INSTITUTE OF THERMOMECHANICS
CZECH ACADEMY OF SCIENCES
Dolejškova 5, 182 00 Praha 8, Czech Republic

Review of the Doctoral Dissertation:

Biomechanics and acoustics of voice production

Author: Hugo Lehoux, MSc

The thesis is composed from the main text of the dissertation (52 pages) supplemented by copies of four publications of the author published in scientific impacted journals. He is noted as the first author in three publications. The four articles occupy in total 57 pages of the thesis.

The topic of the thesis belongs to a multidisciplinary scientific area on the border between medical sciences (physiological view on human voice production), art sciences (singing techniques) and physics (voice acoustic analysis and measurement techniques, especially those for investigation of vocal folds vibration).

The first part of the thesis, written on ca 30 pages, is an excellent overview of the state of art in vocology from different points of view starting with basic principles of voice production, physiology and anatomy of the phonatory system, measurement techniques used by phoniatricians in clinical investigations as well as in voice research, the bases of the source-filter theory of phonation by Fant and the myoelastic theory of voice production according to Titze. The thesis is focused on analysis and understanding of the sound source in the laryngeal part of the vocal tract, i.e. mainly on varies adjustments of the vocal folds and analysis of their vibration by ~~the~~ laryngeal video-endoscopy using a high speed camera and electroglottography method (EGG). Logically, the introduction to the state of the art concentrates predominately on the topics that were studied in separate papers included in the second part of the thesis. Therefore, the author overviews in more detail studies related to the observed frequency jumps during vocal folds' vibration, to the video-kymography used for analysis of mucosal waves detected on vibrating vocal folds, the voice registers, the effects of the subglottic resonances, and arrangements of excised larynx experiments as well as synthesis of kymograms using computational aeroelastic or finite element (FE) model of vocal folds vibration generating mucosal waves.

The main part of the author's original work, published in the two impacted articles (I-II), consists of two experiments performed on excised red deer and human larynges, applying the newly developed anechoic subglottal tract.

The first experiment showed that the origin of the vocal folds vibration is based on the aeroelastic mathematical model of phonation onset when the airflow velocity is above the critical flutter velocity. The prephonatory setting of the vocal folds and the mean level of the subglottic air pressure are in the dominant roles. The effects of acoustic subglottal resonances are of course substantial, but rather of a secondary importance for initiation of vocal folds vibration.

Similar conclusions follow from the second experiment~~s~~ with the anechoic subglottal tract studying the frequency jumps during vocal folds vibration by changing the vocal folds pre-stress in longitudinal direction. No qualitatively important effects in frequency jumps were found in these experiments with and without the subglottic resonances. The experiments showed that the frequency jumps are probably primarily caused by changes in bulging and pre-stressing of the vocal folds rather than by acoustic properties of the subglottal tract. May be more information on changes of the vocal folds vibration before, during and after frequency jumps could be identified by using HS camera, which has not been used in these experiments, however.

Phone (Secretary): (+420) 266 052 021
Phone (Financial Dept): (+420) 266 053 122
EU VAT No: CZ61388998
Email: secr@it.cas.cz
www.it.cas.cz

The third paper, which is probably thematically closer to the music and singing interests of the thesis author, concerns differences between chest and head registers, studied in a professional female singer. The HS camera is used for measuring glottal area waveforms (GAW), and simultaneously the EGG and microphone signals were recorded. Clear differences between chest and head registers were found in the normalized amplitude quotient (NAQ) and closing quotient (CgQ). The quotient CgQ could be closely related to impact stress during vocal folds' contact because it tells about the closing speed of the glottis. I would like to note that the time derivative of the measured glottal area waveform (GAW) could say even more on the vocal folds' loading by impact stress level during the vocal folds' collision. The measured data suggest that the chest-like register offers a higher risk for creation of nodules on the vocal folds due to the lower CgQ than the head-like register. An interesting result is that the closed quotient (CQ) does not clearly distinguish between the chest and head registers, similarly as so called "outsiders" included in the listening tests.
Finally, Hugo Lehoux contributed as one of the co-authors (software development, writing review and editing) in publishing the article (IV) on 3D modeling of vocal folds' vibration with mucosal waves. The simulation results seemed to correspond to vocal fold vibration patterns observed in HS registrations of normophonic and dysphonic speakers. Thus the simulation results may help to identify clinically various voice problems from HS registrations, and increase understanding of the relations between the deviant vocal fold vibration patterns and the resulting acoustic and auditory-perceptual characteristics of dysphonia.

In general, the thesis is well and clearly written, nearly without any errors, misprints and problematic unclear formulations. The thesis is well structured, and the referenced literature is adequate and up-to-date.

The author has published the original approaches developed by him during his PhD. studies in four high quality impacted journals (*Scientific Reports, Journal of Sound and Vibration, Journal of Voice, Journal of Speech Language and Hearing Research*).

The obtained results are novel and important for better knowledge of principal physical processes in voice science and for opening up new pathways for future investigations.

Thus in conclusion I can fully recommend excepting the thesis for the defense and after its successful presentation to award Hugo Lehoux the PhD. title in Biophysics.

Some remarks and notes for discussion:

Page 6_3 – How the "bandwidth", reflecting the damping in the vocal tract, is defined and are there differences in damping of acoustic waves between supraglottal and subglottal spaces?

Page 14_3 – Please, could you specify more clearly that the output voice sound is the result of the linear combination of the source and the filter? Is it really a linear combination?

Page 16_4 – Please, could you specify if the low-frequency band from 0 to 2000 Hz and the high-frequency band from 2000 to 4000 Hz chosen for the spectral balance are possible to use generally?

Page 23, end of paragraph 2.4.1. – How was the vocal fold thickness defined in the experiments with excised larynges and was it possible to detect any effect of vocal fold thickness on the vocal folds vibration amplitude?

Prague, August 3, 2023

Ing. Jaromír Horáček, DrSc.
tel.+420 266 053 125
e-mail: jaromirh@it.cas.cz

KTH
JOHAN SUNDBERG
Professor of Music Acoustics (retired)
Department of Speech Music Hearing
School of Computer Science and Communication
SE-100 44 Stockholm
Mobile: +4670 740 7873, Email jsu@ kth.se

2023-08-10

Review of Hugo Lehoux's doctoral dissertation
Biomechanics and acoustics of voice production

The dissertation consists of an introduction chapter presenting an overview of the functioning of the human voice and summaries of the four articles, which constitute the main part of the work. Of these articles, Lehoux is the first author of three.

The articles, all published in prestigious international journals, concern the phonatory function of the vocal folds, specifically the register phenomenon. The first analyses the role of tracheal resonance, the second is about factors determining the size of the pitch jumps that typically occur at register breaks, and the third concerns voice source differences between the modal and falsetto registers. The fourth article presents a development of a tool for analysis of vocal fold vibration.

As can be expected from articles published in prestigious international journals, all articles in the dissertation keep a high quality. Help from the respective journals' reviewers is not mentioned in any of the articles' *Acknowledgements* sections, but it would be reasonable to assume such help has contributed to the final quality.

The two articles, which concern the phenomenon of sudden shifts in vocal fold vibration frequency, are based on experiments where ex vivo vocal folds are brought to vibration by an airstream through a windpipe with systematically varied acoustic properties. The experimental conditions are carefully controlled, granting reliable results. In one condition, the windpipe is void of resonance and in the remaining conditions it possesses resonances at different frequencies. The results show that the sometimes assumed relevance of tracheal resonance is negligible.

Figure 2 of the first article is quite remarkable. It shows the waveform of the subglottal pressure in the resonance-free windpipe. The waveform is an almost exact upside-down copy of the glottal airflow waveform. This is novel and convincing proof of the reliability of the experimental setup.

- It can be noted that the subglottal pressures used are remarkably high, actually higher than those typically used even by singers. This suggests that the properties of the vocal folds used in the experiment differed from those of humans.

The third article analyses vocal fold vibration parameters in co-author Lisa Popeil, who feels that she can shift between a chest-like and a head-like register over a wide range of phonation frequencies. The article analyses differences in vocal fold vibration parameters between the singer's versions of these registers over a wide range of phonation frequencies. The results reveal a number of differences. In the chest-like register her closed quotient was longer, her vertical phase difference was greater, her normalised amplitude quotient was lower and her glottal closure was quicker. All these differences support the conclusion that the glottal adduction was firmer in the chest-like register. The article invites the following points discussion:

- Subglottal effect is known to have strong effects on the glottal airflow, generally referred to as the voice source. In fact, an increase of this pressure causes all the effects observed in this article. It increases the closed quotient, which would be associated with a larger phase difference between the lower and upper parts of the glottis. Further, it lowers the normalized amplitude quotient, and it increases the maximum flow declination rate, which is related to the glottal closing speed. It is unfortunate that the subglottal pressure was not measured in this investigation.
- As is common in single-subject studies, the analysed examples were assessed in a listening test. Two panels were recruited, one was well informed (the authors) while the other one was not. The results showed that the less qualified listeners were less consistent. This raises an important question. Perceptual tests with uniformed or insufficiently informed listeners are common in the voice research field. Is it fair to conclude that listening tests have a limited value as compared to analyses of physiological parameters?
- This investigation analysed a single singer, who can shift between chest and head registers across a much wider frequency range than other singers. This raises the question if this should be ascribed to special anatomical characteristics or to training.

Some subjects inviting discussion:

Vocal fold contact can happen before the folds stop the glottal airflow completely. What is the relevance of this difference in the analysis of phonation?

On page 17 it is mentioned that sudden stopping of thyroarytenoid activation is *assumed* to create register breaks. Has this not been shown in several studies?

On page 19 it is mentioned that Flanagan's two-mass model could not vibrate without a vocal tract? This is sometimes believed to be the case also with the real vocal folds. Is that correct? Lip trills can be produced without any resonator attached to the mouth.

On page 14 variation of the speed of the mucosal wave is mentioned. This should be relevant to the closed quotient. What could be the reason for this speed difference?

On page 23 it is stated that increase of the vocal fold thickness *can* increase the vertical phase differences, so is it possible that it does not?

As compared with kymograms, high-speed video documentation would offer more information and there are now powerful softwares that can perform automatic analysis. What is the role of the kymogram in the future?

Finally, analyses of source-filter interaction based on digital models has dominated voice research for a long time, while experimental data have been more rare. This dissertation presents carefully planned and carefully realised experiments to elucidate the same phenomenon. This is a welcome example for future voice research. The dissertation shows that Hugo Lehoux possesses the skills and gifts to produce excellent scientific research and thus to carry on the international top quality research arising around his supervisor, professor Jan Ŝvec.

In conclusion I fully recommend that the dissertation be accepted for defense and that Hugo Lehoux be awarded the PhD. title in Biophysics.

KTH as above

Johan Sundberg