to the ventricular-fold vibrations frequently found in our database. Indeed, ventricular vibration was unexpectedly observed in several phonatory gestures (growl, shout, vocal fry, throat singing), even though their microstructural properties lead some authors to rate them as poor biomechanical

oscillators (Agarwal, 2004). Moreover, simulations based on the  $M2 \times 2M$  model showed that ventricular oscillations can settle in the larynx and be sustained by aerodynamic coupling with the vocal folds, providing a sufficient initial ventricular narrowing.

These results are in agreement with the medical observation that the ventricular folds can become an efficient "alternative oscillator" after cordectomy with reconstruction, a surgery in which the ventricular folds are moved to the glottal plane (Bertino, Bellomo, Ferrero, & Ferlito, 2001; Fukuda, Tsuji, Kawasaki, Kawaida, & Sakou, 1990; Guida & Zorzetto, 2007). Indeed, Fukuda et al. (1990) found better vibratory patterns in patients for whom the vocal fold was reconstructed with the ventricular fold than in those who underwent conventional laser treatment. Of 11 patients who underwent surgical procedure, five presented normal voice, five had hoarse voice, and only one patient had rough and hoarse voice.

From this point of view, the precise knowledge of the ventricular-fold dynamics and the understanding of their impact on phonation and vocal output are of obvious interest and should be further explored. The vocal-ventricular aerodynamic interactions will be tested in vitro in a future study, by combining a deformable ventricular-fold replica to a self-oscillating vocal-fold one (Bailly, 2009; Ruty et al., 2007). The perceptual differences implied on the produced sound will be studied by coupling a model of acoustical wave-propagation in the vocal tract to the laryngeal-source models described here.

## Acknowledgments

We gratefully acknowledge the French Ministry of Research and Education, which supported this research through a PhD grant, and the European COST Action 2103 Advanced Voice Function Assessment for travel funds. We would like to thank all the subjects for their valuable help during the experimental procedure. We would also like to thank Joe Wolfe for his suggestions on this work.

## References

- Agarwal, M. (2004). *The false vocal folds and their effect on trans-laryngeal airflow resistance*. Unpublished doctoral dissertation, Bowling Green State University, Bowling Green, Ohio.
- Agarwal, M., Scherer, R. C., & Hollien, H. (2003). The false vocal folds: Shape and size in frontal view during phonation based on laminagraphic tracings. *Journal of Voice*, *17*, 97–113.
- Alipour, F., & Finnegan, E. (2013). On the acoustic effects of the supraglottic structures in excised larynges. *The Journal of the Acoustical Society of America*, *133*, 2984–2992.
- Alipour, F., Jaiswal, S., & Finnegan, E. (2007). Aerodynamic and acoustic effects of false vocal folds and epiglottis in excised larynx models. *Annals of Otolaryngology, Rhinology & Laryngology*, *116*, 135–144.
- Alipour, F., & Scherer, R. (2012). Ventricular pressures in phonating excised larynges. *The Journal of the Acoustical Society of America*, *132*, 1017–1026.
- Bailly, L. (2009). *Interaction entre cordes vocales et bandes ventriculaires en phonation: Exploration in-vivo, modélisation physique, validation in-vitro* [Interaction between vocal folds and ventricular folds in phonation: In-vivo exploration, physical modeling] (Doctoral dissertation). Retrieved from TEL Multidisciplinary Theses Servers (Accession No. tel-00416587).
- Bailly, L., Henrich, N., & Pelorson, X. (2010). Vocal fold and ventricular fold vibration in period-doubling phonation: Physiological description and aerodynamic modeling. *The Journal of the Acoustical Society of America*, *127*, 3212–3222.
- Bailly, L., Pelorson, X., Henrich, N., & Ruty, N. (2008). Influence of a constriction in the near field of the vocal folds: Physical modeling and experimental validation. *The Journal of the Acoustical Society of America*, *124*, 3296–3308.
- Bak-Pedersen, K., & Nielsen, K. O. (1986). Subepithelial mucous glands in the adult human larynx. *Acta Oto-laryngologica*, *102*, 341–352.
- Behrman, A., Dahl, L. D., Abramson, A. L., & Schutte, H. K. (2003). Anterior-posterior and medial compression of the supraglottis: Signs of nonorganic dysphonia or normal postures? *Journal of Voice*, *17*, 403–410.
- Bertino, G., Bellomo, A., Ferrero, F. E., & Ferlito, A. (2001). Acoustic analysis of voice quality with or without false vocal fold displacement after cordectomy. *Journal of Voice*, *15*, 131–140.
- Brücker, C., Triep, M., & Kob, M. (2004). Study of the vortex dynamics in a mechanical model of the vocal folds using particle-image velocimetry. In *Proceedings of the International Conference on Voice Physiology and Biomechanics* (pp. 11–17). Marseille, France.
- Chan, R. W., Fu, M., & Tirunagari, N. (2006). Elasticity of the human false vocal fold. *Annals of Otolaryngology, Rhinology & Laryngology*, *115*, 370–381.
- Chevallier, G., Guilbault, R., Renard, N., & Tran Ba Huy, P. (2008). La voix saturée du chanteur de rock metal, un mécanisme performant [The "saturated" voice of rock singers, a high-performance mechanism]. In *Actes du 64<sup>ième</sup> Congrès de la Société Française de Phoniatrie et des Pathologies de la Communication*. Paris, France.
- Crevier-Buchman, L., Vincent, C., & Hans, S. (2008). Comportements laryngés en voix chuchotée, étude en caméra ultra-rapide [Laryngeal behavior in whispered voice: A study using high speed imaging]. In *Actes du 64<sup>ième</sup> Congrès de la Société Française de Phoniatrie et des Pathologies de la Communication*. Paris, France.
- de Cheveigné, A., & Kawahara, H. (2002). Yin: A fundamental frequency estimator for speech and music. *The Journal of the Acoustical Society of America*, *111*, 1917–1930.
- Drechsel, J. S., & Thompson, S. L. (2008). Influence of supraglottal structures on the glottal jet exiting a two-layer synthetic, self-oscillating vocal fold model. *The Journal of the Acoustical Society of America*, *123*, 4434–4445.
- Esling, J. H., Zeroual, C., & Crevier-Buchman, L. (2007). A study of muscular synergies at the glottal, ventricular and aryepiglottic levels. In J. Trouvain & W. J. Barry (Eds.), *Proceedings of the 16<sup>th</sup> International Congress of Phonetic Sciences* (pp. 585–588). Saarbrücken, Germany: Universität des Saarlandes.
- Finnegan, E. M., & Alipour, F. (2009). Phonatory effects of supraglottic structures in excised canine larynges. *Journal of Voice*, *23*, 51–61.
- Fuks, L., Hammarberg, B., & Sundberg, J. (1998). A self-sustained vocal-ventricular phonation mode: Acoustical, aerodynamic and glottographic evidences. *Kungliga Tekniska Högskolan TMH—Quarterly Progress and Status Report*, *3*, 49–59.
- Fukuda, H., Tsuji, D. H., Kawasaki, Y., Kawaida, M., & Sakou, T. (1990). Displacement of the ventricular fold following cordectomy. *Auris Nasus Larynx*, *17*, 221–228.

- Gauffin, J., Binh, N., Ananthapadmanabha, T. V., & Fant, G.** (1983). Glottal geometry and volume velocity waveform. In D. Bless & J. Abbs (Eds.), *Vocal fold physiology: Contemporary research and clinical issues* (pp. 194–201). San Diego, CA: College-Hill.
- Granqvist, S., & Lindestad, P. A.** (2001). A method of applying Fourier analysis to high-speed laryngoscopy. *The Journal of the Acoustical Society of America*, *110*, 3193–3197.
- Guida, H. L., & Zorzetto, N. L.** (2007). Morphological and histochemical analysis of the human vestibular fold. *International Journal of Morphology*, *25*, 537–543.
- Haji, T., Mori, K., Omori, K., & Isshiki, N.** (1992). Mechanical properties of the vocal fold. *Acta Oto-laryngologica*, *112*, 559–565.
- Henrich, N.** (2006). Mirroring the voice from Garcia to the present day: Some insights into singing voice registers. *Logopedics Phoniatrics Vocology*, *31*, 3–14.
- Henrich, N., d'Alessandro, C., Castellengo, M., & Doval, B.** (2004). On the use of the derivative of electroglottographic signals for characterization of nonpathological phonation. *The Journal of the Acoustical Society of America*, *115*, 1321–1332.
- Hirano, M., Kurita, S., & Nakashima, T.** (1983). Growth, development and aging of human vocal folds, In D. M. Bless & J. M. Abbs (Eds.), *Vocal fold physiology: Contemporary research and clinical issues* (pp. 22–43). San Diego, CA: College-Hill.
- Hollien, H., & Colton, R. H.** (1969). Four laminagraphic studies of vocal fold thickness. *Folia Phoniatrica*, *21*, 179–198.
- Holmberg, E. B., Hillman, R. E., & Perkell, J. S.** (1988). Glottal airflow and transglottal air pressure measurements for male and female speakers in soft, normal, and loud voice. *The Journal of the Acoustical Society of America*, *84*, 511–529.
- Iijima, H., Miki, N., & Nagai, N.** (1992). Glottal impedance based on finite element analysis of two-dimensional unsteady viscous flow in a static glottis. *IEEE Transactions on Signal Processing*, *40*, 2125–2135.
- Ikeda, T., Matsuzaki, Y., & Aomatsu, T.** (2001). A numerical analysis of phonation using a two-dimensional flexible channel model of the vocal folds. *Journal of Biomechanical Engineering*, *123*, 571–579.
- Imagawa, H., Sakakibara, K. I., Tayama, N., & Niimi, S.** (2003). The effect of the hypopharyngeal and supraglottic shapes on the singing voice. In *Proceedings of the Stockholm Music Acoustics Conference* (pp. 51–54). Stockholm, Sweden.
- Ishizaka, K., & Flanagan, J. L.** (1972). Synthesis of voiced sounds from a two-mass model of vocal folds. *The Bell System Technical Journal*, *51*, 1233–1267.
- Kawasaki, A., Fukuda, H., Shiotani, A., & Kanzaki, J.** (2001). Study of movements of individual structures of the larynx during swallowing. *Auris Nasus Larynx*, *28*, 75–84.
- Kimura, K., Sakakibara, K. I., Imagawa, H., Chan, R., Niimi, S., & Tayama, N.** (2002). Histological investigation of the supraglottal structures in humans for understanding abnormal phonation. *The Journal of the Acoustical Society of America*, *112*, 2446.
- Kitajima, K., & Fujita, F.** (1990). Estimation of subglottal pressure with intraoral pressure. *Acta Oto-laryngologica*, *109*, 473–478.
- Kotby, M. N., Kirchner, J. A., Kahane, J. C., Basiouny, S. E., & El-Samaa, M.** (1991). Histo-anatomical structure of the human laryngeal ventricle. *Acta Oto-laryngologica*, *111*, 396–402.
- Kucinschi, B. R., Scherer, R. C., DeWitt, K. J., & Ng, T. T. M.** (2006). Flow visualization and acoustic consequences of the air moving through a static model of the human larynx. *Journal of Biomechanical Engineering*, *128*, 380–390.
- Kutta, H., Steven, P., Kohla, G., Tillmann, B., & Paulsen, F. P.** (2002). The human false vocal folds: An analysis of anti-microbial defense mechanisms. *Anatomy and Embryology*, *205*, 315–323.
- Kutta, H., Steven, P., Varoga, D., & Paulsen, F. P.** (2004). Tff peptides in the human false vocal folds of the larynx. *Peptides*, *25*, 811–818.
- Li, S., Wan, M., & Wang, S.** (2008). The effects of the false vocal fold gaps on intralaryngeal pressure distributions and their effects on phonation. *Science in China Series C: Life Sciences*, *51*, 1045–1051.
- Lindestad, P. A., Blixt, V., Pahlberg-Olson, J., & Hammarberg, B.** (2004). Ventricular fold vibration in voice production: A high-speed imaging study with kymographic, acoustic and perceptual analyses of a voice patient and a vocally healthy subject. *Logopedics Phoniatrics Vocology*, *29*, 162–170.
- Lindestad, P. A., Hammarberg, B., Larsson, H., & Granqvist, S.** (1998). Ventricular fold co-vibration in chronic laryngitis studied with high-speed imaging and acoustic analysis. In *Proceedings of the 24th Congress of the International Association of Logopedics and Phoniatrics* (pp. 205–208). Amsterdam, the Netherlands.
- Lindestad, P. A., Sodersten, M., Merker, B., & Granqvist, S.** (2001). Voice source characteristics in Mongolian “throat singing” studied with high-speed imaging technique, acoustic spectra, and inverse filtering. *Journal of Voice*, *15*, 78–85.
- Loscos, A., & Bonada, J.** (2004). Emulating rough and growl voice in spectral domain. In *Proceedings of the 7th International Conference on Digital Audio Effects (DAFX-04)*. Naples, Italy.
- Lous, N. J. C., Hofmans, G. C. J., Veldhuis, R. N. J., & Hirschberg, A.** (1998). A symmetrical two-mass vocal-fold model coupled to vocal tract and trachea, with application to prosthesis design. *Acta Acustica*, *84*, 1135–1150.
- McGowan, R. S., & Howe, M. S.** (2010). Influence of the ventricular folds on a voice source with specified vocal fold motion. *The Journal of the Acoustical Society of America*, *127*, 1519–1527.
- Miller, J. A., Pereira, J. C., & Thomas, D. W.** (1988). Fluid flow through the larynx channel. *Journal of Sound and Vibration*, *121*, 277–290.
- Moisik, S. R.** (2013). *The epilarynx in speech*. Unpublished doctoral dissertation, University of Victoria, British Columbia, Canada.
- Nasri, S., Jasleen, J., Gerratt, B. R., Sercarz, J. A., Wenokur, R., & Berke, G. S.** (1996). Ventricular dysphonia: A case of the false vocal fold mucosal traveling wave. *American Journal of Otolaryngology*, *17*, 427–431.
- Nomura, H., & Funada, T.** (2007a). Effects of the false vocal folds on sound generation by an unsteady glottal jet through rigid wall model of the larynx. *Acoustical Science & Technology*, *28*, 403–412.
- Nomura, H., & Funada, T.** (2007b). Sound generation by unsteady flow ejecting from the vibrating glottis based on a distributed parameter model of the vocal cords. *Acoustical Science & Technology*, *28*, 392–402.
- Oliveira Rosa, M., & Pereira, J. C.** (2008). Aerodynamic study of three-dimensional larynx models using finite element methods. *Journal of Sound and Vibration*, *311*, 39–55.
- Oliveira Rosa, M., Pereira, J. C., Grellet, M., & Alwan, A.** (2003). A contribution to simulating a three-dimensional larynx model using the finite element method. *The Journal of the Acoustical Society of America*, *114*, 2893–2905.

- Pelorsen, X., Hirschberg, A., Wijnands, A. P. J., & Baillet, H.** (1995). Description of the flow through in-vitro models of the glottis during phonation. *Acta Acustica*, 3, 191–202.
- Pepinsky, A.** (1942). The laryngeal ventricle considered as an acoustical filter. *The Journal of the Acoustical Society of America*, 13, 32–35.
- Pinho, S. M. R., Pontes, P. A. L., Gadelha, M. E. C., & Biasi, N.** (1999). Vestibular vocal fold behaviour during phonation in unilateral vocal fold paralysis. *Journal of Voice*, 13, 36–42.
- Reidenbach, M. M.** (1998). The muscular tissue of the vestibular folds of the larynx. *European Archives of Oto-Rhino-Laryngology*, 225, 365–367.
- Rothenberg, M.** (1992). A multichannel electroglottograph. *Journal of Voice*, 6, 36–43.
- Roubeau, B., Henrich, N., & Castellengo, M.** (2009). Laryngeal vibratory mechanisms: The notion of vocal register revisited. *Journal of Voice*, 23, 425–438.
- Rubin, A. D., Praneetvatakul, V., Gherson, S., Moyer, C. A., & Sataloff, R. T.** (2006). Laryngeal hyperfunction during whispering: Reality or myth? *Journal of Voice*, 20, 121–127.
- Ruty, N., Pelorsen, X., Van Hirtum, A., Lopez Arteaga, I., & Hirschberg, A.** (2007). An in-vitro setup to test the relevance and the accuracy of low-order vocal folds models. *The Journal of the Acoustical Society of America*, 121, 479–490.
- Sakakibara, K. I., Fuks, L., Imagawa, H., & Tayama, N.** (2004). Growl voice in ethnic and pop styles. In *Proceedings of the International Symposium on Musical Acoustics*. Nara, Japan.
- Sakakibara, K. I., Imagawa, H., Konishi, T., Kondo, K., Murano, E. Z., Kumada, M., . . . Niimi, S.** (2001). Vocal fold and false vocal fold vibrations in throat singing and synthesis of khömei. In *Proceedings of the International Computer Music Conference* (pp. 135–138). Havana, Cuba.
- Sakakibara, K. I., Imagawa, H., Niimi, S., & Osaka, N.** (2002). Synthesis of the laryngeal source of throat singing using a 2 × 2-mass model. In *Proceedings of the International Computer Music Conference* (pp. 5–8), Gothenburg, Sweden.
- Sakakibara, K. I., Imagawa, H., Niimi, S., & Tayama, N.** (2004). Physiological study of the supraglottal structure. In *Proceedings of the International Conference on Voice Physiology and Biomechanics*. Marseille, France.
- Serrurier, A., & Badin, P.** (2008). A three-dimensional articulatory model of nasals based on MRI and CT data. *The Journal of the Acoustical Society of America*, 123, 2335–2355.
- Shadle, C., Barney, A. M., & Thomas, D. W.** (1991). An investigation into the acoustics and aerodynamics of the larynx. In J. Gauffin & B. Hammarberg (Eds.), *Vocal fold physiology: Acoustic, perceptual and physiological aspects of voice mechanisms* (pp. 73–82). San Diego, CA: Singular Publishing Group.
- Stager, S.** (2011). The role of the supraglottic area in voice production. *Otolaryngology*, 51. doi:10.4172/2161-119X.S1-001
- Stager, S. V., Bielamowicz, S., Gupta, A., Marullo, S., Regnell, J. R., & Barkmeier, J. M.** (2001). Quantification of static and dynamic supraglottic activity. *Journal of Speech, Language, and Hearing Research*, 44, 1245–1256.
- Stager, S. V., Bielamowicz, S., Regnell, J. R., Gupta, A., & Barkmeier, J. M.** (2000). Supraglottic activity: Evidence of vocal hyperfunction or laryngeal articulation? *Journal of Speech, Language, and Hearing Research*, 43, 229–238.
- Stager, S. V., Neubert, R., Miller, S., Regnell, J. R., & Bielamowicz, S. A.** (2003). Incidence of supraglottic activity in males and females: A preliminary report. *Journal of Voice*, 17, 395–402.
- Svec, J., Horacek, J., Sram, F., & Vesely, J.** (2000). Resonance properties of the vocal folds: In vivo laryngoscopic investigation of the externally excited laryngeal vibrations. *The Journal of the Acoustical Society of America*, 108, 1397–1407.
- Triep, M., & Brücker, C.** (2010). Three-dimensional nature of the glottal jet. *The Journal of the Acoustical Society of America*, 127, 1537–1547.
- Triep, M., Brücker, C., & Schröder, W.** (2005). High-speed PIV measurements of the flow downstream of a dynamic mechanical model of the human vocal folds. *Experiments in Fluids*, 39, 232–245.
- Tsai, V., Celmer, A., Berke, G. S., & Chhetri, D. K.** (2007). Videostroboscopic findings in unilateral superior laryngeal nerve paralysis and paresis. *Otolaryngology—Head and Neck Surgery*, 136, 660–662.
- Vilain, C. E.** (2002). *Contribution à la synthèse de parole par modèle physique. Application à l'étude des voix pathologiques* [Contribution to the physical modeling of speech with application to voice pathologies] (Doctoral dissertation). Retrieved from GIPSA-lab ([www.gipsa-lab.grenoble-inp.fr/~coriandre.vilain/mes\\_documents/These/these\\_cv.pdf](http://www.gipsa-lab.grenoble-inp.fr/~coriandre.vilain/mes_documents/These/these_cv.pdf)).
- Von Doersten, P. G., Izdebski, K., Ross, J. C., & Cruz, R. M.** (1992). Ventricular dysphonia: A profile of 40 cases. *Laryngoscope*, 102, 1296–1301.
- Wilson, J. E.** (1976). Variations of the laryngo-pharynx in singing. *The National Association of Teachers in Singing Bulletin*, 31, 20–22.
- Zanger Borch, D., Sundberg, J., Lindestad, P. A., & Thalen, M.** (2004). Vocal fold vibration and voice source aperiodicity in “dist” tones: A study of a timbral ornament in rock singing. *Logopedics Phoniatrics Vocology*, 29, 147–153.
- Zeroual, C., Esling, J., & Crevier-Buchman, L.** (2005). Physiological study of whispered speech in Moroccan Arabic. In *Proceedings of Interspeech* (pp. 1069–1072). Lisbon, Portugal.
- Zhang, C., Zhao, W., Frankel, S. H., & Mongeau, L.** (2002). Computational aeroacoustics of phonation. Part II : Effects of flow parameters and ventricular folds. *The Journal of the Acoustical Society of America*, 112, 2147–2154.
- Zheng, X., Bielamowicz, S., Luo, H., & Mittal, R.** (2009). A computational study of the effect of false vocal folds on glottal flow and vocal fold vibration during phonation. *Annals of Biomedical Engineering*, 37, 625–642.
- Zheng, X., Mittal, R., Xue, Q., & Bielamowicz, S.** (2011). Direct-numerical simulation of the glottal jet and vocal-fold dynamics in a three-dimensional laryngeal model. *The Journal of the Acoustical Society of America*, 130, 404–415.

## Appendix A

## Nomenclature

Voice tasks<sup>a</sup>

M0	laryngeal mechanism used in vocal fry
M1	laryngeal mechanism used in modal phonation
M2	laryngeal mechanism used in falsetto
M3	laryngeal mechanism used in whistle register

## Models

M2M <sub>vof</sub>	two-mass model of vocal folds
M2M <sub>venf</sub>	two-mass model of ventricular folds
M2 × 2M	model of vocal–ventricular vibrations combining M2M <sub>vof</sub> and M2M <sub>venf</sub>

## Modeling variables

t	time
x	axial direction, from the lower part to the upper part of the vocal folds
z	lateral direction, from the anterior glottal commissure to the posterior one
y	transversal direction, perpendicular to plane (x, z)
A <sub>vof</sub>	glottal space area in plane (y, z) <sup>b</sup>
d <sub>vof</sub>	glottal length along x-direction
h <sub>vof</sub>	vocal-folds aperture along y-direction
h <sub>vof1</sub>	aperture of the M2M <sub>vof</sub> at the first mass location
h <sub>vof2</sub>	aperture of the M2M <sub>vof</sub> at the second mass location
m <sub>vof</sub>	mass parameter of the M2M <sub>vof</sub> model
k <sub>vof</sub>	spring stiffness parameter of the M2M <sub>vof</sub> model
k <sub>cvof</sub>	coupling stiffness parameter of the M2M <sub>vof</sub> model, equal to 0.5 k <sub>vof</sub>
r <sub>vof</sub>	damping parameter of the M2M <sub>vof</sub> model
f <sub>0</sub>	fundamental frequency of glottal oscillations
P <sub>2</sub>	pressure in the laryngeal ventricle
P <sub>0</sub>	subglottal pressure
A <sub>venf</sub>	ventricular space area in plane (y, z) <sup>b</sup>
d <sub>venf</sub>	ventricular-folds length along x-direction
h <sub>venf</sub>	ventricular-folds aperture along y-direction
h <sub>venf1</sub>	aperture of the M2M <sub>venf</sub> at the first mass location
h <sub>venf2</sub>	aperture of the M2M <sub>venf</sub> at the second mass location
m <sub>venf</sub>	mass parameter of the M2M <sub>venf</sub> model
k <sub>venf</sub>	spring stiffness parameter of the M2M <sub>venf</sub> model
k <sub>cvenf</sub>	coupling stiffness parameter of the M2M <sub>venf</sub> model, equal to 0.5 k <sub>venf</sub>
r <sub>venf</sub>	damping parameter of the M2M <sub>venf</sub> model
f <sub>r</sub>	resonance peak frequency of the ventricular folds
Q <sub>r</sub>	quality factor of the ventricular folds mechanical resonance
γ	ratio of the ventricular-folds aperture over the glottal aperture
ξ	geometrical aspect ratio h <sub>venf</sub> <sup>o</sup> /h <sub>venf</sub> <sup>m</sup> , varying between 0 (full ventricular adduction) and 1 (maximal ventricular aperture)
h <sub>ventricle</sub>	ventricle height along y-direction
d <sub>ventricle</sub>	ventricle length along x-direction
h <sub>0</sub>	trachea height along y-direction
ρ	air density
μ	air dynamic viscosity
c <sub>o</sub>	speed of sound

## Modifiers of a variable X(t)

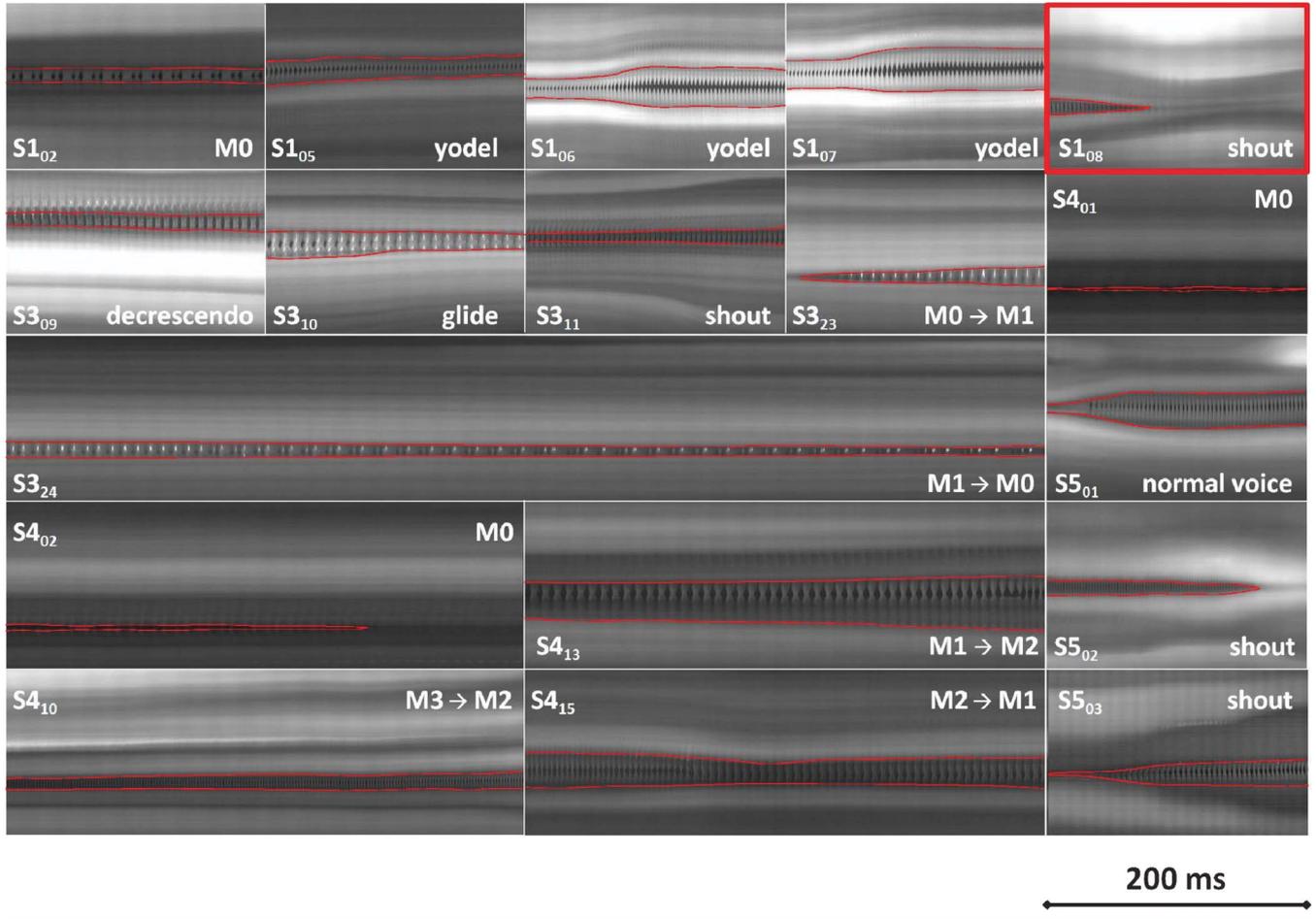
$\tilde{X}$	corresponding real quantity measured from in vivo data <sup>b</sup>
$X^m = \max_t(X)$	maximum value of X over time duration t
mean <sub>t</sub> (X)	mean value of X over time duration t
min <sub>t</sub> (X)	minimum value of X over time duration t
X <sub>n</sub>	normalized quantity defined by ratio X/max <sub>t</sub> (X)
< X >	average value of X along z-direction <sup>2</sup>
X <sup>o</sup>	initial value of X, at instant t = 0 of the simulation
X <sup>e</sup>	cycle-to-cycle equilibrium position of X

<sup>a</sup>See Henrich (2006) and Roubeau et al. (2009) for more details about vocal registers and laryngeal mechanisms. <sup>b</sup>As introduced in Bailly et al. (2010, Section 2a).

Appendix B

Kymographic Illustrations

Figure B1. Kymographic illustration of typical *ventricular slow nonoscillatory motions* identified in the database.



**Figure B2.** Kymographic illustration for all sequences demonstrating a *ventricular fast oscillatory motion*. Red rectangles are the cases chosen as representative of each type of ventricular-fold motion in the numerical study. For each kymogram, the corresponding laryngeal images showing the selected kymographic line, sound, and video sequence are available in the online supplemental materials (<http://hal.archives-ouvertes.fr/hal-00998464>).

