

Experimental and theoretical contribution to the analysis and the modelling of the vocal folds vibration. Anne Bouvet

▶ To cite this version:

Anne Bouvet. Experimental and theoretical contribution to the analysis and the modelling of the vocal folds vibration.. Signal and Image Processing. Université Grenoble Alpes, 2019. English. NNT: 2019GREAT040. tel-02520623

HAL Id: tel-02520623 https://theses.hal.science/tel-02520623

Submitted on 26 Mar 2020

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.



Thèse

pour obtenir le grade de **DOCTEUR DE L'UNIVERSITE GRENOBLE ALPES** Spécialité : **Signal Image Parole Télécoms**

Arrêté ministériel : 25 mai 2016

Présentée par Anne BOUVET

Thèse dirigée par **Xavier PELORSON** et codirigée par **Annemie VAN HIRTUM**

préparée au sein des laboratoires GIPSA-lab et Legi dans l'École doctorale Électronique, Électrotechnique, Automatique, Traitement du Signal (EEATS)

Experimental and theoretical contribution to the analysis and the modelling of the vocal folds vibration

Thèse soutenue publiquement le **29 novembre 2019**, devant le jury composé de :

Madame Anna BARNEY
Professeur des Universités, University of Southampton, Rapportrice
Monsieur Jorge C. LUCERO
Professeur des Universités, Department of Computer Science
University of Brasília, Rapporteur
Monsieur Yves LAPRIE
Directeur de Recherche, CNRS, LORIA, Présdent du jury
Monsieur Isao T. TOKUDA
Professeur des Universités, Department of Mechanical Engineering,
Ritsumeikan University, Examinateur
Monsieur Xavier PELORSON
Directeur de Recherche, CNRS, Legi, Directeur de thèse
Madame Annemie VAN HIRTUM
Chargée de Recherche, CNRS, Legi, Directrice de thèse



UNIVERSITÉ DE GRENOBLE ALPES ÉCOLE DOCTORALE EEATS

Électronique, Électrotechnique, Automatique, Traitement du Signal (EEATS)

THÈSE

pour obtenir le titre de

Docteur en sciences

de l'Université Grenoble Alpes Mention : SIGNAL IMAGE PAROLE TÉLÉCOMS

Présentée et soutenue par

Anne BOUVET

Experimental and theoretical contribution to the analysis and the modelling of the vocal folds vibration

Thèse dirigée par Xavier PELORSON et codirigée par Annemie VAN HIRTUM

préparée aux laboratoires GIPSA-Lab et Legi

soutenue le 29 novembre 2019

Jury :

Anna BARNEY	-	Engineering and Physical Sciences,
		University of Southampton, United Kingdom
Jorge C. LUCERO	-	Dept. of Computer Science,
		University of Brasília, Brazil
Yves LAPRIE	-	Loria, Lorraine University, France
Isao T. TOKUDA	-	Dept. of Mechanical Engineering,
		Ritsumeikan University, Japan
Xavier PELORSON	-	Legi, Grenoble Alpes University, France
Annemie VAN HIRTUM	-	Legi, Grenoble Alpes University, France
	Anna BARNEY Jorge C. LUCERO Yves LAPRIE Isao T. TOKUDA Xavier PELORSON Annemie VAN HIRTUM	Anna BARNEY-Jorge C. LUCERO-Yves LAPRIE Isao T. TOKUDA-Xavier PELORSON Annemie VAN HIRTUM-

Acknowledgements

Je voudrais tout d'abord remercier mes directeurs de thèse pour leur encadrement durant ces 3 ans et plus, de m'avoir fait confiance, envoyée 3 mois au Japon et permis de rencontrer deux fois plus de monde en partant au Legi !

Merci Xavier pour la bienveillance et la pédagogie dont tu fais preuve. De me dire que mon travail était très bon et de repartir avec ma copie intégralement rouge. Et puis, de ne pas m'avoir étripée après m'avoir demandée de décrire mes axes pour la 1500^{me} fois. Et aussi d'être venu m'aider pour les expériences pour que j'arrête de me prendre des poignées de châtaigne avec les maquettes.

Merci Annemie pour toutes les discutions thèse et pas thèse, les idées nulles mais en fait si... pour la relecture phrase par phrase de mon manuscrit et des articles. Merci aussi pour ta patience lors des corrections qui ne se sont, malgré tout, pas terminées en nuit blanche. Merci pour ta bienveillance, le recul et les encouragements.

I also would like to thanks Isao Sensei and all his students and post-PhD for welcoming me in is lab and allowed me to discover their wonderful country.

I thank my jury, Yves Laprie, Anna Barney, Jorge Lucero and Isao Tokuda for agreeing to evaluate my work and for their kind and very positive feedback.

Remercier ma famille, qui n'ont surement toujours pas compris le pourquoi du comment de mon parcours, mais qui acceptent (malgré eux) de ramer à répondre quand on leur demande ce que font leur fille dans la vie (mention spéciale à celle qui à la bonne idée de prendre des notes). Merci d'être restés derrière moi et de m'avoir soutenu dans les moments moins faciles. Merci d'être venu jusqu'à Grenoble pour ma soutenance et d'avoir organisé un pot de thèse tellement super qu'on m'en parle encore aujourd'hui !

Remercier aussi tout particulièrement mes co-bureaux Firas, Andrei et Sophie au Gipsa et Christophe au Legi pour tous les échanges pro et surtout pas pro qui sont toujours un plaisir ! Et puis tous les doctorants, post-doctorants... qui ont répondu présent pour le thé et le jeu de carte de midi au Gipsa et qui répondent toujours pour les "boissons chaudes" de 10h au Legi. Et aussi tous les permanents, personnels techniques et administratifs du Legi avec qui les discussions de midi sont toujours captivantes.

Mention spéciale pour Julie, Vincent et Thomas pour avoir fait tant de kilomètres pour profiter avec moi de ce premier week-end de docteur !

Je voudrais aussi dire un grand merci à Patrick Van de Ponseele et Laurent de Ryck qui m'ont donné la confiance et la motivation de me lancer dans cette aventure.

Merci Maman, Merci Papa, Merci Xavier, Merci Annemie, Merci Jury. Merci copain d'école,

Merci copains Gipsa, Merci copains Legi !

Maintenant je vous laisse j'ai ma partie de Zelda à terminer sur ma Switch et c'est de votre faute !

Bisous,

Anne



Contents

1	Intr	oduction	1
	1.1	Context	1
	1.2	Formulation of the problem and objectives	2
	1.3	Outline of this thesis	3
2	The	vocal folds apparatus	5
	2.1	Phonation system	5
	2.2	Vocal folds structure and oscillation	7
	2.3	Quantification: physical numbers and perturbation	10
	2.4	Studied vocal folds conditions	13
	2.5	Vocal folds replicas	15
	2.6	General set-up	23
	2.7	Summary	25
3	Exte algo	ernal lighting and sensing photoglottography: characterization and MSePGG rithm	29
	3.1	Introduction	30
	3.2	ePGG measurement system	31
	3.3	Mechanical replicas and set-ups	32
	3.4	ePGG signal characterization	35
	3.5	Multi-Signal-ePGG (MSePGG)	39
	3.6	MSePGG results	44
	3.7	Discussion	50

Contents

	3.8	Conclusion	51
4	Influ	nence of level difference due to vocal folds vertical angular asymmetry	53
	4.1	Introduction	54
	4.2	Glottal replicas with imposed asymmetry angle α set up $\ldots \ldots \ldots \ldots$	54
	4.3	Fluid-structure interaction experiments	58
	4.4	Parallel level difference analogy from P_{On}	64
	4.5	Conclusion	67
5	Exp	erimental study of water influence	69
	5.1	Introduction	71
	5.2	Experimental approach	72
	5.3	Results: voice quality features	81
	5.4	Discussion: voice quality features	93
	5.5	Results: complexity analysis	94
	5.6	Discussion: complexity analysis	99
	5.7	Conclusion	100
6	Ana	lytical flow modelling of water influence	103
	6.1	Introduction	104
	6.2	Model approach	105
	6.3	Experimental approach	115
	6.4	Results and discussion	117
	6.5	Conclusion	124
7	Con	clusion and perspectives	125

Bi	bliogi	aphy	141
A	Axes	s and planes	143
B	Cali	bration of sensors	145
	B.1	Pressure sensor	145
	B.2	Optical sensor	146
	B.3	Photo-diode	147
	B.4	Laser transceiver	147
С	Mec	hanical frequency response	151
	C.1	Experimental set up	151
	C.2	Frequency response function	151
	C.3	Results	152
D	Silic	one replica procedure	157
	D.1	M5 casting	157
	D.2	MRI casting	159
	D.3	EPI casting	161
	D.4	Rigid support	164
E	You	ng modulus estimation	167
	E.1	Tensile test	167
	E.2	Results: silicone layers	167
	E.3	Results: latex tubes	172
F	Vert	ical angular asymmetry: leakage area and model	181
	F.1	Calculation of the vertical leakage area	181

Contents

	F.2	Simplified fluid-structure interaction model	182
G	Arti	ficial saliva fluid properties	187
	G.1	Artificial saliva	187
	G.2	Results: density	189
	G.3	Results: kinematic viscosity	189
	G.4	Results: wettability of different replicas surfaces	191
H	Add	itional preliminary results for hydration	193
	H.1	Influence of mean upstream pressure	193
	H.2	Influence of liquid viscosity	194

List of Figures

1.1	Illustration of the voice process.	1
2.1	Illustration of phonation system.	5
2.2	Structure of a human vocal fold: a) realistic view from [26, 27], with intermediate and deep layer forming the vocal ligament and b) simplified illustration	7
2.3	Mechanical response of <i>in-vivo</i> human vocal folds taken from Švec's thesis [28].	8
2.4	Schematical illustration in the medio-frontal plan, of one cycle of vocal folds oscillation. Inspired from [21, 28]	9
2.5	Videostroboscopic image of UVPF. Adapted from Rosen and Simpson	13
2.6	Motion driven rigid (MDR) replica: a) side and b) top view, c) general view in medio-frontal view and d) schema of rigid vocal fold half cylinder (fixed with pressure tap and mobile without)	17
2.7	Illustration of the deformable pressurized latex tube (PLT) replica	18
2.8	Typical amplitude curve of mechanical response for a latex vocal fold tube $(P_{PLT} = 3800 \text{ Pa})$ following the procedure outlined in Appendix C	19
2.9	Typical PLT replica's curves a) threshold pressures and b) oscillation f_0 and me- chanical f_1^M frequencies as a function of internal pressure P_{PLT} for micrometric screws position $o_{mm} = 1 \text{ mm.} \dots \dots$	20
2.10	Silicone vocal folds replicas M5, MRI and EPI in a) the medio-frontal plane b) the transverse plan, c) and d) detailed layers in the medio-frontal plane	21
2.11	Typical amplitude curve of mechanical response for silicone vocal folds replicas, M5, MRI and EPI measured in Appendix C	23
2.12	General set-up.	26
3.1	External photoglottograph (ePGG) principle (mid-coronal plane) of glottal trans- illumination: orientation angle of light source (IR), distance from source d, light sensor (S) at position d_k , trachea area A_t , glottal area $A_g(t)$ and light flux $\Phi(d)$.	30

3.2	Illustration of ePGG fixed subglottal light sensor (S) position and 3 supraglottal source (IR) positions along the front neck surface and source-sensor distances d_k [mm]	32
3.3	Mechanical airway replica with ePGG light sensor (S) and source (IR)	34
3.4	Vocal folds replica a) MDR and b) PLT with subglottal and supraglottal airway replicas to which ePGG sensor (S) and source (IR) are mounted.	35
3.5	Mean ePGG signal as a function of source-sensor distance d for the airway replica with 2 (×) and 3 (+) leather layers	36
3.6	Mean ePGG signal as a function of source orientation.	37
3.7	Mean ePGG signal as a function of static glottal area A_g for a rigid (×) and a deformable (+) mechanical glottal replica.	37
3.8	Correlated time signals of glottal area $A_g(t)$ (dashed line) and scaled ePGG (full line): a) rigid replica ($f_0 = 10$ Hz), b) deformable replica ($f_0 = 129$ Hz)	38
3.9	MSePGG workflow overview.	45
3.10	Deformable mechanical replica: a) ePGG signal portion, b) measured A_g (dashed line) and estimated \hat{A}_g (full line) area. Mean error yields 5.4%	46
3.11	Sustained vowel /a/ pronounced by a human speaker for 3 ($\tilde{K} = 3$) source-sensor distances d : a) ePGG data as a function of time, b) extracted ($\tilde{K} \times N = 9$) ePGG values for MSePGG parameter estimation (symbols) as a function of d and their linear fit (dashed line) for three ($N = 3$) areas ($A_1 > A_2 > A_3$), c) estimated glottal areas $\hat{A}_g(t)$.	48
3.12	Estimated glottal areas \hat{A}_g for consecutive (5 s) utterance of /sa/ by a human speaker.	49
4.1	Imposed asymmetry angle α , glottal gap parameters A , l_A , h_A and geometrical parameters E , e , w and a : a) glottal gap (red triangle) for MRI replica with $\alpha = 20^{\circ}$ ($A \approx 32 \text{ mm}^2$, $l_A \approx 14.2 \text{ mm}$ and $h_A \approx 4.6 \text{ mm}$), b) sagittal view, c) single left vocal fold.	55

List of Figures

٠	
1	\mathbf{v}
	х

4.2	Geometrical characterization of M5, MRI and EPI replicas as a function of asymmetry angle α : a) degree of contact \mathscr{G} (Eq. (4.1)) and critical angles $\alpha_{(I,II),(II,III)}$ (vertical lines), b) glottal gap area A (Eq. (4.2)) and maximum normal prephonatory glottal area (5 mm ² , horizontal line) and c) glottal gap height h_A (Eq. (4.4)) and maximum vertical level difference during phonation (3 mm, horizontal line). Symbols indicate experimentally assessed α values.	56
4.3	Measured upstream pressure time signal of $P_u(t)$ for the EPI replica with $\alpha = 0^\circ$. Oscillation onset P_{On} and offset P_{Off} pressures are indicated (square) as is the steady state oscillatory signal portion ($\overline{P}_u \approx P_{On}$) used for analysis (5 s). For clarity, a zoom of $P_u(t)$ near oscillation onset (left), during steady state oscillation (middle) and near oscillation offset (right) is provided.	59
4.4	Mean (symbols) and standard deviation (error bars) of onset P_{On} and offset P_{Off} pressures as a function of asymmetry angle α : a) M5, b) MRI and c) EPI. Critical angles (vertical lines) indicate contact regime changes between $\mathscr{G} = 1$ (I), $\mathscr{G} \in]0, 1[$ (II) and $\mathscr{G} = 0$ (III).	60
4.5	Mean (symbols) and standard deviation (error bars) of harmonic frequencies (f_N) as a function of asymmetry angle α : a) M5, b) MRI and c) EPI. Critical angles (vertical lines) indicate contact regime changes between $\mathscr{G} = 1$ (I), $\mathscr{G} \in]0, 1[$ (II) and $\mathscr{G} = 0$ (III). Horizontal dashed lines indicate mechanical resonance frequencies $f_{1,2}^M$.	61
4.6	Mean (symbols) and standard deviation (error bars) of total harmonic distortion (THD in [dB]) as a function of asymmetry angle α for all replicas. Critical angles (vertical lines) indicate contact regime changes between $\mathscr{G} = 1$ (I), $\mathscr{G} \in]0,1[$ (II) and $\mathscr{G} = 0$ (III).	62
4.7	Measured upstream pressure $P_u(t)$ for EPI steady state oscillation: a) $\alpha = 0^\circ$ and b) $\alpha = 20^\circ$.	63
4.8	Harmonics power ratio $\mathscr{P}_{f_1}/\mathscr{P}_{f_0}$ as a function of asymmetry angle α for M5, MRI and EPI. Critical angles (vertical lines) indicate contact regime changes between $\mathscr{G} = 1$ (I), $\mathscr{G} \in]0, 1[$ (II) and $\mathscr{G} = 0$ (III)	63
4.9	Mean (symbols) and standard deviation (error bars) of signal-to-noise ratio (SNR in [dB]) as a function of asymmetry angle α for all replicas. Critical angles (vertical lines) indicate contact regime changes between $\mathscr{G} = 1$ (I), $\mathscr{G} \in]0,1[$ (II) and $\mathscr{G} = 0$ (III)	64
4.10	Measured (symbols) $P_{On}(h_{l_A}(\alpha))$ and fitted (full lines) with Eq. (4.6) for all replicas.	67

5.1	Schematic illustration of experimental setup in the medio-frontal plane indicating liquid injection time tag t_L , sprayed liquid volume V_L , upstream pressure $P_u(t)$, glottal area $A_c(t)$. Assessed vocal folds replicas: deformable (silicone and PLT) and forced motion (MDR with constriction pressure $P_c(t)$).	74
5.2	Graduated syringe equipped with a commercial round spray nozzle	75
5.3	Illustration of waveform shape feature vectors for f_0 cycle-to-cycle analysis for pressure $P_1 = P_u \dots \dots$	77
5.4	Illustration of f_N cycle analysis and oscillation regimes: a) normal $f_N = f_0$, b) subharmonic $f_N = f_0/2$.	78
5.5	Deformable vocal folds replicas in parameter space $(f_{0,d}, \overline{P}_u)$ for $V_L = 0$ mL: silicone vocal folds replicas (M5, MRI, EPI: filled rectangles) and PLT vocal folds replica conditions $(P_{PLT}, 'o_{mm})$ for varied P_{PLT} (full line rectangles) or varied o_{mm} (dashed rectangles) from reference condition (2800,0).	82
5.6	Motion driven rigid MDR vocal folds replica: a) measurement, b) $\mathscr{A}(V_L)/\mathscr{A}_d$ from $P_c(t)$, c) $T^t(V_L)/T^t_d$ from $P_c(t)$ and d) $(T^c/T^o)/(T^c/T^o)_d$ from $P_c(t)$. Sub- script <i>d</i> indicates $V_L = 0$ mL. Mean (symbols) and standard deviation (vertical bars) are plotted in b), c) and d). For clarity data for {1,10} Hz are shifted around V_L	83
5.7	Deformable vocal folds replicas: a) measurement, b) $P_u(t \cdot f_N)$ for PLT condition (3300,0) and $V_L \leq 5$, c) $P_u(t \cdot f_N)$ for EPI and $V_L \leq 5$. b,c) Time is normalised with f_N yielding $f_{0(,d)}$ or $f_0/2$. Curves for $V_L \geq 3$ mL are upshifted, $V_L \geq 2$ mL (dotted lines) and $V_L = 0$ mL (full line).	84
5.8	Mean (symbols) and standard deviation (vertical bars) as a function of sprayed water volume V_L : a) normalized period T/T_d and added mass hypothesis (dash-dotted line) with $m_d = \overline{m}_d$ (Eq. (5.10) for $\alpha_L = 1$ and $\alpha_m = 1$) and $m_d = \overline{m}_d \pm \Delta m_d$ (shaded region), b) period perturbation ζ_T . Values from f_0 cycle-to-cycle analysis.	86
5.9	Mean (symbols) and standard deviation (vertical bars) as a function of sprayed water volume V_L : a) normalized amplitude $\mathscr{A}/\mathscr{A}_d$, b) amplitude perturbation $\zeta_{\mathscr{A}}$, c) turbulence intensity $\Delta U_{P_{u,max}}/\overline{U}_{P_{u,max}}$.	87
5.10	Mean (symbols) and standard deviation (vertical bars) as a function of sprayed water volume V_L : a) normalised closed portion $(T^t/T)/(T^t/T)_d$, b) closing-opening asymmetry $(T^c/T^o)/(T^c/T^o)_d$.	88

List of Figures

5.11	Evidence for subharmonic frequency $f_0/2$ as a function of V_L : a) subharmonic- to-fundamental power ratio $1 - P_{f_0/2}/P_{f_0}$, b) probability of <i>T</i> from f_N cycle-to- cycle analysis for PLT condition (2800,0)
5.12	Mean (symbols) and standard deviation (vertical bars) as a function of sprayed water volume V_L : a) normalized period T/T_d , b) period perturbation ζ_T . Values from f_N cycle-to-cycle analysis
5.13	Overall spectral features: a) total harmonic distortion rate THD, b) signal-to- noise ratio SNR
5.14	Illustration of upstream pressure time series $P_u(t \cdot f_N)$ (left) and their phase space representation (right) for volumes $V_L \leq 5$ mL (gray scaling indicating V_L in mL and f_N in Hz) for silicone (MRI) and PLT (condition (2800,0)) vocal folds repli- cas. For clarity time series $P_u(t \cdot f_N)$ are up-shifted (+1500 Pa) with each V_L increment
5.15	Recurrence plots for $V_L \in \{2,4\}$ mL for vocal folds replicas: a,c) PLT condition (2800,0) and b,d) silicone MRI
5.16	Degree of determinism $(\mathscr{D}(V_L))$, recurrence rate $(\mathscr{R}(V_L))$ and their ratio $(\gamma(V_L))$ for vocal folds replicas (symbols) shifted around each V_L value for clarity 97
5.17	Entropy $E(V_L)$ and estimated correlation dimension $D_2(V_L)$ for vocal folds replicas (symbols) shifted around each V_L value for clarity. The frame contains a zoom for $D_2(V_L)$ with $V_L \in \{1,2\}$ mL
6.1	Schematic overview of the vertical ($\beta = 90^{\circ}$) rigid channel with time-varying rectangular constriction (radius 1 cm, width $L = 3.0$ cm), pressure taps and fluid supplies: orientation angle β , gravitational acceleration g , imposed mobile vocal folds frequency f_c , time-varying minimum spacing $h_c(t)$, liquid supply time tag $t_L(t)$, liquid volume V_L , upstream pressure $P_u(t)$ and minimum constriction pressure $P_c(t)$. Flow is supplied along the positive z-direction and liquid is injected along the negative z-direction.
6.2	Void fraction v_f (thick full line) and normalised homogeneous density ρ_h/ρ_L (thin dashed line) as a function of mass quality <i>x</i> for air-water mixture. As a reference $v_f = x$ (diagonal line) for $\rho_L = \rho_G = \rho_h$ is indicated
6.3	Scaling factors $\alpha_{\mu,i}$ in Eq. (6.5) for gas $(i = G)$ and liquid $(i = L)$ ruled flows as a function of <i>x</i> for air-water mixture: $\alpha_{\mu,i} = f(\rho_h/\rho_i)$ (thin dashed line), $\alpha_{\mu,i} = f(\rho_i/\rho_j, x)$ (thick full line) and unity $\alpha_{\mu,i} = 1$ (horizontal dotted line) 109

6.4	Normalised homogeneous mixing viscosity $\mu_h(x)/\mu_L$ from Eq. (6.6) for airwater mixture with (γ_L , γ_G) following: McAdam (gray thick dashed), Cicchitti (black diagonal thick dashed), Lin (black thick full), Dukler (gray thin full), Beattie (black thick dotted). Also $\mu_h(x)/\mu_L$ from Eq. (6.5) with Garcia $(f(\rho_h/\rho_i))$, discontinuous thin dash-dotted) is shown. Note that $\mu_G/\mu_L \approx 0.018$. A zoom for $x \leq 0.01$ is included.	110
6.5	Normalised homogeneous mixing viscosity $\mu_h(x)/\mu_L$ from Eq. (6.6) (McAdam (gray thick dashed) and Cicchitti (black diagonal thick dashed)) and turbulent mixing viscosity μ_h^{τ}/μ_L from Eq. (6.9) (Aung (black thin full)) for air-water mixture.	111
6.6	Multipliers ϕ_L^2 and ϕ_G^2 as a function of Martinelli's parameter X for air-water flow ($\rho_L/\rho_G = 832$) and for different Chisholm's constants: $C_{LG} = 5$ for laminar- laminar (LL) flow (full line), $C_{LG} = 10$ for turbulent-laminar (TL) flow (dashed- dotted line), $C_{LG} = 12$ for laminar-turbulent (LT) flow (dashed line) and $C_{LG} = 20$ for turbulent-turbulent (TT) flow (dotted line).	113
6.7	C_{LG} (thin dashed line) constant Eq. (6.19) and Laplace number (thick full line) Eq. (6.20) for air-water flow ($\rho_L/\rho_G = 832$) as a function of hydraulic diameter D. As a reference constant Chisholm's values for laminar-laminar (LL) flow ($C_{LG} = 5$) and for turbulent-laminar (TL) flow ($C_{LG} = 10$) are shown	114
6.8	Measured data for $f_c = 1$ Hz and $V_L = 4$ mL: (top) $P_u(t)$ (black full line), $P_c(t)$ (gray full line) and $t_L(t)$ (dashed line) with duration of liquid injection Δt_L and (bottom) $h_c(t)$ (full line).	116
6.9	Illustration of minimum constriction pressure P_c during a single period $t \cdot f_c$ ($f_c = 10$ Hz) for $V_L \in \{0, 1, 3, 5\}$ mL for given P_u and h_c (scaled with respect to P_u).	117
6.10	Illustration of upstream pressure P_u and minimum constriction pressure P_c during a single period $t \cdot f_c$ for $f_c \in \{1, 6, 10\}$ Hz and prescribed h_c (scaled with respect to P_u): a) $V_L = 0$ mL and b) $V_L = 0$ mL (black curves) and $V_L = 5$ mL (gray curves). For clarity P_c is shifted down for $f_c \in \{1, 6\}$ Hz	118
6.11	Measured P_c (black full) and best models outcomes for $V_L \in \{1, 2, 3, 4, 5\}$ mL (shifted for clarity) and $f_c \in \{1, 6, 10\}$ Hz during closing and opening $(h_c > 0)$: gas dominated slip (SG) model Eq. (6.12) for $C_{LG} \approx 12$ (thick dotted) or $C_{LG}(z)$ (gray thin full), slip (S) model Eq. (6.16) for $C_{LG} \approx 20$ (thick dashed) and homogenous (H) flow model Eq. (6.6) for Cicchitti parameters (gray thin	
	dash-dotted) and following Aung Eq. (6.9) (thin dotted).	123
A.1	Illustration of the different planes and axes used.	143

B.1	Calibration curves obtained for the sensors of Kulite XCS-093, Endevco 8507C-2 and Endevco 8507C-5 types: pressure measured by a liquid column manometer as a function of the voltage measured by the sensors
B.2	Calibration curve obtained for the optical sensor OPB700: heigth of the aperture of the MDR replica constriction as a function of voltage given by the sensor 147
B.3	Example of calibration curve obtained for the photo-diode BPW34: aperture imposed with the MDR replica constriction as a function of voltage given by the sensor
B.4	Schematic drawing of the electronic circuit allowing to adapt and recover the photo-diode signal
B.5	Setup for the laser calibration. The laser (left) pointing onto the target (right) placed between 60 mm and 180 mm
B.6	Calibration curves obtained for the laser transceiver. Distance as a function of voltage between 60 mm and 180 mm
C.1	Mechanical frequency response set up
C.2	Example mechanical responses silcone replicas EPI, M5 and MRI
C.3	Example of PLT replica mechanical response for $P_{PLT} = 3800$ Pa
C.4	PLT vocal folds mechanical response: a) oscillation amplitude as a function of internal pressure P_{PLT} and frequency f and b) quality factor as function of the inside water pressure P_{PLT} and frequency f
D.1	Silicone products mixed to obtain the different layers of the silicone VF replica 157
D.2	Example of mixing Ecoflex parts A and B and Silicone thinner (ratio <i>M</i> ES 1:4) for the body layer of MRI replica
D.3	Illustration (a) and (b) and picture (c) and (d) of the M5 molds
D.4	Picture of the MRI molds
D.5	Illustration (a), (b), (c) and (d) and picture (e), (f) and (g) of the EPI molds 162
D.6	Illustration a), b) and c) of the sheet placement for each type of mold and d) product used to fix the sheet on the mold

D.7	a) Adhesive silicone and b) rigid supports used to fix silicone VF (M5 and EPI on the left and MRI on the right)	164
E.1	Illustration of tensile test on DS 1:1 specimen (the bottom bite is fixed and the top bite moves in the upward direction).	168
E.2	Silicone VF replicas layers in the medio-frontal plane	168
E.3	Specimen design used for tensile test	170
E.4	Illustration of a) mold and b) specimen	170
E.5	Effect of measured data smoothing using a moving average (MA) of length 10 (MA 10) and 20 (MA 20) data points on $\mathscr{C}(\varepsilon)$.	171
E.6	Measured section areas $\mathscr{S}(symbols)$ as a function of the deformation Δl and their quadratic fitting (full lines).	172
E.7	Comparison of the constrains calculated with the varying section \mathscr{S}_0 and with the section $\mathscr{S}_0=150 \text{ mm}^2$	173
E.8	a) Smoothed $\frac{\mathscr{C}}{\mathscr{T}}(\varepsilon)$ for all mixing ratios. b) Example of charging and decharging for DS $\mathscr{M} = 1:1$ specimen.	174
E.9	Derivatives curves of maximum deformation measurements (Figure E.8(a)) 1	175
E.10	Linear fitting of silicone specimens	176
E.11	Tensile test configuration for a) latex tube with diameter 11 mm and b) latex tube with diameter 28.5 mm	177
E.12	Charge and decharge of each measurement with observation of the relaxation deformation.	179
E.13	Linear fitting of latex tubes	180
F.1	Illustration of the different area use to calculate A with $A = A_b - A_j + Av$ 1	182
F.2	Illustration vertical difference for parallel vocal folds based on [5]	183
G.1	Artisial presentation and composition	188
G.2	Teijin presentation.	189
G.3	Cannon-Fenske viscometer with two measurement spheres S_1 and S_2	190

G.4	Illustration of the wettability of the three types of vocal folds replicas for the different liquids used. Contact angles are indicated
H.1	Influence of upstream pressure for two configurations of the PLT replica: (1900,1) and (2800,1)
H.2	Influence of liquid spraying, distilled water and Artisial, for two configuration of the PLT replica: (1900,1) and (2800,1)

List of Tables

2.1	Summarised values of density and viscosity measured in Apppendix G	15
2.2	Summary of moduli of elasticity \mathscr{E} [kPa] for different molding layer (mixing ratio \mathscr{M}) from literature and estimated in the current study (Appendix E)	22
2.3	Layer properties in human and in silicone vocal folds replicas $(M5^a, MRI^b, EPI^c)$: measured Young modulus \mathscr{E} , ratio \mathscr{M} , layer thickness l_d .	23
2.4	Vocal folds replicas and human [30, 52, 61].	24
2.5	Typical vocal folds physical and perturbation values.	24
2.6	Vocal folds replicas used and main quantities analysed in the next chapters (\checkmark).	25
3.1	Typical values for variables of interest for adults and mechanical replicas [21, 48, 49, 85–87].	33
4.1	Dimensional parameters and critical asymmetry angles $\alpha_{(I,II),(II,III)}$ for all replicas (M5, MRI and EPI).	54
4.2	Minimum and maximum onset pressures relative [%] to the value for $\alpha = 0^{\circ}$ and associated α [°]	59
4.3	Overall decrease of fundamental frequency f_0 relative [%] to the value for $\alpha = 0^\circ$ in the full contact (I), partial contact (II) and no contact (III) regime	62
4.4	Overview for $\alpha = 24.6^{\circ}$ of leakage area <i>A</i> and potential parallel level difference analogy approximations $\Delta E \approx h_A$, $\Delta E \approx h_L$ and $\Delta E \approx h_{l_A}$.	65
4.5	Fit accuracy (R^2 in %) with $\Delta E(\alpha) \approx h_L(\alpha)$ and with $\Delta E(\alpha) \approx h_{l_A}(\alpha)$	66
6.1	Injected liquid volume V_L and duration Δt_L . Estimated flow rate Φ_L during injection.	115
6.2	Summary of best model Eq. (·) and their accuracies (coefficient of determination $R^2 \leq 1$ and mean absolute relative error $\xi \geq 0$): gas dominated slip model (SG), slip model (S) and homogenous model (H) and their parameters.	121

7.1	Typical vocal folds physical and perturbation values
B.1	Pressure sensors calibration
B.2	Sensitivities of the laser transceiver calibration for different materials
C.1	Overview mechanical resonance properties of replicas for $\alpha = 0^{\circ}$
E.1	Different mixing ratio used for the layers of each replica
E.2	Summary of moduli of elasticity & [kPa] found in literature and estimated in Appendix E
E.3	Measured modulus of elasticity $\mathscr E$ and their accuracy $\mathbb R^2$ of each silicone layer 174
E.4	Initial section and relaxation deformation values for each tube
E.5	Measured modulus of elasticity and their accuracy of each latex tube
G.1	Summarised values of density and viscosity
G.2	Size and constants of capillaries used
G.3	Results of viscosity measurements realized for artificial saliva Artisial 190
G.4	Results of viscosity measurements realized for artificial saliva Teijin
G.5	Results of wettability for used liquids

Nomenclature

Acron	iyms	
VF	Vocal fold	
FS	Fluid structure	
FSI	Fluid structure interaction	
UVFP	Unilateral vocal fold paralysis	
FRF	Frequency response function	
MDR	Motion driven replica	
PLT	Pressurized latex tube	
M5	Silicone VF replica based on M5 model	
MRI	Silicone VF replica based on MRI VF images	
EPI	Silicone VF replica based on VF structure with epithelium	
SNR	Signal-to-Noise ratio	dB
THD	Total Harmonic Distortion	dB
MSeP	GG Multi-Signal external Photoglottograph	
L	Liquid	
G	Gaz	
d	Dry configuration	
Physi	cal quantities	
Р	Pressure	Pa
ρ	Density	kg.m ⁻³
μ	Dynamic viscosity	$kg.s^{-1}.m^{-1}$
v	Kinematic viscosity	$m^2.s^{-1}$
Re	Reynolds number	[—]

Sr	Strouhal number	[—]
Ма	Mach number	[—]
E	Young modulus	Pa
P	Signal power	dB
g	Gravitational constant	9.81 m.s^{-2}
Geo	metry Parameters	
A	Area	m ²
A_c	Area at the minimum of the constriction	m ²
A_g	Human glottal area	m ²
A_s	Area at the jet separation position	m ²
A_t	Trachea area	m ²
A_0	Initial constriction aperture	m ²
h_c	Spacing between the two VFs	m
L_z	Inferior-superior length of the VF	m
l_d	Thickness of different layers of VF	m
W	Posterior - anterior width of the VF	m
R_t	Trachea radius	m
L_t	Trachea length	m
Mea	surement Parameters	
P_u	Upstream pressure	Pa
P_{c}	Constriction pressure	Pa
P _{On}	Onset threshold pressure	Pa
P _{Off}	Offset threshold pressure	Pa
t	Time	S
ζ_A	Shimmer	[—]

Nomenclature

ζ_T	Jitter	[—]
Т	Period	S
f_0	Fundamental frequency	Hz
f_{-}	Harmonic frequencies	Hz
f_c	Imposed oscillation frequency for MDR replica	Hz
f^M	Mechanical frequency	Hz
Q	Quality factor	[-]
f_s	Sampling frequency	Hz
P _{PLT}	Internal pressure of the PLT replica	Pa
0 _{mm}	Position of the micrometric screws	[—]
Φ	Volume flow velocity	m ³ /s
V_L	Sprayed liquid volume	mL
Spa	atial Parameters	
Z	Inferior - superior axis	m
Х	Left - right axis	m
У	Posterior - anterior axis	m
x_c, y_c, z	z_c Position of the minimum of the glottal constriction	

xxiii

 x_s, y_s, z_s Position of the jet separation

CHAPTER 1 Introduction

1.1 Context

Human voice is generated by the auto-oscillation of vocal folds. From the physical point of view, as illustrated in Figure 1.1, this oscillation is created by the interaction between the air flow from the lungs (in dark blue) and the soft tissues that constitute the vocal folds (in light blue). This vibration will modulate the air flow and form an acoustic wave (in red) that propagates through the vocal tract (in green). The vocal tract is articulated by several muscles and will modify the spectrum of the acoustic wave to finally form a voiced sound, which is radiated from the lips. Understanding the complex phenomena that are involved in the vibration of vocal folds is essential to be able to reproduce them in the context of speech synthesis by physical models or to be able to provide diagnostic or surgery tools in the case of pathologies. Besides normal phonation [1, 2], physical studies contribute to the understanding of several vocal folds normal and abnormal conditions affecting phonation, such as pathologies related to kystes and nodules [3, 4], vocal folds asymmetries [5–8] or yet the influence of the glottal area geometry [9]. This thesis focuses on the physical study of the fluid-structure interaction in order to considerate the impact of particular parameters that may affect the vibration of vocal folds and the glottal airflow.



(a) Medio-sagittal plane

(b) Schematic overview of the fluid-structure interaction

Figure 1.1 – Illustration of the voice process.

This work was realised as part of the ANR ArtSpeech¹ project in collaboration with the Phonology Phonetics Laboratory and the Lorraine Research laboratory for Computer science and its Applications, which aims to synthesize speech by digital simulation from physical processes. It was partly achieved in collaboration with the Department of Mechanical Engineering of the Ritsumeikan University, as part of the Japan Society for the Promotion of Science Summer Program Grant².

1.2 Formulation of the problem and objectives

In the framework of physical studies on vocal folds auto-oscillation, three main contributions are aimed in this thesis.

Firstly, the observation of the glottal opening – during phonation, breathing, speech production or swallowing – is a constant question in the voice studies, either physical or clinical. In this work it is aimed to measure accurately the glottal area variation. Accurate knowledge of the glottal area is of interest for clinical purposes and is necessary to estimate the glottal flow and pressure forces on the vocal folds. Therefore, the objective is to calibrate the existing device external PhotoGlottoGraph (ePGG) [10–12], so that a relationship is established between measured ePGG signals and the variation of the glottal area $A_g(t)$.

Secondly, an experimental study of unilateral vocal fold paralysis (UVFP) [3] is realised. This pathology results in mechanical and geometrical asymmetry characterised by vocal folds asymmetry in both tension and spatial positioning. On human subjects, the effect of tension and geometry can not be studied independently nor systematically. Therefore, the objective is a systematic study of geometrical changes, mimicking an unilateral vocal fold paralysis, and more in particular to quantify the influence of an angular vertical asymmetry on the auto-oscillation. This study is done in the continuation of a previous study by Tokuda *et al.* [5, 6] on the influence of a vertical level difference for parallel vocal folds.

Thirdly, the influence of a liquid layer on the vocal folds surface is quantified and investigated. It is known that vocal folds are covered by a thin layer of liquid essential for normal phonation [13–18] as it prevents from vocal folds lesions associated with speech [19, 20]. Nevertheless, few physical studies are realised considering the influence of this liquid on the airflow and the fluid-structure interaction of the vocal folds. Therefore, the objective is to realise a first experimental study to quantify the influence of liquid and to consider its effect on the flow model classically used in physical voice production modelling [1, 2].

¹ArtSpeech ANR-15-CE23-0024

²JSPS summer program fellowship SP18205

1.3 Outline of this thesis

This document is organized as follows:

- Chapter 2: Realizes a brief summary about the phonary system, typical anatomical and fluid mechanical numbers, the measurements used to quantify vocal folds movement and the studied conditions are introduced. Then a description of three types of vocal folds replicas used in this thesis is detailed and a general overview of the experimental set up is provided.
- Chapter 3: Presents the calibration algorithm MSePGG and its validation on mechanical replicas for a non-invasive measurement of the glottal area variation with the external PhotoGlottoGraph device.
- Chapter 4: Exposes the consequences of left-right angular vertical vocal folds asymmetry on the auto-oscillation. This pathology is related to the glottal area as it creates air leakage associated with glottal area enlargement.
- Chapter 5: Shows an experimental study of the influence of the presence of a liquid on the vocal folds. Using different vocal folds replica types, the study allows to show the influence on the airflow as well as on the auto-oscillation of the vocal folds. Water is used as a liquid fluid.
- Chapter 6: Is related to Chapter 5 as it validates a theoretical flow model for gas-liquid fluid mixing. It intends to take into account the presence of liquid on vocal folds in a classical theoretical model for voice production. As in Chapter 5 water is used as a liquid fluid.

The vocal folds apparatus

2.1 Phonation system

The phonation system is the whole set of organs that makes it possible to produce speech. These organs are shared with the respiratory, olfactory and part of the digestive systems. As shown in Figure 2.1, using axes and planes defined in Appendix A, the phonation system is divided into three parts: the thorax (in dark blue) with the lungs and trachea, the larynx (in light blue), and the vocal tract (in green).



Figure 2.1 – Illustration of phonation system.

2.1.1 Trachea and lungs

Located in the chest, both lungs have a combined capacity of 4-5 liters for a male subject [21]. Each lung is connected via a bronchial tube to the trachea (length about 12 cm and diameter 2 cm corresponding to tracheal area A_t =3.1 cm²) which is, in turn connected to the larynx as shown in Figure 2.1(a). The primary function of the lungs is breathing, that allows the blood to be fed

with oxygen during inspiration (contraction of the diaphragm which expands the lungs) and the release of carbon dioxide during expiration (relaxation of the diaphragm that allows the elastic contraction of the lungs and the ejection of air).

The airflow rate is very different during speaking than during normal breathing. While the inhalation-to-expiration ratio in normal breathing is around 40% - 60%, most sounds are produced during exhalation, which brings the ratio to around 10% - 90%. It can go up to 5% - 95% for long conversations [21].

During phonation, the lungs are the pressure reservoir that generates and regulates the airflow: typical volume flow $\phi \approx 2 \cdot 10^{-4}$ m³/s and associated subglottal pressure up to 1 kPa or even 2 kPa for loud speech, compared to 100-200 Pa for breathing.

2.1.2 Larynx

The larynx is located in the throat, between the trachea and the pharynx, coloured in light blue in Figure 2.1. It contains the vocal folds which generates the source of many speech sounds, *i.e.* the so-called voiced sounds produced during phonation [21]. The larynx consists of four cartilages (thyroid, cricoid, arytenoid and epiglottis) connected by ligaments and membranes.

Firstly, the larynx with the epiglottis allows food to be redirected to the oesophagus and stomach, covering the airway system. Secondly, the larynx contains, controls and protects the two vocal folds. The vocal folds are connected together at their anterior end to the thyroid cartilage, responsible of the vocal folds tension, and individually to the posterior end to two individual arytenoid cartilages, responsible of the spacing between the vocal folds called the glottis. The vocal folds structure and phonation process is further detailed in Section 2.2.

During normal breathing, the glottis is widely open (glottal area A_g up to 15 mm² [21]) allowing air to pass freely. During speech, the vocal folds configuration reduces the passage of air in the glottis, which leads to their auto-oscillation, detailed in Section 2.2.2. From this auto-oscillation emerges an acoustic wave that propagates in the vocal tract.

2.1.3 Vocal tract

The vocal tract coloured in green in Figure 2.1, is the superior part of the phonation system. It is located in the head, from the larynx to the lips (and nose), over a length of about 17 cm for a male adult and 14 cm for a female adult [21]. The vocal tract consists of the pharynx, the oral cavity and the nasal cavity. When vocal folds vibrate, the acoustic waves they generate pass through the pharynx and propagate into the cavities, which act as acoustic resonators [22, 23]. While the nasal cavity has a fixed geometry, the oral cavity can be modulated by several

articulatory muscles (tongue, lips, etc.) and by the velum at the posterior extremity of the palate, which borders the oral and nasal cavities, allows to close the nasal cavity.

This modulation of the entire vocal tract allows two things [21]. Firstly, to modify the spectral distribution (resonances) of the energy of the sound wave generated by the vocal folds, constituting the voiced sounds [23]. Modulation of the vocal tract geometry is not considered in this work. Instead the vocal tract is approximated by a uniform circular duct with smooth rigid walls. Secondly, it allows to generate the so-called unvoiced sounds such as fricative consonants [24, 25], which are out of the scope of this thesis.

2.2 Vocal folds structure and oscillation

The vocal folds, by their vibration, are the source of the voiced sound of speech. They are placed in the larynx and measure, about 15 mm for a male adult and about 13 mm for a female adult in the posterior - anterior direction and typically 8 mm for a male and 4 mm for a female adult, in the inferior - superior direction [21].

2.2.1 Structure



(a) In-vivo observation

(b) Simplified illustration

Figure 2.2 – Structure of a human vocal fold: a) realistic view from [26, 27], with intermediate and deep layer forming the vocal ligament and b) simplified illustration.

The vocal folds are composed of three major layers [3, 26]: the epithelium, the lamina propria (superficial layer and ligament) and thyroarytenoid muscle as illustrated in Figure 2.2.

The epithelium measures around 0.1 mm and consists of about twenty layers of cells. Below, the superficial layer of lamina propria, also known as Reinke's space, is composed of proteins,
water and loose elastin and collagen fibers. These two layers form the main vibrating part of the vocal folds. Then the ligament, formed by the intermediate and deep layer of the lamina propria, is made of elastin and collagen fibers, which increases in density and rigidity up to the muscle. The deeper layer of the vocal fold is the thyroarytenoid muscle, which forms the body. Finally, each vocal fold is covered with a thin liquid layer, which is the subject of Section 2.4.2.

Švec *et al.* [28] conducted an *in-vivo* study on the mechanical resonances of healthy human vocal folds. This study allowed to identify the first three mechanical resonances of a vocal fold (here the left one) as illustrated in Figure 2.3:

- $f_1^M = 114$ Hz, with an unidentifiable bandwidth,
- $f_2^M = 171$ Hz, with a bandwith $\Delta f_2^M = 44$ Hz,
- $f_1^M = 241$ Hz, with a bandwith $\Delta f_1^M = 45$ Hz.



Figure 2.3 – Mechanical response of *in-vivo* human vocal folds taken from Švec's thesis [28].

2.2.2 Vocal folds oscillation

The oscillation of the vocal folds that produces the voiced sound source is driven by a fluidstructure interaction between the airflow coming from the lungs and the soft tissues of the vocal folds. This phenomenon can only occur if 1) the tissues on the surface of the vocal folds (epithelium and upper layer of the lamina propria) are sufficiently deformable, 2) the vocal folds are sufficiently close together to create a significant cross-sectional area change at the glottal constriction and 3) the difference between the subglottal pressure P_u (upstream of the vocal folds) and the supraglottic pressure (downstream) is large enough to sustain oscillation.

Threshold pressure P_{On} , or onset pressure, is the minimal upstream pressure P_u required to sustain the auto-oscillation of the vocal folds. It is associated with the offset pressure P_{Off} , which corresponds to the minimal upstream pressure before oscillation ceases. Because of the non-linearity of the fluid-structure interaction a hysteresis appears between the threshold pressures so that $P_{On} > P_{Off}$. This hysteresis is typically about 100 - 200 Pa [29, 30].

As seen in Figure 2.4 illustrating one oscillation cycle, at first, air from the lungs creates a pressure difference between the upstream and downstream end of the vocal folds (grey shades) (1 - 2). This difference allows the vocal folds to detach and open (3 - 4). The airflow rate Φ remains constant throughout the phonation system by the principle of mass conservation along the streamwise direction z. This principle states that $\Phi = A(z) \cdot v(z)$ with area A and flow velocity v. When air passes through the glottal constriction A_g so that $A_g \ll A_t$, the velocity v(z) increases and the pressure P(z) decreases (5). This induces the Bernoulli force resulting in vocal folds closure (6 - 8). This cycle is repeated periodically with fundamental frequency f_0 . This frequency depends on the length and elasticity of the vocal folds and varies between 75 Hz and 250 Hz for adult subjects [29].

During oscillations of the vocal folds, non-linearities can be observed. These can be due to asymmetry between vocal folds or posterior - anterior modes [7, 8, 31]. The non-linearities lead to sub-harmonics frequencies, larynx cartilage oscillation and even to chaos.



Figure 2.4 – Schematical illustration in the medio-frontal plan, of one cycle of vocal folds oscillation. Inspired from [21, 28].

As upstream pressure P_u drives vocal folds oscillation, its analysis is essential. The same way the dimension of the glottal constriction and in particular its area A_g determines the pressure forces and therefore its accurate quantification is an ongoing issue.

2.3 Quantification: physical numbers and perturbation

Auto-oscillation can be characterised by non-dimensional numbers derived from physical quantities associated with the underlying fluid-structure interaction. Moreover perturbation of oscillation cycles can occur, for instance due to non-linearities, which is quantified by parameters measuring voice quality. Characteristic non-dimensional numbers (Reynolds, Strouhal and Mach) and perturbation measures (jitter, shimmer, signal-to-noise ratio and total harmonic distortion rate) are defined in the following and typical values observed on human speakers are reported in Table 2.5 where also typical values for vocal folds replicas, discussed later on, are summarised.

2.3.1 Mach number

The Mach number Ma, determines the compressibility of the flow in the glottis. It expresses the ratio between the local velocity of a fluid (here air) and the velocity of sound in the same fluid:

$$Ma = \frac{v}{c} \tag{2.1}$$

where v is the airflow velocity, c is the sound propagation velocity.

For $Ma^2 \ll 1$, the flow is considered as incompressible [32]. For the glottic airflow $v_g \in [10-40]$ m/s [33] and all the vocal folds replicas used, $Ma^2 \approx O(10^{-2})$, which motivates the incompressibility hypothesis for glottal flow.

2.3.2 Reynolds number

The Reynolds number, *Re*, represents the ratio between the inertial forces and the viscous forces of the fluid. It is used to determine if the viscosity of the fluid is negligible in the flow. It also determines the flow regime, laminar or turbulent:

$$Re = \frac{\rho v h}{\mu} \tag{2.2}$$

where ρ is the density of the fluid, v is the velocity of the flow, h a typical flow channel dimension, and μ is the kinematic viscosity coefficient of the fluid, here air.

For low Reynolds values (Re < 2000), the flow can be considered as laminar [32]. For high Reynolds values (Re > 3000) the airflow can be considered as turbulent. For a maximum glottal velocity $v_g = 40$ m/s and using a glottal width $h_g = 1$ mm in the right-left direction, we obtain Re

< 2000. This is characteristic of a laminar flow for which the effects of viscosity are lower than inertia.

2.3.3 Strouhal number

The Strouhal number *Sr* estimates the importance of the unsteadiness of the flow in relation to its inertia. It represents the ratio between the flow velocity and the oscillation velocity of the vocal folds:

$$Sr = \frac{f_0 L_z}{v} \tag{2.3}$$

where f_0 is the oscillation frequency of the vocal folds, L_z the vocal fold length in inferiorsuperior direction and v the velocity of the flow.

For $Sr \ll 1$ the unsteady effects of the vocal folds oscillations are considered as negligible on the airflow [32]. For human and for the vocal folds replicas used in this work, $Sr \approx O(10^{-2})$, as shown in Table 2.5. The flow in the glottis can therefore be considered as quasi-steady during oscillation.

2.3.4 Cycle-to-cycle perturbation: jitter and shimmer

Jitter and shimmer are parameters used to quantify the stability of a periodic signal. In the following chapters, the analysis will be realised for i = 1...N with N = 50 periods.

The jitter analyses the relative cycle-to-cycle variation of the fundamental frequency. It averages the absolute time differences between each consecutive period T divided by the average period of the signal:

$$\zeta_T = \frac{\frac{1}{N-1} \sum_{i=1}^{N-1} |T_i - T_{i+1}|}{\frac{1}{N} \sum_{i=1}^{N} T_i}.$$
(2.4)

The shimmer analyses the relative cycle-to-cycle variation of the signal amplitude \mathscr{A} . It averages the absolute differences of the amplitude of each consecutive period, divided by the average amplitude of the signal:

$$\zeta_{\mathscr{A}} = \frac{\frac{1}{N-1} \sum_{i=1}^{N-1} |\mathscr{A}_i - \mathscr{A}_{i+1}|}{\frac{1}{N} \sum_{i=1}^{N} \mathscr{A}_i}.$$
(2.5)

Consequently, a signal with a stable fundamental frequency will have a low jitter. And a signal with a stable amplitude will have a low shimmer. Typical values for normal speech yield $\zeta_{\mathcal{T}} \approx 0.3\%$ and $\zeta_{\mathcal{A}} \approx 2.7\%$ as reported in Table 2.5 [34].

2.3.5 Signal-to-noise ratio

Other common voice parameters are derived from a harmonic decomposition of the signal under analysis. Signal-to-noise ratio *SNR* (in dB) is computed as the ratio of the summed power of all signal harmonics \mathcal{P}_{signal} based on the first harmonic frequency f_N to the summed power of the remaining noise \mathcal{P}_{noise} :

$$SNR = 10\log_{10}\left(\frac{\mathscr{P}_{signal}}{\mathscr{P}_{noise}}\right).$$
 (2.6)

Consequently, the signal quality is said to improve for increasing SNR since it indicates less noise. Typical values for normal speech yield SNR ≈ 24 dB as reported in Table 2.5 [34].

2.3.6 Total harmonic distortion rate

Total harmonic distortion rate THD (in dB) is computed from the ratio between the summed power of higher harmonic frequencies¹ \mathscr{P}_{harm} and the power of the first harmonic frequency \mathscr{P}_{f_N} :

$$THD = 10\log_{10}\left(\frac{\mathscr{P}_{harm}}{\mathscr{P}_{f_N}}\right).$$
(2.7)

Consequently, the signal quality is said to improve for decreasing THD since it indicates less harmonic distortion. Typical values for normal speech yield THD ≈ -9.9 dB as reported in Table 2.5 [35].

¹Theoretically an infinite number of harmonics is accounted for, but in practice 50 harmonics are used.

2.4 Studied vocal folds conditions

In this work, two particular conditions are studied, which from a physical point of view, are rarely considered in literature. Firstly, the effect of the left-right vertical angular asymmetry in the vocal folds positioning on the auto-oscillation is studied. This condition is relevant to unilateral vocal fold paralysis. Secondly, the presence of liquid (mucus) on the vocal folds surface is considered.

2.4.1 Vertical angular asymmetry



Figure 2.5 – Videostroboscopic image of UVPF. Adapted from Rosen and Simpson.

Vocal folds asymmetry is reported as one of the main causes of glottic insufficiency resulting in dysphonia [3]. Glottic insufficiency is often caused by complete or partial vocal fold paralysis such as unilateral vocal fold paralysis (UVFP). Unilateral vocal fold paralysis is due to a dysfunction of one or several laryngeal nerves and results from a myriad of causes, in 70 % the origin is from malignant tumours, disease or iatrogenic (side effect) post-operative trauma (nerve cut, etc.). unilateral vocal fold paralysis is characterized by vocal fatigue and a breathy voice [3].

Clinical examinations show that unilateral vocal fold paralysis induces left-right vocal folds asymmetries to the shape, tension and positioning. Videostroboscopic imaging (Figure 2.5) reveals that asymmetrical positioning of the vocal folds in the vertical sagittal plane prevents complete glottal closure [3, 36, 37], which is suggested to be the main physical cause of the observed breathy voice. The vocal folds positioning asymmetry results from both the normal and paralysed vocal folds. In order to compensate the reduced movement of the paralysed vocal fold, the normal vocal fold changes its position in the inferior-superior direction. Consequently, a vertical level difference between the left and right vocal fold in Figure 2.5) and the normal vocal fold remains in the transverse plane (left vocal fold in Figure 2.5). This asymmetry changes the initial configuration of the glottis.

Quantitative characterization of this spatial asymmetry from imaging on human subjects is extremely tedious [3, 36, 37]. Nevertheless, the vertical level difference during phonation is reported [37] to yield several millimetres, *i.e.* 1.3 ± 1.5 mm or up to 3 mm. The effect of positioning asymmetry on *in-vivo* vocalization can not be studied independently from tension or shape asymmetry between left and right vocal folds. In that sense, physical studies of auto-oscillation using mechanical replicas are of interest, when aiming a systematic study of the influence of vocal folds positioning asymmetry.

2.4.2 Mucus presence

The surface of human vocal folds, illustrated in Figure 2.2(b), is covered with a thin liquid layer [19, 38].

It is established from observations on human speakers and cadaveric larynges that good surface hydration is essential for phonation. This depends on a combination of environmental factors [13], vocal efforts [14, 15], lifestyle [16, 17] and general health status [18]. One of the physical consequences attributed to surface hydration is to reduce phonotraumatic lesions [19, 20] and to reduce the minimum subglottic pressure required for vocal folds self-oscillation (threshold onset pressure P_{On}) [20, 39–41]. Other self-oscillation properties of the voice are sensitive to hydration such as the fundamental frequency f_0 , the closing quotient, the velocity quotient and cycle-to-cycle disturbances jitter and shimmer [38, 42–44].

Artificial saliva sprays are developed to remedy lack of oral hydration including at the vocal folds surface hydration [13, 45, 46]. Liquid sprays and nebulization techniques are common in (re)hydration studies on human speakers or excised larynges focusing on vocal quality properties listed in the previous paragraph [42, 43, 46]. In this work, distilled water and two artificial salivas (liquid properties determined in Appendix G) are used to study the effect of surface hydration experimentally. These fluids are used in order to mimic the mucus presence on the vocal folds. As a start, distilled water is chosen as a reference since fluid properties are well known. Values for density and viscosity at room temperature are given in Table 2.1. From this table is seen that the water density is similar to the density of vocal folds mucus and of human saliva which yields about 1000 kg/m³ [38, 45]. However, values reported for viscosity (dynamic μ and kinematic v) of distilled water are ten times lower than values observed for mucus or saliva [38, 45]. Therefore, it is proposed to use two artificial saliva products as well. The first artificial saliva is called Artisial (available in French pharmacies) and the second one is called Teijin (available in Japanese pharmacies). These sprays are thus commonly used in the event of a decrease or absence of salivary secretions. Their compositions are detailed in Appendix G. The use of these sprays has the advantage that their densities and viscosities are of the same order of magnitude as human mucus or saliva. The viscosity of Artisial is roughly in the middle of the range observed for human saliva and the viscosity of Teijin is situated at the lower end of human saliva range. In addition it is interesting to note that the use of the three fluids allows to study the influence of viscosity.

	Density ρ [kg.m ⁻³]	Kinematic viscosity v [m ² .s ⁻¹]	Dynamic viscosity μ [kg.m ⁻¹ .s ⁻¹]
Human saliva [45]	~ 1000	\sim [2.75 - 15.51] $\times 10^{-6}$	[2.75 - 15.51]×10 ⁻³
Mucus [38]	~ 1000	\sim [1 - 10] $ imes$ 10 ⁻⁶	$[1 - 10] \times 10^{-3}$
Artisial saliva	1015	7.3×10^{-6}	7.4×10^{-3}
Teijin saliva	1014	4.9×10^{-6}	4.9×10^{-3}
Water	1000	1.00×10^{-6}	1.0×10^{-3}
Air	1.20	1.57×10^{-5}	1.8×10^{-5}

Table 2.1 – Summarised values of density and viscosity measured in Apppendix G.

2.5 Vocal folds replicas

All experimental studies on vocal folds oscillation presented in this work, either to reproduce healthy or pathological conditions, rely on the use of mechanical vocal folds replicas. This way physical parameters potentially influencing vocal folds oscillation can be controlled and the repeatability of experimental results can be verified. Three types of vocal folds replica are assessed. The first replica represents rigid vocal folds (Section 2.5.1) in order to study the influence of a forced oscillation on the airflow. The next two replica types are deformable. Made of pressurised latex (Section 2.5.2) and silicone layers (Section 2.5.3) respectively, they allow the study of the full fluid-structure interaction.

The use of both deformable replica types of the vocal folds structure are of interest to study the fluid-structure interaction because:

- Pressurized latex allows to mimic the variation of vocal folds elasticity (associated to the stiffness or tension). Furthermore, the design of the replica allows to vary the initial glottal area,
- Silicone replicas aim to reproduce the multi-layered physiological vocal fold structure with increasing degree of complexity. The initial glottal area can be varied.

2.5.1 Motion driven rigid: MDR

The motion driven rigid (MDR) vocal folds replica shown in Figure 2.6 is 1:3 up scaled from human vocal folds and consists of two rigid parallel half cylinders (duralium, neutral wettability, see Section G.4), with a radius of 10 mm, an inferior-superior length $L_z = 20$ mm and posterior - anterior width w = 30 mm illustrated in Figure 2.6(d) and summarized in Table 2.4. One cylinder is fixed and the other one is forced into motion by an eccentric motor indicated in Figure 2.6(a) [47, 48]. Sinusoidal movement of the rectangular gap between the two half-cylinders shown in Figure 2.6(b), is imposed with a forcing frequency f_c between 0 Hz and 12 Hz and a constant forcing amplitude between 0 mm (complete closure) up to 1.9 mm.

The minimum imposed time-varying spacing (constriction) $h_c(t)$ ($0 \le h_c \le 1.9$ mm) is assessed using an optical sensor (OPB700, accuracy ± 0.01 mm, calibration shown in Appendix B.2), so that the time-varying constriction area is obtained as $A_c(t) = h_c(t) \cdot w$ with $A_c(t) \le 57$ mm². This maximum value of the constriction area is indicated as $A_{c,max} = 57$ mm² in Table 2.4. A pressure tap is provided in the middle of the fixed cylinder, seen in Figure 2.6(d), to measure the pressure at the constriction level $P_c(t)$ (pressure sensor Kulite XCS093, accuracy ± 5 Pa, calibration given in Appendix B.1).

2.5.2 Deformable pressurized latex tube: PLT

The deformable pressurized latex tube (PLT) replica (up scaled 1:3), shown in Figure 2.7, consists of two latex tubes (neutral wettability, see Section G.4) with diameter 11 mm, thickness 0.2 mm and length 25 mm representing the vocal folds. These tubes envelop a metallic structure and are pressurized (P_{PLT}) by filling it with distilled water. The distilled water is supplied by a water column whose height and hence internal water pressure P_{PLT} can be controlled. The resulting vocal folds replica (inferior-superior length $L_z = 12$ mm, posterior - anterior width w = 25 mm, mass $m_d = 2.1$ g in Table 2.4) is illustrated in Figure 2.7 and detailed in [4, 49]. Both vocal folds are placed face to face in a metallic frame. A central latex tube (diameter 30 mm, thickness 0.2 mm) is placed between the vocal folds and is fixed downstream and upstream of the constriction so that the airflow passes through without leakage. This central latex tube is further connected to an upstream and downstream rigid uniform tube (diameter 25 mm) representing the vocal tract and trachea respectively.

The position of each vocal fold relative to the axis of the central tube and therefore the spacing between the two vocal folds can be modified using four micrometric screws (Mitutoyo 153-101, accuracy 0.01 mm, range 5 mm), which value is denoted o_{mm} . When airflow passes through the gap between both tubes a fluid-structure interaction leads to self-oscillation, in the same way as during human phonation. The time-variation of the glottal-like constriction is observed using a high-speed camera (Motion BLITZ Eosens Cube 7, frame rate 525 and shutter 750 μ s). The



Figure 2.6 – Motion driven rigid (MDR) replica: a) side and b) top view, c) general view in medio-frontal view and d) schema of rigid vocal fold half cylinder (fixed with pressure tap and mobile without).

constriction area $A_c(t)$ (accuracy $\pm 0.002 \text{ mm}^2$) can be extracted from the gathered images as detailed in [49].

It should be noted that the initial constriction area $A_0 = A_c(t = 0)$ is the combined result of the pressure P_{PLT} imposed and the position of the micrometric screws o_{mm} . Therefore, initial conditions of the PLT replica are indicated as (P_{PLT}, o_{mm}) . Typical values of the initial glottal area A_0 varies from 18 mm² to 74 mm² and $A_{c,max} = 100$ mm² as indicated in the Table 2.4.

As P_{PLT} can be varied, the elasticity of each tube can be modified. The elasticity determines mechanical resonance properties: peak frequencies f^M , bandwidths Δf^M and resulting quality factors $Q = \frac{f^M}{\Delta f^M}$. These properties are obtained from frequency response curve detailed in Appendix C. An example is presented in Figure 2.8 for $P_{PLT} = 3800$ Pa and extracted properties for $P_{PLT} \in [1500 - 4500]$ Pa (triples the minimum P_{PLT}) are summarised in Table 2.4. Note that f_1^M almost doubles for 60 Hz to 110 Hz which is in the range of human vocal folds from 62 Hz up to 114 Hz.

Figure 2.9 shows the oscillation threshold pressure P_{On} and P_{Off} (Figure 2.9(a)) and associ-



(a) Schematic illustration (frontal plan)



(b) Top view (transverse plan



(c) Front view (frontal plan)

(d) Side view (sagittal plan)

Figure 2.7 – Illustration of the deformable pressurized latex tube (PLT) replica.



Figure 2.8 – Typical amplitude curve of mechanical response for a latex vocal fold tube ($P_{PLT} = 3800 \text{ Pa}$) following the procedure outlined in Appendix C.

ated oscillation frequencies $f_{0,On}$ and $f_{0,Off}$ (Figure 2.9(b)) as a function of P_{PLT} . As a reference, measured mechanical frequencies $f_1^M(P_{PLT})$ are also indicated in Figure 2.9(b). It is seen that oscillation frequencies $f_{0,On} \approx f_{0,Off}$ increase quasi-linearly with the internal pressure P_{PLT} which is also the case for mechanical frequencies f_1^M . Note that for all P_{PLT} the oscillations frequency f_0 is just above the mechanical frequency f_1^M and in the range characterising male voice.

Pressure curves $P_{On}(P_{PLT})$ and $P_{Off}(P_{PLT})$ show a typical parabolic shape exhibiting a local minimum [1, 2]; in this case yielding $P_{On} = 265$ Pa for $P_{PLT} = 2700$ Pa. Note that this kind of curve shape with a local minimum is also observed on experiments with excised larynges when tension is varied by elongation of vocal folds [29]. As for human vocal folds an hysteresis between P_{On} and P_{Off} of about 100 Pa is observed.

2.5.3 Deformable silicone replicas: M5, MRI and EPI

Silicone replicas approximate the multi-layer human vocal folds structure, shown in Figure 2.2, with different degree of complexity. Vocal folds replicas consisting of an overlay of two, three or four silicone molding layers are considered. These three deformable silicone vocal folds replicas (moderately non-water wettability, see Section G.4), labelled M5, MRI and EPI, are illustrated in Figure 2.10. They are obtained following the procedures proposed in [5, 50, 51] and detailed in Appendix D. In the following, firstly different molding layers are briefly detailed, next the replica compositions are presented and finally oscillation is discussed.

Different molding layers:

Each layer, illustrated in Figure 2.10, has a different thickness l_d , and different Ecoflex(A & B)-to-Silicone (ES) or Dragonskin(A & B)-to-Silicone (DS) mixing ratio \mathcal{M} . \mathcal{M} is given as (ES,DS $x_{E,D}$: y_S) with $x_{E,D}$ the quantity of Ecoflex/Dragonskin A and B compared to y_S the



(b) Auto-oscillation and mechanical frequencies

Figure 2.9 – Typical PLT replica's curves a) threshold pressures and b) oscillation f_0 and mechanical f_1^M frequencies as a function of internal pressure P_{PLT} for micrometric screws position $o_{mm} = 1$ mm.

quantity of silicone. Young moduli \mathscr{E} of all layers are estimated using a tensile test as detailed in Appendix E. From Table 2.2 is seen that estimated moduli have the same order of magnitude as reported in literature [51–56]. Moreover, comparison with values estimated for human vocal folds, given in Table 2.3, show that the Young moduli of molding layers approximate \mathscr{E} values for the thyroarytenoid (vocalis) muscle, the superficial layer of the lamina propria (Reinke's space) and epithelium layer. On the other hand, the Young modulus of the ligament (intermediate and deep layers of the lamina propria) is under estimated since no tension is applied in the deep layer.

Each vocal fold is mounted on a backing layer (ratio \mathcal{M} DS 1:1, thickness 4 mm) in order to attach the replica to its rigid support, *i.e.* region outside of the black frames in Figure 2.10(b).

Silicone replicas composition:

Different multi-layer compositions and geometrical casts are used to mold the M5, MRI and EPI vocal folds replicas, illustrated in Figure 2.10.



(d) Description of the different layers

Figure 2.10 – Silicone vocal folds replicas M5, MRI and EPI in a) the medio-frontal plane b) the transverse plan, c) and d) detailed layers in the medio-frontal plane.

The M5 vocal folds replica, illustrated on the left in Figure 2.10, is a two-layer model reproducing the muscle (l_d =6.4 mm) and the superficial layer (l_d =1.5 mm) (see Table 2.3 and Figure 2.10(c) and 2.10(d)) which is cast using the geometrical so called M5 model [60].

	DS 1:1	ES 1:1	ES 1:2	ES 1:4	ES 1:8
Drechsel et al. (2008) [57]	22	_	4.1	—	_
Riede et al. (2008) [58]	—	—	~ 10	—	—
Pickup et al. (2010) [50]	_	_	8.5	2.9	—
Murray et al. (2010) [59]	~ 21	~ 2.3	~ 1.8	~ 0.3	~ 0.08
Murray et al. (2011) [51]	50	12	—	1.6	0.2
Current study	52.0	21.9	10.4	4.9	0.2*

Table 2.2 – Summary of moduli of elasticity \mathscr{E} [kPa] for different molding layer (mixing ratio \mathscr{M}) from literature and estimated in the current study (Appendix E).

*Value for ES 1:8 is taken from [59]

The geometrical cast used to obtain the MRI vocal folds replica, illustrated in the middle in Figure 2.10, is more complex and realistic as the cast is derived from magnetic resonance imaging data of a human vocal folds [5, 50]. The realistic vocal folds geometry is characterised by a triangular shape (sharp angle 60°) in the transverse plane (top view in Figure 2.10(b)) compared to the rectangular shaped M5 geometrical model. In addition, the MRI vocal folds replica has a three-layer structure: the muscle (l_d =10 mm), the superficial layer (l_d =3.0 mm) (see Table 2.3) and adding a third thin and stiff surface layer representing the epithelium (l_d =0.1 mm) to the two-layer structure of the M5 vocal folds replica. Note that compared to MRI vocal folds replicas presented in literature [5, 50] an epithelium layer is added in this work.

Finally, a four-layered EPI vocal folds replica, illustrated on the right in Figure 2.10, is obtained by inserting an extremely soft deep layer between the muscle (l_d =6.4 mm) and superficial layer (l_d =1.0 mm) of the three-layer structure used for the MRI vocal folds replica [51] (see Table 2.3 and Figure 2.10). The EPI vocal folds replica cast is inspired on the geometrical M5 model so that its geometry is a scaled version of the M5 vocal folds replica.

Silicone replica oscillation:

Mechanical response of the silicone vocal folds replicas is measured following the same procedure as for the PLT replica. The procedure and detailed results are given in Appendix C. Examples of frequency response curve is plotted in Figure 2.11. An overview of the first mechanical oscillation properties (frequency f_1^M , bandwidth Δf_1^M and quality factor Q_1^M) is summarized in Table 2.4, as well as their dimensions and masses. As for the PLT vocal folds replica, when airflow passes through the vocal folds gap, fluid-structure interaction leads to self-oscillation. Note that in contrast to what is observed for the PLT replica, f_1^M is higher than oscillation frequency f_0 , but the magnitude of the difference is the same (about 40 Hz).

2.6. General set-up

Table 2.3 – Layer properties in human and in silicone vocal folds replicas (M5^{*a*}, MRI^{*b*}, EPI^{*c*}): measured Young modulus \mathscr{E} , ratio \mathscr{M} , layer thickness l_d .

	Male adult		Silicone vocal folds replica		
Layer	€ [kPa]	l_d [mm]	\mathscr{E}^* [kPa]	M [-]	l_d [mm]
Muscle	8-29	6.0	$10.4^a \ 4.9^b \ 21.9^c$	ES 1:2 ^{<i>a</i>} ES 1:4 ^{<i>b</i>} ES 1:1 ^{<i>c</i>}	$6.4^a \ 10.0^b \ 6.4^c$
Superficial	2-9	0.6	$4.9^a \ 0.2^{b,c}$	ES 1:4 ^{<i>a</i>} ES 1:8 ^{<i>b</i>,<i>c</i>}	$1.5^a \ 3.0^b \ 1.0^c$
Ligament	10-45	0.8	4.9 ^c	ES 1:4	1.0^{c}
Epithelium	40-60	0.1	$52.0^{b,c}$	DS $1:1^{b,c}$	$0.1^b \ 0.1^c$

^{*a*} M5: muscle and superficial

^b MRI: muscle, superficial and epithelium

^c EPI: muscle, superficial, ligament and epithelium

ES: Ecoflex/Silicone mixure

DS: Dragonskin/Silicone mixture

*: values taken from the current study Table 2.2



Figure 2.11 – Typical amplitude curve of mechanical response for silicone vocal folds replicas, M5, MRI and EPI measured in Appendix C.

2.6 General set-up

The overall set-up is schematically illustrated in Figure 2.12(a) and illustrated with the PLT replica in Figure 2.12(b). Although the elements presented in this general set-up are used throughout this work, their particular features (air supply system, tube length, etc.). In the following the most used set-up is detailed. If elements differ, this will be explicitly mentioned and clarified.

Each vocal folds replica is inserted in an experimental set-up so that glottal-like area $A_c(t)$, *i.e.* the space between right and left vocal folds, is situated in a transverse horizontal plane as for a standing human speaker which is an improvement compared to previous work [1, 2].

	$\begin{vmatrix} L_z^{\star} \\ [mm] \end{vmatrix}$	w [†] [mm]	$A_{c,max}^{\ddagger}$ [mm ²]	$\begin{matrix} A_0^{\triangleleft} \\ [\mathrm{mm}^2] \end{matrix}$	m_d^{\Box} [g]	<i>f</i> 0 [Hz]	$\begin{vmatrix} f_1^M \\ [Hz] \end{vmatrix}$	Δf_1^M [Hz]	$egin{array}{c} Q_1^M \ [-] \end{array}$
MDR	20.0	30.0	60	_	_	1 - 12	—	_	_
PLT	12.0	25.0	100	18-74	2.1	80 - 150	60 - 110	10±3	6 - 20
M5	10.7	17.0	85	<10	1.7	~ 100	152	16	9.5
MRI	10.0	18.0	85	<10	2.1	~ 115	143	10	14.3
EPI	10.2	17.0	85	<10	1.7	~ 85	145	11	13.2
Adult	4 - 8	13 – 25	50	<20	_	75 - 250	62 - 114	[-]	[-]

Table 2.4 – Vocal folds replicas and human [30, 52, 61].

* L_z : inferior-superior vocal folds length

[†] w: posterior - anterior vocal folds width

[‡] upper limit for glottal area $\max(A_c(t)) \leq A_{c,max}$)

 $A_0 = A_c(t=0)$: initial glottal area

 \square *m_d*: vocal fold mass

Table 2.5 – Typical vocal folds physical and perturbation values.

	Re	Sr	P_{On}	P _{Off}	SNR		ζ_T	$\zeta_{\mathscr{A}}$
	[-]	[-]	[Pa]	[Pa]	[[aB]	[ab]	[%]	[%]
MDR	$\propto O(10^3)$	$< 5 \times 10^{-2}$	_	_	20	0	0	0
PLT	$< 9 \times 10^{3}$	$< 5 \times 10^{-2}$	250 - 350	250 - 270	25	-10	1.5 - 3	2.2 - 4.6
M5	$< 20 \times 10^{3}$	$< 2 \times 10^{-2}$	1100	890	40	-20	0.2	0.3
MRI	$< 19 \times 10^{3}$	$< 2 \times 10^{-2}$	980	675	40	-20	0.7	1
EPI	$< 11 \times 10^{3}$	$< 3 \times 10^{-2}$	440	325	40	-10	0.6	0.3
Adult*	700-2000	$\propto O(10^{-2})$	300 - 1000	100 - 500	24	-9.9	0.3	2.7

*: Values for human adult from [2, 30, 34, 35, 62]

A rigid uniform circular channel with diameter 2.5 cm is mounted vertically to the upstream(12 cm trachea inferior to vocal folds, quarter wave resonance 660 Hz) and downstream² (11 cm vocal tract superior to vocal folds half wave resonance 1500 Hz) end of the vocal folds replica.

Continuous steady airflow (density $\rho_G = 1.2 \text{ kg} \cdot \text{m}^{-3}$, dynamic viscosity $\mu_G = 1.8 \times 10^{-5} \text{ Pa} \cdot \text{s}$, temperature $22 \pm 2^{\circ}\text{C}$) is provided along the *z*-direction by a manual valve followed by a pressure regulator (Norgren, 11-818-987) controlled air supply. Air is delivered by an air compressor (Atlas Copco GA5 FF-300-8, GA15 FF-8) connected to an upstream pressure reser-

 $^{^{2}}P_{d} \approx 0$, so the driven pressure difference is approximated as $\Delta P = P_{u}$

voir (volume $\ge 0.22 \text{ m}^3$) filled with acoustic foam in order to avoid parasite acoustic resonances. The air supply chain is illustrated for the PLT replica in Figure 2.12(b).

A pressure transducer (Endevco 8507C-5, accuracy ± 5 Pa, calibrated in Appendix B.1) is positioned in a pressure tap 65 mm upstream of the vocal folds replicas in order to measure upstream pressure P_u . In the case of a motion driven rigid vocal folds replica an additional pressure sensor (Kulite XCS-093, accuracy ± 5 Pa, calibrated in Appendix B.1) is placed at the glottal spacing so that constriction pressure P_c is measured as well. P_c could not be measured for the deformable vocal folds replicas (PLT and silicones). All signals are acquired with a card National Instruments PCI-6225 combined with a National Instruments BNC-2110 input/output card and processed with the Labview software programs with a sample frequency $f_s=10$ kHz. The high speed camera (Motion BLITZ Eosens Cube 7) is placed so that the vocal folds constriction can be observed.

Table 2.6 – Vocal folds replicas used and main quantities analysed in the next chapters (\checkmark).

	MDR replica	PLT replica	Silicone replicas	$P_{(u)}$	A_c
Chapter 3	\checkmark	\checkmark	_		\checkmark
Chapter 4	—	—	\checkmark	\checkmark	\checkmark
Chapter 5	\checkmark	\checkmark	\checkmark	\checkmark	_
Chapter 6	\checkmark	—	_	\checkmark	—

2.7 Summary

In this chapter the human vocal folds apparatus is characterised and simplifications are introduced which will allow a systematic physical study of vocal folds oscillations in the next chapters. Different simplifications let to three types of mechanical vocal folds replicas: MDR, PLT and silicones.

In order to study the fluid-structure interaction, an overview for human vocal folds and simplifications is given concerning:

- Fluid properties in Table 2.1,
- Structural properties including geometrical features of the vocal folds as well as the glottis, mechanical resonances and typical oscillation frequencies are given in Table 2.4,
- A characterisation of the oscillation in terms of physical numbers, threshold pressures and perturbation measures related to voice quality is provided in Table 2.5. Results for replicas



(b) Illustration with PLT replica

Figure 2.12 – General set-up.

are obtained for the interaction between airflow and vocal folds structure and are useful as a reference for the next chapters. Values for a human subject correspond for normal phonation whereas values for replicas are obtained in absence of mentioned conditions: without hydration and without vertical angular asymmetry.

It was shown that the upstream driving pressure P_u is crucial for oscillation as is the glottal-

2.7. Summary

like constriction area $A_{g,c}$. Quantification and characterisation of these two quantities is therefore assessed in the following chapters for different conditions as indicated in Table 2.6. The table also indicates which replicas are used in each chapter:

- In Chapter 3, the focus is on the calibration of the ePGG glottal area measurement device. In order to obtain a quantitative validation of the proposed calibration algorithm replicas allowing an accurate measurement of the constriction area are used, *i.e.* the MDR replica with an optical sensor and the PLT replica with the high speed camera.
- In Chapter 4, the left-right vocal folds vertical angular asymmetry is systematically studied. As only the effect of geometrical changes on auto-oscillation features is of interest, only silicone vocal folds replicas are used as the vertical asymmetry can be imposed without affecting structural properties.
- In Chapter 5, the influence of surface hydration following water spraying on oscillation is studied. In order to obtain general results all replicas are used. Using both the deformable silicone replicas and the deformable PLT replica allows to asses a wide region in the parameter space spanned between driving pressure and auto-oscillation frequency. In addition, the rigid forced oscillation replica MDR allows to study the effect of water spraying on the flow in more detail.
- In Chapter 6, only data obtained for water spraying on the MDR replica are considered in order to validate a theoretical flow model accounting for the presence of air and water following hydration with a spray.

CHAPTER 3

External lighting and sensing photoglottography: characterization and MSePGG algorithm

Based on Bouvet A., Amelot A., Pelorson X., Maeda S., Van Hirtum A., 2019. *External lighting and sensing photoglottography: characterization and MSePGG algorithm*. Biomedical Signal Processing and Control, 51:318-327.

Experiments involving the human speaker were performed at the Phonology Phonetics Laboratory, Paris, France.

Specific nomenclature

PGG	PhotoGlottoGraphy	
ePGG	external PhotoGlottoGraph	
MSePGG	Multi Signal external PhotoGlottoGraph	
IR	InfraRed light	
S	Sensor	
d	Source - Sensor distance	m
d_k	Sensor position	m
A_g	Glottal area (human and replicas)	m^2
A_u	Internal area of plexiglass tube	m^2
U_{ePGG}	ePGG sensor voltage	V
Ι	Light intensity	cd
Φ	Ligth flux	lm
η	Light signal of $A_c=0$	lm

3.1 Introduction

Observation, and further measurement, of the glottal area between the moving vocal folds during breathing, speech production or swallowing, has been for long a major technological challenge. Since Garcia's pioneer experiments using mirrors [63], several different techniques have been developed and optimized. Video recordings using an endoscope coupled with a stroboscopic light or a high speed camera has become a very popular technique despite of the costs of the equipment and the need for extensive post-processing in the case of high-speed recordings [64–68]. This technique might cause discomfort and is invasive due to the insertion of optic devices (through the oral cavity in the case of a rigid endoscope, or through the nasal cavity in the case of a flexible endoscope), and thus a medical environment is required and pronunciation of certain phonemes can be hindered or inhibited. Quantitative area extraction from endoscopic images remains challenging even when stereo-endoscopy or additional devices are used [65, 69–71], also being due to the trade-off between spatial and temporal resolution for image acquisition among others. Non-invasive alternatives are very few [72]. Ultrasound techniques have been tested but lack of spatial resolution [73–75], and therefore ultrasound-based imaging is mostly used for innocuous visualization only [76–80].



Figure 3.1 – External photoglottograph (ePGG) principle (mid-coronal plane) of glottal transillumination: orientation angle of light source (IR), distance from source d, light sensor (S) at position d_k , trachea area A_t , glottal area $A_g(t)$ and light flux $\Phi(d)$.

PhotoGlottoGraphy (PGG) [72, 81] consists of devices for illuminating the glottis and measuring the amount of light that passes between the vocal folds. In its original development, PGG is an invasive technique as it requires the insertion of a light source or of a light sensor through the oral or nasal cavity. In contrast, External PhotoGlottoGraphy (ePGG) [10–12] is a non-invasive technique, both light source and sensor being placed outside of the vocal tract on the exterior of the neck as shown in Figure 3.1. Another difference with the classical PGG is the use of a lightning in the near infrared (IR) instead of visible light. Indeed, IR wavelengths in this range 700-1000 nm are reported to transilluminate large sections of human tissue [82–84]. Given the non-invasive nature of ePGG, this system no longer requires a medical environment and allows to make continuous measurements with as less disturbance as possible, *e.g.* during speech production. Consequently, if a relationship between measured ePGG signals and glottal area variation $A_g(t)^1$ can be established, ePGG is suitable to observe variations of the glottal area non-invasively and continuously regardless of location (medical practice, laboratory, field...) which makes it an interesting technique for many disciplines. Therefore, the aim of this work is to investigate and formalize the relationship between ePGG measurements and glottal area.

3.2 ePGG measurement system

The ePGG system [10–12], shown in Figure 3.1, consists of two main elements: a light source (infrared LED, LSF812N1, wavelength 810 nm, size ≤ 5 mm, beam angle $45 \pm 5^{\circ}$) and a single light sensor (photo-diode, Vishay Semiconductors BP104, peak sensitivity at wavelength 950 nm, size ≤ 3 mm) placed in a holder. Electrical ePGG signals (between 0 V and 5 V) are acquired using a data acquisition card (Data Translation, 16 bit) and software (QuickDaq 7.8.10). In addition, the ePGG signal is amplified linearly prior to acquisition in order to compensate for *e.g.* inter-subject differences affecting light absorption (tissue, skin, *etc.*). Spurious light sources (ambient light, screen, *etc.*) are dimmed during acquisition to ensure the ePGG signal quality. In addition, both light source (IR) and sensor (S) are shielded once their position is fixed.

The light source (IR) is positioned at different supraglottal positions to illuminate the glottis through the surface of the front neck and a sensor (S) is placed at a subglottal position to record the light variation due to vocal folds displacement that modulates glottal area A_g as schematized in Figure 3.1. Concretely for a human subject, in this work, the sensor (S) is fixed in the midsagittal plane above the suprasternal notch and the light source (IR) is positioned somewhere along the mid-coronal plane as shown in Figure 3.2. Distance *d* (Figure 3.1) indicates then the distance between transverse (horizontal) planes containing the source and sensor, respectively. Furthermore, the orientation of the IR source is varied by turning its holder in the mid-coronal plane as depicted in Figure 3.1. As a convention, light emitted towards the sensor (following axis *d*) yields orientation angle 90° and light emitted parallel to the transverse plane (perpendicular to *d*) yields orientation angle 0°. The ePGG signal is sampled at 20 kHz so that the temporal resolution of 0.05 ms is excellent given the frequency range of interest, *i.e.* typical ≤ 250 Hz during phonation (Table 3.1). This temporal resolution is also high in comparison with other

¹In this chapter, the interest is to measure the glottal aperture. For this purpose the term A_g will be used for both humans as well as replicas.

techniques using high-speed imaging [66].



Figure 3.2 – Illustration of ePGG fixed subglottal light sensor (S) position and 3 supraglottal source (IR) positions along the front neck surface and source-sensor distances d_k [mm].

3.3 Mechanical replicas and set-ups

To fully assess the potential of ePGG as a non-invasive measurement of glottal area A_g , the relationship between ePGG and A_g needs to be studied quantitatively as a function of parameters potentially affecting the ePGG signal. Therefore, mechanical replicas of laryngeal airway portions are mounted to an experimental set-up developed to control and measure physical quantities in a reproducible and accurate way.

An overview of variables of interest and their order of magnitudes on human adult subjects and on replicas is given in Table 3.1. Note that the given fundamental frequency range is associated with vocal folds self-oscillation and hence rapid A_g variations during human phonation. Slower A_g variation occurs during other phoneme production, respiration, *etc*.

The upper bound of glottal area A_g for human subjects (Table 3.1) corresponds to the maximum reported for quiet breathing [86, 87]. An average value for normal subjects is disagreed on in literature due to inter subject variability, the difference between inspiration and expiration and the dependence on breathing effort. As an example, average values reported for normal subjects in [86, 87] differ for both inspiration (126±8 mm² versus 217±54 mm²) and expiration (70±7 mm² versus 178±35 mm²).

From trachea radius $R_t \approx 9 \pm 1$ mm (Table 3.1), it follows that trachea area $A_t = \pi R_t^2$ is

Quantity	Symbol	Human	Replicas ^{<i>a,b</i>}
Glottal area	A_g	$\leq 270 \text{ mm}^2$	\leq 250 mm ²
Trachea radius	R_t	10 mm	12.5 mm
Trachea length	L_t	120 mm	150-260 mm
Frequency ^c	f_0	75 - 250 Hz	90-350 Hz
Subglottal pressure ^c	P_u	$\leq 1500 \text{ Pa}$	\leq 1500 Pa

Table 3.1 – Typical values for variables of interest for adults and mechanical replicas [21, 48, 49, 85–87].

^{*a*} Rigid mechanical replica (Section 2.5.1) is scaled 3:1.

^b Deformable mechanical replica (Section 2.5.2) is scaled 1:1.

^{*c*} During self-oscillation of the deformable replica.

approximated as $A_t \approx 254 \pm 60 \text{ mm}^2$. This standard deviation results in $\leq 23\%$ under- or overestimation of A_t , which reflects the limited inter- and intra-subject variability of A_t compared to A_g . Furthermore, A_t is independent from the glottal aperture so that A_t holds for normal subjects as well as subjects suffering from a (vocal folds) pathology affecting the glottal area during respiration, closure or phonation.

To study the variation of the glottal area in a controlled way the two mechanical vocal folds replicas are used, the MDR (Section 2.5.1) one and the PLT (Section 2.5.2) one. This way rapid glottal area variation during vocal folds self-oscillation (PLT replica) as well as slow glottal area variation during glottal abduction or adduction (MDR replica) can be reproduced while the glottal area is known. To mimic the effect of airway walls, a simple mechanical airway replica (Section 3.3.1) is mounted to the subglottal and to the supraglottal side of each vocal folds replica so that the light sensor and source of the ePGG device can be attached on the exterior of the airway replicas.

3.3.1 Mechanical airway wall replica

The mechanical 'airway wall' replica, illustrated in Figure 3.3, consists of a uniform tube (transparent plexiglas, internal area $A_u = 491 \text{ mm}^2$, wall thickness 2 mm, length 150 - 260 mm) which can be covered with layers of lamb leather (layer thickness 0.7 mm) in order to vary wall light absorption, *i.e.* representing differences in human airway wall absorption properties due to wall thickness, wall tissue composition, *etc.* [82–84]. Lamb leather is used since it easily available and allows to vary ePGG voltages within a range pertinent for human subjects [10–12]. Note that it is not attempted to reproduce realistic details of a multi-layered airway tissue since the sensor is sensible to the absorption resulting from all tissue between the source and the sensor

regardless of its composition. Instead, a simple way to vary the wall absorption is proposed.



Figure 3.3 – Mechanical airway replica with ePGG light sensor (S) and source (IR).

3.3.2 Vocal folds replicas

First, the MDR vocal folds replica detailed in Section 2.5.1, is used, as illustrated in Figure 3.4(a), in order to compare the known glottal aperture $A_g(t)$ to the MSePGG signal for slow variation, as glottal aperture movement.

Then PLT replica, shown in Figure 3.4(b) and detailed in Section 2.5.2, is used to simulate fast vocal folds movements. The associated time-variation of the gap is observed using the high-speed camera (525 frames per second and aperture time 750 μ s).

It is noted that absorption properties of airway wall material (mostly lamb leather) and vocal folds (mostly water) are different and that transillumination through the water-filled latex tubes occurs.



(a) MDR replica

(b) PLT replica

Figure 3.4 – Vocal folds replica a) MDR and b) PLT with subglottal and supraglottal airway replicas to which ePGG sensor (S) and source (IR) are mounted.

3.4 ePGG signal characterization

The ePGG system, described in Section 3.2, is assessed on the mechanical replicas described Section 3.3. Since experimental set-ups are equipped to measure the glottal area, the relation-ship between ePGG signal and glottal area can be systematically studied on these replicas as a function of parameters potentially affecting the ePGG signal illustrated in Figure 3.1. In the following, the experimental ePGG signal characterization is presented firstly for static geometrical configurations with constant glottal area in Section 3.4.1 and secondly for dynamic geometrical configurations with time-varying glottal area in Section 3.4.2. During all the experiments, the temperature was maintained at 21.5 ± 1.0 °C.

3.4.1 Static glottal area

In this section, static geometries are considered so that mean ePGG signals are shown. Mean values are derived from consecutive ePGG signal for 3 s and the coefficient of variation yields less than 24%.

Firstly, the effect of the source-sensor distance d (Figure 3.1) on the ePGG signal is sought.

Chapter 3. External lighting and sensing photoglottography: characterization and MSePGG algorithm

The ePGG system is positioned on the mechanical airway replica with the constant area² ($A_u = 491 \text{ mm}^2$, Section 3.3.1) as shown in Figure 3.3. The source-sensor distance *d* is systematically varied by repositioning the source in the range $d \le 200 \text{ mm}$ and the orientation angle is 27°. In addition, in order to mimic the influence of wall tissue thickness, measurements are performed for two (thickness 1.4 mm) and three leather layers (thickness 2.1 mm). Measured mean ePGG signals are plotted in Figure 3.5. The ePGG signal decreases with *d* regardless



Figure 3.5 – Mean ePGG signal as a function of source-sensor distance *d* for the airway replica with 2 (\times) and 3 (+) leather layers.

of wall thickness. Linear fitting of measured ePGG signals in the range $d \le 100$ mm (appropriate for human subjects) and in the range $d \ge 100$ mm (appropriate for dynamic mechanical replicas), results in coefficient of determination $R^2 \ge 98.9\%$. Consequently, a first order linear approximation can be used to characterize the evolution of ePGG signal with source-sensor distance *d*, while the negative slope depends on wall absorption (thickness) and distance *d*. All the remaining experiments are done with 2 layers (thickness 1.4 mm).

Secondly, static geometrical configurations are assessed in order to determine the effect of the source orientation angle in the mid-coronal plane, as in Figure 3.1) on the ePGG signal. The ePGG system is again positioned on the uniform mechanical airway replica, in Figure 3.3, *i.e.* in absence of a glottal constriction (no glottal replica). The source orientation angle is systematically varied from 0° up to 40° and the source-sensor distance is held constant to d = 100 mm. Measured mean ePGG signals are plotted in Figure 3.6.

For orientation angles up to about 15° , the ePGG signal is minimum and only marginally (< 0.3 V) affected by the orientation angle due to the source (IR) half beam angle of $22.5 \pm 2.5^{\circ}$ (Section 3.2). Further increasing the orientation angle above 15° results in a linear ($R^2 = 98.1\%$) increase of the mean ePGG signal. All remaining experiments are done for orientation angle 27° .

Thirdly, static geometrical configurations are performed to determine the effect of glottal area

²Six different cross-section shapes were tested in [88] confirming current results.



Figure 3.6 – Mean ePGG signal as a function of source orientation.

on the ePGG signal (Figure 3.1). The MDR and PLT mechanical replicas are used and a uniform mechanical airway wall replica is attached at each end as shown in Figure 3.4. The source and sensor are positioned on each airway replica (trachea end and vocal tract end) so that the glottal area of the mechanical replica corresponds to the minimum area of the channel portion between source and sensor. The source-sensor distance is d = 150 mm for the rigid and d = 257 mm for the deformable glottal replica. Glottal area A_g is varied in the range 0-55 mm² (rigid) and 20-100 mm² (deformable). Measured mean ePGG signals are plotted in Figure 3.7.



Figure 3.7 – Mean ePGG signal as a function of static glottal area A_g for a rigid (×) and a deformable (+) mechanical glottal replica.

The ePGG signal increases linearly with A_g for the rigid ($R^2 = 99.2\%$) and the deformable ($R^2 = 98.2\%$) replica, indicating that ePGG signal and glottal area relate well using a linear approximation. Note that, in general, the differences in slope and offset in the figure may be due to 1) positioning of the ePGG system (source-sensor distance *d*, orientation angle) and 2) channel wall properties affecting light absorption (thickness, material, *etc.*).

3.4.2 Time-varying glottal area

The correlation between the time-varying ePGG signal and the time-varying glottal area is quantified for the MDR replica (Figure 3.4(a)) with source-sensor distance d = 150 mm and for the PLT replica (Figure 3.4(b)) with source-sensor distance d = 257 mm. For the motor-driven replica, oscillations are imposed at frequencies $f_0 \in \{2, 5, 10, 12\}$ Hz and the glottal area is varied in the range $0 \le A_g \le 40$ mm². For the flow-driven deformable replica self-oscillations are observed for mean subglottal pressures $P_u \in \{500, 570, 720, 840\}$ Pa so that the fundamental frequencies yielded $f_0 \in \{113, 125, 129, 131\}$ Hz and the area varied in the range $20 \le A_g \le 100$ mm².

Typical examples of correlated time signals for slow (MDR) and fast (PLT) vocal folds displacement are plotted in Figure 3.8. The scaled ePGG signal is firstly normalized by accounting for its mean and standard deviation and then scaled to mm^2 by accounting for the mean and standard deviation of the shown A_g .



(b) Deformable, correlation coefficient 90%

Figure 3.8 – Correlated time signals of glottal area $A_g(t)$ (dashed line) and scaled ePGG (full line): a) rigid replica ($f_0 = 10$ Hz), b) deformable replica ($f_0 = 129$ Hz).

Correlation coefficients between ePGG signals and glottal area $A_g(t)$ yield 90% for the rigid and 85% for the deformable glottal folds replica.

Consequently, the ePGG signal and glottal area are correlated during the oscillation. In the

following section, it is aimed to formalize the relationship between ePGG signal and glottal area $A_g(t)$ accounting for the different variables affecting the measured ePGG signal.

3.5 Multi-Signal-ePGG (MSePGG)

In Section 3.4, it was shown that the ePGG signal is mainly determined by 1) the sourcesensor distance, 2) the minimum area of the channel portion between the source and sensor and 3) the measurement condition determined by the combination of wall properties (*e.g.* absorption), environment (*e.g.* light) and ePGG system settings (*e.g.* amplification outlined in Section 3.2) and positioning (*e.g.* orientation angle).

In the following, a Multi-Signal-ePGG (MSePGG) approach is proposed accounting for each of these factors. The underlying model (Section 3.5.1), parameters estimation (Section 3.5.2.1) and their initialization (Section 3.5.2.2) is outlined. The sought relationship between ePGG signal and glottal area $A_g(t)$ is detailed (Section 3.5.3) and an overview of the resulting MSePGG workflow is summarized (Section 3.5.4).

3.5.1 Model

Following the transillumination principle shown in Figure 3.1, ePGG sensor voltage U is proportional to light intensity I at distance d_k from the light source,

$$U(d_k) \propto I(d_k),\tag{3.1}$$

where transmitted light intensity $I(d_k)$ at sensor position d_k is then expressed using light flux Φ as

$$I(d_k) = \oiint_{A_{min}(d_k)} \Phi(d_k) dA, \qquad (3.2)$$

where

$$A_{min}(d_k) = \min_{d \in [0,d_k]} (A(d))$$
(3.3)

is the minimum area encountered by the transmitted light flux between the source and sensor positions. Furthermore, in Section 3.4 it was shown that for the ranges of interest (Table 3.1), the dependence on d and A_{min} can be described using a first order linear approximation. Conse-

quently, light flux $\Phi(d) > 0$ can be approximated by model $\Phi_m(d)$ defined as

$$\Phi_m(d) = \alpha_d d + \beta_d, \tag{3.4}$$

with slope $\alpha_d < 0$ and offset $\beta_d > 0$ (see Section 3.4). From (3.2), $I(d_k)$ is now modeled as

$$I_m(d_k) = A_{min}(d_k) \cdot \Phi(d_k),$$

= $A_{min}(d_k) \cdot (\alpha_d d_k + \beta_d).$ (3.5)

Inserting (3.5) in (3.1) results in modeling the ePGG voltage $U(d_k)$ as $U_m(d_k)$ given by

$$U_m(d_k) = \gamma ((\alpha_d d_k + \beta_d) \cdot A_{min}(d_k)) + \eta$$

= $(\alpha_v d_k + \beta_v) \cdot A_{min}(d_k) + \eta$ (3.6)

where $\eta > 0$ is the signal measured for $A_{min}(d_k) = 0$ and $\gamma > 0$ is the scaling factor of (3.1). For sake of simplicity, let us denote $\alpha_v = \gamma \alpha_d < 0$ and $\beta_v = \gamma \beta_d > 0$. It is worth noting that this latter model only holds for positive coefficient $\alpha_v d_k + \beta_v$: in other words source-sensor distance d_k must satisfy $d_k \leq -\beta_v / \alpha_v$. Additionally, it is noted that $A_{min} = 0$ corresponds to glottal closure for which no direct light is transmitted through the vocal folds although light transmitted due to tissue transillumination might remain. Therefore, η is independent of d and A_{min} so that η reflects solely the measurement condition. Considering now the time-variation of the glottal opening, model (3.6) can be directly extended as

$$U_m(d_k,t) = (\alpha_v d_k + \beta_v) \cdot A_{min}(d_k,t) + \eta.$$
(3.7)

Consequently using model (3.7), extracting area $A_{min}(d_k,t)$ from measured ePGG signals $U(d_k,t)$ reduces to a problem of parameter estimation as assessed in the next section.

3.5.2 Calibration

3.5.2.1 Parameter estimation

Let vector $U = [U(d_1) \dots U(d_K)]^{\dagger} \in \mathbb{R}^K$ denote the concatenation of *K* mean ePGG measurements and .[†] the transpose operator. Vectorizing model (3.6) yields

$$\mathbf{U}_m = [\boldsymbol{\alpha}_v \mathbf{d} + \boldsymbol{\beta}_v] \odot [\mathbf{A}_{\min}] + \boldsymbol{\eta}, \qquad (3.8)$$

where $d = [d_1 \dots d_K]^{\dagger} \in \mathbb{R}^K$ is the sensor position vector, $A_{\min} = [A_{\min}(d_1) \dots A_{\min}(d_K)]^{\dagger} \in \mathbb{R}^K$ is the minimum area vector and \odot is the Hadamard product. It is worth noting that this latter model relies on the K+3 parameters $\{\{A_{\min}(d_k)\}_{1 \le k \le K}, \alpha_v, \beta_v, \eta\}$ that should be estimated from K ePPG measurements $\{U(d_k)\}_{1 \le k \le K}$. So, without additional constraints, model (3.8) is not identifiable. A solution to overcome this inherent problem is to reduce the number of parameters to be estimated by measuring for several distances the ePGG voltage associated with the same A_{\min} .

To simplify the notations, and without loss of generality, let us assume that \widetilde{K} measurements have been performed for each of the *N* different A_{min} values (so that $N \times \widetilde{K} = K$). Consequently, model (3.8) can be recast as

$$\mathbf{U}_{m} = \left[\boldsymbol{\alpha}_{v}\mathbf{d} + \boldsymbol{\beta}_{v}\right] \odot \left[\tilde{\mathbf{A}}_{\min} \otimes \mathbb{1}_{\widetilde{K}}\right] + \boldsymbol{\eta}, \qquad (3.9)$$

where $\tilde{A}_{\min} = [A_{\min_1} \dots A_{\min_N}]^{\dagger} \in \mathbb{R}^N$, $\mathbb{1}_{\tilde{K}} \in \mathbb{R}^{\tilde{K}}$ is the identity vector and \otimes is the Kronecker product. Using such a constraint, the number of parameters is reduced to N + 3, leading thus to an identifiable model as soon as $N + 3 \leq K$.

In practice, the set of parameters $\{\tilde{A}_{\min}, \alpha_{\nu}, \beta_{\nu}, \eta\}$ is estimated by minimizing mean square error J() between the *K* ePGG measurements U and the related ePGG model U_m provided by (3.9):

$$\left\{\hat{\tilde{A}}_{\min},\hat{\alpha}_{\nu},\hat{\beta}_{\nu},\hat{\eta}\right\} = \operatorname{argmin} \begin{array}{c} \tilde{A}_{\min} \ge 0, \, \alpha_{\nu} \le 0, \\ \beta_{\nu} \ge 0, \, \eta \ge 0 \end{array} J\left(\tilde{A}_{\min},\alpha_{\nu},\beta_{\nu},\eta\right)$$
(3.10)

where

$$J(\tilde{A}_{\min}, \alpha_{\nu}, \beta_{\nu}, \eta) = \frac{1}{K} (U - U_m)^{\dagger} (U - U_m).$$
(3.11)

To solve (3.10), a gradient descent iterative method is used [89]. The initialization of this iterative optimization algorithm is discussed in the next section.

3.5.2.2 Initialization

To initialize parameters set $\{\tilde{A}_{\min}, \alpha_{\nu}, \beta_{\nu}, \eta\}$ a three-step procedure is applied.

Firstly, η is initialized as $\eta^{(0)}$, the mean ePGG voltage when no direct light is transmitted. Indeed, this condition is obtained when $A_{min} = 0$, so that (3.7) resumes to $U_m(d_k, t) = \eta$, $\forall (d_k, t)$. Thus, $\eta^{(0)}$ is estimated as the average value of all the measurements for which $A_{min} = 0$. Note that trained subjects can produce glottal closure on instruction, for other subjects closure is mimicked by supplying no excitation light. The effect of the subject instruction on η^0 and resulting parameter and area estimations needs to be assessed in future studies.

Secondly, it is sought to initialize α_v and β_v by measuring ePGG signals $U^{(0)}$ associated with a constant reference area $A_0 > 0$ (detailed in the next section) for at least two different sensor positions $d = [d_1 \dots d_P]^{\dagger} \in \mathbb{R}^P$, with $P \ge 2$ (see³) and $\tilde{A}_{min} = A_0 \mathbb{1}_P \in \mathbb{R}^P$. From model (3.8) and assuming that A_0 is known, one can estimate $\alpha_v^{(0)}$ and $\beta_v^{(0)}$ by the following least mean square (LMS) minimization

$$\alpha_{v}^{(0)}, \ \beta_{v}^{(0)} = \operatorname{argmin} \alpha_{v}, \ \beta_{v} \left\| \mathbf{U}^{(0)} - \mathbf{U}_{m}^{(0)} \right\|_{2}^{2},$$
(3.12)

where

$$\mathbf{U}_m^{(0)} = A_0 \left[\alpha_v \mathbf{d} + \beta_v \right] + \boldsymbol{\eta}^{(0)}$$

Consequently, $\alpha_{\nu}^{(0)}$, $\beta_{\nu}^{(0)}$ are expressed in a closed form as

$$\begin{pmatrix} \boldsymbol{\alpha}_{\nu}^{(0)} \\ \boldsymbol{\beta}_{\nu}^{(0)} \end{pmatrix} = \left(\mathbf{D}^{\dagger} \mathbf{D} \right)^{-1} \mathbf{D}^{\dagger} \frac{\left(\mathbf{U}^{(0)} - \boldsymbol{\eta}^{(0)} \right)}{A_0},$$
 (3.13)

with $\mathbf{D} = [\mathbf{d}, \mathbb{1}_P] \in \mathbb{R}^{P \times 2}$.

Thirdly, from $\{\alpha_v^{(0)}, \beta_v^{(0)}, \eta^{(0)}\}\$ and measured ePGG signals U, the *N* elements of \tilde{A}_{\min} $(\{A_{\min_n}\}_{1 \le n \le N})$ are initialized by a LMS minimization obtained by approximating ePGG measurements U with model (3.9), leading to

$$\forall n, \quad A_{\min_n}^{(0)} = \frac{[\alpha_v^{(0)} \mathbf{d}_n + \beta_v^{(0)}]^{\dagger} (\mathbf{U}_n - \boldsymbol{\eta}^{(0)})}{[\alpha_v^{(0)} \mathbf{d}_n + \beta_v^{(0)}]^{\dagger} [\alpha_v^{(0)} \mathbf{d}_n + \beta_v^{(0)}]}, \quad (3.14)$$

where $U_n \in \mathbb{R}^{\widetilde{K}}$ (resp. $d_n \in \mathbb{R}^{\widetilde{K}}$) is the n-th subvector of U (resp. d) so that $U = [U_1^{\dagger}, \dots, U_N^{\dagger}]^{\dagger}$ (resp. $d = [d_1^{\dagger}, \dots, d_N^{\dagger}]^{\dagger}$).

³Note that $P \ge 2$ is a theoretical criterion. In practice, $P \ge 3$ is used.

3.5.3 Estimation of glottal area $A_g(t,k)$

When the glottis is situated in between the source (IR) and the sensor (S) (Figure 3.1) glottal area $A_g(t,k)$ can be estimated, where $A_g(t,k)$ indicates the glottal area as a function of time t using the signal gathered through the sensor at position d_k . Indeed in this case, minimum area A_{min} encountered by the light flux corresponds to glottal area A_g and thus

$$A_g(t,k) = A_{min}(d_k,t). \tag{3.15}$$

Consequently, glottal area $A_g(t,k)$ is estimated in the same way as $A_{min}(d_k,t)$ outlined in the previous section. This means that firstly the parameter set $\{\tilde{A}_g, \alpha_v, \beta_v, \eta\}$ is estimated using \tilde{K} different sensor positions for N different glottal areas $A_g \in \mathbb{R}^N$ so that the model is identifiable, *i.e.* $N+3 \leq K$ with $K = N \times \tilde{K}$.

Once the parameter set estimation is assessed and assuming that the measurement condition is not altered – *i.e.* unchanged combination of subject, source positioning, environment and hardware settings – it is argued (and shown in Section 3.6) that any additional ePGG measurement $U(d_i,t)$ with a sensor at any source-sensor distance d_i results in glottal area estimation $\hat{A}_g(t,i)$ using (3.7) and substituting $U_m(d_i,t)$ with $U(d_i,t)$ and $\{\alpha_v, \beta_v, \eta\}$ with $\{\hat{\alpha}_v, \hat{\beta}_v, \hat{\eta}\}$ so that

$$\hat{A}_g(t,i) = \frac{U(d_i,t) - \hat{\eta}}{\hat{\alpha}_v d_i + \hat{\beta}_v}.$$
(3.16)

It is worth noting that the estimated area $\hat{A}_g(t,i)$ does not depend on source-sensor distance d_i , but depends on the choice of reference area A_0 during initialization. Setting $A_0 = 1$ results in a normalized area estimation relative to A_0 . When a quantitative area estimation is needed, it is proposed to use trachea area ($A_0 \approx 254 \text{ mm}^2$) or glottal area during quiet inspiration ($A_0 \approx 127 \text{ mm}^2$) as a reference area (Section 3.3, [86]).

Consequently, once MSePGG calibration is done, resulting $\{U(d_i,t)\}_{1 \le i \le I}$ associated with $I \ge 1$ sensors results in I estimations of the minimum area, *i.e.* $\{\hat{A}_g(t,i)\}_{1 \le i \le I}$. Thus a single or multiple sensor positions can be used to estimate the glottal area once calibration is achieved. For I > 1 a set of estimations is obtained, which can be further post-processed, *e.g.* to further optimize the estimation by applying a sensor-selection criteria (*e.g* signal-to-noise ratio) or to further characterize the estimation (*e.g.* determine an uncertainty on the resulting minimum area estimation). Such post-processing is not considered in this work.
3.5.4 Workflow

To summarize, the MSePGG workflow of minimum area estimation is schematized in Figure 3.9. The first two steps – parameter initialization and estimation (Section 3.5.2) – lead to MSePGG calibration in order to account for the specific measurement condition (subject, environment, amplification, *etc.*). Once calibration is achieved the minimum area as a function of time $\hat{A}_{min}(t,i)$ can be estimated for any experimental protocol and for any of $I \ge 1$ sensor positions as long as calibration parameters $\{\hat{\alpha}_v, \hat{\beta}_v, \hat{\eta}\}$ remain valid, *i.e.* measurement conditions are unaltered. Different post-processing strategies can be applied to the resulting I estimations of minimum area.

Concretely, in Section 3.6.2 where the MSePGG approach is applied to measurements on a human speaker in order to estimate the glottal area, the same three sensors positions are considered during all the steps so that $\tilde{K} = 3$, P = 3 and I = 3. Reference area A_0 corresponds to the area during quiet inspiration. It is noted that the choice of A_0 needs to be addressed carefully when an extended study on human subjects is aimed since an error in A_0 might affect estimated parameters α_v and β_v and hence estimated glottal areas $\hat{A}_g(t,i)$ following (3.16). Furthermore, voltages associated with N = 3 minimum areas were determined as the mean values associated with the maximum, minimum and averaged voltage during 16 subsequent periods of vowel /a/. No post-processing is performed so that glottal area estimations $\hat{A}_g(t,i)$ are directly shown.

3.6 MSePGG results

MSePGG outlined in Section 3.5 is applied following the workflow shown in Figure 3.9. Measurements on the deformable mechanical replica (Section 3.6.1) and on a human speaker (Section 3.6.2) are assessed.

3.6.1 Deformable mechanical replica

Time-varying ePGG signals due to time-varying glottal area during self-oscillation are measured on the same deformable mechanical replica and under the same measurement conditions (environmental and instrumental) as described in Section 3.3.2. MSePGG parameter estimation is assessed using data discussed in Section 3.4 following the procedure outlined in Section 3.5.2.1. Resulting values of parameters $\hat{\alpha}_v$, $\hat{\beta}_v$ and $\hat{\eta}$ can be used directly in (3.16) to estimate glottal area \hat{A}_g since estimated parameters ($\hat{\alpha}_v$, $\hat{\beta}_v$ and $\hat{\eta}$) depend solely on the replica and measurement conditions and not on the glottal area (Section 3.5.1).

Mechanical replica properties [49] are set asymmetrical between both vocal folds. The re-



Figure 3.9 – MSePGG workflow overview.



Figure 3.10 – Deformable mechanical replica: a) ePGG signal portion, b) measured A_g (dashed line) and estimated \hat{A}_g (full line) area. Mean error yields 5.4%.

sulting area variation during self-oscillation and hence associated ePGG signal (Figure 3.10(a)) are characterized by a high harmonic distortion rate (58%) due to the presence of the fundamental frequency ($f_0 = 131$ Hz) and the second harmonic ($f_1 = 262$ Hz). High harmonic distortion rates might occur in the case of voice disorders characterized by diplophonia [67, 68, 90, 91].

Estimated (\hat{A}_g) glottal areas by (3.16) and glottal areas (A_g) measured by imaging are shown in Figure 3.10(b). Quantitative comparison between estimated and measured glottal areas results in a mean error of 5.4%. This suggests that the MSePGG approach can be applied to assess the glottal area in the case of a normal glottal variation and in the case of a voice disorder related to diplophonia. In the next section, MSePGG is applied to a human speaker.

3.6.2 Human speaker

In this section, ePGG measurements are obtained on a human speaker (21-yr-old male, Telugu native speaker, source orientation angle 1°) without any voice disorder (self-reported). Measurement of ePGG data were performed while the speaker pronounced a sustained vowel /a/, during quiet respiration and during glottal closure. The light sensor and source are positioned as outlined in Section 3.2 and illustrated in Figure 3.2. The three source positions resulting in the largest source-sensor distances (d = 61 mm, d = 70 mm and d = 81 mm) were assessed as (6.2) is verified to hold: P = 3, $\tilde{K} = 3$ and I = 3. The ePGG measurements for each source-sensor distance during vowel /a/ are shown in Figure 3.11(a).

Next, ePGG values for MSePGG parameters estimation are selected. For all three sourcesensor distances *d*, it is aimed to determine ePGG voltages associated with three different glottal areas (N = 3) during vowel production so that the condition $N + 3 \le K$ (with $K = \tilde{K} \times N$) holds and the procedure for MSePGG parameter estimation (Section 3.5.2.1) can be applied. For each distance *d*, the highest value and lowest value of the range of ePGG voltages in each period are determined.

The mean value of the 16 maxima of 16 consecutive periods is considered as the ePGG voltage associated with the largest (A_1) glottal area during vowel /a/ production and this for each distance d. The same way the mean value of the 16 minima of 16 consecutive periods is considered as the ePGG voltage associated with the smallest (A_3) glottal area and this again for each distance d.

The ePGG voltage corresponding to the average of these extrema for each distance *d* is then associated with a third glottal area (A_2) and $A_3 < A_2 < A_1$ holds. Resulting K ($K = \tilde{K} \times N$ so K = 9) ePGG data associated with each of these areas are plotted in Figure 3.11(b) as a function of source-sensor distance *d*. It is observed that in agreement with (3.6) the ePGG level decreases linearly for each area with respect to source-sensor distance *d*. The linear fits yield zero voltage level at nearly the same source-sensor distance ($d = 89.2 \pm 0.2$ mm) and thus the linear fits intersects near -0.012 V at $U_m(d_k) - \eta = 0$ following (3.6) for all d_k . Therefore, the intersection point's magnitude (0.012 V) provides a rough estimation of the order of magnitude of the estimated ePGG voltage associated with closed glottal area $\hat{\eta} = 0.01$ V.

The extracted ePGG voltages plotted in Figure 3.11(b) (U_{1,2,3}) are then used to initialize MSePGG parameters following the procedure described in Section 3.5.2.2, *i.e.* step 1 of the calibration protocol shown in Figure 3.9. Concretely, at first, $\alpha_v^{(0)}$ and $\beta_v^{(0)}$ are initialized from (3.13) using ePGG voltages associated with maximum glottal opening during quiet inspiration and taking glottal area $A_0 = 127 \text{ mm}^2$ as a known reference area value (Section 3.5.3, [86]). Next, $\alpha_v^{(0)}$, $\beta_v^{(0)}$ and extracted U_{1,2,3} voltages associated with $A_{1,2,3}$ (shown in Figure 3.11(b)) are used in (3.14) in order to initialize $A_1^{(0)} \approx 41 \text{ mm}^2$, $A_2^{(0)} \approx 38 \text{ mm}^2$ and $A_3^{(0)} \approx 34 \text{ mm}^2$. Once all MSePGG parameters are initialized their value is estimated from (3.10) as described in Section 3.5.2.1, *i.e.* step 2 of the calibration protocol shown in Figure 3.9. It is noted that the initialized ($\tilde{A}_{min}^{(0)} = [A_1^{(0)}A_2^{(0)}A_3^{(0)}]^{\dagger}$ resulting from (3.14)) and estimated ($\tilde{A}_{min} = [\hat{A}_1\hat{A}_2\hat{A}_3]^{\dagger}$ resulting from (3.10)) area values are in good agreement ($\leq 1 \text{ mm}^2$) and thus the proposed initialization procedure results in reasonable area values.



Figure 3.11 – Sustained vowel /a/ pronounced by a human speaker for 3 ($\tilde{K} = 3$) source-sensor distances d: a) ePGG data as a function of time, b) extracted ($\tilde{K} \times N = 9$) ePGG values for MSePGG parameter estimation (symbols) as a function of d and their linear fit (dashed line) for three (N = 3) areas ($A_1 > A_2 > A_3$), c) estimated glottal areas $\hat{A}_g(t)$.

3.6. MSePGG results

Once MSePGG parameters are estimated, glottal areas associated with the time-varying ePGG signals shown in Figure 3.11(a) are estimated from (3.16). Estimated glottal areas $\hat{A}_g(t)$ for all three assessed sensor-source distances d are shown in Figure 3.11(c). The mean absolute relative error between between estimated glottal areas obtained from ePGG measurements at different source-sensor distances yields less than 12%. It follows that estimated areas are of the same order of magnitude as they vary in the same range, *i.e.* between 32 mm² and 48 mm². Given that ePGG signals for each source-sensor distance d are measured during a different utterance of the vowel /a/ some variation between the areas is expected. Consequently, the agreement between estimated areas is satisfying. Furthermore, this maximum amplitude of 16 mm² approximates the order of magnitude mentioned in literature for phonation of 12 mm² [71, 92]. It is noted that the slight overestimation of the amplitude is obtained for the ePGG signal at d = 81 mm for which the signal quality is less, illustrating that sensor positioning can be further addressed. The lack of glottal closure during vowel production observed in Figure 3.11(c) is confirmed by endoscopic imaging on the same subject.

Next, the same MSePGG parameters are applied to estimate (3.16) the time-varying glottal area during consecutive utterances of /sa/ (up to 5 s) by the same speaker in the same conditions at d = 61 mm as illustrated in Figure 3.12. Rapid glottal area variations are observed during vowel (/a/) whereas slow variations associated with glottal opening and closing occurs to pronounce sibilant fricative /s/. As expected the glottal area widens during frication noise. It is seen that the temporal resolution allows to study the transition between vowels and fricatives in detail.



Figure 3.12 – Estimated glottal areas \hat{A}_g for consecutive (5 s) utterance of /sa/ by a human speaker.

3.7 Discussion

Results shown in Section 3.6 illustrate that MSePGG provides an estimation of the timevarying minimum area on a mechanical replica and on a human speaker following the workflow summarized in Section 3.5.4. The MSePGG algorithm and workflow provides an elegant and innocuous method relying on 3 parameters to be estimated simultaneously. The MSePGG algorithm avoids dealing with the complexity of the composing tissue layers and anatomy. This way some restrictions related to the use of other techniques mentioned in the introduction are removed. Indeed, the MSePGG algorithm combined with the ePGG device strives to a noninvasive, non-expensive, continuous and quantitative measurement with minimum discomfort for the subject and a straightforward interpretation following an elegant and short calibration protocol at a low computational and data storage cost. Since the ePGG device is external, the danger for infections is minimal and there is no need for medical supervision. The MSePGG algorithm and workflow is a suitable candidate to provide a quantitative metric of the glottal area outside of a medical practice for clinical or non-clinical studies (health care, research, field studies) using ePGG measurements.

Some remarks can be made considering the proposed MSePGG method in Section 3.5. Firstly, it is noted that the accuracy of MSePGG on a human subject can be improved using a sensor array, instead of moving a single sensor to different positions to avoid intra-subject variability as well as to further simplify the parameter estimation protocol. Secondly, it is seen that the same algorithm and workflow can potentially be applied to obtain a quantitative area estimation from invasive PGG measurements as well since PGG relies on transillumination as well. Thirdly, the accuracy of MSePGG estimations is determined against area measurements on mechanical replicas for which accurate area values are available. MSePGG outcome on human subjects is merely done as an illustration since no quantitative glottal area values were available. Therefore, the influence of the parameter initialization procedure for human subjects, *i.e.* combination of instructions and parameter initialisation as well as sensor-source distances, on the parameter and area estimation needs to be assessed in future studies. In addition, thorough validation for human subjects needs to be assessed in future for a large variety of anatomical and clinical conditions. This way the applicability, advantages and limitations of the proposed MSePGG approach can be determined, e.g. to consider the impact of different positions of the epiglottis. It is expected that such studies will lead to further improvement of the MSePGG algorithm and workflow, the ePGG device as well as in term to a standardized protocol with respect to device positioning and settings.

3.8 Conclusion

Following characterization of ePGG measurements on mechanical replicas, the MSePGG algorithm and workflow is proposed in order to provide a quantitative estimation of the timevarying glottal area following a brief calibration protocol exploiting several source-sensor distances. The good quantitative agreement obtained on mechanical replicas (mean error 5.4%) and preliminary observations on a human subject (estimations within 12%) suggests that MSePGG is a promising technique to estimate the glottal area during normal as well as pathological vocal folds configurations. Future research is needed to consider the use of a sensor array ePGG device to fully investigate the MSePGG algorithm and workflow for different human subjects in order to further improve, validate and facilitate MSePGG on human subjects.

Observation of the glottal area was assessed in this chapter. Ongoing development of glottal area observation illustrates the importance of the glottal area for voicing and hence the interest for clinicians. In the following chapter, unilateral vocal fold paralysis is considered. It is a condition affecting the glottal area and it is assessed to which extent oscillation features are affected.

CHAPTER 4

Influence of level difference due to vocal folds vertical angular asymmetry

Based on Bouvet A., Tokuda I., Pelorson X., Van Hirtum A., 2020. *Influence of level difference due to vocal folds angular asymmetry on phonation*. Journal of the Acoustical Society of America, 147 (2).

This work was partly supported by JSPS summer fellowship (SP18205). Experiments were performed in the Department of Mechanical Engineering at Ritsumeikan University, Kyoto, Japan.

Specific nomenclature

α	Angle asymmetry between the two vocal folds	0
E	Largest thickness of vocal folds	m
e	Smallest thickness of vocal folds	m
a	Rotation axis – anterior VF edge distance	m
Α	Triangular leakage gap due to asymmetry	m ²
l_A	Base of the triangular gap	m
h_A	Heigh of the triangular gap	m
G	Degree of contact between vocal folds	[—]
$\alpha_{(I,II)}, \alpha_{(II,III)}$	Critical asymmetry angles	0
I,II,III	Different contact regimes	[—]

4.1 Introduction

Recently [5, 6], the impact of vertical level difference between parallel located vocal folds, hereafter referred to as parallel level difference, was investigated. These studies, however, did not take into account inclination of the paralysed vocal fold, which is physiologically more close to reality as oultined in Section 2.4.1. Therefore, in this work, the influence of inclination of one vocal fold is studied, *i.e.* angular level difference. Concretely, the normal vocal fold is kept in its transverse plane, whereas the paralysed vocal fold is rotated in the sagittal plane so that its posterior edge is moved in the superior direction imposing a left-right vocal folds asymmetry angle (α). Our aim is to systematically study the potential influence of asymmetry angle $\alpha > 0^{\circ}$ on the auto-oscillation of mechanical replicas. There have been numerous studies on left-right asymmetry focusing on other vocal folds properties such as tension and shape [7, 31, 93–96]. In the present study, symmetry of such properties is maintained so as to concentrate only on the angular asymmetry between the left and the right vocal fold.

4.2 Glottal replicas with imposed asymmetry angle α set up

The vocal folds replicas dimensions are summarized in Table 4.1 and depicted in Figure 4.1: largest *E* and smallest thicknesses *e* along the inferior-superior direction and length *w* along the posterior - anterior direction. The right vocal fold (normal) is fixed in the transverse plane, whereas the left vocal fold (paralysed) is rotated so that its posterior edge is lifted in the superior direction resulting in left-right vocal folds asymmetry angle α shown in Fig 4.1. For all glottal vocal folds replicas, the rotation axis along the left-right direction is fixed in the rigid holder at a = 4.5 mm from the anterior edge (Table 4.1).

Experiments are performed for 13 different values of left-right asymmetry angle α ranging from $\alpha = 0^{\circ}$ up to $\alpha = 24.6^{\circ}$. For symmetric angular configuration ($\alpha = 0^{\circ}$), the glottal area in the transverse plane is negligible as the vocal folds are in full contact (Figure 2.10(b)). When α is increased, air leakage occurs through a glottal gap in the medio-sagittal plane since the vocal folds are no longer in full contact as depicted in Figure 4.1 (red dashed triangle). The triangular

Table 4.1 – Dimensional	parameters and	critical	asymmetry	angles	$\alpha_{(I,II),(II,III)}$	for all	replicas
(M5, MRI and EPI).							

Replica	w [mm]	E [mm]	e [mm]	a [mm]	$lpha_{I,II}$ [°]	$\alpha_{II,III}$ [°]
M5	17.0	10.7	2.99	4.5	8.1	29.2
MRI	18.0	10.0	3.00	4.5	7.8	28.8
EPI	17.0	10.2	1.09	4.5	3.0	14.0

gap with area A, base l_A and height h_A appears near the posterior vocal folds edge and extends towards the anterior edge as α increases. In this sense, l_A indicates the length, for which the vocal fold's contact is interrupted. The degree of vocal fold's contact as a function of α is then expressed as $\mathscr{G} \in [0 1]$:

$$\mathscr{G}(\alpha) = \begin{cases} 0 & \text{if } l_A > w, \\ 1 - \frac{l_A(\alpha)}{w} & \text{otherwise.} \end{cases}$$
(4.1)

Three contact regimes (I, II and III) are identified as α increases: I) $\mathscr{G} = 1$ when $l_A = 0$, *i.e.* full vocal fold's contact, II) $1 > \mathscr{G} > 0$ when $0 < l_A < w$, *i.e.* partial vocal fold's contact and III) $\mathscr{G} = 0$ when $l_A = w$, *i.e.* no vocal folds contact. Note that full contact (regime I) implies no air leakage and hence A = 0, whereas, for A > 0, either partial contact (regime II) or no



Figure 4.1 – Imposed asymmetry angle α , glottal gap parameters A, l_A , h_A and geometrical parameters E, e, w and a: a) glottal gap (red triangle) for MRI replica with $\alpha = 20^{\circ}$ ($A \approx 32 \text{ mm}^2$, $l_A \approx 14.2 \text{ mm}$ and $h_A \approx 4.6 \text{ mm}$), b) sagittal view, c) single left vocal fold.

56 Chapter 4. Influence of level difference due to vocal folds vertical angular asymmetry



Figure 4.2 – Geometrical characterization of M5, MRI and EPI replicas as a function of asymmetry angle α : a) degree of contact \mathscr{G} (Eq. (4.1)) and critical angles $\alpha_{(I,II),(II,III)}$ (vertical lines), b) glottal gap area A (Eq. (4.2)) and maximum normal pre-phonatory glottal area (5 mm², horizontal line) and c) glottal gap height h_A (Eq. (4.4)) and maximum vertical level difference during phonation (3 mm, horizontal line). Symbols indicate experimentally assessed α values.

contact (regime III) occurs depending on α . Two critical angles $\alpha_{I,II}$ and $\alpha_{II,III}$ are then defined: $\alpha_{I,II}$ indicating the largest angle, for which full contact occurs (shift from regime I to II) and $\alpha_{II,III}$ indicating the largest angle, for which partial contact occurs (shift from regime II to III).

An analytic expressions for glottal gap area $A(\alpha)$ (more detailed in Appendix F), its associated triangle base length $l_A(\alpha)$, and triangle height $h_A(\alpha)$ follow from trigonometric reasoning using geometrical vocal folds replica parameters w, E, e and a:

$$A(\alpha) = \begin{cases} 0 & \text{if } \tan(\alpha) \left(w + a + \frac{E/2 - e}{\sin(\alpha)} \right) < \frac{E}{2}, \\ \frac{1}{2 \cdot \tan(\alpha)} \cdot \left(\left(w + a + \frac{1}{\sin(\alpha)} \cdot \left(\frac{E}{2} - e \right) \right) \cdot \tan(\alpha) - \frac{E}{2} \right)^2 \\ + \frac{1}{2} \cdot \tan(\alpha) \cdot (E - e) \cdot \left(e - 2 \cdot \tan\left(\frac{\alpha}{2} \right) \cdot (w + a) \right) & \text{otherwise,} \end{cases}$$
(4.2)

and

$$l_A(\alpha) = \begin{cases} 0 & \text{if } \tan(\alpha) \left(w + a + \frac{E/2 - e}{\sin(\alpha)} \right) < \frac{E}{2} & \text{or } A = 0, \\ 2\sqrt{\frac{A(\alpha)}{\sin(2\alpha)}} & \text{otherwise,} \end{cases}$$
(4.3)

so that h_A depicted in Figure 4.1(b) varies as:

$$h_A(\alpha) = \begin{cases} 0 & \text{if } l_A = 0, \\ \frac{2A(\alpha)}{l_A(\alpha)} & \text{otherwise,} \end{cases}$$
(4.4)

and thus ratio $(h_A/l_A)(\alpha) = 2A(\alpha)/l_A^2(\alpha)$.

For each replica (M5, MRI and EPI), $\mathscr{G}(\alpha)$ (Eq. (4.1)), $A(\alpha)$ (Eq. (4.2)) and $h_A(\alpha)$ (Eq. (4.4)) are plotted in Fig 4.2. The curves correspond to analytic expressions, while the symbols indicate asymmetry angles ($\alpha < 25^{\circ}$), at which experiments were carried out. Critical angles $\alpha_{I,II}$ and $\alpha_{II,III}$ are summarized in Table 4.1 and plotted (vertical annotated lines) in Figure 4.2(a). Values of $\mathscr{G}(\alpha)$, $A(\alpha)$ and $\alpha_{(I,II),(II,III)}$ for M5 and MRI are in close range, since their geometrical parameters (w, E, e and a in Table 4.1) are similar. For EPI, on the other hand, e yields about one-third of the value for M5 and MRI so that glottal gap $A(\alpha)$ increases more rapidly and vocal fold's contact degree $\mathscr{G}(\alpha)$ reduces more rapidly than those of M5 and MRI. As a consequence, in the range of our study $\alpha < 25^{\circ}$, all three contact regimes occur for EPI as $\alpha_{II,III}^{EPI} < 25^{\circ}$ (Table 4.1), whereas regime III (no contact, $\mathscr{G} = 0$) does not occur for M5 and MRI, since $\alpha_{II,III}^{MRI,M5} > 25^{\circ}$ holds (Table 4.1).

Considering that initial pre-phonatory glottal gap area for healthy human speakers [97] yields less than 5 mm² (dashed horizontal line in Figure 4.2(b)), it follows that the glottal gap area for $\alpha = 24.6^{\circ}$ is increased by a factor of 8 (for M5) or more (for EPI and MRI). Height h_A

58 Chapter 4. Influence of level difference due to vocal folds vertical angular asymmetry

(Figure 4.2(c)) quantifies the maximum distance between both vocal folds along the inferiorsuperior direction and therefore provides a rough (since *e* is neglected) approximation for the vertical level difference measured during human phonation [37], *i.e.* 1.3 ± 1.5 or up to about 3 mm (dashed horizontal line in Figure 4.2(c)). For $\alpha = 24.6^{\circ}$, the height of $h_A \approx 6$ mm is twice as much as the value reported on humans. Thus, the range of asymmetry angle α studied in our experiments covers the range observed for human phonation. Note that $h_A \leq 3$ mm corresponds to $\alpha \leq 13^{\circ}$ for EPI and to $\alpha \leq 17^{\circ}$ for M5 and MRI. The ratio $(h_A/l_A)(\alpha)$ increases from 0 up to 0.4 for all replicas, indicating that height h_A of the leakage triangle remains less than half of its base l_A for all α .

4.3 Fluid-structure interaction experiments

4.3.1 Experimental approach

In order to reproduce a fluid-structure interaction leading to auto-oscillation, glottal vocal folds replicas are positioned vertically as shown in Figure 4.1(a). A rigid uniform tracheal tube (diameter 16 mm, length 460 mm, acoustic resonance frequency 185 Hz) is attached to the inferior end of the glottal silicone vocal folds replica. Continuous steady airflow (density $\rho_G = 1.2 \text{ kg} \cdot \text{m}^{-3}$, dynamic viscosity $\mu_G = 1.8 \times 10^{-5}$ Pa·s, temperature $24 \pm 2^{\circ}$ C) is provided by an air compressor (Hitachi SC820) connected to an upstream pressure reservoir (volume 0.04 m³) filled with acoustic foam in order to avoid parasite acoustic resonances to which the tracheal tube is mounted [5]. The compressor is equipped with a pressure regulator (10202U, Fairchild, Winston-salem, NC). A pressure transducer (Kyowa PDS-70GA, accuracy ± 5 Pa) is positioned in a pressure tap in the tracheal wall, 185 mm upstream of the glottal vocal folds replica, in order to measure upstream pressure P_u with sampling frequency $f_s = 10$ kHz.

For each vocal folds replica, 13 different asymmetry angles $\alpha < 25^{\circ}$ are imposed as outlined in Section 4.2 and a rigid uniform vocal tract tube of length 170 mm (diameter 16 mm, length 460 mm, acoustic resonance frequency 500 Hz) can be attached airtight along the superior end of the vocal folds replica. The effect of angular asymmetry α on the fluid-structure interaction is characterized by analyzing the measured upstream pressure $P_u(t)$. Upstream pressures associated with auto-oscillation onset (P_{On}) and offset (P_{Off}) are sought. Next, auto-oscillation features are analyzed for 5 s of steady-state upstream pressure signal $P_u(t)$, while the mean upstream pressure $\overline{P}_u \approx P_{On}$ holds as illustrated in Figure 4.3. First harmonic (or fundamental) frequency f_0 is extracted. Second f_1 and third f_2 harmonic frequencies of f_0 are detected when its power (in dB) yields at least half the power of f_0 . The harmonic contents of $P_u(t)$ is then further characterized by the total harmonic distortion rate THD (in dB) describe in Section 2.3.6. The power ratio $\mathcal{P}_{f_1}/\mathcal{P}_{f_0}$ between the second and first harmonics and the signal-to-noise ratio SNR detailed in Section 2.3.5 are quantified as well. All experiments are repeated 3 times so that the average



Figure 4.3 – Measured upstream pressure time signal of $P_u(t)$ for the EPI replica with $\alpha = 0^\circ$. Oscillation onset P_{On} and offset P_{Off} pressures are indicated (square) as is the steady state oscillatory signal portion ($\overline{P}_u \approx P_{On}$) used for analysis (5 s). For clarity, a zoom of $P_u(t)$ near oscillation onset (left), during steady state oscillation (middle) and near oscillation offset (right) is provided.

and standard deviation of the extracted features are plotted. Experiments are performed with and without vocal tract tube. Since their resulting features are similar, only experimental results with vocal tract tube are shown. It follows that the results are not much affected by acoustical coupling with the supra-glottal vocal tract.

4.3.2 Experimental results

Table 4.2 – Minimum and maximum onset pressures relative [%] to the value for $\alpha = 0^{\circ}$ and associated α [°].

Replica	minin	num	maximum		
Replica	[%]	[°]	[%]	[°]	
M5	-10.5	10.7	46.9	24.6	
MRI	-2.0	1.7	29.5	24.6	
EPI	-12.9	11.7	45.0	24.6	

Oscillation onset (P_{On}) and offset (P_{Off}) pressures are plotted as a function of α in Figure 4.4. Associated fundamental (f_0) frequencies and detected second (f_1) and third (f_2) harmonics are shown in Figure 4.5. The minimum and maximum onset pressures $P_{On}(\alpha)$ relative to the value



Figure 4.4 – Mean (symbols) and standard deviation (error bars) of onset P_{On} and offset P_{Off} pressures as a function of asymmetry angle α : a) M5, b) MRI and c) EPI. Critical angles (vertical lines) indicate contact regime changes between $\mathscr{G} = 1$ (I), $\mathscr{G} \in]0, 1[$ (II) and $\mathscr{G} = 0$ (III).

for symmetric angular configuration ($\alpha = 0^{\circ}$) are quantified in Table 4.2. The minimum of the onset pressure (min($P_{On}(\alpha)$)) occurs for $\alpha > 0^{\circ}$: within regime I (full contact, $\mathscr{G} = 1$) for the MRI replica and within regime II (partial contact, $\mathscr{G} \in]0,1[$) for the M5 and EPI replicas. Nevertheless, the amount of decrease is limited to less than 13% for all replicas. The maximum of the onset pressure (max($P_{On}(\alpha)$)) occurs near the largest asymmetry angle 24.6° and the relative



Figure 4.5 – Mean (symbols) and standard deviation (error bars) of harmonic frequencies (f_N) as a function of asymmetry angle α : a) M5, b) MRI and c) EPI. Critical angles (vertical lines) indicate contact regime changes between $\mathscr{G} = 1$ (I), $\mathscr{G} \in]0, 1[$ (II) and $\mathscr{G} = 0$ (III). Horizontal dashed lines indicate mechanical resonance frequencies $f_{1,2}^M$.

increase yields more than 29.5% for all replicas. For all replicas, the onset pressure increases for $\alpha \ge 15^{\circ}$. For $\alpha < 15^{\circ}$, the shape of the $P_{On}(\alpha)$ curves varies between replicas, because of its dependence on the detailed geometry and layer composition of the individual replica. It is observed that more realistic replicas (MRI and EPI) exhibit a more monotonic tendency than the

62 Chapter 4. Influence of level difference due to vocal folds vertical angular asymmetry



Figure 4.6 – Mean (symbols) and standard deviation (error bars) of total harmonic distortion (THD in [dB]) as a function of asymmetry angle α for all replicas. Critical angles (vertical lines) indicate contact regime changes between $\mathscr{G} = 1$ (I), $\mathscr{G} \in]0,1[$ (II) and $\mathscr{G} = 0$ (III).

M5 replica. For all replicas, hysteresis between the onset and offset pressures is observed as expected. The degree of hysteresis is only marginally affected by α in contact regimes I and II (all replicas: M5, MRI and EPI), whereas it is seen to be reinforced in contact regime III (only EPI replica).

Table 4.3 – Overall decrease of fundamental frequency f_0 relative [%] to the value for $\alpha = 0^\circ$ in the full contact (I), partial contact (II) and no contact (III) regime.

Replica	I	II	III
M5	0.3	-10.3	not measured
MRI	-2.4	-19.8	not measured
EPI	-0.5	-22.8	-34.2

For symmetric angular configuration ($\alpha = 0^{\circ}$), the detected oscillation frequencies f_0 and f_1 are closely located to the first and second mechanical resonance frequencies given in Table C.1. This suggests that the auto-oscillation resulting from the fluid-structure interaction is aerodynamically driven. Thereafter, the fundamental frequency f_0 decreases monotonically with α . The overall decrease of f_0 relative to its value for symmetric configuration ($\alpha = 0^{\circ}$) is quantified in Table 4.3 for all three replicas. It follows that the decrease is negligible (<3%) in contact regime I (full contact, $\mathscr{G} = 1$) and becomes more pronounced (> 10%) in regimes II (partial contact, $\mathscr{G} \in]0,1[$) and III (no contact, $\mathscr{G} = 0$). The second harmonic f_1 is detected for all replicas and all asymmetry angles α of regimes I and II. For the EPI replica, the third harmonic f_2 is detected for all α in regime III, indicating that the harmonic power increases with α .

The increased power of harmonics can be confirmed by the total harmonic distortion rate (THD) plotted in Figure 4.6. For all replicas, the THD increases as the asymmetry angle α is in-



Figure 4.7 – Measured upstream pressure $P_u(t)$ for EPI steady state oscillation: a) $\alpha = 0^\circ$ and b) $\alpha = 20^\circ$.



Figure 4.8 – Harmonics power ratio $\mathscr{P}_{f_1}/\mathscr{P}_{f_0}$ as a function of asymmetry angle α for M5, MRI and EPI. Critical angles (vertical lines) indicate contact regime changes between $\mathscr{G} = 1$ (I), $\mathscr{G} \in]0,1[$ (II) and $\mathscr{G} = 0$ (III).

creased. The increase is most prominent in contact regime II ($\mathscr{G} \in]0, 1[$) as it yields about 10 dB for the EPI replica and about 25 dB for the M5 and MRI replicas. The imprint of higher harmonics on the temporal waveform shape of steady state $P_u(t)$ for EPI is illustrated in Figure 4.7 for $\alpha = 0^\circ$ (few harmonics, low THD) and $\alpha = 20^\circ$ (rich harmonic contents, large THD). The ratio between first and second harmonic powers $\mathscr{P}_{f_1}/\mathscr{P}_{f_0}$ is plotted in Figure 4.8. It is seen that the ratio increases with α towards 1 for all replicas in regime II ($\mathscr{G} \in]0, 1[$), indicating that both harmonics contribute almost equally. Figure 4.9 shows that, for all replicas, increased THD is accompanied by a reduced signal-to-noise (SNR) ratio of about 15 dB.

For all three replicas, angular asymmetry within contact regime I (full contact) showed only a minor influence on onset pressures P_{On} , oscillation frequencies f_0 , THD and SNR, which are





Figure 4.9 – Mean (symbols) and standard deviation (error bars) of signal-to-noise ratio (SNR in [dB]) as a function of asymmetry angle α for all replicas. Critical angles (vertical lines) indicate contact regime changes between $\mathscr{G} = 1$ (I), $\mathscr{G} \in]0,1[$ (II) and $\mathscr{G} = 0$ (III).

close to those observed for symmetric angular configuration ($\alpha = 0^{\circ}$). In contact regime II (up to $\alpha \le 15^{\circ}$), the THD and SNR start to deteriorate as α increases. For $\alpha > 15^{\circ}$, all features (P_{On} , f_0 , THD and SNR) are affected by the angular asymmetry.

Some of our observations may elucidate medical findings reported on human patients having unilateral vocal fold paralysis [3]. For instance, reduced SNR due to air leakage in contact regime II and III through the glottal gap is consistent with a breathy voice description. Increased oscillation onset pressure P_{On} in contact regimes II and III may cause vocal fatigue. The decrease in fundamental frequency f_0 , on the other hand, does not follow the unnatural highpitched voice described for unilateral vocal fold paralysis. This might be due to several reasons such as the increased contribution of the observed higher harmonics to the perceived pitch, mechanical property changes in the paralysed vocal fold, which are not considered in the current work or yet compensatory strategies involving supra-laryngeal structures of human speakers. In our studied range of asymmetry angles, aphonia (or oscillation death) could not be reproduced. Finally, it is noted from Figure 4.2(c) that α -ranges, for which $h_A \leq 3$ mm, *i.e.* vertical level difference observed during human phonation [37], extend up to contact regime II for all three replicas: $\alpha \leq 13^\circ$ for EPI and $\alpha \leq 17^\circ$ for both EPI and MRI. Therefore, partial contact (regime II) is of particular importance for phonation with unilateral vocal fold paralysis.

4.4 Parallel level difference analogy from *P*_{On}

In [5, 6], a parallel level difference ΔE was imposed by varying the distance in the inferiorsuperior direction between vocal folds, which are parallelly located on the transverse plane. A theoretical expression of oscillation onset threshold pressure P_{On} as a function of parallel level

Table 4.4 – Overview for $\alpha = 24.6^{\circ}$ of leakage area *A* and potential parallel level difference analogy approximations $\Delta E \approx h_A$, $\Delta E \approx h_L$ and $\Delta E \approx h_{l_A}$.

Replica	$A \text{ [mm^2]}$	<i>h_A</i> [mm]	<i>h</i> _{<i>L</i>} [mm]	h_{l_A} [mm]
M5	44.0	5.8	2.6	2.9
MRI	50.5	6.2	2.8	3.1
EPI	74.2	7.5	4.4	3.7

difference ΔE was derived [5], and detailed in Appendix F, using the following assumptions: small amplitude approximation of the vocal folds oscillation, no acoustical coupling from suband supra-glottal systems and mechanical left-right vocal folds symmetry. Oscillation threshold pressure P_{On} was than predicted to increase with parallel level difference ΔE as

$$P_{On}(\Delta E) = \frac{\tilde{w}\tilde{E}}{\left(\tilde{E} - \Delta E\right)^2},\tag{4.5}$$

with model parameter set $\{\tilde{w}, \tilde{E}\}$. The parameter \tilde{w} accounts for mechanical damping per unit area, mucosal surface wave velocity and pre-phonatory glottal half-width. The parameter \tilde{E} corresponds to the largest vocal fold thickness.

The glottal leakage due to the parallel level difference is thus rectangular, when imposing ΔE . This is in contrast to the triangular leakage area formed by the angular level difference. Consequently, when imposing ΔE , only regime I (full contact) and regime III (no contact) occur for the parallel level difference. When imposing asymmetry angle α , on the other hand, partial contact (regime II) can also happen. Nevertheless, major tendencies observed by increasing asymmetry angle α (Section 4.3.2), such as decrease in fundamental frequency f_0 and overall increase in onset pressure P_{On} , have been also reported when parallel level difference ΔE was increased [5, 6]. Therefore, a parallel level difference analogy is sought expressing angular level difference in terms of an equivalent parallel level difference, *i.e.* to establish an approximation $\Delta E(\alpha)$. While $A(\alpha) = 0$ holds in contact regime I, the approximation yields $\Delta E(\alpha) = 0$. For $A(\alpha) > 0$, in contact regimes II and III, *a-priori* three formulae are considered to describe the sought approximation:

- $\Delta E(\alpha) \approx h_A(\alpha)$ with triangle height $h_A(\alpha)$ given in Eq. (4.4) and depicted in Figure 4.2,
- $\Delta E(\alpha) \approx h_L(\alpha)$ with $h_L(\alpha) = A(\alpha)/L$, *i.e.* the height of the rectangle obtained from glottal gap area $A(\alpha)$ and width L,
- $\Delta E(\alpha) \approx h_{l_A}(\alpha)$ with $h_{l_A}(\alpha) = A(\alpha)/l_A(\alpha)$, *i.e.* the height of the rectangle obtained from glottal gap area $A(\alpha)$ and width $l_A(\alpha)$.

	M5	EPI	MRI
$\Delta E(\alpha) \approx h_L(\alpha)$	85%	84%	74%
$\Delta E(\boldsymbol{\alpha}) \approx n_{l_A}(\boldsymbol{\alpha})$	89%	89%	81%

Table 4.5 – Fit accuracy (R^2 in %) with $\Delta E(\alpha) \approx h_L(\alpha)$ and with $\Delta E(\alpha) \approx h_{l_A}(\alpha)$.

An overview of resulting maxima for h_A , h_L and h_{l_A} associated with the largest asymmetry angle $\alpha = 24.6^{\circ}$ is given in Table 4.4. Maximum values are within the ΔE -range assessed in experimental studies imposing a parallel level difference [5, 6]. It is seen that h_L and h_{l_A} are of the same magnitude (up to 3 mm) as vertical level differences reported during human phonation [37] whereas h_A exceeds this range.

In order to identify the most appropriate approximation for $\Delta E(\alpha)$, Eq. (4.5) is fitted to data sets $P_{On}(h_A(\alpha))$, $P_{On}(h_L(\alpha))$ and $P_{On}(h_{l_A}(\alpha))$, respectively. It is noted that the assumption of mechanical left-right vocal folds symmetry underlying Eq. (4.5) is justified as only asymmetry angle α is varied. The same way, neglecting acoustic coupling is justified since it was argued, seen in Section 4.3 and Figure 4.5, that the oscillation resulting from the fluid-structure interaction is aerodynamically driven.

At first, approximations $\Delta E(\alpha) \approx h_A(\alpha)$ and $\Delta E(\alpha) \approx h_{l_A}(\alpha)$ are assessed. Approximation $\Delta E(\alpha) \approx h_A(\alpha)$ is eliminated since it relies on the triangular leakage area characterizing angular level difference, whereas a rectangular area is inherent to a parallel level difference approach. Moreover $h_A(\alpha)$ exceeds the range of vertical level difference (up to 3 mm) reported for human phonation (Table 4.4). It is noted that the fit accuracy can not distinguish between both approximations. Indeed, Eq. (4.4) implies that $h_A(\alpha) = 2 \times h_{l_A}(\alpha)$, so that from Eq. (4.5) follows that estimated parameter sets are proportional $(\{\tilde{w}, \tilde{E}\}_{h_A} = 2 \times \{\tilde{w}, \tilde{E}\}_{h_{l_A}})$ whereas fitted values are the same so that both approximations result in the same fit accuracy.

Next, Table 4.5 compares fitting accuracies (coefficient of determination R^2) of Eq. (4.5) for approximations $\Delta E \approx h_{l_A}$ and $\Delta E \approx h_L$. For all replicas, the best fit is obtained with approximation $\Delta E \approx h_{l_A}(\alpha)$. This implies that the onset pressure may depend upon the angular level difference as

$$P_{On}(h_{l_A}(\alpha)) = \frac{\tilde{w}E}{\left(\tilde{E} - h_{l_A}(\alpha)\right)^2}.$$
(4.6)

Based on this formula, the fitted curves for $P_{On}(h_{l_A}(\alpha))$ are drawn simultaneously with the experimental data in Figure 4.10. The vocal fold thickness \tilde{E} is estimated as 12 mm for M5, 26 mm for MRI and 15 mm for EPI. Compared to the real values of $E \approx 10$ mm (see Table 4.1), the estimates are of the same order of magnitude as the experimental measurements.



Figure 4.10 – Measured (symbols) $P_{On}(h_{l_A}(\alpha))$ and fitted (full lines) with Eq. (4.6) for all replicas.

4.5 Conclusion

An experimental study is presented on the effect of angular level difference on the autooscillation for three mechanical vocal folds replicas. Spectral features of the oscillations (harmonic frequencies, THD and SNR) as well as upstream threshold pressures (oscillation onset and offset) are analysed. The same tendencies are observed for all three replicas. As the asymmetry angle α is increased so that vocal folds are no longer in full contact (contact regimes II and III), dynamics of the vocal folds replica are altered clearly: SNR decreases, THD increases, oscillation threshold pressures increase, fundamental frequencies decrease and higher harmonics emerge. Geometrical details of each vocal folds replica determine the leakage area and critical asymmetry angles associated with a shift in contact regime. Expressions are given to quantify these features from geometrical vocal fold parameters. Apart from the decrease in fundamental frequency, observed tendencies are in good agreement with those reported in clinical studies on vertical level difference. The same geometrical reasoning and expressions might be applied to characterize the glottal gap during phonation of patients with unilateral vocal fold paralysis. Based on analogy to parallel level difference, a formula is proposed to describe the oscillation onset pressure as a function of the asymmetry angle.

Up to date, few measurements have been reported on level difference for UVPF. Various attempts have been recently made to measure three-dimensional *in-vivo* geometry of the vocal folds, *e.g.*, laser systems [98, 99] and stereo endoscopy [100]. Such advanced measurements may eventually provide more precise information about the left-right angular asymmetry of patients with voice pathology. Geometrical expressions derived in this study ($A(\alpha)$ in Eq. (4.2), $l_A(\alpha)$ in Eq. (4.3), *etc.*) could be applied to characterize vocal folds contact (\mathscr{G} and regime) for patients. This study on the oscillation onset and voice quality could be of great use as a reference for breathy voice, vocal fatigue, and other symptoms of voice disorders.

The influence of the glottal geometry and by extent the glottal area on oscillation features was

68 Chapter 4. Influence of level difference due to vocal folds vertical angular asymmetry

considered in this chapter. In the next chapter, the presence of liquid is experimentally studied as it is rarely accounted for in physical studies whereas it is likely to affect oscillation, as introduced in Section 2.4.2. In order to quantify to which extent the presence of liquid is important, at first oscillation features considered in this chapter are also assessed in the next chapter. In addition the effect of water spraying on non-linearities is quantified.

Experimental study of water influence

Based on:

Bouvet A., Pelorson X., Van Hirtum A., 2020. *Influence of water spraying on an oscillating channel*. Journal of Fluids and Structures, 93:102840(20).

Van Hirtum A., Bouvet A., and Pelorson X., 2019. *Quantifying the oscillation complexity following water spraying with interest for phonation*. Physical Review E, 100:043111(7).

Specific nomenclature

Т	Oscillation period	S
T^t	Total close phase of the period	S
T^c	Closing phase	S
T^{o}	Opening phase	S
A	Amplitude signal	Pa
F_A	Fluid force	Ν
V_L	Sprayed liquid	mL
m	Vibrating vocal fold mass	g
m_{V_L}	Liquid mass on the vocal fold	g
Q_{max}	Volume flow rate	m ³ /s
C_{S}	Separation constant	[-]
$U_{P_u,max}$	Maximum airflow velocity	m/s
t_L	Marker time series	[-]
Δt_L	Marker duration	S
Φ_L	Water injection flow rate	mL/s
U_L	Liquid bulk velocity	m/s
We	Weber number	[-]
t_d	Droplet relaxation time	S
Stk	Stokes number	[-]
$lpha_{m,L}$	Mass factors	[-]
l_2	Euclidean norm	[-]
C_m	Correlation sum	[-]
$R_{i,j}$	Recurrence plot	[-]
D	Degree of determination	[-]
\mathscr{R}	Rate of states	[-]
γ	\mathscr{D} and \mathscr{R} ratio	[-]
Ε	Entropy	[-]
T_d	Oscillation period for $V_L=0$ mL	S

5.1 Introduction

The surface of human vocal folds, illustrated in Figure 2.2(b), is hydrated with a thin (≤ 0.9 mm) liquid layer [19, 38]. It is established from observations on human speakers and on cadaver larynges that good surface hydration may diminish phonotraumatic vocal folds lesions [19, 20] Artificial saliva sprays are developed to remedy vocal folds surface hydration [13, 45, 46] so that liquid sprays and nebulisation techniques are common in (re-)hydration studies focusing on voice properties [42, 43, 46]. One of the physical consequences attributed to surface hydration is to reduce the auto-oscillation onset pressure, *i.e.* the minimum subglottic pressure required to sustain vocal folds auto-oscillation or phonation [20, 39–41]. Other voice auto-oscillation properties are sensitive to hydration such as fundamental frequency f_0 , closing quotient, speed quotient and cycle-to-cycle perturbations jitter and shimmer [38, 42–44]. Nevertheless, different tendencies of these features are described in literature stressing the need for physical studies. So far, few physical studies of the vocal folds instability reported in literature deal with surface hydration.

From a physical point of view, vocal folds auto-oscillation during voiced speech sound production is due to a fluid-structure instability in the larynx following the interaction of airflow coming from the lungs and the enveloping vocal folds tissues. This fluid-structure interaction mechanism can be represented as a deformable vocal folds with one mechanical degree of freedom acted on by the flow while assuming a constant time delay associated with the time needed for the deformation wave to travel along the vocal folds surface [4, 101–103]. The variation of transverse area A(t) associated with each streamwise mass position is then modeled considering the equation of transverse motion (left hand side) which is excited from its transverse rest position A_0 by an acting fluid force F_A (right hand side) so that

$$m\ddot{A}(t) + r\dot{A}(t) + k(A(t) - A_0) \approx F_A(t), \qquad (5.1)$$

with vibrating mass *m*, damping *r* and stiffness *k*. Recently, it was shown using a motion driven rigid vocal folds replica that the glottal pressure drop is affected by water spraying [104]. The mean flow was affected by the air-water mixture during the opening and closing phase of the glottal cycle in different ways leading to an altered opening-closing asymmetry. Therefore, fluid force $F_A(t)$ in Eq. (5.1) is likely to depend on the sprayed water volume which in turn affects the fluid-structure interaction and hence the vocal fold's auto-oscillation.

In this chapter, flow features and their perturbation are aimed to be quantified.

Firstly, features often associated with voice quality are systematically quantified from experiments mimicking water spraying on oscillating vocal folds firstly using the MDR vocal folds replica and secondly using deformable vocal folds replicas. Experiments with the MDR vocal folds replicas allow to systematically quantify flow features within the glottis as the pressure within the constriction is analysed. Quantification is of interest as glottal flow features are directly related to $F_A(t)$ in Eq. (5.1). Experiments with deformable mechanical vocal folds replicas, PLT and silicones, allow to systematically quantify flow features during auto-oscillation as the upstream pressure is analysed. A comparison of tendencies observed on the quantified features for the glottal flow (rigid vocal folds replica) and during auto-oscillation (deformable vocal folds replicas) is made in order to assess that in terms of Eq. (5.1) the effect of water spraying is mainly limited to the flow expressed by $F_A(t)$. Such an assumption is needed for future modelling studies of the fluid-structure interaction.

Secondly, water spraying might affect the fluid-structure interaction's complexity among others due to changes to the mean and fluctuating flow patterns. Therefore, it is also aimed to quantify the influence of water spraying on the complexity of the fluid-structure instability from a time series analysis. The correlation dimension is estimated from a recurrence analysis of the phase space trajectories. Correlation dimension D_2 allows to describe the needed degrees of freedom, *i.e.* the complexity, with a single parameter, which is not the case when quantifying classic voice signal features. D_2 has been used in turbulence studies since the groundbreaking paper by Takens [105] and introduced in human voice analysis studies [106–112]. So far, its use in voice analysis studies is mainly motivated by the development of speech synthesis and recognition technologies where speech is considered as a low dimensional nonlinear model [113–115]. Nevertheless, D_2 is also of interest from an aerodynamic point of view in order to further assess the assumption that turbulence, such as droplet induced turbulence following water spraying, affects not only fricative speech sound production, but might affect voicing as well [116]. Different mechanical vocal folds replicas and different initial pre-oscillation conditions of elasticity and glottal aperture are assessed given previous studies showing the coexistence of different vibration patterns related to voicing and voice register transitions [117].

5.2 Experimental approach

5.2.1 Surface hydration liquid: water

Distilled water (density $\rho_L = 998 \text{ kg} \cdot \text{m}^{-3}$, dynamic viscosity coefficient $\mu_L = 1.0 \times 10^{-3}$ Pa·s, temperature 22 ± 2 °C) is used as hydration liquid fluid. Water has well known properties so that it is a good reference for future studies involving different fluids.

Compared to airway mucus [118, 119] and saliva [45, 120], water has similar density, but a reduced (up to 2000-fold for airway mucus of healthy subjects) dynamic viscosity. Despite the difference in visco-elastic properties and the lack of salinity and of an elastic structural matrix, bidirectional water fluxes through the epithelium play an important role in regulating vocal folds surface mucus (≤ 0.9 mm) hydration since it consists of a superficial gel-like cover with high vis-

cosity and thin ($\approx 15 \,\mu$ m) underlying aqueous layer with low viscosity [38, 121–123]. Therefore, water is an important component of vocal folds surface liquid often used in (de-)hydration studies on human subjects [38, 44, 46] and on cadaver larynges [38, 42, 44]. Nebulization of either sterile (hypotonic) water or physiological saline solutions is reported to result in a non-statistically significant difference as a remedy for laryngeal desiccation treatment [46]. It is important to note that natural human saliva and mucus are non-Newtonian fluids [45, 118–120, 124] whereas water is a Newtonian fluid. Nevertheless, some artificial saliva sprays contain a Newtonian fluid as well [45].

Typical liquid quantities for nebulization with artificial saliva sprays yield up to 4 mL [46]. Therefore, sprayed water volume V_L is varied in the range from 0 (dry, subscript d) up to 5 mL so that the total mass of sprayed water approximates $m_{V_L} = \rho_L \cdot V_L$.

5.2.2 Dynamic mechanical vocal folds replicas

The different dynamic mechanical vocal folds replicas are assessed allowing to experimentally mimic the glottal area variation during human phonation expressed in Eq. (5.1): MDR replica detailed in Section 2.5.1, silicone replicas (M5, MRI and EPI) detailed in Section 2.5.3 and PLT detailed in Section 2.5.2. Both deformable vocal folds replica types are of interest for physical studies. Indeed, deformable silicone (moderately non-water wet wettability) vocal folds replicas allow to represent the multi-layer structure of human vocal fold's with different degree of complexity whereas pressurized latex (neutral wettability) tubes allow to mimic the variation of human vocal fold's elasticity in a systematic way. Concretely, 5 different PLT vocal folds replica conditions are considered – (2300,0), (2800,0), (3300,0), (2800,1) and (2800,2) – so that either o_{mm} or P_{PLT} is deflected from condition (2800,0). Mean upstream pressure $\overline{P_u}$ yielded $250 < \overline{P_u} < 480$ Pa for the PLT vocal folds replica conditions.

It is noted that water is not absorbed by any of the mechanical vocal folds replicas with smooth surface, either rigid or deformable. It is therefore assumed that their structural properties (m, r and k in Eq. (5.1)) are unaffected when in contact with water.

5.2.3 Overall setup and protocol

5.2.3.1 Airflow supply

As detailed in Section 2.6, vocal folds replicas, oriented as depicted in Figure 2.12, are inserted in the overall experimental setup illustrated in Figure 5.1. A rigid (smooth duralium, highly wettable) uniform circular channel with diameter 2.5 cm (constant area 490 mm²) is mounted vertically to the upstream (12 cm trachea inferior to the vocal folds) and downstream



Figure 5.1 – Schematic illustration of experimental setup in the medio-frontal plane indicating liquid injection time tag t_L , sprayed liquid volume V_L , upstream pressure $P_u(t)$, glottal area $A_c(t)$. Assessed vocal folds replicas: deformable (silicone and PLT) and forced motion (MDR with constriction pressure $P_c(t)$).

(11 cm vocal tract superior to the vocal folds) end of the vocal folds replica. The downstream tube length is short compared to a typical vocal tract of a male adult (\approx 18 cm) in order to avoid acoustical coupling during vocal folds oscillation. Note that from $A_{c,max}$ values given in Table 2.4 follows that the area constriction ratio, *i.e.* the ratio between glottal area and constant area of the attached upstream and downstream tubes (490 mm²), is greater than 75% at all times for all vocal folds replicas.

Continuous steady airflow (density $\rho_G = 1.2 \text{ kg} \cdot \text{m}^{-3}$, dynamic viscosity $\mu_G = 1.8 \times 10^{-5} \text{ Pa} \cdot \text{s}$, temperature $22 \pm 2^{\circ}\text{C}$) is provided along the streamwise *z*-direction by a valve (Norgren, 11-818-987) controlled air supply. Air is delivered by an air compressor (Atlas Copco GA5 FF-300-8, GA15 FF-8) connected to an upstream pressure reservoir (volume $\geq 0.22 \text{ m}^3$) filled with acoustic foam in order to avoid parasite acoustic resonances. A pressure transducer (Endevco 8507C-5, accuracy ± 5 Pa) is positioned in a pressure tap upstream of the vocal folds

replicas in order to measure upstream pressure $P_u(t)$. In the case of a motion driven rigid vocal folds replica (MDR in Section 2.5.1) an additional pressure sensor is placed at the glottal spacing so that constriction pressure $P_c(t)$ is measured as well.

During experiments, at first single phase airflow is provided in the streamwise inferiorsuperior direction (positive z-direction in Figure 5.1) so that a steady oscillatory pattern is established prior to water spraying, *i.e.* dry condition for $V_L = 0$ mL. Experiments are performed with the mean upstream pressure set just above the auto-oscillation onset pressure P_{on} , *i.e.* the minimum upstream pressure required to sustain auto-oscillation (see Section 2.7).

Prior to water spraying (for $V_L = 0 \text{ mL}$) a stable oscillation is imposed, which is characterised by mean upstream pressure \overline{P}_u and fundamental frequency $f_{0,d}$. An overestimation of volume flow rate Q_{max} is obtained as

$$Q_{max} = \sqrt{\frac{2\overline{P}_u}{\rho_G}} A_{c,max} \cdot c_s \tag{5.2}$$

with glottal area upper limit $A_{c,max}$ for each vocal folds replica listed in Table 2.4. The separation constant $c_s = 1.3$ defines a geometrical criterion for the time-varying position of flow detachment and jet formation as the lowest streamwise position z_s with $z_s > z_c$ (Figure 5.1) for which area $A_{z_s}(t) \ge c_s \cdot A_c(t)$ [47, 48].

Furthermore, a cycle-to-cycle maximum velocity at the position of jet formation is estimated from feature vector $P_{u,max}$ as

$$U_{P_{u,max}} \approx \sqrt{\frac{2P_{u,max}}{\rho_G}}.$$
(5.3)



Figure 5.2 – Graduated syringe equipped with a commercial round spray nozzle.

5.2.3.2 Water spraying

Distilled water (Section 5.2.1) is injected manually at the downstream end of the channel¹, along the negative z-direction, by emptying a graduated (accuracy 0.5 mL) syringe (diameter 14 mm, maximum volume 11 mL) equipped with a commercial round spray nozzle (diffusion angle $20^{\circ} \pm 1^{\circ}$, diameter $D_n = 0.7 \pm 0.1$ mm, length 10 mm) shown in Figure 5.2, containing a known water volume $1 \le V_L \le 5$ mL. Water is then nebulized homogeneously across the constricted area as illustrated in Figure 5.1. Furthermore, water injection is time-tagged t_L by manually operating an electrical switch at the start and end of injection. The duration of liquid injection Δt_L takes less than 3 s and water injection flow rate ϕ_L yields about $\phi_L = 2$ mL/s [104].

The liquid bulk velocity U_L at the nozzle exit approximates $U_L = 5.2$ m/s and the syringes' driving pressure in the barrel yields about 5.4 kPa. The Weber number, which compares inertia to surface tension,

$$We = \frac{\rho_L U_L^2 D_n}{\sigma} \tag{5.4}$$

for water droplets with diameter D_n issued by the spray yields $We \approx O(10^2)$ with water-air surface tension $\sigma = 72$ mN/m. It follows that the injected liquid jet is expected to break up into much smaller droplets favoring initial mixing with the airflow. The droplet relaxation time

$$t_d = \frac{\rho_L D_n^2}{18\mu_G} \tag{5.5}$$

for the initially injected liquid yields $t_d = 1.5$ s or less as droplets break up and their diameter reduces from D_n . Five seconds portions of pressure signals gathered 5 s after water spraying finished are analysed. This way the impact of surface hydration after using a hydration spray is mimicked. It is verified that no water passes through the vocal folds during oscillation. As oscillation is ongoing it follows that typical airflow velocities U_G through the glottal-like constriction area vary from 0 m/s up to $O(10^1)$ m/s [102]. Therefore, the Stokes number, characterising the behaviour of droplets suspended in the flow,

$$Stk = \frac{t_d U_G}{L_z} \tag{5.6}$$

varies with U_G during a single oscillation cycle leading to mixing of the droplets with the flow in the case of low Stokes numbers (smaller than 1) or surface hydration due to droplet deposition along the upstream pipe (vocal tract) and vocal folds for larger Stokes numbers (larger than 1).

¹Other injection types were tested: continuous injection, drop to drop etc. in [125].

5.2.4 Upstream pressure analysis

5.2.4.1 Auto-oscillation features



Figure 5.3 – Illustration of waveform shape feature vectors for f_0 cycle-to-cycle analysis for pressure $P = P_u$.

The influence of water spraying on the waveform shape is quantified by time tracking period T and associated features within each pressure P. oscillation cycle as illustrated in Figure 5.3 for upstream pressure $P = P_u$: maximum pressure $P_{.max}$, peak-to-peak pressure amplitude \mathscr{A} , closed portion T^t which is composed of a closing T^c and an opening T^o component. Closed portion T^t is defined as the duration of the period for which pressure P. exceeds a threshold value. The threshold value is set to $P_{u,max} - 0.95\mathscr{A}$ for deformable vocal folds replicas when upstream pressure P_u is analysed ($P = P_u$) and to $P_{c,max} - 0.85\mathscr{A}$ for motion driven MDR vocal folds replica when constriction pressure P_c is analysed ($P = P_c$). The part of T^t during which P(t) increases corresponds to closing portion T^c and the remaining decreasing part corresponds to opening portion T^o so that $T^t = T^c + T^o$. Note that the ratio T^c/T^o informs on closing-opening asymmetry within each cycle. Therefore T^c/T^o allows to quantify the imprint of mean flow changes, characterised by a changing closing-opening asymmetry [104].

To identify properties for any of the aforementioned temporal feature vectors, for the sake of simplicity denoted \mathscr{P} in general, its mean value $\overline{\mathscr{P}} = \frac{1}{N} \sum_{i=1}^{N} \mathscr{P}_i$ and its standard deviation

$$\Delta \mathscr{P} = \sqrt{\frac{1}{N} \sum_{i=1}^{N} |\mathscr{P}_i - \overline{\mathscr{P}}|^2}$$
(5.7)

are gathered where N = 50 indicates the used number of periods. The perturbation intensity is then obtained as $\frac{\Delta \mathscr{P}}{\mathscr{P}}$, the ratio between standard deviation $\Delta \mathscr{P}$ and mean value \mathscr{P} . It is noted that when \mathscr{P} represents a velocity time series this ratio yields the turbulence intensity. Then cycle-to-cycle perturbation jitter ζ_T and shimmer $\zeta_{mathcalA}$, signal-to-noise ratio (SNR) and total harmonic distortion (THD), characterised in Section 2.3, are quantified.

5.2.4.2 Oscillation cycle-to-cycle analysis: f_0 and f_N



(b) Subharmonic oscillation cycle regime, $f_N = f_0/2$

Figure 5.4 – Illustration of f_N cycle analysis and oscillation regimes: a) normal $f_N = f_0$, b) subharmonic $f_N = f_0/2$.

Classically, first harmonic frequency f_N corresponds to fundamental frequency f_0 of each normal oscillation cycle given as 1/T where period T corresponds to the time interval between successive waveform maxima: $P_{u,max}(i)$ and $P_{u,max}(i+1)$ as shown in Figure 5.4(a).

An alternative subharmonic oscillation cycle regime is observed when

$$\begin{cases} |P_{u,max}^{i} - P_{u,max}^{i+1}| > 50 \text{ Pa}, \\ |P_{u,max}^{i} - P_{u,max}^{i+2}| \le 50 \text{ Pa}, \end{cases}$$
(5.8)

as depicted in Figure 5.4(b). In this case, first harmonic frequency f_N corresponds to subharmonic frequency $f_0/2$ associated with period *T* between maximum $P_{u,max}^i$ and $P_{u,max}^{i+2}$. The used threshold of 50 Pa corresponds to 10% up to 20% of \overline{P}_u for the PLT vocal folds replica and < 10% for silicone vocal folds replicas (EPI, MRI and M5) (Table 2.5).

In summary, first harmonic frequency f_N of each cycle yields:

$$f_N = \begin{cases} f_0/2, & \text{if } |P_{u,max}^i - P_{u,max}^{i+1}| > 50 \text{ Pa} \\ f_0, & \text{otherwise.} \end{cases}$$
(5.9)

Consequently, besides a cycle-to-cycle analysis associated with f_0 , a cycle-to-cycle analysis associated with f_N can be performed when a subharmonic oscillation cycle regime is detected.

5.2.4.3 Period change with added mass hypothesis

The normal cycle period T, and hence fundamental frequency $f_0 = 1/T$ characterising the auto-oscillation due to the fluid-structure interaction described by Eq. (5.1), is dependent on the mechanical resonance frequency, determined by its elasticity and its mass as $T \propto \sqrt{\frac{m}{k}}$ [4, 102]. The assumption that water spraying does not affect the VF's elasticity implies that k remains constant. This is a reasonable working assumption since water does not penetrate the VF replicas. On the other hand, m increases under the hypothesis that spraying water results in added mass m_L to vibrating mass m. In general $m_L = \alpha_L m_{V_L}$ holds with $\alpha_L \leq 1$. Vibrating mass m is less than vocal folds mass m_d so that $m = \alpha_m m_d$ with $\alpha_m < 1$. In the case that m_L is added to m, normal oscillation period T deflects from its value T_d for $V_L = 0$ mL as:

$$\frac{T}{T_d} = \sqrt{1 + \frac{\alpha_L}{\alpha_m} \frac{m_{V_L}}{m_d}} \quad \text{so that} \quad T \ge T_d.$$
(5.10)

It follows that the added mass hypothesis results in an increase of $\frac{T}{T_d}$. The magnitude of the increase is then estimated setting $\alpha_L = 1$ and $\alpha_m = 1$. This way both m_L and m are overestimated so that their ratio provides a first estimation of the tendency. From Table 2.4 (for $V_L = 0$ mL) is seen that the mean mass of deformable VF replicas yields $\overline{m}_d = 2.02$ g and its standard deviation $\Delta m_d = 0.22$ g (11%).

5.2.4.4 Complexity analysis from time series

The correlation dimension D_2 relies on a phase space reconstruction of the system dynamics from a series of measurements at equally spaced intervals in time [126]. Let $P_u(n)$ with n = 1, ..., N denote the measured upstream pressure time series of length N. A set of *m*-dimensional reconstructed vectors y_i^m is generated using the method of delays as

$$\mathbf{y}_{i}^{m} = \left[P_{u}(i), P_{u}(i+\tau), P_{u}(i+2\tau), \dots, P_{u}(i+(m-1)\tau) \right]^{T},$$
(5.11)

where \cdot^{T} is the transpose operator, τ indicates the time-delay between consecutive samples in the reconstructed space and *m* denotes the embedding dimension so that $N_m = N - \tau(m-1)$ vectors can be reconstructed for each *m* dimension [105, 126, 127]. The distance $d_{i,j}^m = d(y_i^m, y_j^m)$ between each pair of reconstructed vectors is computed as the Euclidean l_2 norm of the difference
vector $\Delta \mathbf{y}_{i,j}^m = \mathbf{y}_i^m - \mathbf{y}_j^m$ so that

$$d_{i,j}^{m} = \sqrt{\sum_{k=1}^{m} \left(\mathbf{y}_{i}^{m}(k) - \mathbf{y}_{j}^{m}(k) \right)^{2}}.$$
(5.12)

The probability of the reconstructed vector pair distance being smaller than a certain threshold r is then given by correlation sum $C_m(r)$

$$C_m(r) = \frac{1}{N_m(N_m - 1)} \sum_{\substack{i,j=1\\i \neq j}}^{N_m} H\left(r - d_{i,j}^m\right)$$
(5.13)

where $H(\cdot)$ indicates the Heaviside function so that H(x) = 1 for $x \ge 0$ and H(x) = 0 for x < 0. For large enough values of r, $d_{i,j} < r$ holds $\forall \{i, j\}$ -combinations so that $C_m(r) = 1$. It follows that for deterministic systems $C_m(r)$ decreases monotonically from 1 towards 0 as r approaches 0. Therefore, assuming $C_m(r) \approx r^{D_2^m}$ results in correlation dimension

$$D_2^m = \lim_{r \to 0} \frac{\log C_m(r)}{\log r}.$$
(5.14)

The slope in the linear region of the $\log C_m(r)$ versus $\log r$ domain is considered as an estimation of D_2^m [126, 128–130]. For increasing embedding dimension m, D_2^m converges to a finite value D_2 . As the system is more complex, *i.e.* as more degrees of freedom are required to describe the system's dynamics, D_2 is larger.

Time-delay τ is estimated as the first local minimum of the mutual information [126, 129] and the required minimal embedding dimension *m* is obtained using the method of false nearest neighbours [126, 131, 132]. Correlation dimension D_2 is then estimated from a recurrence analysis of the N_m reconstructed phase space trajectories of upstream pressure P_u [126, 130, 133]. The recurrence analysis is based on recurrence plots $R_{i,j}$ defined by the summation terms in Eq. (5.13) as

$$R_{i,j} = H(r - d_{i,j}^m).$$
(5.15)

Recurrence plots $R_{i,j}$ represent the number and duration of recurrent states and as such expresses the degree of predictability inherent to the system. For a fully predictable, *i.e.* deterministic, oscillating system states are recurrent as initially neighbouring states remain so over time and $R_{i,j}$ is characterised by uninterrupted diagonal lines. As chaos is introduced in the system initially neighbouring states will eventually diverge in time when the system loses its predictability so that $R_{i,j}$ is characterised by diagonal lines of finite length *l*. The shorter *l* the faster the states diverge and the more the system approaches a chaotic regime. Recurrence plots $R_{i,j}$ are then quantified considering the degree of determinism \mathcal{D} , the recurrence rate of states \mathcal{R} and their ratio $\gamma = \mathscr{D}/\mathscr{R}$ defined as

$$\mathscr{D} = \frac{\sum_{l=l_{min}}^{N_m} l \mathscr{P}(l)}{\sum_{l=1}^{N} l \mathscr{P}(l)},$$
(5.16)

$$\mathscr{R} = \frac{1}{N_m^2} \sum_{i,j=1}^{N_m} R_{i,j},$$
(5.17)

$$\gamma = N_m^2 \frac{\sum_{l=l_{min}}^{N_m} l \mathscr{P}(l)}{\left(\sum_{l=1}^{N_m} l \mathscr{P}(l)\right)^2},\tag{5.18}$$

with $\mathscr{P}(l)$ the histogram of diagonal lines with lengths l and minimum line length $l_{min} \ge 2$. In addition, entropy E is obtained from the probability distribution of the diagonal line lengths p(l) as

$$E = -\sum_{l=l_{min}}^{N_m} p(l) \log_2 p(l).$$
(5.19)

5.3 **Results: voice quality features**

5.3.1 Fluid-structure interaction without liquid spraying: $V_L = 0$ mL

During experiments for $V_L = 0$ mL with deformable vocal folds replicas (PLT and silicone), mean upstream pressure \overline{P}_u is set above the onset threshold pressure so that \overline{P}_u differs for each vocal folds replica as does fundamental frequency $f_{0,d}$. Table 2.4 summarises operating conditions for all vocal folds replicas, deformable (MRI, EPI, M5 and PLT) as well as motion-driven rigid (MDR), for mean upstream pressure \overline{P}_u , fundamental frequency $f_{0,d}$, Reynolds number *Re* (Section 2.3.2) and Strouhal number *Sr* (Section 2.3.3). As a reference, typical values observed for airflow through the glottis of a male adult speaker are also indicated in Table 2.5 [30, 52, 61].

5.3.2 Flow through glottal constriction: $V_L > 0$ mL

Figure 5.6(a) illustrates measured upstream pressure $P_u(t)$, constriction pressure $P_c(t)$, water spraying time tag $t_L(t)$ and imposed minimum spacing $h_c(t)$ for $V_L = 4$ mL and forcing frequency $f_c = 1$ Hz near the instant of water spraying. In [104] it was qualitatively shown that when water was sprayed ($V_L > 0$ mL) viscous two-phase mixing affects $P_c(t)$ near closure, *i.e.* during both the opening and closing phase of the oscillation, which results in increased amplitude \mathscr{A} ,



Figure 5.5 – Deformable vocal folds replicas in parameter space $(f_{0,d}, \overline{P}_u)$ for $V_L = 0$ mL: silicone vocal folds replicas (M5, MRI, EPI: filled rectangles) and PLT vocal folds replica conditions $(P_{PLT}, 'o_{mm})$ for varied P_{PLT} (full line rectangles) or varied o_{mm} (dashed rectangles) from reference condition (2800,0).

extended duration of pressure increase near closure T^t and different flow regimes in the opening and closing phase affecting closing-opening asymmetry T^c/T^o . Quantitative evidence of the effect of V_L on \mathscr{A} , T^t and T^c/T^o with respect to their values for single-phase airflow \mathscr{A}_d , T^t_d and $(T^c/T^o)_d$ (subscript d, $V_L = 0$ mL) is provided in Figure 5.6(b), Figure 5.6(c) and Figure 5.6(d).

Both mean (symbols) as standard deviation (vertical error bars) of \mathscr{A} and T^t augments with V_L up to 10% or more. The standard deviation of T^c/T^o increases with V_L as well whereas mean T^c/T^o first increases $V_L \in \{1,2\}$ mL and then decreases for $V_L > 2$ mL up to > 15% so that opening-closing asymmetry is affected. The increased perturbation points to an increased complexity and fluctuation of P_c , *e.g.* due to irregular P_c spikes associated with droplet splashing as observed in Figure 5.6(a) and discussed in [104].

5.3.3 Fluid-structure interaction with auto-oscillation: $V_L > 0$ mL

5.3.3.1 Measured upstream pressure

A typical measurement of upstream pressure $P_u(t)$ and water spraying tag $t_L(t)$ near the instant of water spraying for a deformable vocal folds replica (PLT) is plotted in Figure 5.7(a). The influence of water spraying on the temporal upstream pressure signal P_u is shown in more detail in Figure 5.7(b) (pressurized tube vocal folds replica PLT for (3300,0): $P_{PLT} = 3300$ Pa, screw opening 0 mm) and in Figure 5.7(c) (silicone vocal folds replica: EPI) where P_u before



Figure 5.6 – Motion driven rigid MDR vocal folds replica: a) measurement, b) $\mathscr{A}(V_L)/\mathscr{A}_d$ from $P_c(t)$, c) $T^t(V_L)/T^t_d$ from $P_c(t)$ and d) $(T^c/T^o)/(T^c/T^o)_d$ from $P_c(t)$. Subscript *d* indicates $V_L = 0$ mL. Mean (symbols) and standard deviation (vertical bars) are plotted in b), c) and d). For clarity data for $\{1, 10\}$ Hz are shifted around V_L .



(c) EPI, $P_u(t \cdot f_N)$ for $V_L \leq 5$ mL

Figure 5.7 – Deformable vocal folds replicas: a) measurement, b) $P_u(t \cdot f_N)$ for PLT condition (3300,0) and $V_L \le 5$, c) $P_u(t \cdot f_N)$ for EPI and $V_L \le 5$. b,c) Time is normalised with f_N yielding $f_{0(,d)}$ or $f_0/2$. Curves for $V_L \ge 3$ mL are upshifted, $V_L \ge 2$ mL (dotted lines) and $V_L = 0$ mL (full line).

 $(V_L = 0 \text{ mL}, \text{ full line})$ and after $(V_L > 0 \text{ mL}, \text{ dotted line})$ water spraying are plotted. Note that for clarity pressure signals P_u for $V_L \ge 3 \text{ mL}$ are upshifted. The first harmonic frequency f_N (Section 5.2.4.2) is used for time normalization and varies with V_L as indicated in the legends. Note that depening on V_L a subharmonic oscillation cycle regime becomes apparent for the PLT replica. The amplitude of P_u is affected although the influence is small for EPI (Figure 5.7(c)) and more prominent for PLT (Figure 5.7(b)).

5.3.3.2 Cycle-to-cycle analysis associated with f_0

The normal oscillation cycle regime, determined by fundamental frequency f_0 , is analysed. Upstream pressure features outlined in Section 5.2.4.1 are determined as a function of V_L in order to assess the impact of water spraying on cycle-to-cycle characteristics associated with $f_{0,d}$, *i.e.* for $V_L = 0$ mL. For clarity, features for different deformable replicas are plotted in an interval centered around each V_L value.

Period *T* :

Mean and standard deviation of period T associated with fundamental frequency f_0 is presented in Figure 5.8 as a function of water volume V_L . Plotted values are normalised with mean period T_d for $V_L = 0$ mL.

Mean values (Figure 5.8(a)) for EPI and MRI show a slight (less than 10%) increase with V_L . For M5 and PLT (P_{PLT} , o_{mm})-conditions mean $T/T_d \approx 1$ so that fundamental frequency f_0 is mostly unaffected except for the condition with the largest screw opening $o_{mm} = 2$ mm in (2800,2) for which *T* increases with $\approx 35\%$ when $V_L \ge 2$ mL. Observed T/T_d do not agree with the gradual and significant (up to 100%) increase predicted using Eq. (5.10) with $\alpha_m = 1$ and $\alpha_L = 1$. It follows that the added mass hypothesis (Section 5.2.4.3) can not explain T/T_d so that observed changes to the oscillation period or frequency are not due to changes of the vibrating mass *m*.

The standard deviation of T (error bars in Figure 5.8(a)) and cycle-to-cycle perturbation ζ_T (Figure 5.8(b)) increase with V_L . The increase is limited (< 10%) for silicone vocal folds replicas EPI, MRI and M5 regardless of V_L and is more pronounced (between 10% and 30%) for $V_L \ge 2$ mL in the case of PLT vocal folds replicas. Overall, ζ_T increases as the PLT vocal folds replica condition ($P_{PLT,o_{mm}}$) becomes more elastic, *i.e.* decreasing $P_{PLT} \in \{2300, 2800, 3300\}$ Pa and $o_{mm} = 0$ mm. No clear tendency is observed with respect to the initial aperture condition o_{mm} .

Amplitude \mathscr{A} and turbulence intensity:

Mean and standard deviation of auto-oscillation amplitude \mathscr{A} as a function of water volume



Figure 5.8 – Mean (symbols) and standard deviation (vertical bars) as a function of sprayed water volume V_L : a) normalized period T/T_d and added mass hypothesis (dash-dotted line) with $m_d = \overline{m}_d$ (Eq. (5.10) for $\alpha_L = 1$ and $\alpha_m = 1$) and $m_d = \overline{m}_d \pm \Delta m_d$ (shaded region), b) period perturbation ζ_T . Values from f_0 cycle-to-cycle analysis.

 V_L are shown in Figure 5.9(a). Amplitude values are normalised with respect to \mathcal{A}_d , *i.e.* mean amplitude for $V_L = 0$ mL. For the PLT replica conditions with limited period increase $(T/T_d$ increase < 10%), a large amplitude increase (between 10% and 120%) and large standard deviations (up to 40%) are observed whereas they are less than 15% for the remaining PLT condition (2800,2). For all silicone vocal folds replicas standard deviations of \mathcal{A} are negligible and the mean value increase is limited to < 10% for MRI and M5 and more pronounced up to 35% for EPI.

Cycle-to-cycle perturbation ζ_A (Figure 5.9(b)) remains less than 2% for all silicone vocal folds replicas and increases up to 60% for PLT replica conditions. The same tendency is observed when considering turbulence intensity $\Delta U_{P_{u,max}}/\overline{U}_{P_{u,max}}$ where $U_{P_{u,max}}$ is estimated for each cycle



Figure 5.9 – Mean (symbols) and standard deviation (vertical bars) as a function of sprayed water volume V_L : a) normalized amplitude $\mathscr{A}/\mathscr{A}_d$, b) amplitude perturbation $\zeta_{\mathscr{A}}$, c) turbulence intensity $\Delta U_{P_{u,max}}/\overline{U}_{P_{u,max}}$.

using Eq. (5.3). The turbulence intensity yields less than 1% for all silicone vocal folds replicas and increases with V_L up to 15% for PLT replica conditions. Overall perturbation measures ζ_A and $\Delta U_{P_{u,max}}/\overline{U}_{P_{u,max}}$ increases as the PLT vocal folds replica condition $(P_{PLT,o_{mm}})$ becomes more elastic (decreasing $P_{PLT} \in \{3300, 2800, 2300\}$ Pa and $o_{mm} = 0$ mm) and as the initial aperture decreases (decreasing $o_{mm} \in \{2, 1, 0\}$ mm and $P_{PLT} = 2800$ Pa).



Figure 5.10 – Mean (symbols) and standard deviation (vertical bars) as a function of sprayed water volume V_L : a) normalised closed portion $(T^t/T)/(T^t/T)_d$, b) closing-opening asymmetry $(T^c/T^o)/(T^c/T^o)_d$.

Closing and opening $T^t = T^c + T^o$:

Mean and standard deviation of T^t/T , expressing the ratio between the duration of the pressure peak associated with closing and opening and period T, is plotted in Figure 5.10. Shown values are normalised by $(T^t/T)_d$ for $V_L = 0$ mL. For silicone replicas T^t/T is unaffected by V_L since $T^t/T \approx 1$ and standard deviations are less than 3%. For PLT conditions it is seen that ratio T^t/T decreases compared to its value for $V_L = 0$ mL while the standard deviation increases to 20%. A different tendency is again observed for PLT condition (2800,2) for which the ratio yields $\approx 10\%$ for $V_L > 0$ mL.

Closing-opening asymmetry T^c/T^o associated with T^t is plotted in Figure 5.10(b). Different tendencies are observed for the silicone vocal folds replicas and PLT replica conditions. For silicone vocal folds replica conditions $T^c \ge 1.2T^o$ and its value either remains (EPI for which $T^c/T^o \approx 2.3$) or increases with V_L (MRI and M5) so that the closing phase dominates the opening phase regardless of V_L . The opposite tendency is found for the PLT vocal folds replica conditions since the range for $V_L = 0$ mL ($T^c/T^o \in [0.7 \ 1.2]$ reduces with V_L to $T^c/T^o \approx 0.8 \pm 0.1$. Consequently, for PLT replica conditions the opening phase is more prominent than the closing phase. Note that these tendencies are qualitatively observed from upstream pressure time signals plotted in Figure 5.7(b) for PLT condition (3300,0) and in Figure 5.7(c) for the silicone EPI vocal folds replica.

5.3.3.3 Cycle-to-cycle analysis associated with f_N

Evidence of subharmonic oscillation regime:

The first harmonic frequency f_N is determined as outlined in Section 5.2.4.2. For silicone vocal folds replicas (EPI, MRI and M5) the first harmonic frequency (Eq. (5.9)) equals the fundamental frequency $(f_N = f_0)$ for all V_L (see *e.g.* Figure 5.7(c) for the EPI vocal folds replica), so that the power ratio between subharmonic frequency $f_0/2$ and fundamental frequency f_0 yields zero, *i.e.* $P_{f_0/2}/P_{f_0} = 0$, and hence $1 - P_{f_0/2}/P_{f_0} = 1$ (or 100%). For PLT vocal folds conditions on the other hand, first harmonic frequency f_N shifts from fundamental frequency f_0 to subharmonic frequency $f_0/2$ as V_L increases, which is illustrated for (3300,0) in Figure 5.7(b). Values of $1 - P_{f_0/2}/P_{f_0}$ for all PLT replica conditions as a function of V_L are plotted in Figure 5.11(a). The ratio decreases from 100% for $V_L = 0$ mL, when no subharmonic f_N shifts from fundamental frequency f_0 to subharmonic $f_0/2$ as V_L increases depends on the assessed PLT replica condition (P_{PLT}, o_{mm}) . Note that since $1 - P_{f_0/2}/P_{f_0} > 0$, it implies that in all cases fundamental frequency f_0 has more energy than subharmonic frequency $f_0/2$.

Period *T* :

Figure 5.11(b) shows the probability distribution of *T* as a function of V_L for PLT condition (2800,0). Period *T* is detected as outlined in Section 5.2.4.2. Two distribution peaks are observed. The first peak associated with $f_N = f_0$ for $T \approx 0.01$ s is dominant for $V_L < 3$ mL and the second peak associated with $f_N = f_0/2$ for $T \approx 0.02$ s is dominant for $V_L > 3$ mL. For $V_L = 3$ mL the maximum probability of both peaks is similar. Furthermore, it is noted that both peaks have similar width indicating that the variation around the mean peak value is similar as well.

A cycle-to-cycle analysis for T/T_d associated with first harmonic frequency f_N is plotted in



(b) T probability for (2800, 0)

Figure 5.11 – Evidence for subharmonic frequency $f_0/2$ as a function of V_L : a) subharmonic-tofundamental power ratio $1 - P_{f_0/2}/P_{f_0}$, b) probability of T from f_N cycle-to-cycle analysis for PLT condition (2800,0).



Figure 5.12 – Mean (symbols) and standard deviation (vertical bars) as a function of sprayed water volume V_L : a) normalized period T/T_d , b) period perturbation ζ_T . Values from f_N cycleto-cycle analysis.

Figure 5.12(a) and associated cycle-to-cycle perturbation ζ_T is shown in Figure 5.12(b). For silicone vocal folds replicas (M5, MRI, EPI) $f_N = f_0$ so that plotted values are similar to those in Figure 5.8(a). For PLT replica conditions first harmonic frequency f_N equals either fundamental frequency f_0 or subharmonic frequency $f_0/2$ as illustrated in Figure 5.11(b) for PLT condition (2800,0). Consequently, mean T/T_d -values gradually increase with V_L towards $T/T_d \approx 2$ expressing the growing weight of subharmonic frequency $f_0/2$ on mean values of T/T_d as is illustrated in Figure 5.11.

Overall, for PLT replica conditions standard deviations of T/T_d (Figure 5.12(a)) and cycleto-cycle perturbations ζ_T (Figure 5.12(b)) for $V_L \ge 1$ mL associated with a f_N cycle-to-cycle analysis are increased compared to those obtained from a f_0 cycle-to-cycle analysis (Figure 5.8). This increase is among others due to the first harmonic frequency f_N shifting between f_0 and $f_0/2$ (see Figure 5.11(b)).



Figure 5.13 – Overall spectral features: a) total harmonic distortion rate THD, b) signal-to-noise ratio SNR.

Overall spectral features THD and SNR:

Changes to the harmonic contents with respect to the lowest first harmonic frequency f_N of upstream pressure P_u with V_L are further quantified by the total harmonic distortion rate THD (Eq. (2.7)) depicted in Figure 5.13(a). For silicone replicas (EPI, MRI and M5) the increase is limited since it is most notable for MRI (\approx 4 dB increase). For PLT replica conditions an overall increase (> 4 dB) with V_L is observed. Nevertheless, the increase is not monotonous since for $0 < V_L \leq 3$ mL the subharmonic frequency $f_0/2$ gradually imprints on the harmonic composition with a rate depending on initial conditions (P_{PLT}, o_{mm}). Overall, it is seen that THD increases with V_L for al vocal folds replicas so that for all vocal folds replicas harmonics are generated with V_L . From Figure 5.13(b) is seen that the signal-to-noise ratio SNR (Eq. (2.6)) decreases as V_L increases for all replicas. The overall decrease yields ≥ 10 dB except for MRI and M5 for which the decrease is limited to ≤ 5 dB. It is noted that for all V_L , SNR-values for silicone vocal folds replicas are greater than the ones for PLT vocal folds replica conditions.

5.4 Discussion: voice quality features

For all deformable vocal folds replicas, increasing V_L in the range $V_L \leq 5$ mL affects the mean waveform shape of the upstream pressure during normal oscillation regime: increased oscillation cycle period T, increased amplitude \mathscr{A} and altered closing and opening portions $T^{c,o,t}$ so that closing-opening asymmetry is affected. The perturbation of this mean waveform shape increases with V_L as well: increased total harmonic distortion rate THD, decreased signal-to-noise ratio SNR and increase of both fast ($\zeta_{T,\mathscr{A}}$) and overall (standard deviations) fluctuations of waveform shape features. Therefore, water spraying affects the normal auto-oscillation regime resulting in changes to the mean waveform shape – increased oscillation period and amplitude, changes to opening and closing portions – and increased perturbation of this mean waveform.

A preliminary qualitative description of the effect of water spraying on the flow through the motion driven MDR replica [104] was confirmed quantitatively in this work for waveform amplitude \mathscr{A} and closing and opening portions $T^{t,c,o}$. Found tendencies on the motion driven vocal folds replica compare well with general normal oscillation cycle tendencies described for deformable vocal folds replicas: changes to mean values (amplitude increase and altered opening and closing portions $T^{t,c,o}$) and increased perturbation. Therefore, current findings support that flow changes following two-phase water-air mixing observed for the motion driven MDR replica do also occur in the case of the deformable vocal folds replicas. Flow changes following twophase water-air mixing will alter pressure forces (F_A right-hand side of Eq. 5.1) on the enveloping structure and hence contribute to observed waveform shape changes. It is noted that current data do not provide evidence that vocal folds mass increases significantly with V_L so that m in the left-hand side of Eq. (5.1) can be assumed independent of V_L and mass changes can not explain observed changes to normal oscillation period T. Therefore, it is of interest for further research to account for flow mixing when aiming to model the fluid-structure interaction with surface hydration and to assess how far such a flow model can explain observed general tendencies of the impact of V_L on the waveform shape observed for $V_L = 0$ mL.

A part from mass *m*, initial conditions related to the elasticity and initial glottal aperture $(r, k, \text{ and } A_0 \text{ left-hand side of Eq. (5.1)})$ as well as imposed mean upstream air pressure \overline{P}_u (*F*_A right-hand side of Eq. (5.1)) determine the fluid-structure interaction. Different deformable vocal folds replicas were used resulting in a considerable variation of these initial conditions. Quantified waveform shape features showed the strong impact of initial conditions on the degree

to which water spraying affects the mean waveform shape and its perturbation. In general, mean and even more so perturbation features are less affected by water spraying for silicone vocal folds replicas (EPI, MRI and M5) than for PLT vocal folds replica conditions. Large differences are observed between different PLT conditions as well. It is noted that the silicone M5 vocal folds replica is only marginally affected by V_L . In future, a more systematic study is needed to quantify the effect of water spraying while varying initial conditions.

A subharmonic oscillation regime is generated by water spraying for $V_L \ge 2$ mL in the case of PLT vocal folds replica conditions and no subharmonic oscillation regime is observed for silicone vocal folds replicas. The generation of a sub-harmonic frequency point to non-linearities introduced following water spraying for the PLT vocal folds replica. These non-linearities are potentially due to perturbations introduced by turbulence, estimated to increases with V_L between 1% and 15% for the PLT vocal folds replica whereas it remains less than 1% for silicone vocal folds replicas. Further evidence is needed to sustain this statement.

General tendencies observed for the normal oscillation regime agree with findings reported in literature concerning the effect of surface hydration on human voice features [38, 42, 44] and on excised larynges [43]: decrease of fundamental frequency f_0 , changes to closing-opening asymmetry and to perturbation measures $\zeta_{A,T}$ as well as signal-to-noise ratio SNR and total harmonic distortion rate THD. Note that the shown influence of initial conditions observed for deformable vocal folds replicas might partly explain the variety of tendencies observed on human speakers.

5.5 **Results: complexity analysis**

Measured time series of upstream pressure $P_u(t)$ (left) and their phase representation (right) following water spraying with volumes $V_L \leq 5$ mL are illustrated in Figure 5.14 for a silicone vocal folds replica (MRI in Figure 5.14(a)) and a PLT vocal folds replica condition ((2800,0) in Figure 5.14(b)). To facilitate comparison between times series, P_u is up-shifted (+1500 Pa) with each V_L increment and time t is normalised with respect to the lowest characteristic oscillation period $1/f_N$ with $f_N(V_L)$ as indicated for each time-trace. For all vocal folds replicas embedding dimension m varies in the range $10 \leq m \leq 25$ and time delay τ varies in the range $15 \leq \tau \leq 35$. Water spraying enhances auto-oscillation for all vocal folds replicas since the lowest characteristic oscillation frequency f_N decreases with V_L and from the phase space trajectories is seen that the oscillation amplitude increases which confirms findings reported in [134]. For silicone vocal folds replicas f_N corresponds to the first harmonic, whereas for PLT vocal folds replica conditions a subharmonic frequency is generated as observed from the period doubling for $V_L \geq 2$ mL in Figure 5.14(b). As a consequence, the phase state representation for the silicone vocal folds replicas exhibits a single oscillation trajectory regardless of V_L associated with a stable deter-



Figure 5.14 – Illustration of upstream pressure time series $P_u(t \cdot f_N)$ (left) and their phase space representation (right) for volumes $V_L \le 5$ mL (gray scaling indicating V_L in mL and f_N in Hz) for silicone (MRI) and PLT (condition (2800,0)) vocal folds replicas. For clarity time series $P_u(t \cdot f_N)$ are up-shifted (+1500 Pa) with each V_L increment.

ministic oscillation regime for each V_L . For PLT vocal folds replica conditions, the generation of a subharmonic oscillation frequency makes the phase trajectory more complicated as single oscillation cycles ($V_L \in \{2,3\}$ mL in Figure 5.14(b)) become less stable until ($V_L \in \{4,5\}$ in Figure 5.14(b)) an inner cycle imprints on the trajectory due to the growth of the subharmonic with V_L so that the trajectory in phase space approaches a chaotic oscillatory regime as divergences of the cycles in phase space can be observed as V_L increases. Recurrence plots (Eq. (5.15)) of the phase space state trajectories are illustrated in Figure 5.15 for $V_L \in \{2,4\}$ mL for the silicone (MRI) and PLT vocal folds replica (condition (2800,0)). For $V_L = 2$ mL both recurrence plots display a regular pattern of diagonal lines reflecting the harmonic properties of the auto-oscillation. For $V_L > 2$ mL, recurrence plots can reveal some degree of loss of determinism illustrated for the PLT vocal folds replica at $V_L = 4$ mL (Figure 5.15(d)) as irregular shorter diagonal lines are observed. The MRI replica on the other hand remains in a deterministic oscillatory regime for $V_L = 4$ mL (Figure 5.15(c)).



Figure 5.15 – Recurrence plots for $V_L \in \{2,4\}$ mL for vocal folds replicas: a,c) PLT condition (2800,0) and b,d) silicone MRI.

Recurrence plot properties $(\mathcal{D}, \mathcal{R} \text{ and } \gamma \text{ defined in Eq. (5.18)})$ are then quantified for all vocal folds replicas as a function of V_L . Degree of determinism $\mathcal{D}(V_L)$, recurrence rate $\mathcal{R}(V_L)$ and their ratio $\gamma(V_L)$ are plotted in Figure 5.16. For $V_L < 2$ mL all vocal folds replicas exhibit a stable oscillatory pattern as $\mathcal{D} > 90\%$ in Figure 5.16(a), *i.e.* recurrence plot with mostly uninterrupted diagonal lines as shown in Figure 5.15(a) and Figure 5.15(b). As V_L increases, silicone vocal folds replicas remain characterised by a stable deterministic pattern (*e.g.* Figure 5.15(c)) so that $\mathcal{D} > 90\%$ for all V_L . PLT vocal folds replica conditions on the other hand approaches chaotic behaviour as V_L increases, so that \mathcal{D} reduces to within the range $70\% < \mathcal{D} < 90\%$ for $V_L \in \{4, 5\}$ mL. Recurrence rates $\mathcal{R}(V_L)$ (Figure 5.16(b)) and ratios $\gamma(V_L)$ (Figure 5.16(c)) further confirm tendencies described for \mathcal{D} . Recurrence rate \mathcal{R} increases as a system approaches a chaotic regime as more states occur compared to the stable deterministic regime. The increase of \mathcal{R} remains negligible ($\leq 2\%$) for silicone vocal folds replicas regardless of V_L whereas it in-



Figure 5.16 – Degree of determinism ($\mathscr{D}(V_L)$), recurrence rate ($\mathscr{R}(V_L)$) and their ratio ($\gamma(V_L)$) for vocal folds replicas (symbols) shifted around each V_L value for clarity.

creases with up to 10% for PLT vocal folds replicas for $V_L \ge 2$ mL. It follows that ratio γ is maximum (high \mathcal{D} and low \mathcal{R}) for $V_L \in \{0, 1\}$ mL since the auto-oscillation result in a deterministic stable oscillation for all vocal folds replicas. For $V_L \ge 2$ mL γ will decrease with a degree and at a rate proper to each vocal folds replica and imposed initial conditions (P_{PLT}, o_{mm}) . The decrease remains limited to < 20% for silicone vocal folds replicas ($6.1 \ge \gamma \ge 5.1$) indicating a deterministic oscillation regime and becomes significant up to < 50% ($5.7 \ge \gamma \ge 2.8$) for some PLT vocal folds replica initial (P_{PLT}, o_{mm}) conditions indicating the loss of stable oscillation.



Figure 5.17 – Entropy $E(V_L)$ and estimated correlation dimension $D_2(V_L)$ for vocal folds replicas (symbols) shifted around each V_L value for clarity. The frame contains a zoom for $D_2(V_L)$ with $V_L \in \{1,2\}$ mL.

Quantified properties $\mathscr{D}(V_L)$, $\mathscr{R}(V_L)$ and their ratio $\gamma(V_L)$ suggest that water spraying ($V_L > 0$ mL) can cause the oscillation pattern to approach a chaotic complex oscillation pattern depending on the vocal folds replica, imposed initial conditions and on V_L . The complexity of the auto-oscillation is quantified by entropy $E(V_L)$ (Eq. (E.1)) and estimated correlation dimension

 $D_2(V_L)$ (Eq. (5.14)) plotted in Figure 5.17. Overall entropy $E(V_L)$ (Figure 5.17(a)) shows a small increase in the range $7.7 \le E(V_L) \le 9.0$ for PLT vocal folds replica conditions whereas it remains of similar magnitude for the silicone vocal folds replicas within the range $7.5 \le E(V_L) \le 8.6$. Nevertheless, $E(V_L)$ tendencies are not sufficiently pronounced to sustain a conclusion. Correlation dimension estimates $D_2(V_L)$ are shown in Figure 5.17(b). For $V_L < 2$ mL, $1.01 < D_2 < 1.12$ holds and increases for $V_L \ge 2$ mL at a rate which differs between vocal folds replicas. For silicone vocal folds replicas the increase is limited to within 10% as $1.01 < D_2(V_L) < 1.12$. For PLT vocal folds replica conditions the increase is more pronounced and yields up to 85% as $1.07 < D_2(V_L) < 2.00$ holds.

5.6 Discussion: complexity analysis

The phase space state and recurrence plot analysis, presented in Section 5.5, shows that estimated correlation dimension \mathscr{D}_2 (increase within 85%) and ratio γ (decrease within 50%) are suitable quantities to express changes to the complexity of the steady state auto-oscillation pattern due to surface hydration following water spraying.

Observed $\gamma(V_L)$ and $D_2(V_L)$ tendencies indicate that water spraying for $V_L \leq 5$ mL does not impact the deterministic oscillation regime for silicone vocal folds replicas whereas the system approaches a chaotic regime for PLT vocal folds replica conditions. This confirms observations reported for the signal-to-noise ratio and the fast cycle-to-cycle perturbations of amplitude and period [134]. The increased overall stability of the auto-oscillation for silicone vocal folds replicas compared to PLT vocal folds replicas is in line with the initial complexity observed for $V_L = 0$ mL as $\gamma(V_L = 0)$ and $D_2(V_L = 0)$ are respectively larger and lower for silicone vocal folds replicas than for PLT replica conditions.

For the PLT replica and $V_L \in \{0, 1\}$ mL recurrence plots exhibit a pattern of diagonal lines so that D_2 is near unity, and γ near its maximum, which is inherent to a deterministic oscillation pattern. For $V_L \ge 2$ mL the oscillation regime becomes more complicated as period doubling is observed, the degree of determinism \mathscr{D} decreases and the recurrence rate \mathscr{R} increases illustrating that the system loses some of its predictability as chaos is introduced in the system. The degree to which the initial stable oscillation pattern is disturbed depends on the imposed initial conditions of aperture ($o_{mm} \in \{0, 1, 2\}$ and $P_{PLT} = 2800$) and elasticity ($P_{PLT} \in \{2300, 2800, 3300\}$ and $o_{mm} = 0$). Estimated \mathscr{D}_2 and γ tendencies suggest that the complexity increases with initial aperture (increase of o_{mm}) and with rigidity (increase of P_{PLT}).

For future research, it is of interest to investigate the influence of initial conditions and the influence of the imposed $\overline{P_u}$ on the initial oscillation complexity for $V_L = 0$ mL and on the complexity increase following hydration by water spraying ($V_L > 0$ mL). This way the influence of initial conditions on the coexistence of different attractors can be further studied [117]. In par-

ticular, the role of aerodynamic turbulence needs further investigation and direct observation of the oscillation might contribute to the understanding of the effect of hydration in terms of oscillation modes [7, 8, 31]. Spraying with other liquids needs to be assessed in order to increase the relevance with respect to vocal folds surface hydration and artificial saliva sprays.

It is noted that the estimated range of $D_2(V_L)$ values $(1 \le D_2 \le 2)$ observed for the mechanical vocal folds replicas is similar to the range reported for human speakers during normal voicing by several researchers [108–112]. This encourages further quantification of D_2 as well as γ in order to inform on the oscillation regime and its complexity.

5.7 Conclusion

In this chapter, the effect of water spraying is quantified.

Concerning voice quality features observed on deformable replicas, the mean waveform shape exhibits decreased fundamental frequency, increased amplitude, changing closing-opening asymmetry and increased perturbation. The degree to which these changes occur depends on initial conditions. These tendencies are consistent with quantified waveform shape changes observed for a MDR vocal folds replica subjected to water spraying. It is argued that, as for the MDR replica, changing flow behaviour due to air-water mixing underlies observed tendencies for the deformable vocal folds replicas. Modelling studies are needed in order to assess to which extent mixing flow can explain current findings. In addition, further experimental research is needed in order to systematically consider the effect of initial replica conditions and their surface wettability, the influence of the used nozzle, quantify the droplet distribution during oscillation, physical mechanisms resulting in observed effects of water spraying on oscillation features and to quantify turbulence generation and dissipation during the oscillatory cycle as well. Current results for water provide a reference for future research involving different fluids with properties in agreement with mucus. For instance a fluid with higher viscosity (like Artisial artificial saliva presented in Appendix G) is assessed for which preliminary results are given in Appendix H.2 and compared to water. In addition, the influence of upstream pressure needs to be considered as shown in H.1 for water spraying.

Furthermore the complexity is analysed. The use of different mechanical vocal folds replicas showed that after water is sprayed ($V_L \leq 5 \text{ mL}$) on a stable auto-oscillation either the deterministic regime is maintained or the oscillation approaches a chaotic regime introduced by period doubling. Phase space state recurrence feature γ , *i.e.* the ratio of the degree of determinism to the recurrence rate of the phase space states, and estimated correlation dimension D_2 are suitable parameters to express the influence of hydration following water spraying on the oscillation regime. Besides initial pre-oscillatory conditions related to the glottal aperture and vocal folds elasticity, also the initial fluid-structure complexity determines the impact of water spraying on the complexity of the oscillation regime. The initial complexity near the onset pressure of oscillation is found to be lower for silicone vocal folds replicas than for PLT vocal folds replica conditions.

The feature analysis presented in this chapter showed the significant effect of water on oscillation. As it is assumed that the presence of water does not alter the vocal folds structural properties changes are assumed to be flow related. Therefore, in the following chapter, a quasi-analytical flow model is proposed and validated. On one hand a flow model will help understanding of effect of water during the oscillation cycle. On the other hand an accurate flow model allows to estimate fluid forces, F_A in the right hand side of Eq. (5.1), needed in order to assess reduced order models of auto-oscillation.

CHAPTER 6

Analytical flow modelling of water influence

Based on Van Hirtum A., Bouvet A., Pelorson X., 2018. *Pressure drop for adiabatic air-water flow through a time-varying constriction*. Physics of Fluids, 30:1-10.

Specific nomenclature

V_L	Injected water volume	
C_{S}	Separation constant	_
$\alpha(z,t)$	Viscous contribution factor to the pressure drop	_
v_f	Voice fraction	[-]
$\dot{P_d}$	Downstream pressure	Pa
μ_h	Homogeneous mixture viscosity	
$ ho_h$	Homogeneous mixture density	kg/m ³
$lpha_{\mu,i}$	Factor homogeneous viscosity calculation	_
i, j	Dominant phase and non-dominant phase	_
r	Binary value determine the dominant fluid	_
γ_i	Scaling factor of homogeneous viscosity	_
τ	Turbulence	
S	Slip coefficient	_
ΔP	Pressure drop	Pa
Χ	Martinelli's parameter	_
C_{LG}	Chisholm's constant	_
σ	Surface tension	N/m

6.1 Introduction

Respiratory flow through the human upper airways and in particular through the glottis, *i.e.* airway constriction formed between both vocal folds within the larynx, is often studied considering single-phase airflow through a uniform rigid channel containing a time-varying constriction as illustrated in Figure 6.1.



Figure 6.1 – Schematic overview of the vertical ($\beta = 90^{\circ}$) rigid channel with time-varying rectangular constriction (radius 1 cm, width L = 3.0 cm), pressure taps and fluid supplies: orientation angle β , gravitational acceleration g, imposed mobile vocal folds frequency f_c , time-varying minimum spacing $h_c(t)$, liquid supply time tag $t_L(t)$, liquid volume V_L , upstream pressure $P_u(t)$ and minimum constriction pressure $P_c(t)$. Flow is supplied along the positive z-direction and liquid is injected along the negative z-direction.

As such, the presence of liquid in the fluid is generally neglected [47, 48, 135, 136]. However, this assumption of single-phase airflow is in contrast with the physiological reality and is reported to affect flow-induced phenomena such as the shown effect of surface (de-)hydration on human voiced speech sound production [44, 137, 138].

When the presence of liquid is accounted for, adiabatic two-phase gas-liquid flow occurs. It

is well established that the mixture viscosity of gas-liquid two-phase flow strongly influences the pressure drop and thus the forces exerted by the flow on the surrounding channel walls driving fluid-structure interactions and associated phenomena such as voiced speech sound production, brass instruments play, whistling, *etc*.

As liquid is added, the viscosity of the gas-liquid mixture increases and hence it is expected that the pressure gradient raises [139]. Nevertheless, studies of adiabatic gas-liquid flow mostly focus on steady flow through uniform channels [140, 141] whereas a channel with time-varying constriction degree is more pertinent for glottal flow during speech production. Although that it has been shown from a dimensional analysis [47, 48, 136] of single-phase airflow under glottal conditions (Reynolds number $Re \leq 5 \times 10^3$, Strouhal number $Sr \leq 0.1$ and Mach number $Ma \le 0.12$) that incompressible, laminar and quasi-steady flow can be considered, the presence of a time-varying constriction and hence pulsating jet might affect the mixture viscosity. Indeed, enhanced mixing is reported for intermittent liquid injection which is particularly the case during and after a deceleration phase as occurs for engine jets [142, 143]. In the case of a sinusoidally time-varying constriction, flow deceleration occurs during the opening phase of the oscillatory cycle. Consequently, besides the effect of gas-liquid flow on the pressure drop, effects of the forcing frequency on mixing and hence on the mixture viscosity need to be studied as well. It is noted that near closure viscous flow effects are shown to provide a major contribution to the pressure gradient when single-phase airflow through a glottal-like constriction is considered [47, 48, 135, 136]. Therefore, the aim of this work is twofold. Firstly, it is sought to provide experimental evidence of the effect of two-phase flow and of the imposed oscillation frequency f_c on pressures P_c measured at the minimum constriction for different upstream punctually injected water volumes V_L (Figure 6.1) ranging from dry to excessive. Secondly, it is sought to accurately model observed pressures P_c applying a quasi-steady one-dimensional flow model approach for assessed (f_c, V_L) cases while accounting for viscous mixing. Since different surface hydration conditions occur for glottal flow (due to smoking, drinking, gurgling, etc.), both gas and liquid ruled viscous mixing are considered.

6.2 Model approach

6.2.1 Pressure drop with single-phase viscous contribution

Single-phase flow of a fluid with density ρ and dynamic viscosity μ through a constricted channel is modeled following the quasi-one-dimensional approach outlined in [48, 144, 145]. The model assumes quasi-steady laminar incompressible pressure driven flow while accounting for boundary layer development due to viscosity within the constricted channel portion. An empirical *ad-hoc* geometrical criterion is used to indicate streamwise position of flow separation and jet formation z_s along the diverging side of the constriction ($z \ge z_c$) as the area corresponding

to $A_s = c_s \cdot A_c$ with separation constant $c_s \ge 1$ and minimum channel area A_c . Concretely, the constant is set to $c_s = 1.13$ given the constriction geometry which is within the range commonly reported in literature [2, 48, 136, 146] for glottal-like geometries ($1.05 \le c_s \le 1.4$).

To overcome the constraint of a horizontal flow channel the model is altered to account for gravitational flow acceleration g for any given channel orientation angle β , with $\beta = 0^{\circ}$ and $\beta = 90^{\circ}$ for a horizontal and vertical channel respectively (Figure 6.1). The pressure distribution within the channel with time-varying area A(z,t) as a function of streamwise position z and time t up to flow separation ($z_0 \le z \le z_s$) becomes:

$$P(z,t) = P_u(t) + \frac{1}{2}\rho\Phi^2(t)\left(\frac{1}{A^2(z_0)} - \frac{1}{A^2(z,t)}\right) + \mu\Phi\int_{z_0}^z \frac{dz}{\alpha(z,t)} + \rho g\sin(\beta)(z_0 - z)$$
(6.1)

The function $\alpha(z,t)$ in the term describing the viscous contribution to the pressure drop (third right-hand term) depends on the channel's cross-section shape within the constriction [144, 145]. For a rectangular cross-section shape with constant width w and varying height h(z,t), a two-dimensional Poiseuille flow assumption can be applied when $w \gg h_c$ in the constricted portion:

$$\alpha(z,t) = -\frac{w \cdot h^3(z,t)}{12}.$$

At flow separation ($z = z_s$ and $A = A_s$) $P(z_s, t) = P_d$, with downstream pressure $P_d = 0$, so that volume flow velocity Φ can be estimated from Eq. (6.1):

$$\Phi(t) = \left[\mu \int_{z_0}^{z_s} \frac{dz}{\alpha(z,t)} + \left\{ \left(\mu \int_{z_0}^{z_s} \frac{dz}{\alpha(z,t)} \right)^2 + 2\rho \left(\left(P_u(t) - P_d \right) + \rho g \sin(\beta)(z_0 - z_s(t)) \right) \\ \cdot \left(\frac{1}{A^2(z_s,t)} - \frac{1}{A^2(z_0)} \right) \right\}^{1/2} \right] \\ \times \left[\rho \left(\frac{1}{A^2(z_s,t)} - \frac{1}{A^2(z_0)} \right) \right]^{-1}$$
(6.2)

Once $\Phi(t)$ is known, P(z,t) along the constricted channel portion is estimated using Eq. (6.1). The term describing the viscous pressure drop contribution in Eq. (6.1) and Eq. (6.2) needs to be reconsidered when a two-phase air-water mixture occurs. This is assessed in the next section.

6.2.2 Two-phase viscous pressure drop contribution of air-water mixture

6.2.2.1 Homogeneous mixture

When no slip is assumed between the two phases of the air-water mixture, there is no velocity difference between the gas (air) and liquid (water) phase so that the fluid is considered homogeneous [139]. The flow is then regarded as a single-phase flow of a fluid having average properties determined by the gas-liquid mixing quality. Consequently, the one-dimensional model presented in Section 6.2.1 can be applied and the two-phase pressure drop is again expressed as the sum of inertial (accelerational or decelerational), frictional and gravitational components Eq. (6.1) when averaged mixture properties, density $\rho = \rho_h$ and dynamic viscosity $\mu = \mu_h$, can be determined in terms of the gas (subscript $-_G$) and liquid (subscript $-_L$) properties.

The homogeneous density is estimated as:

$$\rho_h = \left(\frac{x}{\rho_G} + \frac{1-x}{\rho_L}\right)^{-1} \tag{6.3}$$

in terms of the mass quality x ($0 \le x \le 1$) so that homogeneous void (or gas) fraction $v_f = \frac{\Phi_G}{\Phi_G + \Phi_L}$ becomes:

$$v_f = \frac{1}{1 + \left(\frac{1 - x}{x}\right) \left(\frac{\rho_G}{\rho_L}\right)}$$
(6.4)

with $0 \le v_f \le 1$. Consequently, $v_f = 0$ and x = 0 when $\Phi_G = 0$ (pure liquid flow) and $v_f = 1$ and x = 1 when $\Phi_L = 0$ (pure gas flow). From Figure 6.2 is seen that air-water flow at room temperature (density ratio $\rho_L/\rho_G \approx 832$) is gas dominated $v_f \ge 0.99$ ($\Phi_L \le 0.01 \cdot \Phi_G$ and $\rho_h \approx$ ρ_G) for $x \ge 0.1$ compared to $x \ge 0.99$ for an equal density mixture ($\rho_L/\rho_G \approx 1$). It is observed that for x < 0.01 void fraction v_f decreases rapidly as ρ_h/ρ_L increases at the same rate so that the contribution of the liquid to the mixture properties gains importance and will finally dominate $v_f \le 0.5$ ($\Phi_L \ge 0.5 \cdot \Phi_G$ and $\rho_h \ge 0.5 \cdot \rho_L$) as x and hence v_f further reduces towards 0.

Many averaging methods have been described in literature to determine homogeneous mixture viscosity μ_h . The most straightforward approximations [147–149] assume that the flow is liquid dominated. In order to account for liquid as well as gas phase dominated flow, it is proposed to reformulate these models more generally so that μ_h is estimated based on the viscosity of a single phase, either gas (i = G) or liquid (i = L), as:

$$\mu_h = \mu_i \cdot \alpha_{\mu,i},\tag{6.5}$$



Figure 6.2 – Void fraction v_f (thick full line) and normalised homogeneous density ρ_h/ρ_L (thin dashed line) as a function of mass quality x for air-water mixture. As a reference $v_f = x$ (diagonal line) for $\rho_L = \rho_G = \rho_h$ is indicated.

with scaling factor $\alpha_{\mu,i}$. To account for the dominant phase *i* as well as a non-unity density ratio between the phases, scaling factor $\alpha_{\mu,i} = \frac{\rho_h}{\rho_i}$ is introduced following [149]. Following [148], scaling factor $\alpha_{\mu,i} = \left[1 - x^{(r)}(1-x)^{(1-r)}\left(\frac{\rho_i}{\rho_j} - 1\right)\right] (\forall r \in \{0,1\})$ is proposed with *j* indicating the non-dominant phase, *i.e.* j = L for i = G and j = G for i = L, and exponent r = 0 for gas dominated flow and r = 1 for liquid dominated flow. This way μ_i is scaled explicitly by mass quality x and single phase density ratio ρ_i/ρ_i . For liquid dominated flow (r = 1) the original expression given in [148] is retrieved. As expected, $\alpha_{\mu,i} = 1$ holds for pure liquid (x = 0) or pure gas (x = 1) flow. So that following [147] for single phase dominated flows a constant unity scaling factor ($\alpha_{\mu,i} = 1$) is assumed. Used scaling factors $\alpha_{\mu,i}$ in Eq. (6.5), *i.e.* $f(\rho_h/\rho_i)$ (thin dashed line), $f(\rho_i/\rho_i, x)$ (thick full line) and unity constant (*cte* = 1, horizontal dashed line), as a function of x are illustrated in Figure 6.3 for air-water flow. It is verified that scaling factor $\alpha_{\mu,i=L}$ decreases from 1 ($\alpha_{\mu,i=L} \leq 1$) for liquid ruled flows so that mixture viscosity μ_h reduces compared to μ_L as gas is added. On the other hand scaling factor $\alpha_{\mu,i=G}$ increases from 1 ($\alpha_{\mu,G} \ge 1$) as liquid is added for gas ruled flow so that mixture viscosity μ_h augments from μ_G . Decreasing (for liquid ruled) and increasing (for gas ruled) scaling factor $\alpha_{\mu,i}$, and hence increasing and decreasing mixture viscosity following Eq. (6.5), with increased air-water mixing aims to reduce the viscosity difference ($\mu_L \gg \mu_G$) between pure water ($\mu_L \approx 1.0 \times 10^{-3}$ Pa·s) and pure air ($\mu_G \approx 1.8 \times 10^{-5}$ Pa·s) flow. Nevertheless, estimating mixture viscosity μ_h by scaling the viscosity of the gas or liquid phase Eq. (6.5) looses its motivation as mixing increases since the underlying assumption of single-phase dominated flow ebbs away.

To obtain μ_h estimations which are also valid in the transition zone between liquid and gas single-phase ruled flows, μ_h models using the viscosity of both phases, *i.e.* μ_L and μ_G , are proposed in literature [150–156]. A common approach is to express mixing viscosity μ_h (exponent



Figure 6.3 – Scaling factors $\alpha_{\mu,i}$ in Eq. (6.5) for gas (i = G) and liquid (i = L) ruled flows as a function of *x* for air-water mixture: $\alpha_{\mu,i} = f(\rho_h/\rho_i)$ (thin dashed line), $\alpha_{\mu,i} = f(\rho_i/\rho_j, x)$ (thick full line) and unity $\alpha_{\mu,i} = 1$ (horizontal dotted line).

s = 1) or its reciprocal (s = -1) as a sum of single-phase viscosities or their inverses weighted by scaling factors γ_L (for μ_L) and γ_G (for μ_G) as:

$$\mu_h^{(s)} = \mu_L^{(s)} \cdot \gamma_L + \mu_G^{(s)} \cdot \gamma_G + \gamma_0, \quad \forall s \in \{-1, 1\},$$
(6.6)

with offset $\gamma_0 = 0$ when not explicitly stated. The earliest – and most common [157] – models defined scaling factors as mass averages the same way as for ρ_h in Eq. (6.3), *i.e.* $\gamma_G = x$ and $\gamma_L = 1 - x$ according to McAdams *et al.* [150] (for s = -1) and [151] (for s = 1). This scaling is further refined to account besides the mass quality also for densities of the phases which involves void fraction v_f following Beattie *et al.* [153] (for s = 1) since $\gamma_G = v_f$ and $\gamma_L = (1 - 1)$ v_f)(1+2.5 v_f) and kinematic viscosities of the flow phases following Dukler *et al.* [152] (for s = 1) since $\gamma_G = \rho_h x / \rho_G$ and $\gamma_L = \rho_h (1 - x) / \rho_L$ to which Fourar *et al.* [155] (for s = 1) added $\gamma_0 = 2\sqrt{x(1-x)\mu_G\mu_L/\rho_G\rho_L}$, which is maximum for x = 0.5 and reduces symmetrical towards 0 at limiting values $x \in \{0, 1\}$ so that μ_h is increased for intermediate x compared to values obtained for $\gamma_0 = 0$. Furthermore, [154] (for s = -1) optimized (for data in 0 < x < 0.25) scaling factors provided by McAdam et al. [150] by adding an exponent greater than unity to the mass averaging so that resulting μ_h increases as $\gamma_G = x^{1.4}$ and $\gamma_L = (1 - x^{1.4})$. It is noted that Eq. (6.6) satisfies limiting conditions $\mu_h = \mu_L$ and $\mu_h = \mu_G$ for pure liquid (x = 0) and pure gas flow (x = 1), which is not the case for Eq. (6.5). Homogeneous air-water mixture viscosity μ_h from Eq. (6.6) normalized by μ_L obtained for different scaling factors γ_L and γ_G (and $\gamma_0 = 0$) are plotted in Figure 6.4 as a function of x. It is seen that μ_h estimated using Eq. (6.6) decreases continuously as gas concentration (and hence x) is raised. Nevertheless important differences in the mixing viscosity are observed depending on the scaling factors and offset extending (Beattie, Dukler) or shortening (Cicchitti, Lin) the range of gas dominated flow compared to values obtained by applying Eq. (6.5) with scaling factor $f(\rho_h/\rho_i)$ which is in close approximation with McAdam's curve for gas dominated flow and with Dukler's curves for liquid dominated flow. It was noted that Dukler's ($\gamma_0 = 0$) and Fourar's ($\gamma_0 > 0$) curves overlap so that Fourar's curve is not considered and $\gamma_0 = 0$ holds.



Figure 6.4 – Normalised homogeneous mixing viscosity $\mu_h(x)/\mu_L$ from Eq. (6.6) for air-water mixture with (γ_L , γ_G) following: McAdam (gray thick dashed), Cicchitti (black diagonal thick dashed), Lin (black thick full), Dukler (gray thin full), Beattie (black thick dotted). Also $\mu_h(x)/\mu_L$ from Eq. (6.5) with Garcia ($f(\rho_h/\rho_i)$, discontinuous thin dash-dotted) is shown. Note that $\mu_G/\mu_L \approx 0.018$. A zoom for $x \le 0.01$ is included.

Mixing viscosity μ_h from Eq. (6.6) or Eq. (6.5) neglects turbulence. Nevertheless, in the case of gas-liquid flows the homogeneous flow assumption is most suitable to describe flow with bubbles or droplets which induces turbulence. Therefore, it is sought to model μ_h while accounting for turbulence in addition to the fluid characteristics of the phases and their concentration. Turbulent mixing viscosity μ_h^{τ} for liquid dominated homogeneous flow [158] is extended to single-phase ruled homogeneous flow in general as:

$$\mu_{h}^{\tau} = \mu_{h} + \mu_{\tau},$$

= $\gamma_{\tau} \cdot \alpha_{\mu,i} \cdot \mu_{i}, \text{ with } 1 < \gamma_{\tau}$ (6.7)

with turbulent viscosity μ_{τ} , subscript $i \in \{G, L\}$ indicating as before the dominating phase, scaling factor $\alpha_{\mu,i}$ as defined in Eq. (6.5) and introducing turbulent scaling factor $\gamma_{\tau} > 1$. For homogeneous flow, it is shown [158, 159] that γ_{τ} depends on the bubble/droplet size so that $1.1 \leq \gamma_{\tau} \leq 2$. Aung *et al.* [156] proposed to apply the same approach of single-phase ruled flow to flows with intermediate gas and liquid concentrations so that following their approach Eq. (6.6) becomes:

$$\left(\boldsymbol{\mu}_{h}^{\tau}\right)^{(s)} = \boldsymbol{\gamma}_{\tau}^{(s)} \left(\boldsymbol{\mu}_{L}^{(s)} \cdot \boldsymbol{\gamma}_{L} + \boldsymbol{\mu}_{G}^{(s)} \cdot \boldsymbol{\gamma}_{G} + \boldsymbol{\gamma}_{0}\right), \quad \forall s \in \{-1, 1\}$$

$$(6.8)$$

which can be applied for all void fractions v_f . A common value [156] yields $\gamma_{\tau} = 1.7$ so that turbulent mixing viscosity μ_h^{τ} is increased with 70% compared to mixing viscosity μ_h for same scaling factors { $\gamma_L, \gamma_G, \gamma_0$ }. Concretely, Aung [156] proposed to use maximized viscosity μ_h^{τ} regardless *x* by applying Eq. (6.7) with $\alpha_{\mu,L} = 1$ for liquid ruled flow [147] ($0 \le x < x_{thres}$) and Eq. (6.8) with { $\gamma_L, \gamma_G, \gamma_0$ } following Cicchitti [151] elsewhere ($x_{thres} \le x \le 1$):

$$\mu_h^{\tau} = 1.7 \cdot \mu_L, \qquad 0 \le x < x_{thres} \mu_h^{\tau} = 1.7 \left(x \cdot \mu_G + (1 - x) \mu_L \right), \qquad x_{thres} \le x \le 1$$
(6.9)

with concentration threshold $0 < x_{thres} < 1$. Turbulent mixing viscosity μ_h^{τ} for air-water flow obtained from Eq. (6.9) with $x_{thres} = 0.1$ is illustrated in Figure 6.5. For intermediate *x*-values $(x \approx 0.5) \ \mu_h^{\tau}$ yields by definition from 1.7 up to 25 times μ_h from Eq. (6.6) with $\{\gamma_L, \gamma_G, \gamma_0\}$ taken following Cicchitti [151] and McAdams [150], respectively, so that the contribution of turbulence is most notable for intermediate *x*-values.



Figure 6.5 – Normalised homogeneous mixing viscosity $\mu_h(x)/\mu_L$ from Eq. (6.6) (McAdam (gray thick dashed) and Cicchitti (black diagonal thick dashed)) and turbulent mixing viscosity μ_h^{τ}/μ_L from Eq. (6.9) (Aung (black thin full)) for air-water mixture.

6.2.2.2 Separated flow model for air-water mixture

When the assumption of no slip between the phases is dropped, velocities of both phases might be different so that the homogeneous flow model is no longer valid and Eq. (6.4) becomes:

$$v_f = \frac{1}{1 + S\left(\frac{1 - x}{x}\right)\left(\frac{\rho_G}{\rho_L}\right)} \tag{6.10}$$

with *S* indicating the slip between both phases. In order to apply the one-dimensional model approach in Section 6.2.1, the viscous contribution to the pressure drop is modeled considering both phases as separated streams independently of the flow regime [139]. The slip is then accounted for using empirical liquid hold up correlations and empirical relations describing the frictional interaction between the phases. Slip can be accounted for following two-phase friction multipliers ϕ_L^2 and ϕ_G^2 proposed in a landmark paper by [160] which relates frictional two-phase pressure drop $\Delta P_{f,M}$ and frictional pressure drops $\Delta P_{f,i\in\{L,G\}}$, which would exist if the gas (i = G) or liquid (i = L) phase is assumed to flow alone:

$$\Delta P_{f,M} = \phi_L^2 \cdot \Delta P_{f,L}, \tag{6.11}$$

$$\Delta P_{f,M} = \phi_G^2 \cdot \Delta P_{f,G}. \tag{6.12}$$

Multipliers $\phi_L^2, \phi_G^2 \ge 1$ can be determined following fitted correlations in [161]:

$$\phi_L^2 = 1 + \frac{C_{LG}}{X} + \frac{1}{X^2},\tag{6.13}$$

$$\phi_G^2 = 1 + C_{LG}X + X^2, \tag{6.14}$$

with Martinelli's parameter *X* defined as a decreasing function of *x*:

$$X = \left(\frac{1-x}{x}\right)^{0.9} \left(\frac{\rho_G}{\rho_L}\right)^{0.5} \left(\frac{\mu_L}{\mu_G}\right)^{0.1},\tag{6.15}$$

and Chisholm's constant C_{LG} depending on the liquid-gas flow regime: $C_{LG} = 5$ for laminarlaminar flow, $C_{LG} = 10$ for turbulent-laminar flow, $C_{LG} = 12$ for laminar-turbulent flow and $C_{LG} = 20$ for turbulent-turbulent flow. Multipliers as a function of X for different Chisholm's constants are shown in Figure 6.6.

It is seen that gas flow occurs for $X \ll 1$ ($\phi_L^2 \to \frac{1}{X^2}$, $\phi_G^2 \to 1$), liquid flow for $X \gg 100$ ($\phi_L^2 \to 1$, $\phi_G^2 \to X^2$) and two-phase flow for intermediate X-values ($\phi_L^2 \approx f(X)$, $\phi_G^2 \approx f(X)$).



Figure 6.6 – Multipliers ϕ_L^2 and ϕ_G^2 as a function of Martinelli's parameter X for air-water flow $(\rho_L/\rho_G = 832)$ and for different Chisholm's constants: $C_{LG} = 5$ for laminar-laminar (LL) flow (full line), $C_{LG} = 10$ for turbulent-laminar (TL) flow (dashed-dotted line), $C_{LG} = 12$ for laminar-turbulent (LT) flow (dashed line) and $C_{LG} = 20$ for turbulent-turbulent (TT) flow (dotted line).

Different C_{LG} values are reported on as shown correlations provide a one-parameter data-fitting problem for C_{LG} in an attempt to reflect the effect of flow regime and channel geometry (*e.g.* cross-section shape, hydraulic diameter, orientation angle) with more accuracy [94, 141, 162–168]. The meaning of constant C_{LG} is seen writing

$$\Delta P_{f,M} = \Delta P_{f,L} + C_{LG} \left(\Delta P_{f,L} \cdot \Delta P_{f,G} \right)^{1/2} + \Delta P_{f,G}, \qquad (6.16)$$

indicating that the interfacial contribution to the total frictional two phase pressure drop is:

$$\Delta P_{\mathscr{I},M} = C_{LG} \left(\Delta P_{f,L} \cdot \Delta P_{f,G} \right)^{1/2} \tag{6.17}$$

so that C_{LG} weights the mean of the contribution of pure gas and liquid to $\Delta P_{\mathscr{I},M}$. Furthermore, C_{LG} can be derived analytically [169] under certain flow assumptions. For homogeneous air-

water flow $C_{LG} \approx 28.6$ which agrees well with Chisholm's value for turbulent-turbulent flow $(C_{LG} = 20)$. For $C_{LG} = 0$, $\Delta P_{\mathscr{I},M}$ is the sum of the pressure drops of the single-phase flows corresponding to laminar plug flow. A continuous expression of C_{LG} for air-water flow in vertical channels is proposed by [170]:

$$C_{LG} = 21 \Big(1 - \exp(-0.319D) \Big),$$
 (6.18)

and later refined for adiabatic gas-liquid flow by Zhang et al. [94] as:

$$C_{LG} = 21 \Big(1 - \exp(-0.674/Lp) \Big),$$
 (6.19)

as a function of Laplace number *Lp*:

$$Lp = \frac{\sqrt{\sigma/g(\rho_L - \rho_G)}}{D},$$
(6.20)

with surface tension σ and hydraulic diameter *D*. For air-water flow at temperature $T \approx 22^{\circ}$ C and assuming $D \leq 3$ mm, $\sigma \approx 72$ mN/m (at temperature $T = 22 \pm 2^{\circ}$ C), it follows from Eq. (6.20) that Lp > 0.9 so that Eq. (6.19) $C_{LG} \leq 11$ holds as plotted in Figure 6.7. Therefore from Eq. (6.19) follows that C_{LG} increases with *D* and is in fair agreement with Chisholm's value for laminarlaminar flow ($C_{LG} = 5$) for $D \approx 1.1$ mm and for turbulent-laminar flow ($C_{LG} = 10$) for $D \approx$ 2.6 mm. Note that D < 3 mm corresponds to micro-scale channels based on the classification of [171] using 3 mm as the threshold.



Figure 6.7 – C_{LG} (thin dashed line) constant Eq. (6.19) and Laplace number (thick full line) Eq. (6.20) for air-water flow ($\rho_L/\rho_G = 832$) as a function of hydraulic diameter *D*. As a reference constant Chisholm's values for laminar-laminar (LL) flow ($C_{LG} = 5$) and for turbulent-laminar (TL) flow ($C_{LG} = 10$) are shown.

6.3 Experimental approach

Experiments are performed by using MDR vocal folds replica to impose a known timevarying constriction within a vertical rigid uniform circular channel (internal diameter 25 mm) as schematically illustrated in Figure 6.1. Sinusoidal movement of the rectangular gap between both half cylinders is imposed with forcing frequency $f_c \in \{1, 6, 10\}$ Hz.

As detail in Section 2.5.1 the time-varying minimum area along the channel is obtained as $A_c(t) = h_c(t) \cdot L$. Consequently, during experiments hydraulic diameter yields $D \le 2$ mm and area constriction ratio \mathscr{R} is sinusoidally varied between 100% and 95%.

Continuous steady airflow (density $\rho_G = 1.2 \text{ kg} \cdot \text{m}^{-3}$ and dynamic viscosity $\mu_G = 1.8 \times 10^{-5} \text{ Pa} \cdot \text{s}$, temperature $T = 22 \pm 2^{\circ}\text{C}$) is provided along the positive *z*-direction (Figure 6.1) by a valve controlled air supply. Pressure transducers (Kulite XCS-093) are positioned in pressure taps upstream and at the minimum spacing so that upstream pressure P_u and minimum constriction pressure P_c are measured). Air supply is set so that in absence of liquid mean upstream pressure \overline{P}_{up} yields 1136 ± 30 Pa regardless of f_c .

From a dimensional analysis follows that airflow through the time-varying constriction results in non-dimensional numbers characterizing airflow through the glottis of a male adult [2, 47, 48, 135, 136]: $Re \sim O(10^3)$, $Sr \sim O(10^{-2})$, $Ma \sim O(10^{-1})$. Non-dimensional numbers are either of the same order of magnitude or smaller when water is considered. In addition, the aspect ratio h_c/L of the time-varying constriction yields $h_c/L \ll 1$, as is observed for human subjects, which motivates the quasi-one-dimensional flow model approach. Consequently, flow assumptions underlying the model approach outlined in Section 6.2.1 remain valid when considering airflow through the time-varying constriction shown in Figure 6.1.

Distilled water (density $\rho_L = 998 \text{ kg} \cdot \text{m}^{-3}$ and dynamic viscosity $\mu_L = 1.0 \times 10^{-3} \text{ Pa} \cdot \text{s}$, temperature $T = 22 \pm 2^{\circ}\text{C}$) is injected manually at the downstream end of the channel by emptying a graduated (accuracy 0.1) syringe equipped with a spray nozzle (diffusion angle $20^{\circ} \pm 1^{\circ}$, diameter 1 mm) containing a known volume $V_L \in \{1, 2, 3, 4, 5\}$ mL. Liquid is supplied homogeneously along the constricted area (Figure 6.1). Furthermore, liquid injection is time-tagged t_L

V_L [mL]	1.0 ± 0.1	2.0 ± 0.1	3.0 ±0.1	4.0 ± 0.1	5.0 ±0.1
Δt_L [s]	0.86 ± 0.17	1.02 ± 0.10	1.47 ± 0.22	1.66 ± 0.27	2.61 ± 0.41
Φ_L [mL/s]	1.11 ± 0.11	1.98 ± 0.19	2.26 ± 0.22	2.15 ± 0.21	1.95 ± 0.19

Table 6.1 – Injected liquid volume V_L and duration Δt_L . Estimated flow rate Φ_L during injection.
by manually operating an electrical switch. The mean duration of liquid injection Δt_L increases quadratically (coefficient of determination $R^2 = 0.97$) with V_L and its overall value is less than 3 s regardless of V_L (Table 6.1). Liquid supply volume flow rate Φ_L (Table 6.1) is then approximated as the ratio $V_L/\Delta t_L$. The overall order of magnitude yields $\Phi_L \approx 1.8 \pm 0.6$ mL/s. All signals are sampled using a sampling frequency of 10 kHz. No water leakage was observed along the upstream or downstream channel end.



Figure 6.8 – Measured data for $f_c = 1$ Hz and $V_L = 4$ mL: (top) $P_u(t)$ (black full line), $P_c(t)$ (gray full line) and $t_L(t)$ (dashed line) with duration of liquid injection Δt_L and (bottom) $h_c(t)$ (full line).

As an example, Figure 6.8 illustrates measured data for $f_c = 1$ Hz and $V_L = 4$ mL as a function of time t near the instant of liquid injection. Imposed sinusoidally varying $h_c(t)$ is indicated at the bottom and measured pressures $P_u(t)$ and $P_c(t)$ are plotted at the top. Time-tag t_L corresponds to an impulse whose width Δt_L indicates the duration of liquid injection so that pure air flow occurs before the impulse and air-water mixture is observed following the impulse onset. During each period of the shown signals upstream pressure P_u varies sinusoidally so that its maximum is reached when h_c is small and minimum P_u corresponds to large aperture h_c . Next, it is seen that P_c reaches its maximum near closure ($h_c = 0 \text{ mm}$) whereas P_c is negative during most of the open phase ($h_c > 0$ mm). Although these general tendencies are observed for each period of the plotted signals, *i.e.* before, during or after liquid injection, close observation of Figure 6.8 shows some important changes to P_c and to a less extent to P_u following liquid injection. One striking feature characterizing P_c (and which is found to a less extent on P_u) is the irregular appearance of P_c -spikes. This suggests that turbulence is induced due to the presence of droplets, which are visually observed upstream from the constricted channel portion following liquid injection for all assessed V_L . Differences observed comparing airflow and air-water mixture flow data are further detailed in Section 6.4.1. To avoid transitional phenomena immediately following liquid injection $(t_L > 0)$, in the remainder of this work data are extracted 6 s after liquid injection is terminated.

6.4 **Results and discussion**

6.4.1 Experimental observations

Examples of measured data are presented in order to provide a qualitative description of the effect of liquid injection for all assessed (f_c, V_L) .



Figure 6.9 – Illustration of minimum constriction pressure P_c during a single period $t \cdot f_c$ ($f_c = 10 \text{ Hz}$) for $V_L \in \{0, 1, 3, 5\}$ mL for given P_u and h_c (scaled with respect to P_u).

The impact of V_L on measured minimum constriction pressure P_c is illustrated in Figure 6.9 for h_c imposed with forcing frequency $f_c = 10$ Hz. P_u varies less than 4% for all assessed V_L so that a single P_u -curve is plotted. During each period, gap h_c evolves through the following consecutive phases: open, closing, closed, opening and again open. In general, the effect of V_L on P_c increases with increasing V_L . Besides irregular P_c -spikes due to the splashing of droplets throughout the closing phase, differences are observed for small apertures occurring near and during closure corresponding to $h_c \leq 0.5$ mm or constriction ratio $\Re \geq 98\%$. During the closing phase this is most notable when h_c goes to zero, *i.e.* just prior to complete closure, when P_c decreases more with increasing V_L (e.g. 38% more for $V_L \ge 3$ mL) whereas at the beginning of the opening phase the accumulation of liquid during the closed phase results in an increase of P_c followed by a delayed steep decrease which therefore occurs for larger h_c . The extent to which described phenomena are observed varies from period to period, but in general increases with increased V_L . Note that droplet splashing near the gap (and hence spiking) is limited at the start of the opening phase since accumulated liquid downstream from the gap is likely to be evacuated by pulsating jet formation and associated starting vortex and shear-layer role-up at the onset of the opening phase.

The effect of forcing frequency f_c on pressure measurements is illustrated for $f_c \in \{1, 6, 10\}$ Hz in Figure 6.10 for $V_L = 0$ mL and $V_L = 5$ mL.

Without liquid injection ($V_L = 0$ mL in Figure 6.10(a)) the effect of f_c variation on measured pressures P_u and P_c is most notable when h_c reaches towards zero, *i.e.* at the end of the closing



Figure 6.10 – Illustration of upstream pressure P_u and minimum constriction pressure P_c during a single period $t \cdot f_c$ for $f_c \in \{1, 6, 10\}$ Hz and prescribed h_c (scaled with respect to P_u): a) $V_L = 0$ mL and b) $V_L = 0$ mL (black curves) and $V_L = 5$ mL (gray curves). For clarity P_c is shifted down for $f_c \in \{1, 6\}$ Hz.

phase as $h_c \leq 0.15 \text{ mm} \ (\Re \geq 99\%)$, during complete closure $(h_c = 0 \text{ mm})$ and at the start of the opening phase while $h_c \leq 0.43 \text{ mm} \ (\Re \geq 98\%)$. When $f_c = 1 \text{ Hz}$, P_u reflects the inverse tendencies imposed on h_c . This way, at first P_u reaches a minimum associated with maximum aperture, next P_u increases monotonously while h_c decreases during the closing phase until maximum P_u is maintained during the closed phase and afterwards P_u decreases monotonously to its minimum value as h_c increases during the opening phase. When f_c is increased $(f_c \in \{6, 10\} \text{ Hz})$ it is seen that P_u is no longer a monotonous function during the distinct phases, since P_u fluctuates around the maximum value observed for $f_c = 1 \text{ Hz}$ following an increase (overshoot) just before complete closure, which causes P_u to fluctuate during the closed phase and start of the opening phase. The magnitude of the overshoot, resulting fluctuation and affected h_c -range within each period increases with f_c , e.g. overshoot yields 6% for $f_c = 6 \text{ Hz}$ and 18% for $f_c = 10 \text{ Hz}$. As a consequence for $f_c = 1 \text{ Hz}$, $P_u = f(h_c)$ so that P_u and associated P_c during closing and opening phases are mirror-images of one another. For increased forcing frequency $f_c \in \{6, 10\}$ it is ev-

ident that this mirror-symmetry between closing and opening phase does no longer exist for P_u and hence for P_c . The rate of P_c increase at the end of the closing phase is seen to decrease as P_u -overshoot increases with f_c .

During the subsequent closed phase and the start of the opening phase, the rate of P_c variation is dictated by the fluctuation of P_u . During the closed phase, this implies that P_c increases to a maximum at a rate which is either slowed down or accelerated as the amplitude of the P_u fluctuation decreases or increases, respectively. The same way, during the start of the opening phase, the decrease of P_c is decelerated since P_u reduces more slowly. In addition, the minimum P_c value of the opening phase is greater than the minimum observed during the closing phase, *i.e.* increase of 14% for $f_c = 6$ Hz and 25% for $f_c = 10\%$. Consequently, the main impact of increasing f_c for $V_L = 0$ mL lies in the breakdown of mirror-symmetry between the closing and opening phase for P_u and hence P_c .

To assess the effect of f_c on air-water mixture flow ($V_L > 0$ mL) curves measured for $V_L \in \{0,5\}$ mL and $f_c \in \{1,6,10\}$ Hz are plotted together in Figure 6.10(b). As before, P_u is unaffected as V_L is varied so that a single curve is plotted for each f_c . To enhance clarity, P_c curves for $f_c = 6$ Hz and $f_c = 1$ Hz are shifted by extracting a constant offset of 800 Pa and 1600 Pa, respectively. It is seen that for $V_L > 0$ mL, irregular spikes in P_c are superposed on a smooth curve which varies from the one observed for $V_L = 0$ mL to an extent determined by f_c . Hence the effect of $V_L > 0$ mL on the curves observed for $V_L = 0$ mL is two-fold: on one hand irregular spikes occur due to the presence of droplets splashing near the constriction and on the other hand the change of mixing flow properties during the period affects the smooth curve observed for $V_L = 0$ mL.

It is noted that more spikes are seen as f_c reduces since the physical time corresponding to one period is greater (varies between 1 s and 0.1 s for $f_c = 1$ Hz and $f_c = 10$ Hz, respectively) and hence more droplets impact near the constriction during the closing and opening phase. During the closed phase and start of the opening phase, it is seen that P_c increases so that effects pointed out for $V_L = 0$ mL are emphasized which suggests that a change in mixing fluid properties occurs compared to the closing phase induced by the accumulation of liquid or increased contribution of liquid to the mixture upstream from the constriction. A pulsed jet and subsequent decelerating and starting vortex is formed while the gap starts to widen at the beginning at the opening phase accompanying increased water-air mixing and subsequent upstream liquid evacuation so that eventually mixing properties become similar to the ones characterizing the closing phase. Note that increased mixing following jet deceleration during the opening phase is in agreement with findings for a pulsating jet [142, 143] although more research is needed to investigate the contribution of each of the underlying mechanisms. For $f_c = 1$ Hz on the other hand the effect of jet mixing is less so that properties are re-established more rapidly within the period. Consequently, the main impact of $V_L > 0$ mL is to increase P_c during the closed and opening phase onset due to a change in mixing fluid properties so that the mirror-asymmetry of P_c between closing and opening phase is enforced as is observed for all frequencies. In addition, for $V_L > 0$ mL the impact of droplets causes irregular spiking of P_c . Therefore, observed curves are the result of a complex interaction between different phenomena.

6.4.2 Two-phase flow models behavior

Experimental observations provide indications that properties of the mixing fluid and flow regime during an oscillation cycle can be altered firstly by liquid droplet impact along the constriction and secondly by liquid accumulation, mixing and removal upstream of the constricted portion following closure. Obviously, these mechanisms contribute to the changes of liquid and gas concentration in the fluid and might affect the flow type as well. In the following, it is sought if and to which extent quasi-one-dimensional models presented in Section 6.2 allow to explain observed pressure drop tendencies, *i.e.* estimate $P_c(t)$ for known $P_u(t)$ and $h_c(t) > 0$ mm for all assessed (f_c, V_L) cases. From the description given in Section 6.2 and Section 6.4.1 it is *a-priori* expected that homogeneous mixing models for gas dominated flow, whether or not corrected for turbulence due to the spatial distribution of droplets, can be used during most of the open phase, closing phase and opening phase. Near closure, the homogeneous mixing assumption and assumption of gas dominated flow might be less suited so that a separated flow model is *a-priori* motivated.

Model accuracies between N measured (P_c) and modeled (\hat{P}_c) values for each (f_c , V_L) case are objectively expressed considering dimensionless coefficient of determination $R^2 \leq 1$ and mean absolute relative error $\xi \geq 0$, given as:

$$R^{2} = 1 - \frac{\sum_{i=1}^{N} \left(P_{c}(i) - \hat{P}_{c}(i) \right)^{2}}{\sum_{i=1}^{N} P_{c}^{2}(i)},$$
(6.21)

$$\xi = \frac{1}{N} \sum_{i=1}^{N} \left| \frac{P_c(i) - \hat{P}_c(i)}{P_c(i)} \right|.$$
(6.22)

It follows that model outcomes are most accurate when R^2 reaches a maximum (nearest to 1) or when ξ reaches a minimum (nearest to 0). Both R^2 and ξ are considered since experimentally observed instantaneous P_c -spikes associated with liquid droplets which are not modeled, *e.g.* observed for $f_c = 1$ Hz and $V_L = 5$ mL (Figure 6.10(b)), might deteriorate R^2 whereas ξ is less affected. On the other hand, since ξ is relative to the measured value, experimentally observed P_c values in the vicinity of 0 ($P_c \approx 0$) might deteriorate ξ whereas R^2 is less affected.

Resulting R^2 and ξ for most accurate models are summarised in Table 6.2 for the closing as well as the opening phase. Note that the open phase is not assessed since viscous flow effects can be neglected. For each case ($f_c \in \{1, 6, 10\}$ Hz) the accuracy of the single-phase airflow model ($V_L = 0$ mL) is given as a reference. As a general tendency, it is found that R^2 and ξ accuracies

	$f_c = 1 \text{ Hz}$						
	Closing			Opening			
V_L [mL]	Model	R^2	۳ζ	Mod	R^2	ξ	
0	air flow	0.83	1.4	air flow	0.88	0.53	
1	(6.6) H (Dukler)	0.85	1.6	(6.6) H (Dukler)	0.89	0.46	
2	(6.12) SG ^{<i>a</i>,<i>b</i>}	0.89	0.91	(6.12) SG ^{<i>a</i>,<i>b</i>}	0.89	0.42	
2	(0.0) H (Dukler)	0.82	2.4	(0.0) H (Dukler)	0.85	1.9	
3	(0.12) SG ^{4,a} (6.6) H (Dukler)	0.91 0.77	0.78	(0.12) SG ^{(n,n)}	0.85	1.5	
5	(6.12) SG ^{<i>a</i>,<i>b</i>}	0.85	1.7	(6.12) SG ^{<i>a</i>,<i>b</i>}	0.81	0.00	
4	(6.6) H (Dukler)	0.45	2.4	(6.6) H (Dukler)	0.72	1.56	
	(6.12) SG ^{<i>a</i>,<i>b</i>}	0.49	1.7	(6.12) SG ^{<i>a</i>,<i>b</i>}	0.72	1.47	
5	(6.6) H (Dukler)	0.06	1.7	(6.6) H (Dukler)	0.17	2.02	
	(6.12) $SG^{a,b}$	0.10	1.3	(6.12) SG ^{<i>a</i>,<i>b</i>}	0.16	2.0	
			f_c	= 6 Hz			
	Closin	lg	6	Openi	ng	(1	
V_L [mL]	Model	R^2	ξ	Mod	R^2	ξ	
0	air flow	0.84	0.70	air flow	0.70	0.70	
1	(6.12) SG ^{<i>a,b</i>}	0.87	0.46	$(6.16) S^{a,a}$	0.90	0.43	
•	(6.12) SG ^c	0.90	0.22	water flow	0.85	0.73	
2	(6.12) SG ^{<i>a</i>,<i>b</i>}	0.93	0.34	(6.16) S ^{<i>u</i>,<i>u</i>}	0.83	0.60	
2	(0.12) SO (6.12) SC a,b	0.92	0.47	$(6, 16)$ $\mathbf{S}^{a,d}$	0.77	0.40	
3	(6.12) SG ^c	0.85	0.31 0.34	water flow	0.90	0.49	
4	(6.12) SG ^{<i>a</i>,<i>b</i>}	0.96	0.26	$(6.16) S^{a,d}$	0.85	0.55	
-	(6.12) SG ^c	0.97	0.26	water flow	0.81	0.68	
5	(6.12) $SG^{a,b}$	0.80	0.84	(6.16) $S^{a,d}$	0.80	1.2	
	(6.12) SG ^c	0.71	1.1	water flow	0.86	0.51	
	$f_c = 10 \text{ Hz}$						
	Closing			Opening National D2 5			
V_L [mL]	Model	R^2	ς	Mod	<i>R²</i>	<u>ζ</u>	
0	air flow	0.54	0.40	air flow	0.73	0.83	
I	(0.12) SG ^{a,o} (6.12) SG ^c	0.84	0.29	(0.0) H (Cicchitti) (6.0) H (Aung)	0.62	1.2	
	(0.12) 50	0.92	0.23	(0.9) II (Aulig) (6.16) $S^{a,d}$	0.23	0.53	
2	(6.12) SCa,b	0.82	0.25	(0.10) S ((6.6) H (Ciashitti)	0.91	1.0	
2	(6.12) SG ^c	0.82 0.77	0.35	(6.0) H (Cleciniti) (0.32)		2.7	
	(0.12) 50	0.77	0.57	$(6.16) S^{a,d}$	0.86	1.0	
3	(6.12) SG ^{<i>a</i>,<i>b</i>}	0.54	4.5	(6.6) H (Cicchitti)	0.85	0.31	
2	(6.12) SG ^c	0.62	4.5	(6.9) H (Aung)	0.83	1.1	
				(6.16) S ^{<i>a</i>,<i>d</i>}	0.49	1.5	
4	(6.12) $SG^{a,b}$	0.91	0.24	(6.6) H (Cicchitti) 0.73		12	
	(6.12) SG ^c	0.96	0.19	(6.9) H (Aung)	0.44	17	
				$(6.16) \mathbf{S}^{a,d}$	0.88	7.2	
5	(6.12) SG ^{<i>a</i>,<i>b</i>}	0.51	1.4	(6.6) H (Cicchitti)	0.81	0.52	
	(6.12) SG ^c	0.62	1.3	(6.9) H (Aung)	0.76		
				(6.16) S ^{<i>a</i>,<i>a</i>}	0.60	0.92	

Table 6.2 – Summary of best model Eq. (\cdot) and their accuracies (coefficient of determination $R^2 \le 1$ and mean absolute relative error $\xi \ge 0$): gas dominated slip model (SG), slip model (S) and homogenous model (H) and their parameters.

 $\frac{||}{C_{LG}} = \frac{||}{C_{LG}} (6.19), \text{ constant along } z: C_{LG} = \max \left(C_{LG}(Lp(z)) \right).$

^b alternatively Eq. (6.12) with Chisholms's constant value: $C_{LG} \approx 12$ (LT flow). ^c $C_{LG}(z)$ Eq. (6.19), varying along z: $C_{LG} = C_{LG}(Lp(z))$. ^d alternatively Eq. (6.16) with Chisholms's constant value: $C_{LG} \approx 20$ (TT flow).

for $V_L > 0$ mL are similar to those obtained for $V_L = 0$ mL ($R^2 \ge 0.70$ and $\xi \le 1.4$) so that the quasi-analytical approach can be applied to water-air flow with the same accuracy as obtained for single phase airflow. Next, selected models indicated in Table 6.2 are discussed in more detail for all (f_c , V_L) cases.

For $f_c = 1$ Hz, model accuracies (Table 6.2) observed during closing and opening are alike which is consistent with the mirror symmetry observed on the pressure distribution for all assessed V_L . Most accurate model outcomes are obtained for gas dominated homogeneous flow Eq. (6.6) with Dukler's parameters $\gamma_{i \in \{L,G\}}$ so that $\mu_h = v_f \mu_G + (1 - v_f) \mu_L$, and for gas dominated slip flow Eq. (6.12) for constant C_{LG} (label SG in Figure 6.11(a)) with either $C_{LG} = \max \left(C_{LG}(Lp(D(z)))\right)$ according to Eq. (6.19) or alternatively $C_{LG} = 12$, *i.e.* Chisholm's constant for laminar-turbulent (LT) liquid-gas flow. Near closure the gas dominated slip model slightly outperforms the homogeneous flow model as is most notable during the closing phase. This finding holds for all assessed (f_c, V_L) cases. Note that for $V_L \in \{4, 5\}$ mL, R^2 (Table 6.2) decreases (from 0.89 to 0.10) due to the presence of P_c spikes (Figure 6.11(a))) whereas $\xi \approx 1.70$ remains.

As for $f_c = 1$ Hz, gas dominated slip flow Eq. (6.12) (SG) for $C_{LG} \approx 12$ provides accurate P_c estimations during the closing phase for $f_c = \{6, 10\}$ Hz. According to Eq. (6.19) C_{LG} is set either constant $C_{LG} = \max (C_{LG}(Lp(D(z))))$ (label SG $[C_{LG} \approx 12]$ in Figure 6.11(b) and Figure 6.11(c)) or varying along z as $C_{LG}(z) = (C_{LG}(Lp(D(z))))$ (label SG $[C_{LG}(z)]$ in Figure 6.11(b) and Figure 6.11(c) and Figure 6.11(c)) due to varying hydraulic diameter D(z). From Table 6.2 and modeled data curves (SG) illustrated in Figure 6.11(b) and Figure 6.11(c) is seen that varying $C_{LG}(z)$ gains accuracy as f_c increases from $f_c = 6$ Hz to $f_c = 10$ Hz.

Gas dominated slip model (SG) is not suitable within the opening phase when $f_c \in \{6, 10\}$ Hz since for all V_L more accurate P_c estimations are obtained using slip (S) model Eq. (6.16) with either constant C_{LG} set to max $(C_{LG}(Lp(D(z))))$ according to Eq. (6.19) or alternatively $C_{LG} \approx 20$, *i.e.* Chisholm's constant for turbulent-turbulent (TT) liquid-gas flow. From Table 6.2 follows that slip model is in close agreement with single-phase liquid flow and thus that the flow is no longer gas dominated. For $f_c = 10$ Hz and $V_L \in \{3,4,5\}$ mL the model accuracy is further improved by considering homogeneous mixing model (H) with parameters $\gamma_{i \in \{L,G\}}$ according to Cicchitti ($\mu_h = x\mu_G + (1-x)\mu_L$) or homogeneous turbulent flow using Eq. (6.9) (Aung) indicating increased mixture viscosity either generated by increased droplet-induced turbulence or/and jet-related mixing.

Objectively selected liquid-gas models reflect experimental findings and thus the influence of f_c and V_L on water-air mixing and the degree to which mixture viscosity impacts the flow. General tendencies are as follows.

During the closing phase and regardless of (f_c, V_L) , gas dominated slip flow (SG) Eq. (6.12) provides an accurate flow model when C_{LG} is set in accordance with Eq. (6.19) or alternatively



Figure 6.11 – Measured P_c (black full) and best models outcomes for $V_L \in \{1, 2, 3, 4, 5\}$ mL (shifted for clarity) and $f_c \in \{1, 6, 10\}$ Hz during closing and opening ($h_c > 0$): gas dominated slip (SG) model Eq. (6.12) for $C_{LG} \approx 12$ (thick dotted) or $C_{LG}(z)$ (gray thin full), slip (S) model Eq. (6.16) for $C_{LG} \approx 20$ (thick dashed) and homogenous (H) flow model Eq. (6.6) for Cicchitti parameters (gray thin dash-dotted) and following Aung Eq. (6.9) (thin dotted).

 $C_{LG} = 12$ in accordance with laminar-turbulent liquid-gas flow. It is seen that since the hydraulic diameter varies along the channels longitudinal *z*-axis Eq. (6.19) results in overall constant C_{LG} as $C_{LG} = \max (C_{LG}(Lp(z)))$ or C_{LG} is defined locally as $C_{LG}(Lp(z))$. It is seen that this model approach is also suitable during the opening phase for $f_c = 1$ Hz. For imposed oscillation frequencies greater than 1 Hz ($f_c \in \{6, 10\}$ Hz) the mixing is enhanced during the opening phase so that the flow is no longer gas dominated and slip model Eq. (6.16) gains accuracy. Again C_{LG} can be set according to Eq. (6.19) so that it is either constant $C_{LG} = \max (C_{LG}(Lp(z)))$ or locally varying as $C_{LG}(Lp(z))$. When the forced oscillation frequency is further increased to $f_c = 10$ Hz and more liquid is injected $V_L \ge 3$ mL the homogenous turbulent flow model becomes most accurate since it captures the increased range during which P_c is increased due to enhanced mixing.

6.5 Conclusion

Experimental observations of the pressure within a time-varying constricted portion of a vertical channel show the combined influence of injected water volume $V_L \in \{0, ..., 5\}$ mL and imposed oscillation frequency $f_c \in \{1, 6, 10\}$ Hz on water-air mixing and hence on the viscous contribution to the pressure drop. Experimental observations and selected flow models both suggest that the flow remains gas dominated during the closing phase for all assessed (f_c, V_L) whereas during the opening phase both f_c and V_L affect mixing and hence the mixture viscosity. For $f_c = 1$ Hz the flow remains gas dominated during the opening phase for all V_L , whereas for $f_c > 1$ Hz, the contribution of water to the mixture increases for all $V_L > 0$ mL so that the flow is no longer gas dominated. Moreover, for $f_c = 10$ Hz and $V_L \ge 3$ mL mixing increases so that the flow becomes homogeneous and turbulent. Further research is necessary to investigate and quantify droplet properties and their distribution and the reciprocal effect on jet and vortex formation and dissipation mechanisms affecting the mixing for different (f_c, V_L) . In this study, geometrical and flow parameters were inspired on flow through the human glottis. With respect to glottal flow, it is seen that the common quasi-one-dimensional steady flow model approach can be extended to water-air flow with the same accuracy when viscous mixing is accounted for and differences observed between the closing and opening phase. It is of interest to further validate selected flow models for different constriction shapes, for self-oscillating deformable glottal replica's as well as to study liquids other than water, *i.e.* either Newtonian like contained in some artificial saliva sprays or non-Newtonian like natural saliva. Furthermore, it is of interest to further investigate the role of squeezing flow due to the constricted wall motion on pressure driven channel flow.

Conclusion and perspectives

In this thesis, several aspects of the physics of vocal folds oscillation have been studied. The context of this study was presented in a first part, summarising the phonatory system with the vocal folds physiology and oscillation features. Then the different physical numbers and the perturbation quantification variables were presented as well as the used experimental set-up. It was shown that experiments allow to mimic order of magnitudes characterising the fluid-structure interaction for adult human speakers, under normal health conditions. In addition, the use of different types of replicas allows to span a large region in the parameter space of flow and structure variables ensuring the generality of the results presented briefly hereafter.

This thesis aimed three contributions from which the following conclusions are formulated.

Firstly, the MSePGG algorithm and workflow were proposed in order to calibrate an existing ePGG device containing a single light source and sensor. The device parameters were quantified (distance, angle, etc.) to develop the algorithm. The MSePGG algorithm uses measurements obtained from several source-sensor distances for a quantitative estimation of the glottal area as function of time. The algorithm is quantitatively validated in real time on two vocal folds replicas (MDR and PLT) with a mean error of 5.4%. Then calibration was illustrated for measurements on a human subject. It is concluded that the proposed algorithm encourages future use and development of this measurement technique. This is of interest as it allows a continuous and non-invasive measurement which does not need medical supervision and has the advantage of a high time resolution compared to other techniques like high speed video.

Secondly, in the continuation of Tokuda *et al.* [5, 6], who studied the vertical level difference for parallel vocal folds, the effect of the angular vertical asymmetry, as occurring in the case of an unilateral vocal fold paralysis, on the vocal folds oscillation is studied. Since the vocal folds are connected together at their anterior end, imposing an angular asymmetry results in a more realistic geometrical representation of this pathological condition. The experimental work is realised on three mechanical vocal folds silicone replicas while systematically increasing the angular asymmetry. Results show that the progressive loss of full initial contact between the two vocal folds, *i.e.* resulting in an increasing glottal leakage, affects their auto-oscillation properties (harmonic frequencies, THD and SNR) and the threshold pressures at oscillation onset and offset. The oscillation frequency decreases and sustaining oscillation becomes more difficult as seen from the increasing threshold pressures. In addition, perturbation parameters often associated with voice quality increase as well. Finally, a vertical level difference model is fitted allowing to predict the general tendency of increasing oscillation onset pressure for large asymmetry angles associated with loss of full contact.

Thirdly, the effect of water spraying was experimentally investigated on the glottal airflow and on the vocal fold auto-oscillation. Three types of vocal fold replicas were used. Forced oscillation allowed to show the influence of water on the glottal airflow itself. It was found that airflow during opening and closing phases is affected as viscous water-air mixing dominates the opening phase. During auto-oscillation it was observed from systematic feature quantifications, for all deformable replicas, that the waveform shape is altered, the period of oscillation increases, perturbations, like SNR or THD, increase and non-linearities are enhanced. This experimental study stresses the need for further studies considering the presence of liquid, like water, during auto-oscillation. Although this work is a first study and different methodologies could be used, current results encourage further research on these effects which are rarely considered in physical studies so far.

Fourthly, a theoretical flow model was proposed and validated accounting for air-water presence during oscillation. In agreement with experimental findings described in the previous paragraph, the model takes into account different air-water mixing during the oscillation phases. As it is a quasi-analytical flow model it is suitable to be implemented in a reduced order fluid-structure interaction model of vocal fold oscillation. The model needs yet to be validated on data obtained on the deformable replicas.

Threshold pressures at auto-oscillation onset and offset as well as perturbation quantities are considered throughout this thesis. A summary of features often considered in voice quality studies (oscillation frequency, threshold pressures, SNR, etc.) is given in Table 7.1.

It is seen that the conditions studied, vocal folds asymmetry and water hydration, have a significant effect on the vocal folds oscillation. In general, as the condition parameter (asymmetry angle α or water volume V_L) increases, threshold pressure increases, the oscillation frequency decreases and perturbations, like SNR, tend to increase up to 45% for UVFP and with 15% up to 45% for the hydration. However, for small condition parameter values, perturbations can be negligible, like for small water volumes ($V_L < 2$ mL), or even decrease as is the case for small vertical asymmetry angles α for which no glottal leakage occurs. As these conditions are very different in nature, on one hand related to the geometry and on the other hand related to the fluid, it is seen that these common perturbation measures are of equal magnitude illustrating the difficulty of identifying markers enabling to relate obtained values to a single condition.

The following perspectives can be formulated.

It is of interest to further validate the MSePGG algorithm firstly on healthy subjects and

eventually on subjects suffering from different vocal folds pathologies. Such measurements will likely lead to further refinement and improvement of the MSePGG algorithm. Moreover, it is obvious from the proposed algorithm that accuracy of measurements on human subjects would benefit from an improvement of the ePGG device with multiple sensors. Such a prototype is developed in [88].

As unilateral vocal fold paralysis has more symptoms than just the vertical angular asymmetry considered in this thesis, the current study needs to be continued. Indeed, an horizontal positioning asymmetry is also observable, as is an asymmetry of mechanical parameters between both vocal folds. Moreover, the vocal folds shape is likely influenced. The vertical angular asymmetry can lead to non-linearities in the vocal folds vibration which can be studied. Besides these physical changes due to unilateral vocal fold paralysis the used experimental methodology can be elaborated. For instance, high speed imaging of the vocal folds auto-oscillation allows to

	f_0	P _{On}	P_{Off}	SNR	THD	ζ_T	$\zeta_{\mathscr{A}}$
	[Hz]	[Pa]	[Pa]	[dB]	[dB]	[%]	[%]
Adult	75 - 250	300 - 1000	100 - 500	24	-9.9	0.3	2.7
MDR	1 - 10	_	_	20	0	0	0
PLT	80-150	250 - 350	250 - 270	25	-10	1.5 - 3	2.2 - 4.6
M5	≈ 100	1100	890	40	-20	0.2	0.3
MRI	≈ 115	980	675	40	-20	0.7	1
EPI	pprox 85	440	325	40	-10	0.6	0.3
Adult•	$92 \leftarrow 222$	_	_	$8 \leftarrow 17$	_	0.3 ightarrow 0.8	$4 \rightarrow 19$
Adult*	167.7 ± 66.6	_	—	3.7 ± 5.6	—	4.4 ± 3.6	9.9 ± 5.1
Global tendencies for increasing initial glottal area $A > 0$							
M5	$123 \leftarrow 135$	$1000 \rightarrow 1600$	$750 \rightarrow 1300$	$25 \leftarrow 46$	$-25 \rightarrow 0$	$\leqslant 2$	$\leqslant 8$
MRI	$115 \leftarrow 138$	$1000 \rightarrow 1300$	800 ightarrow 1100	$25 \leftarrow 45$	$-10 \rightarrow 0$	≤ 3	≤ 10
EPI	$85 \leftarrow 105$	380 ightarrow 650	300 ightarrow 400	$25 \leftarrow 40$	$-6 \rightarrow -3$	≤ 4	≤ 15
Global tendencies for increasing water spraying volume $V_L \ge 1$							
MDR	1 - 10	—	—		0	0	
PLT	≈ 50	_	_	$15 \leftarrow 25$	$-20 \rightarrow -2$	5 ightarrow 25	4 ightarrow 8
M5	pprox 97	_	—	$37 \leftarrow 43$	≈-15	$\leqslant 2$	< 2
MRI	$105 \leftarrow 110$	_	—	$31 \leftarrow 36$	$-17 \rightarrow -15$	$\leqslant 2$	< 2
EPI	$77 \leftarrow 80$		_	$25 \leftarrow 45$	\approx -7	$1 \rightarrow 5$	$0.3 \rightarrow 3$
	Adult MDR PLT M5 MRI EPI Adult• Adult• Adult* M5 MRI EPI MDR PLT M5 MRI EPI	$ \begin{array}{ c c c c c c c c c c c c c c c c c c c$	$ \begin{array}{ c c c c c c c c c c c c c c c c c c c$	$ \begin{array}{ c c c c c c c c c c c c c c c c c c c$	$ \begin{array}{ c c c c c c c c c c c c c c c c c c c$	$ \begin{vmatrix} f_0 \\ [Hz] \end{vmatrix} P_{On} P_{Off} \\ [Pa] \end{vmatrix} SNR THD \\ [dB] [dB] \\ [$	$ \begin{array}{ c c c c c c c c c c c c c c c c c c c$

Table 7.1 – Typical vocal folds physical and perturbation values.

*Unilateral vocal fold paralysis

•: Pinho2012 [172]

*: Schindler2008 [173]

analyse the vibrating structure in more detail.

The current study showed quantitative evidence of the impact of water on main oscillation features. Therefore, more research is needed either experimentally, numerically or theoretically. The use of different liquids with different viscosity is an example of an experimental study directly related to this work. It would allow not only to detail the influence of viscosity on quantified features, but, in addition, it is of interest to further validate the theoretical flow model. From this point of view, the use of different artificial saliva sprays, for which fluid properties are determined in Appendix G, is motivated and first results are shown in Appendix H. Moreover, it has a direct relationship with fluids used in clinical practice. Besides viscosity, non-Newtonian liquids need to be assessed as human vocal folds mucus is non-Newtonian. Next, the influence of the experimental procedure needs to be addressed in more detail: influence of the used nozzle, characterisation of droplet distribution, liquid supply method, observation of flow patterns, etc. Finally, new replicas allowing liquid exchange with the flow have to be developed.

Concerning the theoretical flow model, it would be interesting to consider predictions on human speakers (*e.g.* after an artificial saliva spray was used) as it assumes that structural properties are unaltered and it does not account for fluid transport between the flow and the vocal folds structure. Furthermore, the theoretical model can be expanded as for instance the presence of a liquid bridge between the vocal folds can be accounted for.

Bibliography

- 1. Ruty, N. *Modèles d'interactions fluide parois dans le conduit vocal* PhD thesis (Institut National Polytechnique de Grenoble, Grenoble, France, 2007) (cit. on pp. 1, 2, 19, 23).
- Cisonni, J. Modélisation et inversion d'un système complexe de production de signaux acoustiques PhD thesis (Institut National Polytechnique de Grenoble, Grenoble, France, 2009) (cit. on pp. 1, 2, 19, 23, 24, 106, 115).
- 3. Rosen, C. & Simpson, C. *Operative techniques in laryngology* 3–8 (Springer-Verlag, 2008) (cit. on pp. 1, 2, 7, 13, 14, 64).
- Luizard, P. & Pelorson, X. Threshold of oscillation of a vocal folds replica with unilateral surface growths. *Journal of the Acoustical Society of America* 141, 3050–3058 (2017) (cit. on pp. 1, 16, 71, 79).
- Tokuda, I. & Shimamura, R. Effect of level difference between left and right vocal folds on phonation: Physical experiment and theoretical study. *Journal of the Acoustical Society of America* 142, 482–492 (2017) (cit. on pp. 1, 2, 19, 22, 54, 58, 64–66, 125, 157, 182, 183).
- 6. Shimamura, R. & Tokuda, I. Experimental study on level difference between left and right vocal folds. *Acoust Sci & Tech* **38**, 264–267 (2017) (cit. on pp. 1, 2, 54, 64–66, 125).
- 7. Steinecke, I. & Herzel, H. Bifurcations in an asymmetric vocal-fold model. *Journal of the Acoustical Society of America* **97**, 1874–1884 (1995) (cit. on pp. 1, 9, 54, 100).
- Mergell, P., Herzel, H. & Titze, I. R. Irregular vocal-fold vibration—High-speed observation and modeling. *Journal of the Acoustical Society of America* 108, 2996–3002 (2000) (cit. on pp. 1, 9, 100).
- 9. Wu, L. & Zhang, Z. A computational study of vocal fold dehydration during phonation. *IEEE Transactions on Biomedical Engineering* **64**, 2938–2948 (2017) (cit. on p. 1).
- Honda, K. & Maeda, S. Glottal-opening and airflow pattern during production of voiceless fricatives: A new non-invasive instrumentation. *Journal of Acoustic Society of America 123*, 3738 (2008) (cit. on pp. 2, 30, 31, 33).
- Vaissière, J., Honda, K., Amelot, A., Maeda, S. & Crevier-Buchman, L. Multisensor platform for speech physiology research in a phonetics laboratory. *Journal of the Physical Society of Japan* 14, 65–77 (2010) (cit. on pp. 2, 30, 31, 33).
- Kim, H., Honda, K. & Maeda, S. ePGG, airflow and acoustic data on glottal opening in Korean plosives in International Congress of Phonetic Sciences (Glasgow, UK, 2015), 4 (cit. on pp. 2, 30, 31, 33).

- Sivasankar, M., Carroll, T., Kosinski, A. & Rosen, C. Quantifying the effects of altering ambient humidity on ionic composition of vocal fold surface fluid. *Laryngoscope* 123, 1275–1278 (2013) (cit. on pp. 2, 14, 71).
- 14. Solomon, N. & DiMatta, M. Effects of a vocally fatigue task and systemic hydration on phonation pressure threshold. *Journal of Voice* **14**, 341–362 (2000) (cit. on pp. 2, 14).
- 15. Santana, E., Masson, M. & Araujo, T. The effect of surface hydration on teacher's voice quality: an intervention study. *Journal of Voice* **31**, 5–11 (2017) (cit. on pp. 2, 14).
- 16. Erickson-Levendoski, E. & Sivasankar, M. Investigating the effects of caffeine on phonation. *Journal of Voice* **25**, 215–219 (2011) (cit. on pp. 2, 14).
- 17. Sivasankar, M. & Erickson-Levendoski, E. Influence of obligatory mouth breathing during realistic activities, on voice measures. *Journal of Voice* **26**, 9–13 (2012) (cit. on pp. 2, 14).
- Fisher, K., Ligon, J., Sobecks, J. & Roxe, D. Phonatory effects of body fluid removal. *Journal of Speech, Language, and Hearing Research* 44, 354–367 (2001) (cit. on pp. 2, 14).
- Gartner-Schmidt, J. Voice Therapy for the treatment of voice disorders 4th ed. (eds Bailey, B., Johnson, J. & Newlands, S.) 1388 (Lippincott Williams & Wilkins, Philadelphia, USA, 2006) (cit. on pp. 2, 14, 71).
- Verdolini, K., Titze, I. & Fernell, A. Dependence of phonatory effort on hydration level. *Journal of Speech, Language, and Hearing Research* 37, 1001–1007 (1994) (cit. on pp. 2, 14, 71).
- O'Shaughnessy, D. Speech Communication Human and Machine (Addison-Wesley Publishing Company, 1987) (cit. on pp. 5–7, 9, 33).
- 22. Blandin, R., Hirtum, A. V., Pelorson, X. & Laboissière, R. The effect on vowel directivity patterns of higher order propagation modes. *Journal of Sound and Vibration* **432**, 621–632 (2018) (cit. on p. 6).
- 23. Blandin, R. *Theoretical and experimental study of vocal tract acoustics* PhD thesis (Université Grenoble Alpes, 2016) (cit. on pp. 6, 7).
- Yoshinaga, T., Hirtum, A. V., Nozaki, K. & Wada, S. Influence of the Lip Horn on Acoustic Pressure Distribution Pattern of Sibilant /s/. *Acta Acustica united with Acustica* 104, 145–152 (2018) (cit. on p. 7).
- Pont, A., Guasch, O., Baiges, J., Codina, R. & van Hirtum, A. Computational aeroacoustics to identify sound sources in the generation of sibilant /s/. *International Journal for Numerical Methods in Biomedical Engineering* 35, e3153 (2019) (cit. on p. 7).
- 26. Sataloff, R. T. *et al.* Laryngeal mini-microflap: A new technique and reassessment of the microflap saga. *Journal of Voice* **9**, 198–204 (1995) (cit. on p. 7).
- 27. Hirano, M. & McCormick, K. R. *Clinical Examination of Voice by Minoru Hirano* **4**, 1273–1273 (Acoustical Society of America, 1986) (cit. on p. 7).

- 28. Svec, J. G. *On vibration properties of human vocal folds* PhD thesis (University of Groningen, 2000) (cit. on pp. 8, 9).
- Berry, D. A., Titze, I. R., Story, B. H. & Herzel, H. Bifurcations in excised larynx experiments. *Journal of the Acoustical Society of America* 98, 2930–2930 (1995) (cit. on pp. 9, 19).
- Plant, R., Freed, G. & Plant, R. Direct measurement of onset and offset phonation threshold pressure in normal subjects. *Journal of the Acoustical Society of America* 116, 3640–3646 (2004) (cit. on pp. 9, 24, 81).
- Neubauer, J., Mergell, P., Eysholdt, U. & Herzel, H. Spatio-temporal analysis of irregular vocal fold oscillations: Biphonation due to desynchronization of spatial modes. *Journal* of the Acoustical Society of America **110**, 3179–3192 (2001) (cit. on pp. 9, 54, 100).
- 32. White, F. Viscous Fluid Flow 2nd (McGraw-Hill, New York, 1991) (cit. on pp. 10, 11).
- 33. Pelorson, X., Hirschberg, A., van Hassel, R. R., Wijnands, A. P. J. & Auregan, Y. Theoretical and experimental study of quasisteady-flow separation within the glottis during phonation. application to a modified two-mass model. *Journal of the Acoustical Society of America* **96**, 3416–3431 (1994) (cit. on p. 10).
- Teixeira, J. P. & Fernandes, P. O. Jitter, Shimmer and HNR Classification within Gender, Tones and Vowels in Healthy Voices. *Proceedia Technology* 16, 1228–1237 (2014) (cit. on pp. 12, 24).
- 35. Childers, D. G. & Lee, C. K. Vocal quality factors: Analysis, synthesis, and perception. *Journal of the Acoustical Society of America* **90**, 2394–2410 (1991) (cit. on pp. 12, 24).
- Hong, K. H. & Jung, K. S. Arytenoid appearance and vertical level difference between the paralyzed and innervated vocal cords. *Laryngoscope* **111**, 227–232 (2001) (cit. on pp. 13, 14).
- Oyamada, Y., Yumoto, E., Nakano, K. & Goto, H. Asymmetry of the Vocal Folds in Patients With Vocal Fold Immobility. *Archives of Otolaryngology–Head & Neck Surgery* 131, 399 (2005) (cit. on pp. 13, 14, 58, 64, 66).
- Leydon, C., Sivasankar, M., Falciglia, D., Atkins, C. & Fisher, K. Vocal fold surface hydration: a review. *Journal of Voice* 23, 658–665 (2009) (cit. on pp. 14, 15, 71, 73, 94, 187).
- 39. Jiang, J., Ng, J. & Hanson, D. The effects fo rehydration on phonation in excised canine larynges. *Journal of Voice* **83**, 1536–1552 (1999) (cit. on pp. 14, 71).
- 40. Roy, N., Tanner, K., Gray, S., Blomgren, M. & Fisher, K. An evaluation of the effects of three laryngeal lubricants on phonation threshold pressure (PTP). *Journal of Voice* **17**, 331–342 (2003) (cit. on pp. 14, 71).

- 41. Leydon, C., Wroblewski, M., Eichorn, N. & Sivasankar, M. A meta-analysis of outcomes of hydration intervention on phonation threshold pressure. *Journal of Voice* **24**, 637–643 (2010) (cit. on pp. 14, 71).
- 42. Ayache S. Ouaknine, M., Dejonkere, P., Prindere, P. & Giovanni, A. Experimental study of the effects of surface mucus viscosity on the glottic cycle. *Journal of Voice* **18**, 107–115 (2004) (cit. on pp. 14, 71, 73, 94).
- 43. Dollinger, M., Grohn, F., Berry, D., Eysholdt, U. & Luegmaira, G. Preliminary results on the influence of engineered artificial mucus layer on phonation. *Journal of Speech, Language, and Hearing Research* **57**, 637–647 (2014) (cit. on pp. 14, 71, 94).
- 44. Alves, M., Kruger, E., Pillay, B., van Lierde, K. & van der Linde, J. The effect of hydration on voice quality in adults: a systematic review. *Journal of Voice* **In Press** (2018) (cit. on pp. 14, 71, 73, 94, 104).
- 45. Preetha, A. & Banerjee, R. Comparison of artificial saliva substitutes. *Trends in Biomaterials & Artificial Organs* **18**, 178–186 (2005) (cit. on pp. 14, 15, 71–73, 187).
- Tanner, K. *et al.* Comparing nebulized water versus saline after laryngeal desiccation challenge in Sjögren's Syndrome. *Laryngoscope* 123, 2787–2792 (2013) (cit. on pp. 14, 71, 73).
- 47. Vilain, C. *et al.* Experimental validation of a quasi-steady theory for the flow through the glottis. *Journal of Sound and Vibration* **276**, 475–490 (2004) (cit. on pp. 16, 75, 104, 105, 115).
- Cisonni, J., Van Hirtum, A., Pelorson, X. & Willems, J. Theoretical simulation and experimental validation of inverse quasi one-dimensional steady and unsteady glottal flow models. *Journal of the Acoustical Society of America* 124, 535–545 (2008) (cit. on pp. 16, 33, 75, 104–106, 115).
- 49. Van Hirtum, A. & Pelorson, X. High-speed imaging to study an auto-oscillating vocal fold replica for different initial conditions. *International Journal of Heat and Fluid Flow* 9 (2017) (cit. on pp. 16, 17, 33, 44).
- 50. Pickup, B. & Thomson, S. Flow-induced vibratory response of idealized versus magnetic resonance imaging-based synthetic vocal fold models. *Journal of the Acoustical Society of America* **128**, 124–129 (2010) (cit. on pp. 19, 22, 157, 167, 169).
- Murray, P. & Thomson, S. Vibratory responses of synthetic, self-oscillating vocal fold models. *Journal of the Acoustical Society of America* 132, 3428–3438 (2012) (cit. on pp. 19, 20, 22, 157, 167, 169).
- 52. Hirano, M., Kurita, S. & Nakashima, T. in (eds Bless, D. & J.H., A.) 22–43 (College-Hill Press, 1983) (cit. on pp. 20, 24, 81).
- 53. Alipour, F. & Titze, I. Elastic models of vocal fold tissues. *Journal of the Acoustical Society of America* **90**, 1326–1331 (1991) (cit. on p. 20).

- 54. Min, Y., Titze, I. & Alipour, F. Stress-Strain response of the human vocal ligament. *Annals of Otology, Rhinology & Laryngology* **104,** 563–569 (1995) (cit. on p. 20).
- Chan, R., Fu, M., Young, L. & Tirunagari, N. Relative contributions of collagen and elastin to elasticity of the vocal fold under tension. *Annals of Biomedical Engineering* 35, 1471– 1483 (2007) (cit. on p. 20).
- Smith, S. & Thomson, S. Effect of inferior surface angle on the self-oscillation of a computational vocal fold model. *Journal of the Acoustical Society of America* 131, 4062–4075 (2012) (cit. on p. 20).
- 57. Drechsel, J. S. & Thomson, S. L. Influence of supraglottal structures on the glottal jet exiting a two-layer synthetic, self-oscillating vocal fold model. *Journal of the Acoustical Society of America* **123**, 4434–4445 (2008) (cit. on pp. 22, 167, 169).
- 58. Riede, T., Tokuda, I. T., Munger, J. B. & Thomson, S. L. Mammalian laryngseal air sacs add variability to the vocal tract impedance: Physical and computational modeling. *Journal of the Acoustical Society of America* **124**, 634–647 (2008) (cit. on pp. 22, 167, 169).
- 59. Murray, P. & Thomson, S. Synthetic, multi-layer, self-oscillating vocal fold model fabrication. *Journal of Visualized Experiments* **58**, e3498 (2011) (cit. on pp. 22, 157, 167, 169).
- 60. Scherer, R. *et al.* Intraglottal pressure profiles for a symmetric and oblique glottis with a divergence angle of 10 degrees. *Journal of the Acoustical Society of America* **109**, 1616–1630 (2001) (cit. on p. 21).
- 61. Mobashir, M., Mohamed, A., Quriba, A., Anany, A. & Hassan, E. Linear measrurements of vocal folds and laryngeal dimensions in freshly excised human larynges. *Journal of Voice* **32**, 525–529 (2018) (cit. on pp. 24, 81).
- 62. Lucero, J. C. & Koenig, L. L. Time normalization of voice signals using functional data analysis. *Journal of the Acoustical Society of America* **108**, 1408–1420 (2000) (cit. on p. 24).
- 63. Garcia, M. Observations on the human voice. *Proc. Royal Society of London* **7**, 399–410 (1855) (cit. on p. 30).
- Lohscheller, J., Eysholdt, U., Toy, H. & Döllinger, M. Phonovibrography: mapping highspeed movies of vocal fold vibrations into 2-D diagrams for visualizing and analyzing the underlying laryngeal dynamics. *IEEE Transactions on Medical Imaging* 27, 300–309 (2008) (cit. on p. 30).
- 65. Wurzbacher, T. *et al.* Calibration of laryngeal endoscopic high-speed image sequences by an automated detection of parallel laser line projections. *Medical image analysis* **12**, 300–317 (2008) (cit. on p. 30).

- Deliyski, D. & Hillman, R. State of the Art Laryngeal Imaging: Research and Clinical Implications. *Current Opinion in Otolaryngology & Head and Neck Surgery* 18, 147–152 (2011) (cit. on pp. 30, 32).
- 67. Aichinger, P. *et al.* Comparison of an audio-based and a video-based approach for detecting diplophonia. *Biomedical Signal Processing and Control* **31**, 576–585 (2017) (cit. on pp. 30, 46).
- 68. Aichinger, P. *et al.* Tracking of Multiple Fundamental Frequencies in Diplophonic Voices. *IEEE Transactions on Audio, Speech, and Language Processing* **26**, 330–341 (2018) (cit. on pp. 30, 46).
- Sawashima, M. & Miyazaki, S. Stereo-fiberscopic measurement of the larynx: a preliminary experiment by use of ordinary laryngeal fiberscopes. *Ann. Bull. RILP* 8, 1–12 (1974) (cit. on p. 30).
- 70. Sawashima, M. *et al.* Measurements of the vocal fold length by use of stereoendocope a preliminary study. *Ann. Bull. RILP* **15**, 9–16 (1981) (cit. on p. 30).
- Imagama, H., Sakakibara, K., Tokuda, I., Otsuka, M. & Tayama, N. Estimation of glottal area function using stereo-endoscopic high-speed digital imaging in Proc. Interspeech (2010), 1–4 (cit. on pp. 30, 49).
- 72. Baken, R. *Clinical measurement of speech and voice* (Allyn and Bacon, USA, pp. 197-240, 1987) (cit. on p. 30).
- 73. Hertz, C., Lindström, K. & Sonesson, B. Ultrasonic recording of the vibrating vocal folds. *Acta Otolaryngologica* **69**, 223–230 (1970) (cit. on p. 30).
- Halmet, S. & Reid, J. Transmission of ultrasound through the larynx as a mean of determining vocal-fold activity. *IEEE Transactions on Biomedical Engineering* 19, 34–37 (1972) (cit. on p. 30).
- 75. Kaneko, T. *et al. Ultrasonic observation of vocal folds vibration* (Univ. of Tokyo, USA, pp. 107-117, 1981) (cit. on p. 30).
- Friedman, E. Role of ultrasound in the assessment of vocal cord function in infants and children. *Annals of Otology, Rhinology & Laryngology* **106**, 199–209 (1997) (cit. on p. 30).
- 77. Arutti, A. & Poumayrac, M. Larynx ultrasonography: an alternative technique in the evaluation of the aero-digestive crossroad. *Rev. Imagenol.* **14**, 30–36 (2010) (cit. on p. 30).
- Parangi, S. Editorial: translarygeal vocal cord ultrasound: ready for prime time. *Surgery* 159, 67–69 (2016) (cit. on p. 30).
- 79. Krausert, C. *et al.* Mucosal wave measurement and visualization techniques. *Journal of Voice* **25**, 395–405 (2011) (cit. on p. 30).

- Masood, M., Huang, B., Goins, A. & Hackman, T. Anatomical factors affecting the use of ultrasound to predict vocal fold motion: a pilot study. *American Journal of Otolaryngol*ogy **39**, 413–417 (2018) (cit. on p. 30).
- 81. Sonesson, B. A method for studying the vibratory movements of the vocal folds. *Journal of Laryngology & Otology* **73**, 732–737 (1959) (cit. on p. 30).
- 82. Delpy, D. & Cope, M. Quantification in tissue near-infrared spectroscopy. *Philosophical Transactions of the Royal Society B* **352**, 649–659 (1997) (cit. on pp. 31, 33).
- 83. Lister, T., Wright, P. & Chappell, P. Optical properties of human skin. *Journal of Biomedical Optics* **17**, 1–15 (2012) (cit. on pp. 31, 33).
- Jacques, S. Optical properties of biological tissues: a review. *Phys. Med. Biol.* 58, R37–R61 (2013) (cit. on pp. 31, 33).
- 85. Burnard West, J. *Bio-engineering aspects of the lung, Volume 3* (ed Dekker, M.) (Marcel Dekker, INC., 1977) (cit. on p. 33).
- 86. Brancatisano, T., Collett, P. & Engel, L. Respiratory movements of the vocal cords. *Journal of Applied Physiology* **54**, 1269–1276 (1983) (cit. on pp. 32, 33, 43, 47).
- 87. Scheinherr, A. *et al.* Realistic glottal motion and airflow rate during human breathing. *Medical Engineering & Physics* **37**, 829–839 (2015) (cit. on pp. 32, 33).
- 88. Cambez, M., Cruaud, N., Eddiza, A., Fabregettes, O. & Wauthier, J. *Mesurer la voix* tech. rep. (PHELMA, Institut National Polytechnique de Grenoble, 2018) (cit. on pp. 36, 127).
- 89. Nocedal, J. & Wright, J. Numerical Optimization (Springer, 2006) (cit. on p. 41).
- 90. Aichinger, P. *Diplophonic voice: definitions, models and detection* PhD thesis (Graz University of Technololgy, Austria, 2014) (cit. on p. 46).
- 91. Aichinger, P. *et al.* Towards Objective Voice Assessment: The Diplophonia Diagram. *Journal of Voice* **31**, 1–10 (2017) (cit. on p. 46).
- 92. Stevens, N. K. The acoustic/articulatory interface. *Acoustical Science and Technology* **26**, 410–417 (2005) (cit. on p. 49).
- 93. Ishizaka, K. & Isshiki, N. Computer simulation of pathological vocal-cord vibration. *Journal of the Acoustical Society of America* **60**, 1193–1198 (1976) (cit. on p. 54).
- Zhang, W., Hibiki, T. & Mishima, K. Correlations of two-phase frictional pressure drop and void fraction in mini-channel. *International Journal of Heat and Mass Transfer* 53, 453–465 (2010) (cit. on pp. 54, 113, 114).
- 95. Zhang, Z., Kreiman, J., Gerratt, B. R. & Garellek, M. Acoustic and perceptual effects of changes in body layer stiffness in symmetric and asymmetric vocal fold models. *Journal of the Acoustical Society of America* **133**, 453–462 (2013) (cit. on p. 54).

- Lucero, J. C., Schoentgen, J., Haas, J., Luizard, P. & Pelorson, X. Self-entrainment of the right and left vocal fold oscillators. *Journal of the Acoustical Society of America* 137, 2036–2046 (2015) (cit. on p. 54).
- 97. Zañartu, M. *et al.* Modeling the effects of a posterior glottal opening on vocal fold dynamics with implications for vocal hyperfunction. *Journal of the Acoustical Society of America* **136**, 3262–3271 (2014) (cit. on p. 57).
- Luegmair, G., Mehta, D. D., Kobler, J. B. & Dollinger, M. Three-Dimensional Optical Reconstruction of Vocal Fold Kinematics Using High-Speed Video With a Laser Projection System. *IEEE Transactions on Medical Imaging* 34, 2572–2582 (2015) (cit. on p. 67).
- Semmler, M. *et al.* 3D Reconstruction of Human Laryngeal Dynamics Based on Endoscopic High-Speed Recordings. *IEEE Transactions on Medical Imaging* 35, 1615–1624 (2016) (cit. on p. 67).
- Sommer, D. E. *et al.* Estimation of inferior-superior vocal fold kinematics from high-speed stereo endoscopic datain vivo. *Journal of the Acoustical Society of America* 136, 3290–3300 (2014) (cit. on p. 67).
- 101. Titze, I. The physics of small-amplitude oscillation of the vocal folds. *Journal of the Acoustical Society of America* **83**, 1536–1552 (1988) (cit. on pp. 71, 182).
- 102. Ruty, N., Pelorson, X., Van Hirtum, A., Lopez, I. & Hirschberg, A. An in-vitro setup to test the relevance and the accuracy of low-order models of the vocal folds. *Journal of the Acoustical Society of America* **121**, 479–490 (2007) (cit. on pp. 71, 76, 79).
- 103. Avanzini, F. Simulation of vocal fold oscillation with a pseudo-one-mass physical model. *Speech Communication* **50**, 95–108 (2008) (cit. on p. 71).
- 104. Van Hirtum, A., Bouvet, A. & Pelorson, X. Pressure drop for adiabatic air-water flow through a time-varying constriction. *Physics of Fluids* **30**, 101901 (2018) (cit. on pp. 71, 76, 77, 81, 82, 93).
- 105. Takens, F. in (ed Warwick) 1–366 (Springer-Verlag, 1981) (cit. on pp. 72, 79).
- Mende, W., Herzel, H. & Wermke, K. Bifurcations and chaos in newborn infant cries. *Physics Letters A* 145, 418–424 (1990) (cit. on p. 72).
- Herzel, H. & Wendler, J. Evidence of chaos in phonatory samples. *EUROSPEECH*, 263–266 (1991) (cit. on p. 72).
- 108. Jiang, J., Zhang, Y. & McGiligan, C. Chaos in voice, from modeling to measurement. *Journal of Voice* **20**, 2–17 (2006) (cit. on pp. 72, 100).
- 109. Awan, S., Roy, N. & Jiang, J. Nonlinear dynamic analysis of disordered voice: the relationship between the correlation dimension (D2) and pre-/post treatment change in perceived dysphonia severity. *Journal of Voice* **25**, 285–293 (2010) (cit. on pp. 72, 100).

- Choi, S., Zhang, Y., Jiang, J., Bless, D. & Welham, N. Nonlinear dynamic-based analysis of severe dysponia in patients with vocal fold scar and sulcus vocalis. *Journal of Voice* 26, 566–576 (2012) (cit. on pp. 72, 100).
- 111. Lin, L., Zhang, Y., Calawerts, W. & Jiang, J. Vibratory dynamics of four types of excised larynx phonations. *Journal of Voice* **30**, 649–655 (2016) (cit. on pp. 72, 100).
- 112. Liu, B., Polce, E., Sprott, J. & Jiang, J. Applied chaos level test for validation of signal conditions underlying optimal performance of voice classification methods. *Journal of Speech, Language, and Hearing Research* **61**, 1130–1139 (2018) (cit. on pp. 72, 100).
- Banbrook, M., McLaughlin, S. & Mann, I. Speech characterization and synthesis by nonlinear methods. *IEEE Transactions on Audio, Speech, and Language Processing* 7, 1–17 (1999) (cit. on p. 72).
- 114. Behrman, A. Global and local dimensions of vocal dynamics. *Journal of the Acoustical Society of America* **105**, 432–443 (1999) (cit. on p. 72).
- Zhang, Y., Jiang, J. & Wallace, S. Comparison of nonlinear dynamic methods and perturbation methods for voice analysis. *Journal of the Acoustical Society of America* 118, 2551–2560 (2005) (cit. on p. 72).
- Maragos, P. & Potamianos, A. Fractal dimensions of speech sounds: computation and application to automatic speech recognition. *Journal of the Acoustical Society of America* 105, 1925–1932 (1999) (cit. on p. 72).
- 117. Tokuda, I. T., Horáček, J., Švec, J. G. & Herzel, H. Comparison of biomechanical modeling of register transitions and voice instabilities with excised larynx experiments. *The Journal of the Acoustical Society of America* **122**, 519–531 (2007) (cit. on pp. 72, 99).
- Hwang, S., Litt, M. & Forsman, W. Rheological properties of mucus. *Rheologica Acta* 8, 438–448 (1969) (cit. on pp. 72, 73).
- 119. Schwarz, W. The rheology of saliva. *Journal of Dental Research* **66**, 660–664 (1987) (cit. on pp. 72, 73).
- 120. Lai, S., Wang, Y., Wirtz, D. & Hanes, J. Micro- and macrorheology of mucus. *Advanced Drug Delivery Reviews* **61**, 86–100 (2009) (cit. on pp. 72, 73).
- 121. Phillips, J., Wong, L. & Yeates, D. Bidirectional transpotchelial water transport: measurement and governing mechanisms. *Biophysical Journal* **76**, 869–877 (1999) (cit. on p. 73).
- 122. Fisher, K., Telser, A., Phillips, J. & Yeates, D. Regulation of vocal fold transepithelial water fluxes. *Journal of Applied Physiology* **91**, 1401–1411 (2001) (cit. on p. 73).
- 123. Sivasankar, M. & Fisher, K. Vocal fold epithelial response to luminal osmotic perturbation. *Journal of Speech, Language, and Hearing Research* **50**, 886–898 (2007) (cit. on p. 73).
- 124. Miri, A. Mechanical characterization of vocal fold tissue: a review study. *Journal of Voice* **28,** 657–666 (2014) (cit. on p. 73).

125.	Bertrand, E. Optmisation et utilisaton de une maquette du système phonatoire humair
	tech. rep. (PHELMA, Institut National Polytechnique de Grenoble, 2017) (cit. on p. 76).

- 126. Kantz, H. & T., S. *Nonlinear time series analysis* (Cambridge University Press, 2003) (cit. on pp. 79, 80).
- 127. Robinson, J. A topological delay embedding theorem for infinite-dimensional dynamical systems. *Nonlinearity* **18**, 2135–2143 (2005) (cit. on p. 79).
- 128. Grassberger, P. & Procaccia, I. Characterization of strange attractors. *Physical Review Letter* **50**, 347–349 (1983) (cit. on p. 80).
- 129. Fraser, A. & Swinney, H. Independent coordinates for strange attractors from mutual information. *Physical Review A* **33**, 1134–1140 (1986) (cit. on p. 80).
- Lee, J. & Verleysen, M. Nonlinear dimensionality reduction 308 (Springer, 2007) (cit. on p. 80).
- Kennel, M., Brown, R. & Abarbanel, H. Determining embedding dimension for phasespace reconstruction using a geometrical construction. *Physical Review A* 45, 3403–3411 (1992) (cit. on p. 80).
- 132. Hegger, R. & Kantz, H. Improved false nearest neighbor method to detect determinism in time series data. *Physical Review E* **60**, 4970–4973 (1999) (cit. on p. 80).
- 133. Eckmann, J., Kamphorst, S. & Ruelle, D. Recurrence plots of dynamical systems. *Europhysics letters* **4**, 973–977 (1987) (cit. on p. 80).
- 134. Bouvet, A., Pelorson, X. & Van Hirtum, A. Influence of water spraying on an oscillating channel. **Revised** (2019) (cit. on pp. 94, 99).
- Deverge, M. *et al.* Influence of collision on the flow through in-vitro rigid models of the vocal folds. *Journal of the Acoustical Society of America* **114**, 1–9 (2003) (cit. on pp. 104, 105, 115).
- Van Hirtum, A., Cisonni, J. & Pelorson, X. On quasi-steady laminar flow separation in the upper airways. *Communications in Numerical Methods in Engineering* 25, 447–461 (2009) (cit. on pp. 104–106, 115).
- 137. Verdolini, K., Titze, I. & Druker, D. Changes in phonation threshold pressure with induced conditions of hydration. *Journal of Voice* **4**, 142–151 (1990) (cit. on p. 104).
- Fujiki, R., Chapleau, A., Sundarrajan, A., McKenna, V. & Sivasankar, M. The interaction of surface hydration and vocal loading on voice measures. *Journal of Voice* **31**, 211–217 (2017) (cit. on p. 104).
- 139. Hanratty, T. *Physics of gas-liquid flows* 333 (Cambridge University Press, 2013) (cit. on pp. 105, 107, 112).
- 140. Awad, M. & Muzychka, Y. Effective property models for homogeneous two-phase flows. *Experimental Thermal and Fluid Science* **33**, 106–113 (2008) (cit. on p. 105).

- Xu, Y., Fang, X., Su, X., Zhou, Z. & Chen, W. Evaluation of frictional pressure drop correlations for two-phase flow in pipes. *Nuclear engineering and design* 253, 86–97 (2012) (cit. on pp. 105, 113).
- 142. Muculus, M. Entrainment waves in decelerating transient turbulent jets. *Journal of Fluid Mechanics* **638**, 117–140 (2009) (cit. on pp. 105, 119).
- 143. Grosshans, H., Szasz, R. & Fuchs, L. Enhanced liquid-gas mixing due to pulsating injection. *Computers & Fluids* **107**, 196–204 (2015) (cit. on pp. 105, 119).
- Wu, B., Van Hirtum, A. & Luo, X. Pressure driven steady flow in constricted channels of different cross section shapes. *International Journal of Applied Mechanic* 5, 1–19 (2013) (cit. on pp. 105, 106).
- 145. Van Hirtum, A. Analytical modeling of constricted channel flow. *Mechanics Research Communications* 83, 53–57 (2017) (cit. on pp. 105, 106).
- 146. Cisonni, J., Van Hirtum, A., Luo, X. & Pelorson, X. Experimental validation of quasione-dimensional and two-dimensional steady glottal flow models. *Medical & Biological Engineering & Computing* 48, 903–910 (2010) (cit. on p. 106).
- 147. Owen, W. Two-phase pressure gradient. *ASME Int. Developpement Heat Transfer Part II*, 363–368 (1961) (cit. on pp. 107, 108, 111).
- 148. Davidson, W. *et al.* Studies of heat transmission through boiler tubing and pressures from 500-3300 Lbs. *Journal of Fluids Engineering* **65**, 553–591 (1943) (cit. on pp. 107, 108).
- 149. Garcia, F. *et al.* Power law and composite power law friction factor correlations for laminar and turbulent gas-liquid flow in horizontal pipelines. *International Journal of Multiphase Flow* **29**, 1605–1624 (2003) (cit. on pp. 107, 108).
- McAdams, W., Woods, W. & Heroman, L. Vaporization inside horizontal tubes IIbenzene-oil mixtures. *Transactions of ASME* 64, 193–200 (1942) (cit. on pp. 108, 109, 111).
- Cicchitti, A., Lombaradi, C., Silversti, M., Soldaini, G. & Zavattarlli, R. Two-phase cooling experiments - pressure drop heat transfer burnout measurements. *Energia Nucleare* 7, 407–425 (1960) (cit. on pp. 108, 109, 111).
- 152. Dukler, A., Moye, W. & Cleveland, R. Frictional pressure drop in two-phase flow. *AIChE Journal* **10**, 38–51 (1964) (cit. on pp. 108, 109).
- 153. Beattie, D. & Whalley, P. Simple two-phase frictional pressure drop calculation method. *International Journal of Multiphase Flow* **8**, 83–87 (1982) (cit. on pp. 108, 109).
- Lin, S., Kwok, C., Li, R., Chen, Z. & Chen, Z. Local frictional pressure drop during vaporization for R-12 through capillary tubes. *International Journal of Multiphase Flow* 17, 95–102 (1991) (cit. on pp. 108, 109).

155.	Fourar, M. & Bories, S. Experimental study of air-water two-phase flow through a fracture
	(narrow channel). International Journal of Multiphase Flow 21, 621-637 (1995) (cit. on
	pp. 108, 109).

- Aung, N. & Yuwono, T. Evaluation of mixture viscosity models in the prediction of twophase flow pressure drops. *ASEAN Journal on Science and Technology for Development* 29, 115–128 (2012) (cit. on pp. 108, 110, 111).
- 157. Collier, J. & Thome, J. *Convective boiling and condensation* 640 (Claredon Press, 1994) (cit. on p. 109).
- 158. Kashinsky, O. Experimental study of laminar bubbly flows in a vertical pipe. *Experiments in Fluids* **15**, 308–314 (1992) (cit. on p. 110).
- 159. Sato, Y. & Sekoguchi, K. Liquid velocity distribution in two-phase bubbly flow. *International Journal of Multiphase Flow* **2**, 79–95 (1975) (cit. on p. 110).
- Lockhart, P. & Martinelli, R. Proposed correlation of data for isothermal two-phase twocomponent flow in pipes. *Chemical Engineering Progress* 45, 39–48 (1949) (cit. on p. 112).
- Chisholm, D. A theoretical basis for the Lockhart-Martinelli correlation for two-phase flow. *International Journal of Heat and Mass Transfer* 10, 1767–1778 (1967) (cit. on p. 112).
- Lee, H. & Lee, S. Pressure drop correlations for two-phase flow within horizontal rectangular channels with small heights. *International Journal of Multiphase Flow* 27, 783–796 (2001) (cit. on p. 113).
- 163. Zhao, T. & Bi, Q. Pressure drop characteristics of gas-liquid two-phase flow in vertical miniature triangular channels. *International Journal of Heat and Mass Transfer* 44, 2523– 2534 (2001) (cit. on p. 113).
- Chung P.M.Y.and Kawaji, M., Kawahara, A. & Shibata, Y. Two-phase flow through square and circular microchannels effect of channel geometry. *Journal of Fluids Engineering* 126, 546–552 (2004) (cit. on p. 113).
- Fries, D., Trachsel, F. & von Rohr, P. Segmented gas-liquid flow characterization in rectangular microchannels. *International Journal of Multiphase Flow* 34, 1108–1118 (2008) (cit. on p. 113).
- 166. Muzychka, Y. & Awad, M. Asymptotic generalizations of the Lockhart-Martinelli method for two phase flows. *Journal of Fluids Engineering* **132**, 1–12 (2010) (cit. on p. 113).
- Li, X. & Hibiki, T. Frictional pressure drop correlation for two-phase flows in mini and micro single-channels. *International Journal of Multiphase Flow* **90**, 29–45 (2017) (cit. on p. 113).

- Trela, M., Kwidzinski, R. & Lackowski, M. Generalization of Martinelli-Nelson method of pressure drop calculation in two-phase flows. *E3S Web of Conferences* 13, 02006 (Jan. 2017) (cit. on p. 113).
- 169. Wallis, G. *One Dimensional Two Phase Flow* 408 (McGraw-Hill, New York, 1969) (cit. on p. 113).
- Mishima, K. & Hibiki, T. Some characteristics of air-water two-phase flow in small diameter vertical tubes. *International Journal of Multiphase Flow* 22, 703–712 (1996) (cit. on p. 114).
- 171. Kandlikar, S. Fundamental issues related to flow boiling in minichannels and microchannels. *Experimental Thermal and Fluid Science* **26**, 389–407 (2002) (cit. on p. 114).
- Pinho, C. M. R., Jesus, L. M. T. & Barney, A. Aerodynamic measures of speech in unilateral vocal fold paralysis (UVFP) patients. *Logopedics Phoniatrics Vocology* 38, 19–34 (2012) (cit. on p. 127).
- 173. Schindler, A. *et al.* Vocal improvement after voice therapy in unilateral vocal fold paralysis. *Journal of Voice* **22**, 113–118 (2008) (cit. on p. 127).
- 174. Pascal, J.-C. Vibration et acoustique 1 & 2 (2007 2009) (cit. on p. 151).
- 175. Smooth-On. Smooth-on (ed Smooth-On) https://www.smooth-on.com/ products/(2020) (cit. on pp. 169, 174).
- 176. Abdallah, W. *et al.* Fundamentals of wettability. *Oil Field Review* **44-61** (2007) (cit. on p. 191).

APPENDIX A

Axes and planes

In order to describe the point of view of the different figures, the medical planes and axes used, illustrated in Figure A.1, are:

Medical planes:

- Medio-frontal plane
- Transverse plane
- Medio-sagittal plane

- Axes:
 - Posterior anterior
 - Inferior superior
 - Right left



Figure A.1 – Illustration of the different planes and axes used.

APPENDIX B

Calibration of sensors

Measurement are performed using sensors that convert physical quantities such as pressure or aperture height into an electrical voltage recorded by the digital acquisition system. The response of these sensors are susceptible on experimental conditions (position, atmospheric pressure, brightness, etc.). A calibration is required to convert the measured voltage back into physical parameters. In this works calibration is performed fitting a linear model to measured data.

The calibration accuracy is expressed with the coefficient of determination R^2 (%):

$$R^2 = 1 - \frac{\sigma^2}{\sigma_{data}^2},\tag{B.1}$$

which expresses how well the calibration model fits experimental data with sample variance σ_{data}^2 about its mean value and σ^2 the sample variance of the model residuals. Accuracy increases as R^2 approaches 100 %.

The voltage signals from the sensors presented below are conditioned by a National Instruments SCXI-1121 pre-processing unit. Their digitisation is carried out with an acquisition card National Instruments PCI-6225 combined with a National Instruments BNC-2110 input/output card, the sampled signals are then processed with the National Instruments Labview software programs.

B.1 Pressure sensor

Pressure measurements are conducted with piezoresistive pressure transducers type Kulite XCS-093 and Endevco 8507C-2 and 8507C-5 types. The calibration of these sensors is performed against a liquid column manometer Kimo MG80 and TJ300 (accuracy \pm 5 Pa). Measurements are made for flows in a tube of uniform cross-section (diameter 25 mm) with a constriction at its end to obtain pressures higher than atmospheric pressure. The uniform section contains a wall pressure tap (diameter 0.4 mm). Figure B.1 shows a typical example of a calibration curve for three sensors representing all types of sensors used. All these sensors have linear responses ($R^2 > 99\%$) but have different gains as summarized in Table B.1.

Endevco	Sensibility [Pa/V]	<i>R</i> ² [%]	Kulite	Sensibility [Pa/V]	$\begin{bmatrix} R^2 \\ [\%] \end{bmatrix}$
2 type n°0	788	99.74	n°0	3351	99.98
2 type n°1	251	99.95	n°3	2498	99.99
2 type n°3	294	99.98	n°5	2455	99.55
5 type n°1	564	99.98			

Table B.1 – Pressure sensors calibration.



Figure B.1 – Calibration curves obtained for the sensors of Kulite XCS-093, Endevco 8507C-2 and Endevco 8507C-5 types: pressure measured by a liquid column manometer as a function of the voltage measured by the sensors.

B.2 Optical sensor

An OPTEK OPB700 optical sensor is used to measure the imposed opening of the constriction of the rigid MDR vocal folds replica. The calibration of this sensor is performed with different known minimum constriction heights of the replica. This height is adjusted by placing metal lamellae of known thickness (0 up to 1 mm with accuracy 0.05 mm) between the two vocal folds. Figure B.2 shows the calibration curve of this sensor. The sensor position must be adjusted so that the minimum and maximum opening is in the linear response area of the sensor. During this calibration, a sensitivity of 0.347 mm/V, with accuracy $R^2 = 99.51$ % is obtained.



Figure B.2 – Calibration curve obtained for the optical sensor OPB700: height of the aperture of the MDR replica constriction as a function of voltage given by the sensor.

B.3 Photo-diode

For the mechanical response measurements, the aperture of vocal folds replicas is measured using an optical system using a photo-diode for its light sensitivity (wavelength 635 nm). The calibration of the photo-diode is carried out from the imposed aperture at the minimum of the constriction of the MDR replica. As for the optical sensor, the imposed aperture is varied by placing metal lamellae whose thickness is known between the vocal folds. Figure B.3 shows the calibration curve of the photo-diode (potentiometer = 100 Ω) whose response is linear with a sensitivity of -9.31 mm/V, $R^2 = 99.82$ %.

To be able to recover its signal, the photo-diode is mounted on an electronic circuit detailed in Figure B.4. The potentiometer allows to adapt the linear response zone according to the brightness of the laser (wavelength 635 nm). The negative gain is explained because the photodiode signal is connected to the negative input of the operational amplifier and is compared to 0 V connected to the positive input.

B.4 Laser transceiver

A Panasonic HL-G112-A-C5 laser transceiver (wavelength 655 nm) is calibrated by displacing a steady target between 60 mm and 180 mm (sensor measurement range) with a step of 10 mm. This is illustrated in Figure B.5. The calibration is repeated for different materials (steel, latex and silicone: taken from different VF replicas) to ensure that the calibration is not



Figure B.3 – Example of calibration curve obtained for the photo-diode BPW34: aperture imposed with the MDR replica constriction as a function of voltage given by the sensor.



Figure B.4 – Schematic drawing of the electronic circuit allowing to adapt and recover the photodiode signal.

influenced by the material of the reflecting surface.

Figure B.6 shows the linear calibration curves ($R^2 > 99\%$) of the transceiver and Table B.2 the sensibility obtained for each material. These are close enough to be considered as the same at 12 mm/V.



Figure B.5 – Setup for the laser calibration. The laser (left) pointing onto the target (right) placed between 60 mm and 180 mm.

Table B.2 – Sensitivities of the laser transceiver calibration for different materials.

	Sensibility [mm/V]	<i>R</i> ² [%]
Steel	12.05	100.0
Latex	12.02	99.99
Silicone	11.94	99.99



Figure B.6 – Calibration curves obtained for the laser transceiver. Distance as a function of voltage between 60 mm and 180 mm.

Mechanical frequency response

Deformable PLT and silicone vocal folds replicas (Section 2.5.2 and 2.5.3) are structures capable of auto-oscillation due to a FS interaction. The resulting oscillation features depend on their mechanical properties. These properties can be determined from the frequency analysis of the mechanical frequency response measurement outlined in this appendix.

C.1 Experimental set up

A shaker (Modalshop K2007E) equipped with a steel bar (length $l_{bar} = 230$ mm and diameter $d_{bar} = 3.7$ mm given a cross section $S_{bar} =$ first resonance 3×10^6 Hz¹) is positioned so that the free bar end is in contact with the VF replica. A laser (wavelength 635 nm, beam width 5 mm), two lens (L₁ concave focal 100 mm and L₂ convex focal 100 mm) and photo-diode (BPW34) (accuracy 0.05 mm, calibration in Appendix B.3) are aligned along the inferior and superior side of the replica respectively as illustrated in Figure C.1. Sinusoidal sweep signals from 60 Hz to 300 Hz with duration 30 s are fed to the shaker in order to excite the structure. An accelerometer (PCB 353B18 with signal conditioner PCB 482A21) is attached to the mobile bar to measure the applied excitation x(t). The photo-diode (BPW34) is used to quantify the varying amount of laser light passing through the replica during excitation given the response y(t) of the replica to the excitation. Note that alternatively the laser tranceiver (calibration in B.4) can be used.

C.2 Frequency response function

Frequency response functions (FRF) H(f) = X(f)/Y(f), representing the amplitude ratio of the Fourier transforms of response $X(f) = \mathscr{F}(x(t))$ to excitation $Y(f) = \mathscr{F}(y(t))$, are estimated in order to characterise mechanical resonances. Frequency f^M associated with maximum peak amplitude |H(f)|, phase change $\phi_{H(f)}$ and -3 dB frequency bandwidth Δf^M

 $^{{}^{1}}f^{M} = \frac{1}{2\pi} (\beta_{n} l_{bar})^{2} \sqrt{\frac{\mathscr{E}_{bar} I}{\rho_{bar} S_{bar} l_{bar}^{4}}}$ with fixed-free factor $\beta_{n} l_{bar} = 1.87510407$ [174], $\mathscr{E}_{bar} = 200 \times 10^{9}$ Pa, $I = \frac{\pi d_{bar}^{4}}{32}$ and $\rho_{bar} = 8000$ kg/m³
$(|H(f)| > |H(f^M)| - 3 \text{ dB})$ are extracted. The resonance quality factor Q^M is obtained as $Q^M = f^M / \Delta f^M$.

C.3 Results

Mechanical response measurements are presented below for silicone replicas and the PLT replica. Mechanical resonance features are summarized in Table C.1.



(a) Top view

(b) Side view



Figure C.1 – Mechanical frequency response set up.

first resonance peak				second resonance peak			
Replica	f_1^M [Hz]	Δf_1^M [Hz]	Q_1^M [-]	f_2^M [Hz]	Δf_2^M [Hz]	Q_{2}^{M} [-]	
PLT	60 - 110	10 ± 3	5-18	90 - 125	9±3	6-20	
M5	152	16	9.5	222	unidentifiable	unidentifiable	
MRI	143	10	14.3	219	unidentifiable	unidentifiable	
EPI	145	11	13.2	222	unidentifiable	unidentifiable	

Table C.1 – Overview mechanical resonance properties of replicas for $\alpha = 0^{\circ}$.

C.3.1 Silicone replicas

Measured data plotted in Figure C.2, are realised for the three silicone replicas, M5, MRI and EPI. Properties extracted for the two first mechanical resonances are given in Table C.1.

The coherence² associated with the FRF is above 90 % for frequencies up to 250 Hz for the three replicas. So the results are considered from 60 Hz to 250 Hz.



Figure C.2 – Example mechanical responses silcone replicas EPI, M5 and MRI.

²Ratio between cross-spectrum and auto-spectrum of signal spectral density $C = \frac{|S_{yx}|^2}{S_{xx}S_{yy}}$

C.3.2 PLT replica

Measured data are plotted in Figure C.3 for $P_{PLT} = 3800$ Pa. Experiments are performed for varying rigidity as the internal water pressure P_{PLT} is imposed in the range from 1500 Pa up to 4500 Pa. The first mechanical frequency f_1^M increases quasi-linearly with P_{PLT} as shown in Figure C.4. The quality factor given in Figure C.4(b) seems to increase with the P_{PLT} as well but not monotonously and is difficult to measure for particular P_{PLT} values. Properties extracted for the two first mechanical resonances are given in Table C.1.

The coherence associated with the FRF is above 90 % for frequencies up to 200 Hz. The results for PLT vocal folds replica are considered from 60 Hz to 200 Hz.



Figure C.3 – Example of PLT replica mechanical response for $P_{PLT} = 3800$ Pa.



Figure C.4 – PLT vocal folds mechanical response: a) oscillation amplitude as a function of internal pressure P_{PLT} and frequency f and b) quality factor as function of the inside water pressure P_{PLT} and frequency f.

APPENDIX D

Silicone replica procedure

The M5, MRI and EPI replicas are used during the experiments. These are made from multi layers mixing silicone thinner and two parts A&B Ecoflex 03-00, Smooth-On, Inc, Easton, PA or two parts A&B Dragon Skin 10 Fast FX Pro, Smooth-On, Inc, Easton, PA, illustrated in Figure D.1. The procedure follows the processes given in [5, 50, 51, 59] and is detailed in this appendix. In this appendix the muscle is referred to as the body of the VF and the superficial layer as the cover.



(a) Ecoflex part A (yellow) and B (blue) (b) Dragon Skin part A (yellow) and

B (blue)

(c) Silicone thinner

Figure D.1 – Silicone products mixed to obtain the different layers of the silicone VF replica.

D.1 M5 casting

1st step: Body

Use the mold marked M5 Body (Figure D.3(a) and D.3(c)).

1. Apply thin plastic sheet (40 mm \times 150 mm) in the mold as illustrated in Figure D.6(a) helped with HIVAC-G (Figure D.6(d)),



Figure D.2 – Example of mixing Ecoflex parts A and B and Silicone thinner (ratio \mathcal{M} ES 1:4) for the body layer of MRI replica.



Figure D.3 – Illustration (a) and (b) and picture (c) and (d) of the M5 molds.

- 2. Mix Ecoflex A, B and Silicone thinner with *M* 1:1:2, (A = 2 g, B = 2 g, Thinner = 4 g)
- 3. Extract air bubbles with the vacuum bell (Figure D.2(d)),
- 4. Fill the mold until the middle line,
- 5. Let dry at least 5 hours.

2nd step: Backing

- Mix Dragon Skin A, B and Silicone thinner with *M* 1:1:1, (A = 3 g, B = 3 g, Thinner = 3 g)
- 2. Extract air bubbles with the vacuum bell,
- 3. Fill the upper line of the mold above body part,
- 4. Let dry at least 3 hours,
- 5. Remove the molding from the mold and the plastic sheet.

3rd step: Cover

Use the mold marked M5 Cover (Figure D.3(b) and D.3(d)).

- 1. Apply thin plastic sheet (40 mm \times 150 mm) in the mold helped with HIVAC-G,
- Mix Ecoflex A, B and Silicone thinner with *M* 1:1:4, (A = 2 g, B = 2 g, Thinner = 8 g)
- 3. Extract air bubbles with the vacuum bell,
- 4. Fill the mold until the middle line,
- 5. Add the Body/Backing layers paying attention to air bubbles,
- 6. Let dry at least 15 hours,
- 7. Remove the replica from the mold and the plastic sheet.

D.2 MRI casting

1st step: Body

Use the molds marked MRI Body left and right (Figure D.4(a)).

1. Apply thin plastic sheet (18 mm × 100 mm) in the mold as illustrated in Figure D.6(b) helped with HIVAC-G (Figure D.6(d))





- 2. Mix Ecoflex A, B and Silicone thinner with *M* 1:1:4, (A = 2 g, B = 2 g, Thinner = 8 g)
- 3. Extract air bubbles with the vacuum bell,
- 4. Fill the molds until the middle line,
- 5. Let dry at least 12 hours.

2nd step: Backing

- Mix Dragon Skin A, B and Silicone thinner with *M* 1:1:1, (A = 3 g, B = 3 g, Thinner = 3 g)
- 2. Extract air bubbles with the vacuum bell,
- 3. Fill the upper line of the molds above body part,
- 4. Let dry at least 3 hours,
- 5. Remove the moldings from the molds and the plastic sheet.

3rd step: Cover

Use the molds marked MRI Cover left and right (Figure D.4(b)).

- 1. Apply thin plastic sheet (18 mm × 100 mm) in the mold as illustrated in Figure D.6(b) helped with HIVAC-G (Figure D.6(d)),
- 2. Mix Ecoflex A, B and Silicone thinner with *M* 1:1:8, (A = 1 g, B = 1 g, Thinner = 8 g)

- 3. Extract air bubbles with the vacuum bell,
- 4. Fill the molds until the middle line,
- 5. Add the Body/Backing layers paying attention to the left/right side and the air bubbles,
- 6. Let it dry during 45 minutes in the oven at $80^{\circ}C$,
- 7. Let the moldings cool down,
- 8. Remove the moldings from the molds and the plastic sheet really carefully.

4th step: Epithelium

- Mix Dragon Skin A, B and Silicone thinner with *M* 1:1:1, (A = 2 g, B = 2 g, Thinner = 2 g)
- 2. Extract air bubbles with the vacuum bell,
- 3. Cover the Body/Backing/Cover layers,
- 4. Let dry at least 3 hours,
- 5. Repeat the procedure.

D.3 EPI casting

1st step: Body

Use the molds marked EPI Body (Figure D.5(a) and D.5(e)).

- 1. Apply thin plastic sheet (17 mm × 100 mm) in the mold as illustrated in Figure D.6(c) helped with HIVAC-G (Figure D.6(d)),
- 2. Mix Ecoflex A, B and Silicone thinner with *M* 1:1:1, (A = 2 g, B = 2 g, Thinner = 2 g)
- 3. Extract air bubbles with the vacuum bell,
- 4. Fill the molds until the middle line,
- 5. Let dry at least 10 hours.



Figure D.5 – Illustration (a), (b), (c) and (d) and picture (e), (f) and (g) of the EPI molds.

2nd step: Backing

- Mix Dragon Skin A, B and Silicone thinner with *M* 1:1:1, (A = 3 g, B = 3 g, Thinner = 3 g)
- 2. Extract air bubbles with the vacuum bell,
- 3. Fill the upper line of the molds above body part,
- 4. Let dry at least 3 hours,
- 5. Remove the moldings from the molds and the plastic sheet.

3rd step: Ligament

Use the molds EPI ligament (Figure D.5(b) and D.5(f)).

- 1. Apply thin plastic sheet (17 mm × 100 mm) in the mold as illustrated in Figure D.6(c) helped with HIVAC-G (Figure D.6(d)),
- 2. Mix Ecoflex A, B and Silicone thinner with *M* 1:1:4, (A = 2 g, B = 2 g, Thinner = 8 g)
- 3. Extract air bubbles with the vacuum bell,
- 4. Fill the molds until the middle line,

- 5. Add the Body/Backing layers paying attention to the air bubbles,
- 6. Let it dry at least 12 hours,
- 7. Remove the moldings from the molds and the plastic sheet really carefully.

4th step: Cover

Use the molds marked EPI Cover (Figure D.5(c) and D.5(g)).

- 1. Apply thin plastic sheet (17 mm × 100 mm) in the mold as illustrated in Figure D.6(c) helped with HIVAC-G (Figure D.6(d)),
- 2. Mix Ecoflex A, B and Silicone thinner with *M* 1:1:8, (A = 1 g, B = 1 g, Thinner = 8 g)
- 3. Extract air bubbles with the vacuum bell,
- 4. Fill the molds until the middle line,
- 5. Add the Body/Backing/Ligament layers paying attention to the air bubbles,
- 6. Let dry it during 45 minutes in the oven at $80^{\circ}C$,
- 7. Let the moldings cool down
- 8. Remove the replica from the mold and the plastic sheet really carefully.

5th step: Epithelium

- Mix Dragon Skin A, B and Silicone thinner with *M* 1:1:1, (A = 2 g, B = 2 g, Thinner = 2 g)
- 2. Extract air bubbles with the vacuum bell,
- 3. Cover the Body/Backing/Ligament/Cover layers,
- 4. Let dry at least 3 hours,
- 5. Repeat the procedure



(d) HIVAC-G

Figure D.6 – Illustration a), b) and c) of the sheet placement for each type of mold and d) product used to fix the sheet on the mold.

D.4 Rigid support

Then each VF replica must be mounted on a rigid support as shown in Figure D.7(b).



(a) Sil-Poxy adhesive silicone

(b) Rigid supports

Figure D.7 – a) Adhesive silicone and b) rigid supports used to fix silicone VF (M5 and EPI on the left and MRI on the right).

- 1. Cut the unnecessary parts of the mold,
- 2. Use Sil-Poxy adhesive silicone (Figure D.7(a)) to fix the mold in the support (Figure D.7(b)),
- 3. Ajust the position to be symmetrical,
- 4. Let dry the glue for 5 minutes,

- 5. Add Sil-Poxy to fix the boundaries,
- 6. Let dry the glue for 5 minutes,
- 7. Put talcum powder.

Young modulus estimation

The modulus of elasticity or Young's modulus \mathscr{E} , expressed in Pa, is an intrinsic property of an elastic solid material. It relates the relative linear deformation ε to the charge or tensile stress \mathscr{C} required to deform the material with a cross section \mathscr{S} as:

$$\mathscr{E} = \frac{\mathscr{C}}{\varepsilon \cdot \mathscr{S}},\tag{E.1}$$

with deformation $\varepsilon = \frac{\Delta l - l_0}{l_0}$, l_0 corresponds to the initial length of the material and Δl its elongation for the corresponding charge.

E.1 Tensile test

The tensile test (Figure E.1) consists in measuring the deformation Δl in *mm* as a function of the charge \mathscr{C} in *N* applied on a specimen of the material. Young's modulus can be calculated using Eq. (E.1). For this purpose, the INSTRON 3369 electro-mechanical press was used with an INSTRON 2530-445/71212 50kN charge sensor (accuracy 0.16% of the reading) in the ENSE3 school of Grenoble.

Tensile test is applied to silicone VF layer specimens as outlined in Section E.2 and to latex tube samples (VF and central tube) as detailed in Section E.3.

E.2 Results: silicone layers

Silicone VF replicas (M5, MRI and EPI) are made up from 2 to 4 different layers of silicone mixtures (Ecoflex and Silicone thinner (ES) or Dragonskin and Silicone thinner (DS)), illustrated in Figure E.2. The manufacturing procedure is detailed in Appendix D and mixing ratios for the different layers used in the silicone replicas are given in Table E.1. The modulus of elasticity is measured for the mixing ratio of each layer using a tensile test outlined in Section E.1. Moduli have already been reported in literature [50, 51, 57–59]. Values from literature and estimated in



Figure E.1 – Illustration of tensile test on DS 1:1 specimen (the bottom bite is fixed and the top bite moves in the upward direction).

the current study are summarised in Table E.2 and match well for all assessed mixing ratios and hence for all layers.



Figure E.2 – Silicone VF replicas layers in the medio-frontal plane.

E.2.1 Measurement protocol

For each ratio (except $\mathcal{M}=1:8$), 2 specimens are tested. The specimens design is shown in Figure E.3. Each specimen has initial length $l_0 = 80$ mm, initial cross section $\mathcal{S}_0 = 150 \text{ mm}^2$ and total volume 23715 mm³. They were obtained using a 3D printed (ABS-P430 printer) mold

Ratio M	M5	MRI	EPI
DS 1:1	Backing	Backing Epithelium	Backing Epithelium
ES 1:1	_	—	Body
ES 1:2	Body	_	_
ES 1:4	Cover	Body	Ligament
ES 1:8	_	Cover	Cover

Table E.1 – Different mixing ratio used for the layers of each replica.

Table E.2 – Summary of moduli of elasticity \mathscr{E} [kPa] found in literature and estimated in Appendix E.

	DS 1:1	ES 1:1	ES 1:2	ES 1:4	ES 1:8	DS 1:0	ES 1:0
Drechsel et al. (2008) [57]	22	-	4.1	-	-	_	_
Riede et al. (2008) [58]	_	_	~ 10	_	_	_	_
Pickup et al. (2010) [50]	_	_	8.5	2.9	_	_	_
Murray et al. (2010) [59]	~ 21	~ 2.3	~ 1.8	~ 0.3	~ 0.08	_	_
Murray et al. (2011) [51]	50	12	_	1.6	0.2	_	_
Smooth-on [175]	_	_	—	—	—	151.7	68.9
Current study	52.0	21.9	10.4	4.9	0.2*	151.7 [†]	68.9^{\dagger}

*Value for ES 1:8 is taken from [59] as of 0.2 kPa

[†]Value for DS 1:0 and ES 1:0 is taken from the manufacturer Smooth-on: Dragon skin 10 fast and <u>Ecoflex 00-30</u>

(Figure E.4(a)). An example of a specimen is shown in Figure E.4(b). For each specimen, measurements are performed from $\Delta l = 0$ mm to different maximum elongation values:

- 1. $\Delta l = 25$ mm at a deformation rate of 1 mm/s,
- 2. $\Delta l = 50 \text{ mm at } 1 \text{ mm/s},$
- 3. $\Delta l = 100 \text{ mm}$ at 1 mm/s,
- 4. $\Delta l = 150 \text{ mm}$ at 2 mm/s.

The charge \mathscr{C} (up to 8 N) and the longitudinal deformation Δl are measured with a sampling frequency of 10 Hz. For each maximum elongation, the cross section \mathscr{S} at the center of the specimen is measured. Note that the deformation yields $\varepsilon \in [0 2]$.



Figure E.3 – Specimen design used for tensile test.



(a) 3D printed specimen mold

(b) Picture of the DS 1:1 specimen.

Figure E.4 – Illustration of a) mold and b) specimen.

E.2.2 Analysis

First, due to the small charges (up to 8 N), with respect to the range of the force sensor (50 kN), the measured charge signal is very noisy. A smoothing is therefore applied with a sliding average over 10 (MA 10) and 20 (MA 20) measuring data points (Figure E.5). This corresponds to an average of 1 s and 2 s of signal respectively, which implies a variation on 1 mm and 2 mm for elongation up to 25 mm, 50 mm and 100 mm and a variation of 2 mm and 4 mm for the elongation up to 150 mm.

Then, the central section \mathscr{S} of the specimen was measured for 5 elongation values: 0 mm, 25 mm, 50 mm, 100 mm and 150 mm (accuracy $\pm 0.25 \text{ mm}^2$). From these measurements, a quadratic regression was performed to determine the evolution of the section at each measurement point shown in Figure E.6. It is seen that the overall section reduction is about 38 % of \mathscr{S}_0 . The modulus of elasticity (Eq. (E.1)) depends on section \mathscr{S} , so that constraints \mathscr{C}/\mathscr{S} and $\mathscr{C}/\mathscr{S}_0$ (with \mathscr{S} the variable section depending on the deformation shown in Figure E.6 and \mathscr{S}_0 the initial section (150 mm²)) is shown in Figure E.7. It is observed that the curve plotted with a variable section is more linear than with the initial section. The analyses are therefore based on a constraint calculated according to the variable section with the deformation \mathscr{C}/\mathscr{S} .



Figure E.5 – Effect of measured data smoothing using a moving average (MA) of length 10 (MA 10) and 20 (MA 20) data points on $\mathscr{C}(\varepsilon)$.

The maximum deformation is obtained for $\Delta l = 150$ mm, which results in $\varepsilon = 1,88$. On the curves (Figure E.8(a)) DS $\mathcal{M} = 1 : 1$, ES $\mathcal{M} = 1 : 1$ and ES $\mathcal{M} = 1 : 2$ and their derivatives (Figure E.9), a first linear part is observable for $\varepsilon \in [0 \ 0.5]$ and a non-linear part beyond. Finally



Figure E.6 – Measured section areas $\mathscr{S}(\text{symbols})$ as a function of the deformation Δl and their quadratic fitting (full lines).

for $\varepsilon \ge 1.24$, a decrease of the derivative curve due to a sliding of the specimen in the tensile bites (≤ 2 mm) is observable and data in this range cannot be taken into account. The analyses are therefore made on the linear zone of the deformation: $\varepsilon \in [0 \ 0.5]$. For the curve ES $\mathcal{M} = 1:4$ (Figure E.8(a)), the linear part is extended to the maximum deformation measured ($\varepsilon = 1.875$). The analyses are then extended to $\varepsilon \in [0 \ 1.5]$. Charging and decharging gives the same result as illutrated in Figure E.8(b).

Then, a linear fitting (Figure E.10) of the slope $\frac{\mathscr{C}}{\mathscr{S}}(\varepsilon)$ is calculated for each measurement, without and with smoothing (MA 10 and MA 20), in order to obtain the slope coefficient, corresponding to the modulus of elasticity \mathscr{E} . The fit accuracy (coefficient of determination \mathbb{R}^2) is calculated. The average of all maximum elongation values for each mixing is shown in Table E.3. Except for the mixing ES $\mathscr{M} = 1 : 4$ for which the average is realised for the two last largest elongation values ($\Delta l = [0 - 100]$ mm and $\Delta l = [0 - 150]$ mm). The best accuracy ($\mathbb{R}^2 > 90 \%$) is obtained for maximum smoothing (MA 20), so that these values are reported in Table E.2.

For completeness, the density is estimated (Table E.3) as the mass (accuracy 0.01 g) of the specimens divided by their volume (23715 mm^3).

E.3 Results: latex tubes

The latex is a natural rubber and tubes are used in the PLT replica for each VF and for the central tube. In the current study, four tubes are analysed with two thicknesses (0.2 mm and

0.3 mm) and two diameters (\emptyset =11 mm used for the VF and \emptyset =28.5 mm for the central tube). Latex is provided by Piercan company.



Figure E.7 – Comparison of the constrains calculated with the varying section \mathscr{S} and with the section $\mathscr{S}_0=150 \text{ mm}^2$.



Figure E.8 – a) Smoothed $\frac{\mathscr{C}}{\mathscr{F}}(\varepsilon)$ for all mixing ratios. b) Example of charging and decharging for DS $\mathscr{M} = 1:1$ specimen.

Table E.3 – Measured modulus of elasticity & a	and their accuracy R ²	of each silicone lay	yer.
--	-----------------------------------	----------------------	------

Data		Moving average		Moving average		Density	Density	
ratio <i>M</i>			10 data points		20 data points		measured	calculated $^{\vee}$
	€ [kPa]	$R^2 [\%]$	ℰ [kPa]	R^2 [%]	€ [kPa]	R^2 [%]	$[kg.m^{-3}]$	$[kg.m^{-3}]$
DS 1:1	51.6	96.2	51.6	99.5	52.0	99.6	1021	1040.0
ES 1:1	21.8	82.6	21.7	97.3	21.9	98.0	991.4	1034.8
ES 1:2	10.3	50.2	10.2	88.6	10.4	92.8	990.9	1020.5
ES 1:4	4.90	47.6	4.89	89.3	4.90	93.2	988.0	1006.7
ES 1:8	_	—	_	—	—	-	—	995.8

^{\diamond}Values ρ_{eq} calculated for $\mathcal{M} = r_E : r_S$ from densities $\rho_{E,S}$ given by the manufacturer [175] as

$$\frac{1}{\rho_{eq}} = \frac{\frac{2 \cdot r_E}{\rho_E} + \frac{r_S}{\rho_S}}{2 \cdot r_E + r_S}$$

with $\rho_{Dragonskin} = 1072.9 \text{ kg.m}^{-3}$, $\rho_{Ecoflex} = 1064.6 \text{ kg.m}^{-3}$ and $\rho_{Silicone} \in [940 - 1000] \text{ kg.m}^{-3}$, measured $\rho_{Silicone} = 980 \text{ kg.m}^{-3}$.



Figure E.9 – Derivatives curves of maximum deformation measurements (Figure E.8(a)).

E.3.1 Measurement protocol

Tubes with a diameter of 11 mm are simply fixed, bent in order to have two layers of latex as shown in Figure E.11(a). The tubes with a diameter of 28.5 mm are fixed bent so as to have four layers to be able to be held entirely in the tensile bites, illustrated in E.11(b).

Each tube has initial length $l_0 = 80$ mm and an initial section given in Table E.4. For each tube, measurements consist of a charge followed by a decharge from $\Delta l = 0$ mm to different maximum elongation values:

- 1. $\Delta l = 25$ mm at a deformation rate of 1 mm/s,
- 2. $\Delta l = 50 \text{ mm at } 1 \text{ mm/s},$
- 3. $\Delta l = 75 \text{ mm at } 1 \text{ mm/s},$

The charge \mathscr{C} (up to 15 N) and the longitudinal deformation Δl are measured with a sampling frequency of 10 Hz.



Figure E.10 – Linear fitting of silicone specimens.

E.3.2 Analysis

As seen in Section E.2, the raw signal is noisy so that smoothing is applied with again a sliding average over 10 (MA 10) and 20 (MA 20) data points. This corresponds to an average of 1 s and 2 s of signal respectively, which implies a variation of 1 mm and 2 mm. Due to the small initial section (< 30 mm²) and thickness (≤ 0.3 mm) the variation of the section with the applied charge could not be measured. So that the section \mathscr{S} is taken constant to the initial section \mathscr{S}_0 ,



Figure E.11 – Tensile test configuration for a) latex tube with diameter 11 mm and b) latex tube with diameter 28.5 mm.

i.e. $\mathscr{S} = \mathscr{S}_0$.

A relaxation deformation is observed in Figure E.12 and values are given in Table E.4. Relaxation consists of the delayed (non-instantaneous) irreversible deformation of the material under stress, below the yield strength of the material. Due to this phenomenon and the unknown evolution of the section, the linear part of the curve $\frac{\mathscr{C}}{\mathscr{T}_{\ell}}(\varepsilon)$ is reduced to $\varepsilon \in [0 - 0.2]$ as illustrated in Figure E.13.

Thickness	Ø	Section	Relaxation deformation				
[mm]	[mm]	$[mm^2]$	25 mm [%]	50 mm [%]	75 mm [%]		
0.2	11	10.4	2.0	4.0	5.7		
0.3	11	6.91	2.5	3.4	4.0		
0.2	28.5	17.9	1.5	2.0	3.4		
0.3	28.5	26.9	1.2	2.0	3.0		

Table E.4 – Initial section and relaxation deformation values for each tube.

As in Section E.2 the linear fitting of the slope $\frac{\mathscr{C}}{\mathscr{F}}(\varepsilon)$ is calculated to obtain the modulus of elasticity \mathscr{E} , with its fit accuracy (coefficient of determination R²). The best accuracy (R² > 97 %) is obtained for maximum smoothing (MA 20), so that these values are reported in Table E.5, with the estimated density.

Table E.5 – Measured modulus of elasticity and their accuracy of each latex tube.

Tube		Data		Moving average		Moving average		Density	
Thickness	Ø			10 data points		20 data points		Density	
[mm]	[mm]	€ [kPa]	$R^2 [\%]$	€ [kPa]	R^2 [%]	€ [kPa]	R^2 [%]	[kg.m ⁻³]	
0.2	11	1067.8	88.7	1065.9	98.2	1091.1	97.9	1635	
0.3	11	1157.4	95.0	1152.2	99.1	1159.2	99.3	1244	
0.2	28.5	1510.4	98.7	1507.7	99.4	1514.2	99.3	1324	
0.3	28.5	1121.7	98.7	1119.6	99.4	1130.8	99.1	930.7	



Figure E.12 – Charge and decharge of each measurement with observation of the relaxation deformation.



Figure E.13 – Linear fitting of latex tubes.

APPENDIX F

Vertical angular asymmetry: leakage area and model

F.1 Calculation of the vertical leakage area

In Section 4.2 the VF level difference asymmetry creates a triangle area leakage *A*. Its dimensions can be calculated from the known dimensions *w*, E, e and *a* as a function of the angular level difference α as illustrated in Figure F.1. To simplify the calculation, it was divided into four steps: 1) Equation (F.1), 2) Equation (F.2), 3) Equation (F.3) and 4) Equation (F.4) = Eq. (F.1) - Eq. (F.2) + Eq. (F.3).

$$A_b = \frac{1}{2 \cdot \tan(\alpha)} \cdot \left(\tan(\alpha) \cdot \left(w + a + \frac{1}{\sin(\alpha)} \cdot \left(\frac{E}{2} - e \right) \right) - \frac{E}{2} \right)^2$$
(F.1)

$$A_{j} = \frac{1}{2} \cdot \tan(\alpha) \cdot \left(\frac{E}{2} - e + \tan\left(\frac{\alpha}{2}\right) \cdot (w+a)\right)^{2}$$
(F.2)

$$A_{\nu} = \frac{1}{2} \cdot \tan(\alpha) \cdot \left(\frac{3 \cdot E}{2} - 2 \cdot e - \tan\left(\frac{\alpha}{2}\right) \cdot (w+a)\right)^2$$
(F.3)

$$A(\alpha) = \begin{cases} 0 & \text{if} \quad \tan(\alpha) \left(w + a + \frac{E/2 - e}{\sin(\alpha)} \right) < \frac{E}{2}, \\ \frac{1}{2 \cdot \tan(\alpha)} \cdot \left(\left(w + a + \frac{1}{\sin(\alpha)} \cdot \left(\frac{E}{2} - e \right) \right) \cdot \tan(\alpha) - \frac{E}{2} \right)^2 \\ + \frac{1}{2} \cdot \tan(\alpha) \cdot (E - e) \cdot \left(e - 2 \cdot \tan\left(\frac{\alpha}{2} \right) \cdot (w + a) \right) & \text{otherwise,} \end{cases}$$
(F.4)



(a) Areas A and known parameters shown in Section 4.2.



Figure F.1 – Illustration of the different area use to calculate A with $A = A_b - A_j + A_v$.

F.2 Simplified fluid-structure interaction model

Next, the simplified fluid-structure interaction model introduced in [5] is based on the delayed one mass model or mucosal wave model proposed in [101]. The resulting equation (F.14) for the onset pressure as a function of the level difference motivates the fitting in Section 4.4.

In this section Δd indicates the vertical level difference for parallel vocal folds with thickness d, as shown in Figure F.2.

 $\begin{cases} \text{Right vocal fold} \\ -\frac{d+\Delta d}{2} \leqslant z \leqslant \frac{d-\Delta d}{2} \end{cases}$ Left vocal fold $-\frac{d-\Delta d}{2} \leqslant z \leqslant \frac{d+\Delta d}{2} \end{cases}$



Figure F.2 – Illustration vertical difference for parallel vocal folds based on [5].

For the glottis:

$$-\frac{d+\Delta d}{2} \leqslant z \leqslant \frac{d+\Delta d}{2} \tag{F.5}$$

and

$$A_g(z,t) = w[2x_0 + x_R(z,t) + x_L(x,t)]$$
(F.6)

with A_g the glottal area, w the vocal fold width, x_0 the prephonatory glottal half-width and x_R and x_L the time dependent tissue displacement of the left and right vocal folds.

$$\tau = \frac{d - \Delta d}{2} \cdot \frac{1}{v_{mw}} \tag{F.7}$$

 τ is the time delay within which the mucosal surface wave travels with velocity v_{mw} .

Assumption 1: τ is small.

$$x_{RL}(t+\tau) = x_{R,L}(t) \pm \tau x_{RL}(t)$$
(F.8)

so the glottal area become:

$$A_{1,2} = w[2x_0 + x_R(t) + x_L(t) \pm \tau \{\dot{x_R}(t) + \dot{x_L}(t)\}]$$

with A_1 the glottal entry area and A_2 the glottal exist area.

Assumption 2: Only Bernoulli principle is taken into account (neglect *e.g.* acoustic loads).

$$\begin{cases} \phi \approx A_s \sqrt{\frac{2\Delta P_{tot}}{\rho}} \\ \frac{1}{2}\rho \phi^2 + P = Constant \end{cases}$$

with $A_s = A_2$ and $\Delta P_{tot} = P_s$. So:

$$\Rightarrow \frac{1}{2}\rho \frac{\phi^{2}}{A_{2}^{2}} + 0 = \frac{1}{2}\rho \frac{\phi^{2}}{A_{z}^{2}} + P(z)$$
$$\Rightarrow P(z) = \frac{1}{2}\rho \phi^{2} \left(\frac{1}{A_{2}^{2}} - \frac{1}{A_{z}^{2}}\right) = P_{s} \left[1 - \left(\frac{A_{2}}{A_{z}}\right)^{2}\right]$$

$$P(z) = \begin{cases} P_s & z \in \left[-\frac{1}{2}(d + \Delta d), -\frac{1}{2}(d - \Delta d)\right] \\ P_s \left(1 - \left(\frac{A_2}{A_z}\right)^2\right) & z \in \left[-\frac{1}{2}(d - \Delta d), \frac{1}{2}(d - \Delta d)\right] \\ 0 & z \in \left[\frac{1}{2}(d - \Delta d), \frac{1}{2}(d + \Delta d)\right] \end{cases}$$
(F.9)

Average of intra-glottal pressure at the left vocal fold:

$$P_g = \frac{1}{d} \int_{-\frac{1}{2}(d-\Delta d)}^{\frac{1}{2}(d+\Delta d)} P(z) dz$$
(F.10)

Assumption 3: Glottis surface is linear.

$$\frac{\delta A}{\delta z} = \frac{A_2 - A_1}{d - \Delta d} \tag{F.11}$$

In the glottis:

$$P_g = \frac{1}{d} \int P(z) dz$$

$$P_{g} = \frac{1}{d} \int_{A_{1}}^{A_{2}} P_{s} \left(1 - \left(\frac{A_{2}}{A(z)} \right)^{2} \right) \frac{d - \Delta d}{A_{2} - A_{1}} dA$$

$$P_{g} = \frac{1}{d} P_{s}(d - \Delta d) \left(1 - \frac{A_{2} - A_{1}}{A_{1}} \right)$$
(F.12)

Here:

$$A_{1} = w [2x_{0} + x_{R}(t) + x_{L}(t) + \tau (\dot{x}_{R}(t) + \dot{x}_{L}(t))]$$

$$A_{2} = w [2x_{0} + x_{R}(t) + x_{L}(t) - \tau (\dot{x}_{R}(t) + \dot{x}_{L}(t))]$$

$$\tau = \frac{d - \Delta d}{2v_{mw}}$$

From Equation F.12:

$$P_g = \frac{d - \Delta d}{d} P_s \frac{A_1 - A_2}{A_1}$$

$$P_g = \frac{d - \Delta d^2}{2} P_s \frac{1}{v_{mw}} \frac{\dot{x_R} + \dot{x_L}}{2x_0 + x_R + x_L + \frac{d - \Delta d}{2v_{mw}}} (\dot{x_R} + \dot{x_L})$$

Assumption 4: Left-right vocal folds symmetry $x_L(z,t) = x_R(z,t)$

Assumption 5: Approximation of P_g by its linearisation around $x_{L,R} = 0$ and $x_{L,R} = 0$

$$P_g = \frac{d + \Delta d}{d} P_s \frac{2\tau(\dot{x_R} + \dot{x_L})}{2x_0 + x_R + x_L}$$

So using the mucosal wave model

$$M\ddot{x_R} + B\dot{x_R} + Kx_R = P_g \tag{F.13}$$

with *M* the mass, *B* the damping, *K* the stiffness per unit area of the vocal fold medial surface.

By applying assumption 4 and 5 this becomes:

$$M\ddot{x_R} + B\dot{x_R} + Kx_R = \frac{d - \Delta d}{d} P_s \left[\tau \frac{\dot{x_R} + \dot{x_L}}{x_0} \right].$$

$$\Rightarrow M\ddot{x}_{R} + B\dot{x}_{R} + Kx_{R} = \frac{(d - \Delta d)^{2}}{dx_{0}} P_{s} \frac{2}{2v_{mw}} \dot{x}_{R}$$
$$\Rightarrow M\ddot{x}_{R} + \left\{ B - \frac{(d - \Delta d)^{2}}{d} \frac{P_{s}}{v_{mw}x_{0}} \right\} \dot{x}_{R} + Kx_{R} = 0$$

Oscillation growths when damping coefficient is negative: $B - \frac{(d - \Delta d)^2}{d} \frac{P_s}{v_{mw}x_0} \leq 0$ Oscillation threshold pressure is given as:

$$P_s \geqslant \frac{d}{(d - \Delta d)^2} B v_{mw} x_0 \tag{F.14}$$

- 1. P_{thres} is proportional to phonatory glottal length x_0 , and it is inverse proportional to the vocal fold thickness d,
- 2. P_{thres} increases with imposed vertical level difference as $P_s \approx \frac{1}{(d \Delta d)^2}$.

So when the level difference is zero: $\Delta d = 0$, $P_s \ge Bv_{mw}x_0/d$. When the level difference is positive: $\Delta d > 0$: $P_s \ge \frac{d}{(d-\Delta d)^2}Bv_{mw}x_0$.

Given

$$P_s = \alpha \frac{\tilde{d}}{(\tilde{d} - \Delta d)^2}$$

with fitting parameter $\{\alpha, \tilde{d}\}$.

Artificial saliva fluid properties

During the measurements of the influence of liquids on VF, two artificial saliva were used (Artisial and Teijin). In order to be able to compare their properties with human saliva and distilled water, their density ρ and kinematic viscosity v were measured. The dynamic viscosity μ follows from the relationship $v = \frac{\mu}{\rho}$. An overview of these fluids properties for human saliva, mucus, air and water and values determined for the two artificial saliva in this study is provided in Table G.1.

	Density ρ [kg.m ⁻³]	Kinematic viscosity v [m ² .s ⁻¹]	Dynamic viscosity μ [kg.m ⁻¹ .s ⁻¹])
Human saliva [45]	~ 1000	\sim [2.75 - 15.51] $ imes 10^{-6}$	$[2.75 - 15.51] \times 10^{-3}$
Mucus [38]	~ 1000	\sim [1 - 10] $ imes$ 10 ⁻⁶	$[1 - 10] \times 10^{-3}$
Artisial saliva	1015	7.328×10^{-6}	7.44×10^{-3}
Teijin saliva	1014	4.884×10^{-6}	4.95×10^{-3}
Water	1000	1.00×10^{-6}	1.00×10^{-3}
Air	1.20	1.57×10^{-5}	1.85×10^{-5}

Table G.1 – Summarised values of density and viscosity.

In addition, the wettability of these fluids for surfaces of the vocal folds replicas is determined in Section G.4.

G.1 Artificial saliva

The properties of two artificial salivae are determined.

G.1.1 Artisial

This solution (Figure G.1), which can be found in French pharmacies, is used in the event of a decrease or absence of salivary secretions.
Active ingredients: Potassium chloride, Sodium chloride, Magnesium chloride, Calcium chloride, Dipotassium phosphate, Monopotassium phosphate.

Excipients: Sorbitol, Methyl parabenzoate, Carboxymethylcellulose sodium, Purified water, Sterile nitrogen.



Figure G.1 – Artisial presentation and composition.

G.1.2 Teijin Saliva

This solution (Figure G.2), which can be found in Japanese pharmacies, is also used in the event of a decrease or absence of salivary secretions.

Active ingredients: Sodium chloride, Potassium chloride, Hydrated calcium chloride, Magnesium chloride, Dipotassium phosphate.

Excipients: Carmelose sodium, D-sorbitol, sodium benzoate, sorbic acid, sodium hydroxide, carbon dioxide.

	<u>Composition</u>			
世でで	Active ingredient (per 50 g)	Sodium chloride	42.2 mg	
		Potassium chloride	60.0 mg	
上部の青印と噴射成合わせてご使用下む		Calcium chloride hydrated	7.3 mg	
		Magnesium chloride	2.6 mg	
_人 工唾液 サリベート		Dipotassium phosphate	17.1 mg	
	Additive	Sodium Carmelose		
ΤΕΙͿΙΝ		D-sorbitol		
50g		sodium benzoate		
		sorbic acid		
NERSER 教Jアーマ株式会社		sodium hydroxide		
^{東京都千代} 田区置が開3-2-1		carbon dioxide (propellant))	
(a) Product		(b) Composition		

Figure G.2 – Teijin presentation.

G.2 Results: density

Density measurement were realised collecting 10 ml of liquid with a micro-pipette (accuracy \pm 0.1 mL) and weighed with a precision balance (accuracy \pm 0.01 g). The process is repeated ten times to obtain the average density and its standard deviation (<1% of the averaged value). The results are obtained for 22°C:

- Artisial: $\rho = 1015 \pm 6.16 \text{ kg.m}^{-3}$
- Teijin: $\rho = 1014 \pm 3.63 \text{ kg.m}^{-3}$

The average density for both artificial salivae is reported in Table G.1.

G.3 Results: kinematic viscosity

Viscosity measurements were realised using a Cannon-Fenske viscometer using the principal of capillarity. The measurement consists of measuring the filling times t_1 and t_2 of the liquid studied in both measuring spheres (S₁ and S₂), illustrated in Figure G.3 and multiply each by the constant k_1 for time t_1 and the constant k_2 for time t_2 . Setting $v_1 = t_1 \times k_1$ and $v_2 = t_2 \times k_2$ gives

two estimates of kinematic viscosity v of the liquid.

In this case, two viscometers with different capillary size, detailed in Table G.2, were used. Results for the two artificial saliva are respectively given in Table G.3 and G.4: averaged values \overline{v} and standard deviation $\sigma(v)$ (<2% of the averaged value). Averaged values are reported in Table G.1. The reported dynamic viscosity is obtained as: $\mu = \overline{v} \times \rho$.



Figure G.3 – Cannon-Fenske viscometer with two measurement spheres S_1 and S_2 .

N°	Capillary ref.	Ø Capillary [mm]	k_1 [m ² .s ⁻²]	k_2 [m ² .s ⁻²]
1	75	0.54	8.419×10^{-9}	6.500×10^{-9}
2	100	0.63	1.606×10^{-8}	1.201×10^{-8}

Table G.2 – Size and constants of capillaries used.

Table G.3 – Results of viscosity measurements realized for artificial saliva Artisial.

N°	$\begin{vmatrix} t_1 \\ [s] \end{vmatrix}$	t ₂ [s]	$\begin{vmatrix} v_1 \\ [m^2.s^{-1}] \end{vmatrix}$	v_2 [m ² .s ⁻¹]	$\begin{bmatrix} \overline{\nu} \\ [m^2.s^{-1}] \end{bmatrix}$	$\frac{\sigma(v)}{[m^2.s^{-1}]}$
1 2	879 457	1139 597	$\begin{array}{ c c c c c } 7.400 \times 10^{-6} \\ 7.339 \times 10^{-6} \end{array}$	$\begin{array}{ c c c c c c c c c c c c c c c c c c c$	7.328 ×10 ⁻⁶	1.100×10^{-7}

N°	<i>t</i> ₁ [s]	t ₂ [s]	v_1 [m ² .s ⁻¹]	v_2 [m ² .s ⁻¹]	$[m^2.s^{-1}]$	$\sigma(v)$ [m ² .s ⁻¹]
1 2	581 309	734 409	$\begin{array}{ } 4.891 \times 10^{-6} \\ 4.963 \times 10^{-6} \end{array}$	$\begin{array}{c c} 4.771 \times 10^{-6} \\ 4.912 \times 10^{-6} \end{array}$	4.884 ×10 ⁻⁶	8.119×10^{-8}

Table G.4 – Results of viscosity measurements realized for artificial saliva Teijin.

G.4 Results: wettability of different replicas surfaces

Wettability is the degree of spreading of a liquid phase on a solid surface. It is quantified by the contact angle observable between the surface and the inside droplet as shown in Figure G.4. From 0° to 180° , angles are divided into five different ranges which define the degree of wettability from strongly non-water wet (oil wet) to strongly water wet [176]. The wettability is determined for the three types of vocal folds replicas, for water and for the two artificial salivas as presented in Table G.5.

Table G.5 – Results of wettability for used liquids.

	MDR replica	PLT replica	Silicone replicas
Distilled water	Neutral wet	Neutral wet	Moderately non-water wet
Artisial saliva	Neutral wet	Neutral wet	Moderately non-water wet
Teijin saliva	Neutral wet	Neutral wet	Moderately non-water wet



(c) Teijin saliva

Figure G.4 – Illustration of the wettability of the three types of vocal folds replicas for the different liquids used. Contact angles are indicated.

Appendix H

Additional preliminary results for hydration

Immediate perspectives of experimental results presented in Chapter 5 are:

- influence of the mean upstream pressure $\overline{P_u}$ for water spraying,
- influence of liquid fluid viscosity using Artisial artificial saliva in comparison with water,

on quantified features.

Preliminary results for these two perspectives are presented hereafter using the PLT replica for two configurations:

- $P_{PLT} = 1900 \text{ Pa}, o_{mm} = 1 \text{ mm},$
- $P_{PLT} = 2800 \text{ Pa}, o_{mm} = 1 \text{ mm}.$

H.1 Influence of mean upstream pressure

The influence of the mean upstream pressure $\overline{P_u}$ is measured starting from the onset pressure (≈ 140 Pa) in absence on water ($V_L = 0$ mL) up to 300 Pa. The pressure increment is 25 Pa. Quantified features obtained without water are compared with those for $V_L = 4$ mL of distilled water. As in Chapter 5, 5 s portions are analysed. In the case that water is sprayed, the analysed pressure time trace is again taken 5 s after spraying was finished where the effect of water praying is stabled.

Results are shown in Figure H.1.



Figure H.1 – Influence of upstream pressure for two configurations of the PLT replica: (1900,1) and (2800,1).

H.2 Influence of liquid viscosity

The influence of the liquid viscosity is assessed by comparing quantified features with different liquid spray fluids: distilled water and Artisial artificial saliva. Fluid properties are presented in Appendix G. It is seen that the viscosity is increased with a factor 7 for Artisial.





Figure H.2 – Influence of liquid spraying, distilled water and Artisial, for two configuration of the PLT replica: (1900,1) and (2800,1).

Résumé — La production de la voix humaine est générée par l'auto-oscillation des cordes vocales , due à l'interaction entre le flux d'air venant des poumons et la structure élastique des cordes vocales. Le but de cette thèse est de réaliser une étude expérimentale et théorique permettant de mieux comprendre et de modéliser ce phénomène et certaines de ses perturbations. Premièrement, l'algorithme du MSePGG est proposé pour la calibration d'un dispositif non-invasif de mesure in vivo de l'air glottique. Cet algorithme est validé sur des répliques de cordes vocales et illustré pour des mesures sur un locuteur. Deuxièmement, les cordes vocales sont recouvertes par une fine couche de liquide essentielle à la phonation. Une approche expérimentale est proposée afin d'étudier l'influence de la présence de ce liquide sur des répliques de cordes vocales. Elle démontre que la pulvérisation d'eau a un impact sur les paramètres basiques de la voix et sur leurs perturbations. Un modèle théorique simplifié tenant compte de la présence de l'air et de l'eau est ensuite proposé et validé. Troisièmement, l'effet de l'asymétrie angulaire verticale des cordes vocales, dans le cas d'une paralysie unilatérale, sur l'interaction fluide-structure est évalué expérimentalement. Il est observé que la perte initiale de contact des cordes vocales entraîne une variation importante des caractéristiques de phonation et de leurs variations. Un modèle théorique simple est adapté afin de prédire l'augmentation de la pression de seuil de l'oscillation des cordes vocales. Pour des applications cliniques futures, les résultats obtenus suggèrent la poursuite du développement du système MSePGG et illustrent les multiples causes potentielles de perturbation de la voix.

Mots clés : Cordes vocales, Phonation, Interaction fluide-structure, Pathologies vocales, Mesures expérimentales *in-vitro*, Modélisation théorique.

Abstract — The production of the human voice is generated by vocal folds auto-oscillation , due to the interaction between the air flow coming from the lungs and the elastic structure of the vocal folds. The purpose of this thesis is to realise an experimental and theoretical study in order to improve the understanding and modelling of this phenomenon and some of its perturbations. Firstly, the MSePGG algorithm is proposed for the calibration of a non-invasive device for in vivo glottal area measurements. The algorithm is validated on mechanical replicas and illustrated for measurements on a human speaker. Secondly, the vocal folds are covered by a thin layer of liquid, essential for phonation. An experimental approach is proposed to systematically study the influence of the presence of liquid on vocal fold replicas . Water spraying is shown to alter basic voice parameters as well as their perturbation. A simplified theoretical flow model accounting for the presence of both air and water is proposed and validated. Thirdly, the effect of vertical vocal fold angular asymmetry, as occurring in the case of unilateral vocal fold paralysis, on the fluid-structure interaction is experimentally assessed. It is found that loss of full vocal folds contact leads to important variation on phonation features and their variations. A simple theoretical model is shown to fit the increase of auto-oscillation onset threshold pressure with asymmetry angle. For future clinical applications obtained results suggest the further development of the MSePGG device and illustrate the multiple potential causes of voice perturbation.

Keywords: Vocal folds, Phonation, Fluid-structure interaction, Vocal pathologies, Experimental *in-vitro* measurements, Theoretical modelling.