euronoise

Vocal folds and ventricular bands in interaction: comparison between 'in vivo' measurements and theoretical predictions

Lucie Bailly[a], Nathalie Henrich[a], Xavier Pelorson[a] and Joël Gilbert[b]

[a]Département Parole & Cognition, GIPSA-lab, 46, avenue Félix Viallet, 38031 Grenoble Cedex, France

[b]Laboratoire d'Acoustique de l'Université du Maine, Avenue Olivier Messiaen, 72085 Le Mans, France

lucie.bailly@gipsa-lab.inpg.fr

Period-doubling occurrences have been found during singing phonations such as Mongolian *Kargyraa* throat singing or Sardinian *A Tenor*e Bassu singing. The combined vibrations of vocal folds and ventricular bands have been observed during the production of such low-pitch bass-type sounds. The present study aims at better understanding the physical interaction between the ventricular-bands vibration and the vocal-folds self-sustained oscillations.
In this paper, the vibratory properties of both vocal folds and ventricular bands in interaction are analysed on a professional singer by means of acoustical, electroglottographic signals and synchronized glottal images obtained by high-speed cinematography.
Using the detected glottal and ventricular areas, the aerodynamic behaviour of the laryngeal system is simulated using a simplified aerodynamic modelling previously validated *in-vitro* on vocal-folds and ventricular-bands replica. An estimate of the subglottal pressure along with the ventricular aperture extracted from the *in-vivo* data allows a theoretical prediction of the glottal aperture. The *in-vivo* measurements of the glottal aperture are then compared to the simulated estimations. The influence of the subglottal pressure is also discussed.

1 Introduction

The vocal folds are considered to be the main vibratory source of voiced sounds in human speech and singing. The ventricular bands, also called "false vocal folds", are two mucous structures located just above the vocal folds, located superior to the laryngeal ventricle (see Fig.1).

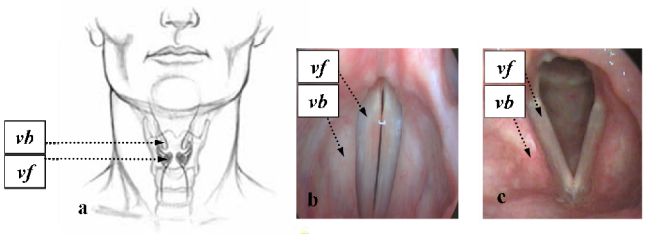

Fig.1 **a**. Frontal view of the human larynx, **b**. Picture of the larynx during phonation, **c**. during breathing (*in-vivo* laryngoscopic recordings).
(***vb***: ventricular bands, ***vf***: vocal folds).

The ventricular bands are not commonly involved as a vibrating structure during normal phonation. Yet, although their physical properties (high viscosity and low stiffness) are different from those of biomechanical oscillators such as the vocal folds [1], their vibration has been observed during certain throat-singing productions found in Asian vocal cultures [2, 3, 4, 5] and Mediterranean traditional polyphonies [6]. In this latter case, the vibratory motion is associated with a period-doubling phenomenon (octave jump below the original tone), symptomatic in voice pathology but used as a phonatory performance in singing [2, 4, 5, 6]. The origin of the ventricular-bands vibration is still poorly understood. *In-vivo* laryngeal examinations support the idea of a strong physical interaction between the ventricular bands and the vocal folds [2, 3, 4]. Yet the physical nature of this interaction is unknown.

The present study intends to characterize this ventricular motion and its correlation with the glottal oscillations, focusing on a possible aerodynamic effect underlying this interaction. What are the physiological correlates associated to this singing phonation? What is the aerodynamic impact of the ventricular vibration on the vocal-folds self-oscillations?

The aerodynamic aspects of the vocal-ventricular interaction have barely been investigated. [2] measured higher oesophageal pressures when phonation switches from modal phonation to the vocal-ventricular mode. Previous experimental investigations dealing with *in-vitro* set-ups have provided an initial insight into the influence of the ventricular bands on the glottal airflow [7, 8, 9, 10]. Using a rigid non-oscillating replica combining vocal folds and ventricular bands, [9] observed an increase or a decrease of the translaryngeal airflow resistance depending upon the laryngeal geometry. The glottal jet curvature due to the Coanda effect, typically observed through *in-vitr*o static replicas [8, 10], decreases with the presence of the ventricular bands [7, 10]. Flow visualization by [10] showed a downstream step of the glottal separation point in the presence of the ventricular bands, inducing a conservation of the flow laminar properties over a longer distance. A significant pressure recovery associated with a reattachment of the jet-flow is observed in [8].

Extending previous theoretical and *in-vitro* investigations presented in [11], this study explores the vibratory properties of both vocal-folds and ventricular-bands in interaction *in-vivo* via the analysis of acoustic, electroglottographic signals and synchronized glottal images obtained by high-speed cinematography on a professional singer. After a quantitative description of the co-oscillations, the aerodynamic behaviour of the laryngeal system is simulated using a simplified aerodynamic modeling and input parameters directly deduced from *in-vivo* measurements. This will open a discussion on the understanding of the physics behind the vocal-ventricular interaction and the resulting sound.

2 Materials and methods

2.1 Aerodynamic modeling

The basic aerodynamic impact implied by the presence of the ventricular bands in the laryngeal tube on the pressure distribution and the vocal-folds self-oscillations is extensively described in [11].

A schematic representation of the laryngeal geometry considered in this investigation is given in Fig.2 along with all relevant parameters of the aerodynamic study. h_{vf} (respectively h_{vb}, $h_{ventricle}$) corresponds to the aperture of the vocal folds (respectively the ventricular bands, the ventricle). A_{vf} (respectively A_{vb}) refers to the cross-sectional area separating the vocal folds (respectively the ventricular bands). In our theoretical approximation, those areas are

rectangular: $A_* = W_* \times h_*$ ($* \in \{vb, vf\}$), W_* being the depth of the section along the z direction (see Fig. 2).
Three coupled subsystems are considered to model the airflow dynamics through the larynx:
- the pressure drop on the inlet of the glottis: $\Delta P_{vf} = P_0 - P_{s1}$,
- the emerging jet evolving in the ventricle, with a dissipation of kinetic energy: $\Delta P_{jet} = P_{s1} - P_2$,
- the pressure drop at the bands: $\Delta P_{vb} = P_2 - P_{s3}$.

The influence of the ventricular-bands presence on the glottal flow and resulting pressure drop has been demonstrated [11].

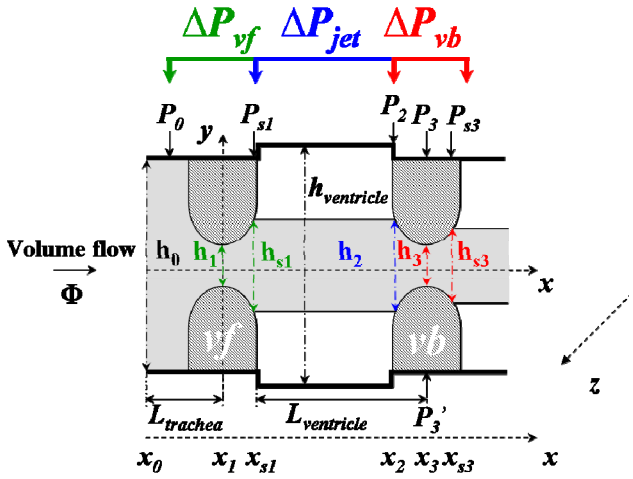

Fig.2 Geometrical sketch of the laryngeal channel and relevant quantities of the aerodynamic study (***vb***: ventricular bands, ***vf***: vocal folds).

A distributed two-mass model of the vocal folds accounting for the additional pressure drop downstream to the glottis owed to the ventricular bands has been proposed, as shown in Fig. 3, and validated on *in-vitro* set-ups of the human larynx [11]. This model is controlled by a set of mechanical (mass (m), stiffness (k, k_c), damping (r)), aerodynamic (including P_0) and geometrical (including $L_{ventricle}$ and h_{vb}) input parameters.

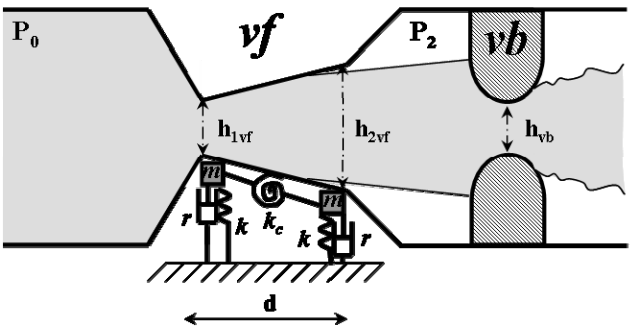

Fig.3 Sketch of the two-mass model of the vocal-folds combined to the theoretical flow description of the ventricular-bands influence used in this study (***vb***: ventricular bands, ***vf***: vocal folds).

All theoretical aspects used in the followings directly refer to this simplified model of phonation, applied under the assumptions of a semi-empirical Liljencrant's flow separation model, a "turbulent" jet-flow geometrical expansion in the ventricle, a dissipation ΔP_{jet} neglected, and a quasi-steady Bernoulli flow dynamics description [11]. Viscous losses are taken into account in the airflow equations of the two-mass model.

2.2 *In-vivo* experimental procedure

The experiment was conducted at the University Medical Center Hamburg-Eppendorf in the Department of Voice, Speech and Hearing Disorders. A professional male singer (MW, age 37) was recorded while performing a specific growly phonation perceptually similar to Asian throat-singing. *In-vivo* investigation of this phonation was carried out using acoustical and electroglottographic measurements, together with clinical examination by high-speed cinematography.

For high-speed imaging, a rigid endoscope (Wolf 90° E 60491) with a continuous light source (Wolf 5131) providing a sufficient illumination of the vocal folds during recording was inserted into the oral cavity. A digital black-and-white CCD camera (Richard WOLF, HS-Endocam 5560) with a resolution of 256x256 pixels was employed to record the high-speed picture sequence. The camera frame rate was selected to 2000 frames/s, which allowed recording sequences of approximately 4 s. The audio signal was captured by a microphone (Wolf 5052.801) placed at the end of the endoscope. An electroglottograph (EG2 [12]) was placed on the singer's neck to get information about glottal period and vocal folds' contact area [13].

2.3 *In-vivo* data analysis procedure

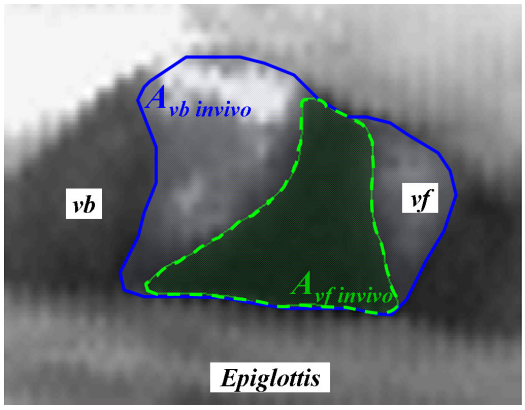

Fig.4 Illustration of a typical high-speed and the detected contours. $A_{vb\ invivo}$ (respectively $A_{vf\ invivo}$) refers to the cross-sectional area between the ventricular bands (respectively the vocal folds) (***vb***: ventricular bands, ***vf***: vocal folds).

High-speed images are resized to focus on the area of interest and smoothed using a bicubic scaling filter (radius 4). Figure 4 illustrates the high-speed imaging processing based on a contours extraction of the vocal folds (dashed line) and the ventricular bands (solid line). These contour plots are shaped as Bezier curves and manually edited on each high-speed image following a method described in [14]. The cross-sectional areas at the maximum constriction of the ventricular bands ($A_{vb\ invivo}$) and of the vocal folds ($A_{vf\ invivo}$) are directly derived from the extracted contours, by means of a triangulation algorithm [14].

Note that in that particular case, the investigated phonatory gesture comes along with a strong narrowing of the laryngeal structures, occasionally combined with a noticeable aryepiglottic constriction. This laryngeal

configuration makes it difficult to detect the effective ventricular-bands and vocal-folds contours with a high accuracy. On the sample illustrated in Fig.4 for instance, both areas $A_{vb\ invivo}$ and $A_{vf\ invivo}$ deduced from the detected contour plots may be underestimated, comparatively to the physiological reality. Yet, as the supraglottic constriction beyond the ventricular bands does not vary much during a typical processed sequence, the method is expected to provide a good estimate of the dynamical evolution of the detected areas.

3 Results and discussion

In the followings, if X is a quantity, X_n refers to the corresponding normalized quantity, such as: $X_n=X/\max(X)$. DEGG corresponds to the time derivative of the electroglottographic (EGG) signal.

3.1 Physiological observations

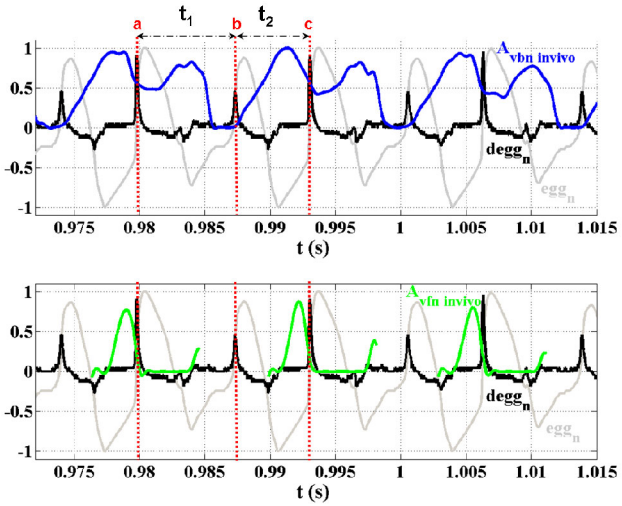

Fig.5 Normalized EGG (egg_n) and DEGG ($degg_n$) signals in function of time measured during growly phonation, along with synchronized and normalized areas $A_{vbn\ invivo}$ (up) and $A_{vfn\ invivo}$ (down) derived from the high-speed images. Shot instants (a, b, c) are plotted as dashed vertical lines.

Figure 5 displays an example of EGG and DEGG signals selected from the growly voice recordings, along with the synchronized cross-sectional areas at the maximum constriction of the ventricular bands, $A_{vbn\ invivo}$, and of the vocal folds, $A_{vfn\ invivo}$. Those areas are deduced from the high-speed images by the contours detection method before-mentioned. Note that no detection of $A_{vfn\ invivo}$ is possible when the ventricular-bands motion hide the glottal aperture, which explains the incomplete signal plotted on Fig.5.

A periodic alteration of both the EGG and DEGG signals is observed, illustrating a period-doubling pattern already detailed in [15]. Yet, whether or not this alteration reflects an additional ventricular motion, a change in glottal contact or both, is still an open question. Which motions are captured by the EGG and DEGG signals? Shot instants a, b, c on Fig.5 provide further insights into this matter. It is shown that at instants a and c, the ventricular bands are coming closer to each other but not touching, while the vocal folds are closing, implying an EGG amplitude rise and a DEGG amplitude "positive" peak. Thus, EGG and DEGG variations are exclusively associated to glottal motions in these cases. At shot instant b, the ventricular bands are already in contact. Data on vocal-folds behaviour are thus not available. Yet, the observed sudden DEGG amplitude peak is not correlated to any sudden changes in ventricular motion; in fact, the adduction of the ventricular bands precedes this instant and is not reflected by a DEGG amplitude peak, although concurrent to a modification of the egg_n signal. Therefore, this suggests that the DEGG amplitude peak at b corresponds only to the closure of the vocal folds. These observations support the thesis that though the EGG may combine information about ventricular and glottal contact areas, the DEGG "positive" peaks exclusively reflect glottal closing states. Consequently, t_1 and t_2 (see Fig.5) define two consecutive glottal cycles durations.

What can be deduced then about the correspondence between ventricular-bands motion and vocal-folds vibrations? Fig.5 confirms quantitatively the results qualitatively demonstrated or speculated in [15] thanks to a kymographical analysis of the high-speed images: a contact between the ventricular bands is periodically observed every two glottal cycles. The vocal folds remain open during part of the ventricular-bands contact. Their closing precedes ventricular-bands opening and occurs at the maximum of ventricular-bands contact. The vocal-folds speed of contact related to the DEGG glottal closing peaks is noticeably reduced under the influence of a ventricular-bands contact downstream. Two following glottal cycles do not have the same duration ($t_1 \neq t_2$) and the duration of the glottal cycle is extended with a ventricular contact ($t_1 > t_2$).

3.2 Theoretical simulations

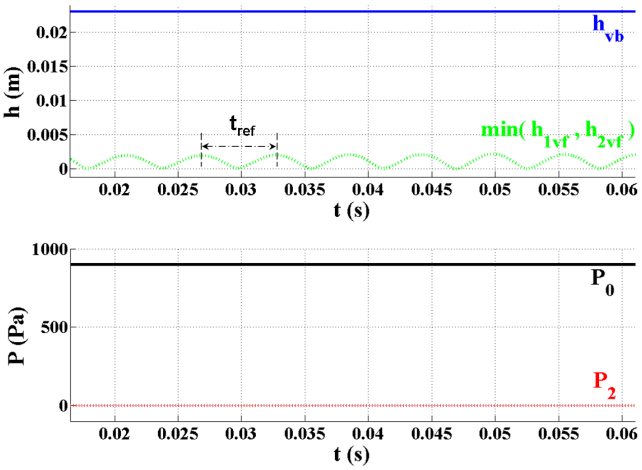

Fig.6 Two-mass simulations of $min(h_{1vf}, h_{2vf})$ and P_2 in function of time. Configuration of reference with $h_{vb} = h_{ventricle}$, $P_0 = 900$Pa. $1/t_{ref} = 170$Hz.

Figure 6 and Figure 7 present theoretical simulations obtained using a two-mass model and accounting for the aerodynamic interaction between the vocal folds and the ventricular bands, as described in 2.1. For all simulations discussed in this section, $L_{ventricle} = 7.5$mm, $h_{ventricle} = 23$mm, $d = 4$mm, $W_{vf} = W_{vb} = 15$mm, $P_0 = 900$Pa and $m = 0.05$g (see Fig.2 and Fig.3). The initial glottal aperture $min(h_{1vf}, h_{2vf})$ is fixed to 0.8mm. These quantities are in agreement with physiological observations [9, 11]. The stiffness input parameter k is chosen so that two consecutive glottal cycles

have the same duration in both the measured *in-vivo* data and the simulation when the ventricular bands are vibrating. In this study, since the corresponding fundamental frequency, $2/t_0$ (see Fig.7), is 150Hz, we choose $k = (4\pi/t_0)^2.m = 44\text{N.m}^{-1}$. The coupling constant k_c is arbitrarily set to 22N.m^{-1}. Finally, the damping input parameter r is fitted so that the two-mass model could simulate stable self-sustained glottal oscillations with complete closure of the vocal folds in the absence of ventricular folds ($r = 1.35.10^{-3}\text{N.s.m}^{-1}$). Figure 6 illustrates this latter configuration. No pressure recovery is predicted in such a case ($P_2 = 0$) and the two-mass model of the vocal folds oscillates regularly with a fundamental frequency $1/t_{ref}$ of 170Hz (see Fig.6).

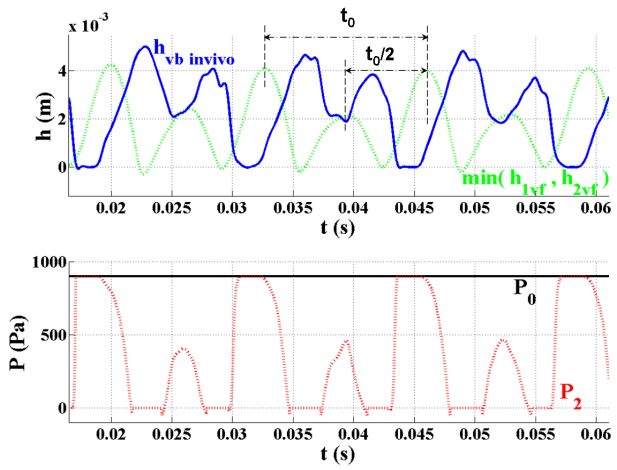

Fig.7 Two-mass simulations of $min(h_{1vf}, h_{2vf})$ and P_2 in function of time. $h_{vb} = h_{vb\ invivo} = A_{vb\ invivo}/W_{vb}$, $P_0 = 900\text{Pa}$. $1/t_0 = 75\text{Hz}$.

Figure 7 presents the pressure P_2 and the glottal aperture $min(h_{1vf}, h_{2vf})$ predicted when a ventricular motion h_{vb}, derived from the high-speed images, is imposed. In this simulation, h_{vb} is approximated by the ratio $A_{vb\ invivo}/W_{vb}$, which comes to $h_{vbmax}.A_{vbn\ invivo}$, h_{vbmax} being the maximum ventricular aperture observed during the phonatory gesture. Because, in this study, we had no possibilities to calibrate the *in-vivo* measurements and consequently no access to this quantity, we chose $h_{vbmax} = 5\text{mm}$, in accordance with physiological data reported in [9, 11].

Under these conditions, a significant pressure recovery P_2 is predicted, even reaching the driving pressure P_0 each time the ventricular bands are in contact. Thus, the pressure drop at the glottis, ΔP_{vf} is affected and the vocal-folds oscillations are altered. A constant alteration of the glottal aperture amplitude between two consecutive vocal-folds oscillations is observed. Therefore, the simulation exhibits a period-doubling phenomenon at the glottis: a second frequency appears in the vocal-folds vibratory pattern, which equals to the fundamental frequency of the ventricular-bands oscillation and of the resulting sound, $1/t_0 = 75\text{Hz}$ (see Fig.7). Figure 8 compares the time derivative of the simulated glottal apertures for the configuration with the ventricular motion extracted from the *in-vivo* measurements (see Fig.7) and for the configuration without any influential ventricular motion (see Fig.6). This result suggests that the aerodynamic interaction between the vocal folds and the ventricular bands does alter the glottal vibrations in terms of frequency and amplitude in the simulations, which may play an important role in the period-doubling phenomenon.

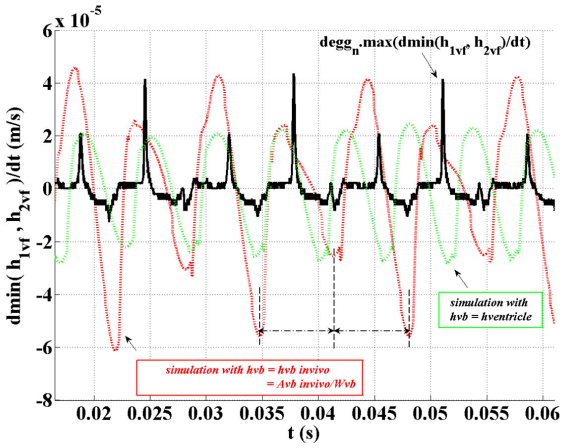

Fig.8 Comparison between $dmin(h_{1vf}, h_{2vf})/dt$ simulated by the two-mass model for two different input values for the ventricular-bands aperture h_{vb}:
$h_{vb} = h_{ventricle}$ and $h_{vb} = A_{vb\ invivo}/W_{vb}$. $P_0 = 900\text{Pa}$.
Superposition of the signal $degg_n.max(dmin(h_{1vf}, h_{2vf})/dt)$.

Yet, in this case, the simulation does not reproduce accurately the frequency difference between two consecutive glottal cycles during the period-doubling phenomenon, as observed *in-vivo* (see section 3.1 and [15]). Besides, it is shown on Fig.8 that an inconsistent phase shift remains between the simulated glottal motions and the measured DEGG signal.

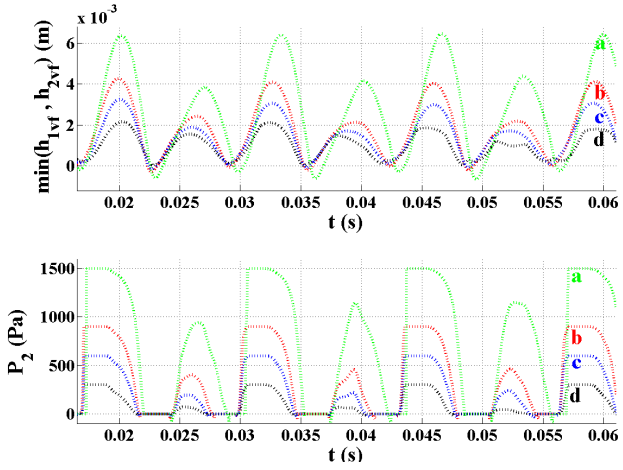

Fig.9 Comparison between $min(h_{1vf}, h_{2vf})$ and P_2 simulated by the two-mass model for four different input values for the driving pressure P_0:
a. $P_0 = 1500\text{Pa}$ **b.** $P_0 = 900\text{Pa}$ **c.** $P_0 = 600\text{Pa}$ **d.** $P_0 = 300\text{Pa}$

The influence of the ventricle length, $L_{ventricle}$, has been tested on the simulations $min(h_{1vf}, h_{2vf})$ and P_2. For $L_{ventricle}$ ranging from 3.5mm to 10mm, which are in the range of physiologically plausible values [9, 11], little differences have been observed on both the amplitude and the frequency of the vocal-folds oscillations. Those differences do not account for the phase shift observed during the *in-vivo* measurements. The driving pressure, P_0 has also been parametrically varied to test the impact on the simulated quantities, as illustrated in Fig.9. Four values of P_0 have been tested (300Pa, 600Pa, 900Pa, 1500Pa), in accordance with physiological data [11]. The amplitude of the vocal-folds oscillations is strongly affected by this parameter. Case **a.** illustrated in Fig.9 shows that the parameter P_0 can also imply a frequency difference between two consecutive glottal cycles.

4 Conclusion

In this study, the laryngeal dynamics of a specific growly phonation produced by a professional singer using period-doubling as a musical performance is explored. This phonation is perceptually similar to Mongolian Kargyraa, Tibetan voice or Sardinian bassu singing and involves the vibrations of the ventricular bands. *In-vivo* investigation of this phonation has been carried out using acoustical and electroglottographical measurements, together with high-speed cinematography. A quantification of the laryngeal dynamics is proposed, extracting glottal and ventricular areas from the high-speed images. These data are used as input parameters of a simplified model of phonation accounting for the aerodynamic interaction between the vocal folds and the ventricular bands.

This study confirms in a quantitative way some results already suggested in [15]. From the processed *in-vivo* data, it is demonstrated that although the EGG signal may combine information about ventricular and glottal contact areas, the DEGG signal exclusively reflects glottal vibratory behaviour. A contact between the ventricular bands is periodically observed every two glottal cycles. Two following glottal cycles do not have the same duration and the duration of the glottal cycle is extended with a ventricular contact. From the theoretical simulations, it is shown that the aerodynamic impact of the ventricular bands motion can induce a constant alteration of the vocal-folds vibration amplitude between two consecutive glottal cycles. Thus, the simulation exhibits a period-doubling phenomenon at the glottis. The phase shift observed between two glottal periods may only be predicted by the theoretical model by changing the specific values of the driving pressure.

Acknowledgments

The authors gratefully acknowledge Pierre Badin for his very helpful contribution to the contours detection method used to process the data. They also acknowledge Frank Müller, Anna Katharina Licht, Markus Hess and Mal Webb for their precious help during the experimental procedure. This work has been supported by a PhD grant from the French Ministry of Research and Education.

References

[1] T. Haji, K. Moir, K. Omori, N. Isshiki, "Mechanical properties of the vocal fold: Stress- Strain studies", *Acta Otolaryngologica (Stockh)* 112, 559-565 (1992)

[2] L. Fuks, B. Hammarberg, J. Sundberg, "A self-sustained vocal-ventricular phonation mode: Acoustical, aerodynamic and glottographic evidences", *KTH TMH-QPSR* 3, 49-59 (1998)

[3] P. A. Lindestad, M. Sodersten, B. Merker, S. Granqvist, S., "Voice source characteristics in Mongolian 'throat singing' studied with high-speed imaging technique, acoustic spectra, and inverse filtering", *J. Voice* 15 (1), 78-85 (2001)

[4] K. I. Sakakibara, T. Konishi, K. Kondo, E. Z. Murano, M. Kumada, H. Imagawa, S. Niimi, "Vocal fold and false vocal fold vibrations and synthesis of khoomei", *Proc. of ICMC* 135-138 (2001)

[5] K. I. Sakakibara, H. Imagawa, S. Niimi, N. Tayama "Physiological study of the supraglottal structure", *Proc. ICVPB* (2004)

[6] N. Henrich, B. Lortat-Jacob, M. Castellengo, L. Bailly, X. Pelorson, "Period-doubling occurences in singing: the 'bassu' case in traditional Sardinian 'A Tenore' singing", *Proc. ICVPB* (2006), Tokyo

[7] C. Shadle, A. Barney, D. Thomas, "An investigation into the acoustics and aerodynamics of the larynx", *Vocal Fold Physiology: Acoustics, Perceptual and Physiological Aspects of Voice Mechanisms*, 73-82 (1991)

[8] X. Pelorson, J. Liljencrants, B. Kroeger, "On the aeroacoustics of voiced sound production", *Proc. 15th ICA*, Trondheim, Norway (1995)

[9] M. Agarwal, "The false vocal folds and their effect on translaryngeal airflow resistance", *Ph.D. thesis*, Bowling Green State University, OH (2004)

[10] B. R. Kucinschi, R. C. Scherer, K. J. DeWitt, T. T. M. Ng, "Flow visualization and acoustic consequences of the air moving through a static model of the human larynx", *Transactions of ASME* 128, 380-390 (2006)

[11] L. Bailly, X. Pelorson, N. Henrich, N. Ruty, "Influence of a constriction in the near field of the vocal folds: Physical modeling and experimental validation", *J. Acoust. Soc. Am.* (revised)

[12] M. Rothenberg, "A multichannel electroglottograph", *Journal of Voice* 6, 36–43 (1992)

[13] N. Henrich, C. d'Alessandro , M. Castellengo, B. Doval, "On the use of the derivative of electroglottographic signals for characterization of nonpathological phonation", *J. Acoust. Soc. Am.* 115(3) 1321-1332 (2004)

[14] A. Serrurier, P. Badin, "A three-dimensional articulatory model of nasals based on MRI and CT data", *J. Acoust. Soc. Am.*, 123(4), 2335-2355. (2008)

[15] L. Bailly, N. Henrich, M. Webb, F. Müller, A.K. Licht, M. Hess, "Exploration of vocal-folds and ventricular-bands interaction in singing using high-speed cinematography and electroglottography," *Proc. 19th ICA, Madrid* (2007)