SINGING REGISTER DIFFERENCES IN VOCAL FOLD OSCILLATIONS OBSERVED ACROSS THREE OCTAVES THROUGH LARYNGEAL HIGH-SPEED VIDEOENDOSCOPY

Faculty of Science

Hugo Lehoux[1], Lisa Popeil[2], Jan G. Švec[1]

[1]Department of Biophysics, Faculty of Science, Palacký University in Olomouc, Olomouc, Czechia

[2]Voiceworks, Los Angeles, California, USA

hugo.lehoux01@upol.cz; lisapopeil@mac.com; jan.svec@upol.cz

INTRODUCTION

- Singing voice registers are mostly based on perceptual sensations; objective evidence has been insufficient.
- The two basic singing voice registers are a) modal/chest (M) and b) head/flutey (H) register.
- Usually, the M register is associated with lower pitches than the H register [1], but a transition regions exist where both registers can be produced at identical pitch.
- In this study we investigated a professional singer (co-author LP) who claimed to be able to produce both M and H registers at every pitch throughout the range of 3 octaves using different laryngeal mechanisms.
- The aim was to verify this claim and identify relevant parameters that could objectively distinguish the two registers.

Assumptions, hypotheses:

- M versus H register: more thyroarytenoid (TA) muscle activity [2], thicker/more bulged vocal folds (VF), more membraneous adduction, possible differences in cartilaginous adduction. The expected measurable differences are listed in Table 1.

Expected laryngeal change (M versus H register)	Measurable parameter
- Higher TA muscle activity - Thicker/more bulged VF - More membraneous adduction	- Larger vertical phase differences (PD) - Larger vibratory amplitudes of the lower VF margin - Higher closing speed - Lower normalized amplitude quotient (NAQ) - Larger closed quotient (CQ) in anterior and middle glottis

Table 1. Expected measurable differences between the M and H registers.

METHODS

Experimental setup (Fig. 1)

- We observed and recorded the VF vibrations with a high-speed (HS) camera (Photron, FASTCAM Mini AX100 54, at 7200 and 13600 fps) connected to a rigid endoscope (K.Storz, 8707 DA, 90°).
- We simultaneously acquired audio signals at 48 kHz sampling rate:
  - Radiated sound with a camera-mounted condenser omnidirectional microphone at about 21 cm distance.
  - VF contact area with an electroglottograph (EGG, Glottal Enterprise EG2-PC).
  - Synchronization signal from the camera.

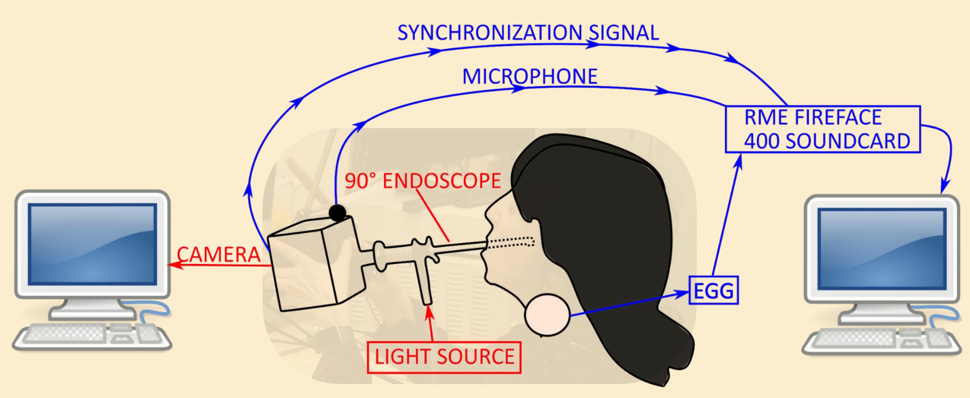

Figure 1. Diagram of the setup used in this study.

Experimental procedure

- Short phonations on 37 pitches were produced from C3 = 131 Hz to C6 = 1047 Hz, first in M then in H (Fig.2).

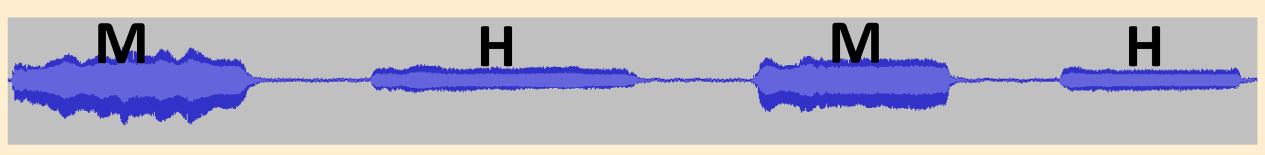

Figure 2. Microphone signal of one recording. The singer alternated the M and H registers at the same pitch

HS Video Processing (Fig. 3 & Fig. 4)

Glottis segmentation from HS video frames was done using the Glottal Analysis Tools (GAT) software developed at the University of Erlangen [3]. Digital kymograms (DKG) were obtained from custom scripts.

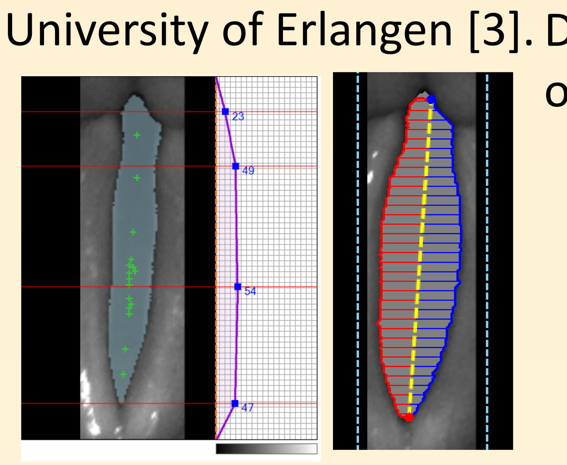

Figure 3. Segmentation of the glottis using GAT software.

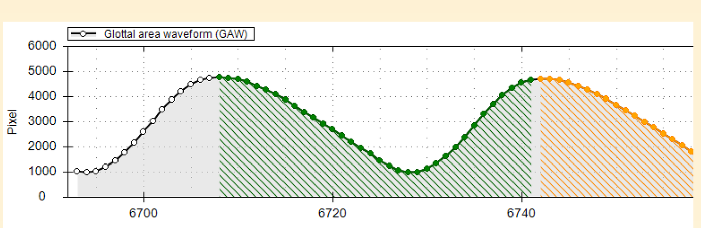

Figure 4. Glottal area waveform (GAW) in pixels derived from the segmented high-speed video recordings.

METHODS (cont.)

- Segmented images were used to obtain phonovibrograms (PVG) [3] mapping the VF edge distance from the glottal midline.
- The CQs were determined from anterior, middle and posterior parts of glottis by analysing PVG.
- The NAQ was derived from GAW:

$$NAQ = \frac{Max(glot.\ area) - Min(glot.\ area)}{Maximum\ Area\ Declination\ Rate * Period}$$

- The PD between the lower and upper margins of the VF were measured from DKG, using a sinusoid fitting method (Jiang et al. [4].)

Listening test

- Listening tests were performed to verify whether M and H registers are perceptually recognizable
- 10 Listeners rated the recordings on a continuous scale from "chest-like" to "head-like" using the Visor software from S. Granqvist [5].
- The participants repeated the test with an interval of at least one week between the tests.

RESULTS

Listening test (Fig. 5)

On average, the samples were correctly identified in 72% of the cases. There was a trend of lowered confidence (below 70%) around pitches D3, Bb3-F4, F5, and Ab5-C6.

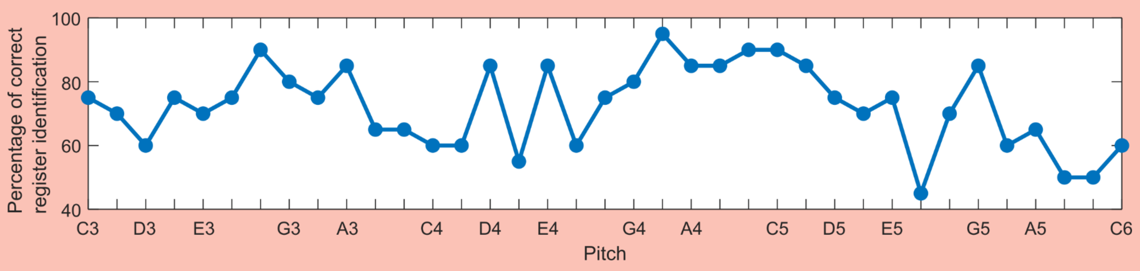

Figure 5. Results of the listening tests.

Visual analysis of the HS videos, DKG and EGG signal

- From C3 to G4, the HS videos revealed a visible stronger membraneous adduction in M. The EGG signals and the DKG revealed a visible longer VF contact in M.
- From C4 to G4, a difference in arytenoids adduction was visible (see Fig. 6). Outside this frequency region there were no clear visual differences in the arytenoids adduction.
- Above G4, no clear visual differences were found.

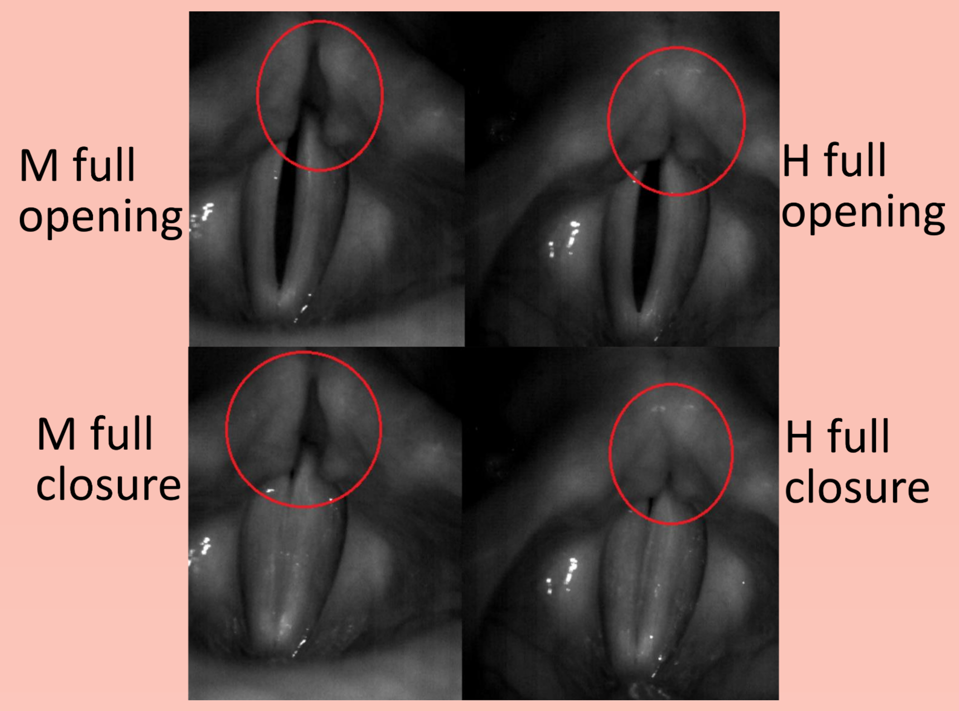

Figure 6. Selected frames from high-speed video recordings, for the F4 pitch. Notice the differences in cartilaginous adjustment (encircled) between M (left) and H (right).

Objective quantitative analysis

Closed quotient (CQ – Fig. 7):

- Up to G4, the CQ was consistently always smaller in H at all three positions
- At the anterior glottis, the CQ seemed smaller in H above B4, but the difference was smaller than below G4. No clear difference was observed between Gb4 and B4.
- At the middle and posterior glottis, no clear difference was visible between M and H above G4.

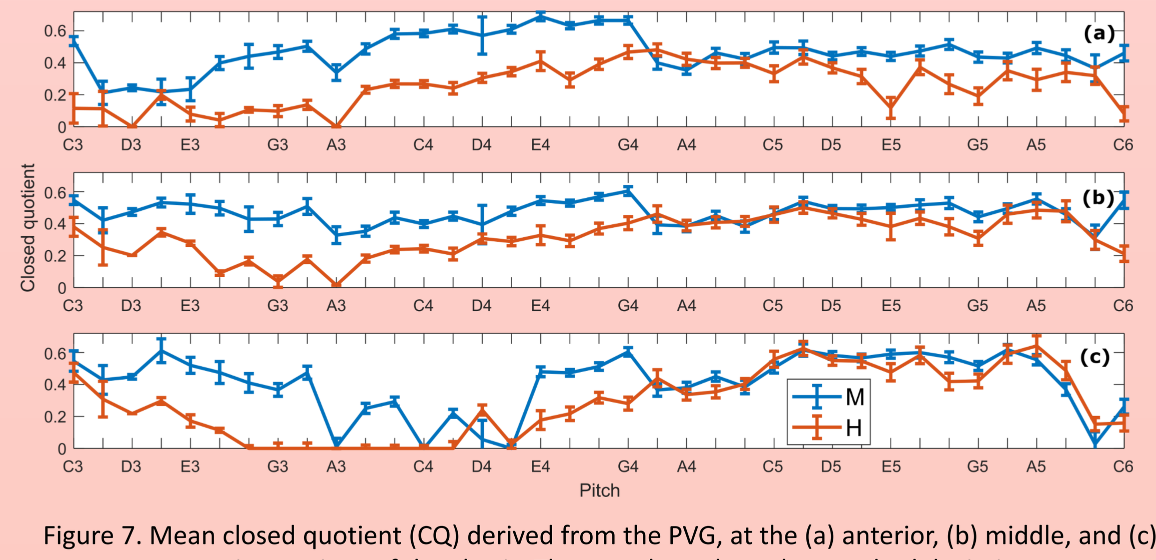

Figure 7. Mean closed quotient (CQ) derived from the PVG, at the (a) anterior, (b) middle, and (c) posterior sections of the glottis. The error bars show the standard deviation.

Vertical Phase Differences (PD):

- The estimated vertical PD were larger in M up to G4. No consistent difference was visible above G4 (Fig. 8).

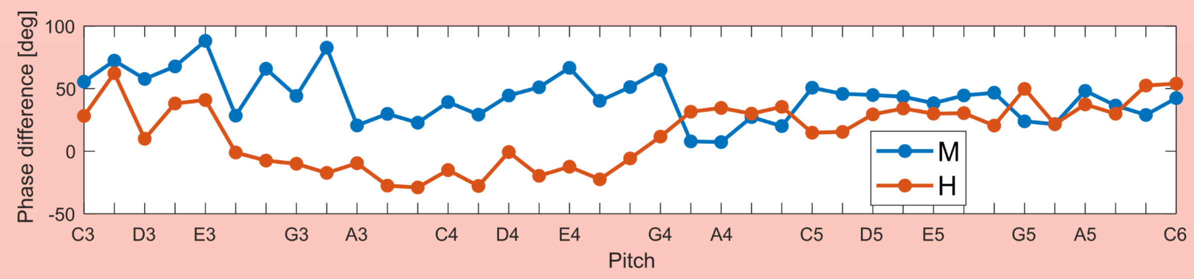

Figure 8. Estimated phase difference between the VF lower and upper margins.

Normalized Amplitude Quotient – NAQ (Fig. 9):

- The NAQ was consistently higher in H across the whole range.

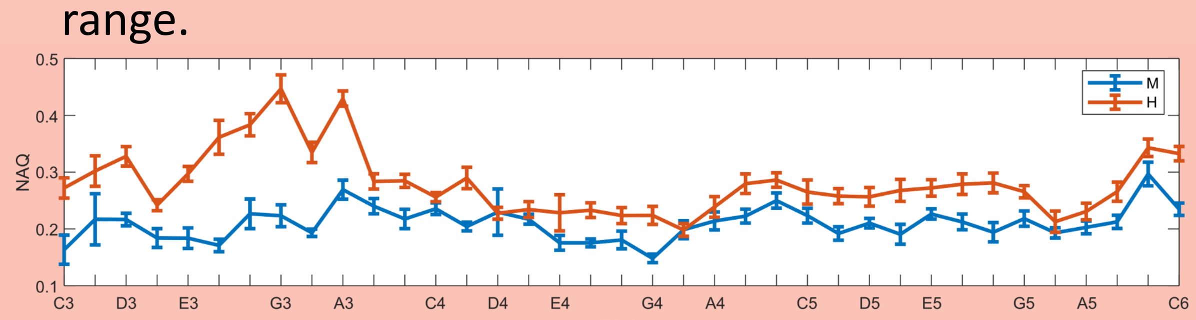

Figure 9. Mean NAQ calculated from the GAW, for every pitch. The error bars show the standard deviation.

EGG signal (Fig. 10):

- The RMS value of the EGG signal was consistently higher in M across the whole range, indicating larger vocal fold contact area in M. The difference was smallest at the medium pitches.

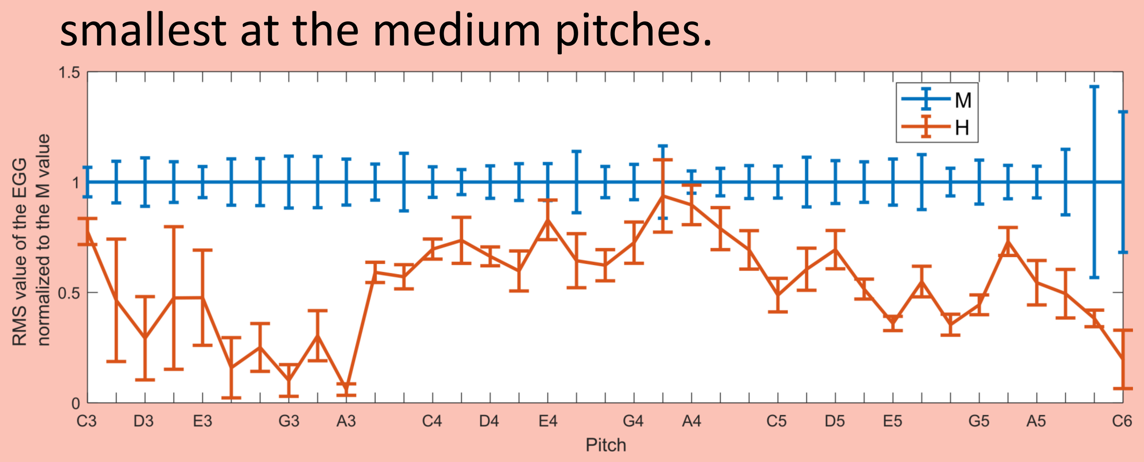

Figure 10. Mean RMS value of the EGG signal normalized by the value in M. The error bars show the standard deviation.

DISCUSSION AND CONCLUSION

- The singer was able to produce two different laryngeal voice qualities across the range of 3 octaves.
- Perceptually, these two voice qualities were correctly identified with the intended voice registers in 72% of the cases.
- Up to G4, all investigated parameters (PD, CQ, NAQ, EGG RMS amplitude) supported the hypothesis of thicker and more adducted vocal folds in the M register.
- Above G4, two of these parameters still showed consistent differences between the two registers: the NAQ was lower in M, the EGG RMS amplitude was higher in M.
- Above B4, the anterior and middle CQ were higher in M, but no clear difference was visible on the posterior CQ.

Limitations:

- This study was done with a **single** singer. Currently we are not aware of any singer who has the abilities to produce the two different registers across such a wide pitch range.
- The maximum frame rate was 13600 fps which is rather low for very high pitches (~ 1000 Hz).

Acknowledgements

This study was supported by the Czech Science Foundation (GACR) project no. 19-04477S.

REFERENCES

[1] H. Hollien, "On Vocal Registers," *J Phonetics,* vol. 2, pp. 125-143, 1974.

[2] M. Hirano, W. Vennard, and J. Ohala, "Regulation of register, pitch and intensity of voice. An electromyographic investigation of intrinsic laryngeal muscles," *Folia phoniatrica*, vol. 22, 1970.

[3] J. Lohscheller, U. Eysholdt, H. Toy, and M. Döllinger, "Phonovibrography: Mapping High-Speed Movies of Vocal Fold Vibrations Into 2-D Diagrams for Visualizing and Analyzing the Underlying Laryngeal Dynamics," IEEE TRANSACTIONS ON MEDICAL IMAGING, vol. 27, 2008.

[4] J. J. Jiang, Y. Zhang, M. P. Kelly, E. T. Bieging, and M. R. Hoffman, "An automatic method to quantify mucosal waves via videokymography," *The Laryngoscope,* vol. 118, pp. 1504-1510, 2008.

[5] S. Granqvist, "Visor: Visual sort and rate", http://www.tolvan.com/index.php?page=/visor/visor.php, last accessed 24.10.2020