Modeling and validation of auto-oscillation onset in a constricted tube with application to phonation

Annemie Van Hirtum*, Xavier Pelorson

Gipsa-lab, CNRS UMR 5216, Grenoble University, 11 rue des Mathematiques (BP46), 38402 Grenoble, France

**ABSTRACT**

An experimental study of the auto-oscillation onset is performed for airflow in a rigid tube containing a constriction downstream from a deformable portion as a function of the constriction degree. A quasi-one dimensional laminar flow model in combination with a reduced order mechanical model (symmetric two mass model) provides a physical fluid–structure interaction model of the deformable portion. Mechanical, geometrical and flow model parameters are chosen to match the experimental setup. Modeled as well as experimental results show that a severe constriction (≥ 80%) at first hinders (≥ 89%) and eventually inhibits (≥ 95%) auto-oscillation. Constrictions of different severity occur naturally in voiced speech sound production (phonation) due to articulation. The current study provides quantitative evidence of the role of the vocal tract constriction degree as a control parameter for phonation (voiced speech sound production) since increasing the constriction degree decreases the vocal folds oscillation frequency (decrease by 25%) and increases the minimum pressure needed to initiate oscillation (increase by 80%).

© 2015 Elsevier Ltd. All rights reserved.

1. Introduction

Normal speech production involves a series of successive transitions between open (e.g. neutral vowel or schwa) and obstructed vocal tract configurations (e.g. oral occlusive or stop) so that the corresponding area constriction ratio varies between 0 (no constriction) and 1 (total closure).

Since the pioneering work by the late 1960s (Lisker and Abramson, 1964, 1971), a large amount of ongoing literature reports on the crucial role of the vocal tract configuration in laryngeal and articulatory adjustment for voicing and devoicing (see e.g. Ohala and Riordan, 1980; Westbury, 1983; Bickley and Stevens, 1986; Lofqvist et al., 1995; Svirsky et al., 1997; Koenig et al., 2008; Pinho et al., 2012). As a result, semi-occlusives are an established tool for voice training and semi-occluded vocal tract exercises are commonly used in speech therapy (see e.g. Laukkanen et al., 2008; Cielo et al., 2013).

In contrast to the cited papers on phonetic properties of the phonological voicing contrast, physical and mathematical studies aiming to understand the possible effect of the vocal tract configuration on the vocal folds dynamics are mostly limited to the impact of acoustical coupling for a uniform open vocal tract configuration (Laje et al., 2001; Zhang et al., 2006; Zanartu et al., 2007; Lucero et al., 2012) or to vocal tract acoustics with (Davies et al., 1993) or without (Arela and Guasch, 2013; Blandin et al., 2015) accounting for convective flow effects. Therefore, the aim of this work is to contribute to the modeling and experimental validation of the impact of a vocal tract constriction on the outcome of a physical phonation...
model, i.e. auto-oscillation of the vocal folds. Such a model approach allows us to express key phonation parameters, such as the pressure threshold at phonation onset, as a function of a limited number of physiologically meaningful parameters to which the vocal tract constriction degree is added. This is a necessary step towards a more extensive study of laryngeal or articulatory adjustment from a physical point of view.

In the following, the impact of a vocal tract by varying constriction degree downstream from a glottal replica (Ruty et al., 2007; Ruty, 2007) on vocal folds auto-oscillation onset is systematically explored using an experimental setup to measure pertinent physical quantities. It was shown (Ruty et al., 2007; Ruty, 2007) that the glottal replica was capable to reproduce phonation pressure thresholds (200–1000 Pa) and auto-oscillation frequencies (100–180 Hz) relevant to the ones observed on human speakers. The streamwise length of a uniformly constricted segment is held constant at 20 mm which corresponds to an order of magnitude of a vocal tract constriction between the tongue and the hard palate (Daniloff et al., 1980; Stevens, 2000). The tube length downstream from the glottal replica has a length of 50 cm which is larger than a human vocal tract (typically 18 cm Daniloff et al., 1980; Stevens, 2000). This is done to avoid secondary noise sources. A simple flow model based on Euler equations is applied to describe the pressure distribution within the vocal tract and to model the influence of the vocal tract constriction on the glottal pressure drop driving phonation. The flow model is then applied to a physical model of speech sound production (van Hirtum et al., 2014) in order to address the impact of a vocal tract constriction from basic fluid dynamical principles for a given vocal tract constriction degree. Acoustical coupling with a downstream pipe representing the supraglottal vocal tract is accounted for in the phonation model. The possible influence of the flow on the wavenumber (Davies et al., 1993) is neglected in the applied model approach as well as the influence of acoustic energy losses along the downstream pipe (Atig et al., 2004; Guilloteau et al., 2014). Moreover, a short upstream pipe is considered in both the model as the experiments. Hence the potential impact of acoustical coupling with an upstream pipe representing the subglottal trachea (Zhang et al., 2006) is not accounted for as well.

2. Method

2.1. Flow model

The onset/offset of vocal fold auto-oscillation is governed by the pressure drop across the glottis. From classical flow studies, it is easily understood that the presence of a constriction in the vocal tract downstream from the glottis alters the pressure distribution within the vocal tract and hence the glottal pressure drop. A simplified vocal tract geometry is schematized in Fig. 1. It consists of a uniform channel with cross-sectional area $A_t$ placed between two constrictions – one at the glottis containing the vocal folds and the other one between the tongue and the hard palate (Daniloff et al., 1980; Stevens, 2000). The tube length downstream from the glottal replica has a length of 50 cm which is larger than a human vocal tract (typically 18 cm Daniloff et al., 1980; Stevens, 2000). This is done to avoid secondary noise sources. A simple flow model based on Euler equations is applied to describe the pressure distribution within the vocal tract and to model the influence of the vocal tract constriction on the glottal pressure drop driving phonation. The flow model is then applied to a physical model of speech sound production (van Hirtum et al., 2014) in order to address the impact of a vocal tract constriction from basic fluid dynamical principles for a given vocal tract constriction degree. Acoustical coupling with a downstream pipe representing the supraglottal vocal tract is accounted for in the phonation model. The possible influence of the flow on the wavenumber (Davies et al., 1993) is neglected in the applied model approach as well as the influence of acoustic energy losses along the downstream pipe (Atig et al., 2004; Guilloteau et al., 2014). Moreover, a short upstream pipe is considered in both the model as the experiments. Hence the potential impact of acoustical coupling with an upstream pipe representing the subglottal trachea (Zhang et al., 2006) is not accounted for as well.

![Fig. 1. Schema of glottal and vocal tract geometry enveloping airflow in the streamwise direction (x): pressure $P$ (subglottal $= P_{sub}$, oral $= P_{oral}$, outlet $= P_{out}$), pressure drop $\Delta P$ (glottal $= \Delta P_s$, total $= \Delta P_t$), cross sectional area $A$ (subglottal $= A_u$, glottal flow separation $= A_l$, downstream end of glottal constriction $= A_{l1}$, unconstructed vocal tract $= A_t$ and uniform vocal tract constriction $= A_{t1}$).](image-url)
these assumptions as well as conservation of volume flow rate \( Q \) along the streamwise direction \( x \), i.e. \( dQ/dx = 0 \), integration of the streamwise momentum equation results in three contributions to the total pressure drop \( \Delta P_{\text{tot}} = \Delta p_k + \Delta p_v + \Delta p_i \) for the channel geometry depicted in Fig. 1 at time \( t \):

\[
\Delta p_k = Q(t)^2 \rho \left( \frac{1}{A_{0i}(t)^2} + \frac{1}{A_{1i}(t)^2} \right),
\]

\[
\Delta p_v = -Q(t)12\mu w^2 \left( \int_{x_0}^{x_0} \frac{dx}{A(x,t)^3} + \int_{x_1}^{x_1} \frac{dx}{A(x,t)^3} \right),
\]

\[
\Delta p_i = -\rho \frac{\partial}{\partial t} \left( Q(t) \left( \int_{x_0}^{x_0} \frac{dx}{A(x,t)} + \int_{x_1}^{x_1} \frac{dx}{A(x,t)} \right) \right),
\]

with air density \( \rho = 1.2 \text{ kg/m}^3 \), dynamic viscosity of air \( \mu = 1.8 \times 10^{-5} \text{ Pa s} \) and subscripts as defined in Fig. 1. The first term \( \Delta p_k \) accounts for kinetic pressure losses due to spatial flow acceleration through a channel with varying streamwise area following Euler’s equation for an ideal inviscid steady pressure driven flow while assuming \( (A_0, A_1) \approx (A_m, A_s) \). The second term \( \Delta p_v \) accounts for viscous pressure losses within constricted channel portions following the lubrication approximation of the Navier–Stokes equations for a flow in a two-dimensional channel. The third term \( \Delta p_i \) accounts for flow inertia following Euler’s equation for an ideal inviscid unsteady incompressible pressure driven flow \( (dQ/dx = 0) \). It is seen that for an unconstricted channel the second term at the right-hand sides of Eqs. (1)–(3) can be neglected. The same way (3) vanishes for steady flow.

The outlined flow model is then applied to determine the forces exercised by the flow on a symmetric two-mass vocal folds model outlined in Section 2.2 in order to model the impact of the vocal tract constriction degree \( (1 - A_{1i}/A_s) \) on vocal folds auto-oscillation onset.

### 2.2. Symmetric two-mass vocal folds model

The vocal folds \( (x_0 \leq x \leq x_0) \) mechanics are modeled as a symmetrical low order model in which each vocal fold is represented by two identical masses \( m = m_1 = m_2 \) at positions \( x_m \) and \( x_m \) covered by three massless plates as illustrated in Fig. 2 (Lous et al., 1998; Ruty, 2007; Ruty et al., 2007; van Hirtum et al., 2014). Each of the vocal folds is modeled as a reduced spring–mass–damper system with 2 degrees of freedom driven by the pressure difference \( \Delta P = P_{\text{sub}} - P_{\text{oral}} \) across the masses as illustrated in Fig. 2. As in Fig. 1 the subscripts refer to upstream \( (u) \), unconstriicted vocal tract \( (t) \), subglottal \( (\text{sub}) \) and oral \( (\text{oral}) \). The glottal area \( A(x,t) \) is then approximated as a piecewise linear function of \( x \) determined by the constant upstream \( (A_u) \) and downstream \( (A_s) \) channel area and the time-varying areas at the streamwise position of the masses \( (A_{m_1} \) and \( A_{m_2} ) \):

\[
A_{i-1,i}(x,t) = \frac{A_i(t) - A_{i-1}(t)}{x_i - x_{i-1}}(x - x_{i-1}),
\]

with streamwise position subscripts \( i \in \{ m_1, m_2, 0 \} \) and corresponding \( i-1 \in \{ 0, m_1, m_2 \} \). The applied models for glottal airflow, vocal folds mechanics and acoustic interaction with a vocal tract downstream from the glottis are severe simplifications of the fluid–structure interaction in the larynx during human sound production.

Following the assumptions outlined in Section 2.1, the anterior–posterior \( y \)-dimension is neglected and the time-varying area within the glottis yields \( A(x,t) = h(x,t) \times w \) using the assumption of a fixed glottal width \( w \) and \( x_0 \leq x \leq x_0 \). The moving position of flow separation along the diverging portion of the glottis is taken into account using a geometrical criterion \( A_0(t) = 1.2 \times \min(A(x,t)) \) defining the position of flow separation \( x_0 \) in the range \( x_{m_1} < x_0 \leq x_0 \) (van Hirtum et al., 2009). At each time instant, the volume flow rate \( Q(n+1) \) is determined from the applied subglottal pressure \( P_{\text{sub}} \) using the flow model outlined in Section 2.1. Using this flow model, the total pressure drop between the channel inlet and outlet \( \Delta P_{\text{tot}} \) yields a quadratic equation of volume flow rate \( Q \). Indeed, from Eqs. (1)–(3), it follows that the contributions to the total

![Fig. 2. Schematic representation of a deformable vocal fold modeled as a symmetrical reduced two mass model with \( m_1 = m_2 = m \) (Lous et al., 1998; Ruty, 2007; van Hirtum et al., 2014) constituting the glottal constriction for \( x_0 \leq x \leq x_0 \) as depicted in Fig. 1.](image)
pressure loss $\Delta P_{tot}$ depend either quadratically or linearly on the volume flow rate $Q$ at time step $t = n + 1$ since

$$\Delta p_k = \alpha_k Q(n+1)^2 \quad \text{with} \quad \alpha_k = \frac{\rho}{2} \left( \frac{1}{A_k(n)^2} + \frac{1}{A_{z_k}(n)^2} \right),$$

(5)

$$\Delta p_x = \alpha_x Q(n+1) \quad \text{with} \quad \alpha_x = -12 \mu \omega^2 \left( \int_{x_0}^{x_{t+1}} \frac{dx}{A(x,n)}^2 + \int_{x_1}^{x_{t+1}} \frac{dx}{A(x,n)^2} \right).$$

(6)

$$\Delta p_i = \alpha_i Q(n+1) + C_i \quad \text{with} \quad$$

(7)

$$\alpha_i = -\frac{\rho}{\Delta t} \left( \int_{x_0}^{x_{t+1}} \frac{dx}{A(x,n)} A(x,n-1) + \int_{x_1}^{x_{t+1}} \frac{dx}{A(x,n)} A(x,n-1) \right),$$

(8)

$$C_i = \frac{\rho Q(n)}{\Delta t} \left( \int_{x_0}^{x_{t+1}} \frac{dx}{A(x,n)} + \int_{x_1}^{x_{t+1}} \frac{dx}{A(x,n)} \right),$$

(9)

where $\Delta t$ indicates the time increment. Consequently, the following quadratic equation is solved for volume flow rate $Q(n+1)$ using $\Delta P_{tot} \approx P_{sub}$ as illustrated in Fig. 1:

$$\alpha_k Q^2 + (\alpha_x + \alpha_i) Q + (C_i - P_{sub}) = 0.$$  

(10)

Once the volume flow rate $Q(n+1)$ is estimated, the pressure drops $\Delta P_0$ and $\Delta P_1$ follow directly by combining the first and second right-hand side terms, respectively, of Eqs. (1)–(3). The pressure distribution in the glottis up to flow separation ($x_0 \leq x \leq x_0$) is then obtained as

$$P(x,n+1) = P_{sub} - \frac{\rho}{2} Q(n+1)^2 \left( \frac{1}{A(x,n)^2} \right) + 12 \mu \omega^2 Q(n+1) \int_{x_0}^{x_{t+1}} \frac{dx}{A(x,n)}^2$$

$$+ \frac{\rho}{\Delta t} Q(n+1) \int_{x_0}^{x_{t+1}} \frac{dx}{A(x,n)} A(x,n-1) - \frac{\rho}{\Delta t} Q(n) \int_{x_0}^{x_{t+1}} \frac{dx}{A(x,n)}.$$  

(11)

Downstream from glottal flow separation, the pressure distribution in the glottis for $x_{0b} < x \leq x_{0f}$ yields $P(x,n+1) = P_{oral}(n+1)$ with $P_{oral}(n+1) = P(x_{0b},n+1)$. The pressure distribution is then used to determine the forces exercised by the flow on the two-mass vocal fold model.

The acoustic model applied to the vocal tract assumes plane wave propagation so that vocal tract acoustics is represented as a transmission line of $N$ hard-walled sections of fixed length $L_k$ and uniform cross-sectional area $A_k = w \times h_k$ for $k = 1, ..., N$. In each section the acoustic pressure $p_{ac,k}$ and particle velocity $u_{ac,k}$ are then given as

$$p_{ac,k} = p^+ (x - ct) + p^- (x + ct),$$

(12)

$$u_{ac,k} = \frac{1}{\rho k} (p^+ (x - ct) - p^- (x + ct)),$$  

(13)

with $c$ denoting the speed of sound in air. Following continuity of acoustic pressure and flow at the junctions between sections $k$ and $k+1$ the reflection coefficient at the junction is $r_k = (A_{k+1} - A_k)/(A_{k+1} + A_k)$ so that for $\Delta t = L_k/c$

$$p_{ac,k+1}^- (n) = \beta_k p_{ac,k+1}^+ (n-1) + r_k p_{ac,k}^- (n),$$

(14)

$$p_{ac,k}^- (n) = -r_k p_{ac,k+1}^+ (n-1) + \beta_k p_{ac,k+1}^- (n),$$

(15)

holds propagation constants $\beta_k = 1 - r_k$ and $\beta_k = 1 + r_k$. Coupling of the acoustic vocal tract model with the vocal-fold model is assessed by imposing continuity of volume flow and pressure at the glottal exit following the description outlined in Lous et al. (1998) and Ruty (2007):

$$p_{ac,oral}^+ (n) = \frac{1}{2} \left( \rho c \frac{Q(n)}{A_1} + P_{oral}(n-1) \right),$$

(16)

$$p_{ac,oral}^- (n) = p_i^- (n-1),$$

(17)

$$p_{ac,oral}^+ (n) = p_{ac,oral}^+ (n-1),$$

(18)

$$p_{ac,oral}^- (n) = r_1 p_i^+ (n) + \phi_1 p_i^- (n-1),$$

(19)

with $p_{ac,oral}^+ (n) = p_{ac,oral}^+ (n) + p_{ac,oral}^- (n)$ being the acoustic pressure at the glottis. The pressure difference at the glottis then becomes $\Delta P_0 = P_{sub} - P_{oral} - p_{ac,oral}$. The acoustic model is completed by considering a low frequency approximation of the impedance of a radiating piston (Ruty, 2007).

The symmetrical mechanical model shown in Fig. 2 describes the movement of the two masses perpendicular to the flow along the z direction (Lous et al., 1998; Ruty, 2007; Ruty et al., 2007; van Hirtum et al., 2014). As before, a rectangular glottal
Therefore, in the following the constriction degree is denoted as \( \Phi \). The mechanical glottal replica (width \( w = 25 \text{ mm} \) and constant height \( h_g = 25 \text{ mm} \) by means of a pressure regulator (Norgren type 11-818-987) and manual valve situated upstream from the constriction at the channel end (length \( l = 25 \text{ mm} \) and constant width \( w = 25 \text{ mm} \)). Whenever collision is detected, the values of \( K \) and \( R \) are increased to \( K_{\text{crit}} = 4K \) and \( R_{\text{crit}} = R + 2\sqrt{K_{\text{crit}}m} \), respectively. The two masses have the same mechanical parameters \( K \) and \( m \). With these notations the mechanical model is written as two coupled equations (Lous et al., 1998; Ruty, 2007; Ruty et al., 2007; van Hirtum et al., 2014):

\[
\frac{d^2 A_{m1}}{dt^2} + \frac{RA_{m1}}{\Delta} + K(1+\gamma)A_{m1} - \gamma K A_{m2} = F_1 (A_{m1}, A_{m2}, \Delta P_0),
\]

\[
\frac{d^2 A_{m2}}{dt^2} + \frac{RA_{m2}}{\Delta} + K(1+\gamma)A_{m2} - \gamma K A_{m1} = F_2 (A_{m1}, A_{m2}, \Delta P_0),
\]

with initial areas in the absence of flow \( A_{m1}^0 \) and \( A_{m2}^0 \) and \( F_1 = w \int_{x_1}^{x_2} \lambda(x)p(x) \, dx \), \( F_2 = w \int_{x_1}^{x_2} \lambda(x)p(x) \, dx \) expressing the forces exerted by the fluid on the masses along the \( z \)-axis on the first and second mass, respectively where \( 0 \leq \lambda(x) \leq 1 \) since only part of the pressure exerted by the fluid is taken into account in the forces. From Eqs. (20) and (21), it is seen that - when neglecting damping - stability or instability resulting in auto-oscillation onset depends on the balance between stabilizing stiffness \( K \) and pressure force related to the pressure difference \( \Delta P_0 \) (Ruty et al., 2007; van Hirtum et al., 2014).

Finite difference discretization is applied to Eqs. (20) and (21) (Lous et al., 1998; Ruty, 2007):

\[
\frac{A_{m1}(n+1) + A_{m1}(n-1) - 2A_{m1}(n)}{\Delta t^2} + \frac{RA_{m1}(n+1) - A_{m1}(n)}{\Delta t} + \frac{K(1+\gamma)A_{m1} - \gamma K A_{m2}}{m} = \frac{1}{m}F_1 (A_{m1}, A_{m2}, \Delta P_0),
\]

\[
\frac{A_{m2}(n+1) + A_{m2}(n-1) - 2A_{m2}(n)}{\Delta t^2} + \frac{RA_{m2}(n+1) - A_{m2}(n)}{\Delta t} + \frac{K(1+\gamma)A_{m2} - \gamma K A_{m1}}{m} = \frac{1}{m}F_2 (A_{m1}, A_{m2}, \Delta P_0),
\]

which expresses \( A_{m1}(n+1) \) and \( A_{m2}(n+1) \) as a function of their values at previous time steps \( n \) and \( n-1 \), mechanical model parameters \( m, K, R, \gamma \) and the \( z \)-components of the forces exerted by the fluid flow \( F_1 \) an \( F_2 \). Concretely, the pressure forces are computed as detailed in Lous et al. (1998) and Ruty (2007).

2.3. Experimental setup

In order to experimentally assess the impact of a vocal tract constriction on auto-oscillation onset, experiments are performed using a mechanical replica enveloping two constrictions as depicted in Fig. 3. Air is supplied by a compressor (Copco GA7). The compressor is connected to a pressure reservoir of 0.75 m³ (‘lung’ replica). Pressure can be provided from a few Pa up to 4000 Pa by means of a pressure regulator (Norgren type 11–818–987) and manual valve situated upstream from the pressure reservoir. A mechanical glottal replica is mounted to the reservoir to which a downstream channel (vocal tract replica) is attached. The mechanical glottal replica (width \( w = 25 \text{ mm} \)) contains two latex tubes filled with water representing two vocal folds (Ruty et al., 2007). Mechanical properties of the vocal folds can be modified by changing the water pressure \( P_{\text{water}} \) inside the latex tubes from 3000 Pa up to 6000 Pa by means of a water column. A uniform channel of length \( L = 50 \text{ cm} \), height \( 25 \text{ mm} \) and width \( 25 \text{ mm} \) is attached to the glottal replica. The height \( h_l \) of a rectangular uniform constriction at the channel end (length \( 20 \text{ mm} \) and constant width \( w = 25 \text{ mm} \)) was changed either steadily or following a prescribed motor-driven motion at 0.4 Hz. Besides the unconstricted case \( h_c = h_l \), height \( h_c \) was varied (and then held steady) from 0.8 mm up to 4.1 mm, corresponding to 3% up to 16% of the uniform downstream pipe area so that the constriction degree \( (1 - h_c/h_l) \) ranges from 84% to 97% (Table 1). Note that for a rectangular constriction with constant width the vocal tract constriction degree is \( 1 - A_{12}/A_1 \), which is the main variable in the flow model (1)–(3), reduces to \( 1 - h_c/h_l \). Therefore, in the following the constriction degree is denoted \( 1 - h_c/h_l \) since it is an explicit function of the experimentally varied variable \( h_c \). The pressure immediately upstream from the glottal replica \( P_{\text{sub}} \) and the pressure immediately upstream.
from the constricted portion $P_t$ are measured by pressure sensors (Kulite, XCS-0.93-0.35-Bar-G) positioned in the pressure taps indicated in Fig. 3. The volume flow rate is not measured since the main model parameter is the upstream pressure $P_{sub}$ from which the volume flow rate is estimated as expressed in (10) and the oscillation is driven by the pressure difference across the vocal folds replica following (21).

From Fig. 4, it is seen that for an unconstricted channel $(1 - h_c/h_l = 0 \%)$ the required upstream pressure to sustain auto-oscillation, corresponding to the phonation threshold pressure, is minimal $(P_{sub,0\%} = 313 \text{ Pa})$ for $P_{water} \approx 5100 \text{ Pa}$ and the corresponding oscillation frequency yields $f_{osc,0\%} = 176 \text{ Hz}$. Results reported in the next section are all obtained for $P_{water} = 5100 \text{ Pa}$ and corresponding mechanical model parameters (initial aperture area, stiffness and damping for $P_{water} \approx 5100 \text{ Pa}$) of the glottal replica are taken from Ruty et al. (2007) where these values are derived experimentally for the same glottal replica using the same setup.

3. Results and discussion

Fig. 5 illustrates vocal tract pressure $P_t$ normalized by mean upstream pressure $P_{sub,\text{mean}}$ observed using the experimental setup depicted in Fig. 3 while varying constriction degree $1 - h_c/h_l$ sinusoidally (motor-driven constriction at 0.4 Hz) between 93% and 99.8%. Auto-oscillation of the vocal folds replica is observed when the constriction degree decreases while auto-oscillation ceases when the constriction degree increases. In addition, the simulated pressure $P_t$ estimated using the symmetric two-mass vocal folds model (van Hirtum et al., 2014) with the flow model outlined in Section 2.1 is plotted as well. Modeled data provide a qualitative approximation of the experimental observations including the observed tendency of oscillation onset and offset as a function of constriction degree. Consequently, the experimental setup as well as the model is capable to reproduce ‘voicing–devoicing’ sequences solely by varying the constriction degree within the vocal tract since all other control parameters are held constant during both experiment ($P_{sub,\text{mean}}$ and $P_{water}$) and simulation ($P_{sub,\text{mean}}$, stiffness, damping, initial aperture area).

The impact of a constriction on auto-oscillation behavior of the glottal replica is systematically studied for static vocal tract configurations with constant constriction degree $1 - h_c/h_l$ while the upstream pressure $P_{sub}$ is gradually increased in the same way (to a maximum value $P_{sub,\text{max}} \approx 545 \text{ Pa}$). Measured pressure values ($P_{sub}$ and $P_t$) for constriction degrees $89\%$ ($h_c = 2.7 \text{ mm}$) and $95\%$ ($h_c = 1.2 \text{ mm}$) are illustrated in Fig. 6 as a function of time and frequency. For a constriction

Table 1
Overview of experimentally assessed static vocal tract constriction heights ($h_c$) and associated constriction degrees ($1 - h_c/h_l$).

<table>
<thead>
<tr>
<th>$h_c$ (mm)</th>
<th>25a</th>
<th>41</th>
<th>2.7</th>
<th>2.4</th>
<th>1.9</th>
<th>1.6</th>
<th>1.2</th>
<th>0.8</th>
</tr>
</thead>
<tbody>
<tr>
<td>$1 - h_c/h_l$ (%)</td>
<td>0a</td>
<td>84</td>
<td>89</td>
<td>90</td>
<td>92</td>
<td>94</td>
<td>95</td>
<td>97</td>
</tr>
</tbody>
</table>

a Unconstricted channel, i.e. $h_c = h_l$ in Fig. 3.

Fig. 4. Measured upstream pressures $P_{sub}$ at auto-oscillation onset ($\star$) and offset ($+$) for an unconstricted channel ($h_c = h_l$) as a function of vocal folds internal water pressure $P_{water}$. A minimum pressure $P_{sub,0\%} = 313 \text{ Pa}$ is observed for $P_{water} \approx 5100 \text{ Pa}$ (dashed vertical line).

Fig. 5. Illustration of measured (data) and modeled (model) vocal tract pressure $P_t$ normalized by the mean value of the upstream pressure $P_{sub,\text{mean}}$ (left axis) for a sinusoidally (motor-driven constriction at 0.4 Hz) varied constriction degree $1 - h_c/h_l$ (right axis) as a function of time $t$. 
degree of 89%, the pressure signal $P_l$ (Fig. 6a) starts to oscillate (onset) around its mean value of 0 Pa as the upstream pressure $P_{sub}$ is increased and oscillation stops (offset) as the upstream pressure is decreased. For a constriction degree of 95% no oscillation is observed and the mean value of $P_l$ is proportional to the upstream pressure $P_{sub}$ indicating that the pressure drop $P_{sub} - P_l$ is reduced compared to the case of 89% and is therefore no longer sufficient to sustain auto-oscillation which supports the flow model expressed in Section 2.1.

Fig. 7 shows measured auto-oscillation onset and offset pressures $P_{sub}$ and associated auto-oscillation frequencies $f_{osc}$ normalized by their values observed in the absence of a constriction ($P_{sub,0\%}$ and $f_{osc,0\%}$) as a function of constriction degree $1 - h_c/h_l$ (Table 1). The upstream pressure $P_{sub}$ (Fig. 7a) needed to sustain auto-oscillation onset and offset increases from values observed in the absence of a constriction ($1 - h_c/h_l \leq 90\%$) to almost twice this value ($1 - h_c/h_l \approx 94\%$) until auto-oscillation is no longer observed ($1 - h_c/h_l \geq 95\%$). The measured oscillation frequency $f_{osc}$ (Fig. 7b) decreases with 10% up to 25% as the constriction degree $1 - h_c/h_l$ increases until no auto-oscillation is observed. Consequently, the presence of a

Fig. 6. Illustration of time–frequency properties of measured pressure signals for two constriction degrees $1 - h_c/h_l = 89\%$ ($h_c = 2.7$ mm) and $1 - h_c/h_l = 95\%$ ($h_c = 1.2$ mm): (a) $P_{sub}$ and $P_l$ normalized by $P_{sub,\max} \approx 545$ Pa as a function of time $t$, and (b) spectrogram of $P_l$.

Fig. 7. Influence of constriction degree ($1 - h_c/h_l$) on auto-oscillation features: (a) normalized onset pressure $P_{sub}$, (b) normalized auto-oscillation frequency $f_{osc}$. No auto-oscillation is observed in the shaded area. Ratio 1 (horizontal line) corresponds to no constriction (subscript 0%).
constriction affects the measured minimum upstream pressure needed to sustain auto-oscillation and to a less extent it affects the associated oscillation frequency.

Modeled values of upstream pressure $P_{sub}$ and oscillation frequency $f_{osc}$ at oscillation onset are indicated in Fig. 7 as a function of constriction degree ($1 - h_c/h_l$) as well. Modeled and measured values agree qualitatively and exhibit the same tendency: increase of onset pressure $P_{sub}$ associated with a decrease of oscillation frequency $f_{osc}$ as the constriction degree ($1 - h_c/h_l$) increases until no oscillation is observed for constriction degrees $> 95\%$. Quantitative agreement between measured and predicted onset pressures $P_{sub}$ (Fig. 7a) is within $< 20\%$ for constriction degrees ($1 - h_c/h_l$) up to $90\%$. For larger constriction degrees ($> 90\%$), modeled values underestimate the onset pressure with $40\%$ indicating that additional losses occur which are not accounted for in the flow model. The error in the predicted oscillation frequencies $f_{osc}$ (Fig. 7b) is independent of constriction degree ($1 - h_c/h_l$) since modeled values systematically overestimate measured oscillation frequencies ($\leq 10\%$).

Experimental and modeled data provide evidence that a severe constriction (degree $> 80\%$) in the tube downstream from the glottal replica affects phonation features since at oscillation onset the required pressure $P_{sub}$ as well as the associated oscillation frequency $f_{osc}$ are influenced. Consequently, current data show that the constriction degree is a phonation control parameter since all other control parameters are held constant. Experimentally two major control parameters are related to the glottal replica: the internal water pressure $P_{water}$ (determining the stiffness applied in the two-mass model describing the vocal folds mechanics) and the initial glottal aperture height (and hence glottal area). Increasing those glottal control parameters will increase the oscillation frequency $f_{osc}$ of a glottal replica (Ruty et al., 2007; Cisonni et al., 2011) whereas increasing the constriction degree has the opposite effect. Consequently, the combination of different control parameters might contribute to the understanding of different control strategies for phonation onset parameters $P_{sub}$ and $f_{osc}$ observed on human speakers, such as laryngeal adjustment affecting the initial glottal area (Bickley and Stevens, 1986; Lofqvist et al., 1995; Koenig et al., 2008; Pinho et al., 2012) or oral cavity enlargement reducing the glottal pressure drop (Lisker and Abramson, 1964, 1971; Ohala and Riordan, 1980; Westbury, 1983; Swirsky et al., 1997).

In addition, taking vocal tract parameters (such as the constriction degree) into account when modeling phonation might in the long run lead to the understanding of a universal (cross-language and cross-speaker observed phenomena) intrinsic fundamental frequency, i.e. a lower intrinsic fundamental frequency (up to $25 Hz$) for low vowels (such as /a/) compared to high vowels (such as /i/) (Crandall, 1925; Whalen and Levitt, 1995). Note that $25 Hz$ corresponds to variation of $f_{osc}$ ($10\%$) observed when using the experimental setup to vary the constriction degree. Nevertheless, no conclusion can be formulated based solely on the current study since besides the constriction degree also the streamwise location of the constriction varies during vowel production. In addition, the effect of viscosity in constricted channel portions is modeled (2) relying on the assumption of two-dimensional flow. This assumption is motivated for the assessed experimental replica and supported by observations of the glottal aperture during normal phonation of human speakers (Svec et al., 2000). However, the cross section shape of constricted portions due to articulation is highly variable (Daniloff et al., 1980) and the assumption of two-dimensional flow for the vocal tract constriction is therefore only a first approximation. In order to improve the model outcome for human articulation, the viscous term (2) related to viscous losses within the vocal tract constriction can be adapted to account for the cross section shape of the vocal tract constriction as proposed in (van Hirtum et al., 2014) without changing the model approach.

### 4. Conclusion

Experimental and modeled data show that increasing the constriction degree ($> 80\%$) results in an increase of the minimum upstream pressure required to sustain oscillation (with $80\%$) and a decrease of the associated oscillation frequency (with $25\%$) until oscillation stops for constriction degrees greater than $95\%$. Predicted and measured oscillation frequencies match within $10\%$ for all assessed constriction degrees. Predicted and measured upstream pressure values at oscillation onset match to within $20\%$ for constriction degrees $< 90\%$ and the accuracy decreases (discrepancy between $20\%$ and $40\%$) for constriction degrees $> 90\%$. Therefore, the constriction degree downstream of a deformable section in a rigid tube is an important parameter affecting auto-oscillation of the deformable portion. The current study quantifies the impact of a single constriction on the oscillation threshold pressure and associated oscillation frequency and identifies a range of constriction degrees for which the effect is major (constriction degrees $> 80\%$). Moreover, the applied model approach is capable to capture the impact of the constriction degree. Therefore, the applied model approach can be used in future studies to investigate the mechanisms underlying the onset of instability leading to auto-oscillation. When applied to phonation, the current study provides strong evidence that the constriction degree of the vocal tract needs to be taken into account in physical models of phonation and might in term contribute to the understanding of phenomena such as intrinsic fundamental frequency and the role of laryngeal or articulatory adjustment strategies for phonation.

The tube length in the current experiments was set to $50\ cm$. It is of interest to expand current findings to shorter tube lengths as well as to differ the streamwise position of the constricted tube segment. Moreover, it is of interest to vary geometrical features of the constriction such as its shape and its length.
Acknowledgements

This work was partly supported by EU-FET Grant (EUNISON 308874).

References


