INTERSPEECH 2023
20-24 August 2023, Dublin, Ireland

A Relationship Between Vocal Fold Vibration and Droplet Production

Tsukasa Yoshinaga[1,2], *Takayuki Arai*[3], *Akiyoshi Iida*[1]

[1]Toyohashi University of Technology, Japan [2]Osaka University, Japan
[3]Sophia University, Japan

yoshinaga@me.tut.ac.jp, arai@sophia.ac.jp, iida@me.tut.ac.jp

Abstract

While some aerosol droplets causing airborne transmissions are argued to be produced by vocal fold vibrations, the detailed production mechanisms were unclear due to the difficulty of direct observation. In this study, by using a transparent acrylic vocal fold model and high-speed imaging, we observed vocal fold vibrations that produce droplets of artificial mucus to clarify the relationship between the vocal fold vibration and the droplet production. The vocal fold model was set on a lung model which has a manual diaphragm. The artificial mucus in between the vocal folds was lighted up by a laser sheet. The results showed that droplets were produced when the sound amplitudes were decreased or the fundamental frequency became unstable. We observed that mucus drops attached to the middle of the vocal fold wall splashed and formed a droplet that flew from the vocal fold wall. This suggests that the shear forces of turbulent airflow passing on the mucus mainly produced the droplets.

Index Terms: speech production, vocal folds, COVID-19, high-speed imaging, droplet

1. Introduction

Since the outbreak of COVID-19 in 2020, many researchers have investigated the transmission mechanism and infectiousness of aerosol droplets. The importance of airborne transmission of COVID-19 has been widely recognized, and the risk of spreading infectious aerosols through speaking and singing has been assessed depending on the situation [1]. In the airborne transmission, disease is transmitted by inhaling aerosol droplets exhaled by an infected person.

There are three possible mechanisms for the generation of aerosol droplets [2]: the aerosol generation by bursts of bronchiole fluid film; the aerosols produced by shear-flow forces in large bronchi; and the droplets formed by the opening and closing of vocal folds in the larynx. The third type of aerosol is produced not only by coughing but also by speech production, and the amount of aerosol production in the speech was reported to be much larger than that in normal breathing [3-4]. Furthermore, it was found that the amount of aerosol varies depending on the type of phonemes [5]. However, it is still unclear how the aerosol droplets are produced from the opening and closing of the vocal folds. Although many studies have been conducted to visualize the aerosol emissions during speech utilizing laser sheets or computational simulation for showing the effectiveness of face masks [6-8], the source of the aerosol droplets in the larynx and upper airways is still unclear. Recently, an attempt was made to measure aerosol in the larynx using a vocal fold replica oscillated by a motor [9]. Investigating how the aerosols are emitted from the vocal folds and how the risk of infection is increased by talking and singing will be useful in developing countermeasures against various diseases transmitted through aerosol droplets.

In this study, to clarify the relationship between the vocal fold vibration and aerosol droplet production, we observed artificial mucus put on a reed-type artificial vocal fold made of transparent acrylic plates by using a high-speed camera. By analyzing simultaneous recordings of sounds with high-speed images, the characteristics of vocal fold vibration during the droplet generation are elucidated.

2. Materials and Methods

2.1. Vocal fold model

In this study, we used a reed-type vocal fold model made of transparent acrylic plates to observe the droplets in between the vibrating folds. The reed-type vocal folds were first proposed by von Kempelen for his speaking machine around 1780 [10]. A reed is fixed to a retainer as shown in Fig. 1. An airflow passing through the channel between the reed and retainer vibrates the reed. The pulsating airflow produces a sound similar to that of human vocal folds, and the sound amplitudes decay with −6 dB/oct. up to 4 kHz [11]. Due to the easy fabrication procedures and reproducibility compared to silicone vocal fold models, the reed-type vocal folds are still used for voice production studies [12-13].

To observe droplet productions in between the reed and retainer, we made larger vocal folds with a longitudinal length of 40 mm compared to the usual vocal folds with a length of 30 mm. The reed has a width of 10.5 mm and a thickness of 0.3 mm. The reed tip was cut to close along the retainer edges. The average oscillation frequency is approximately 84 Hz, and this fundamental frequency f_o is slightly lower than the usual male voice of f_o ~ 120 Hz. However, f_o = 84 Hz is within the vocal range of a professional bass singer (C2 ~ F4: 65.4 Hz ~ 349 Hz) [14]. The artificial mucus was put in the retainer outlet which has a 10.5 mm semicircle, and airflow was exhaled from the lung model to produce the droplet.

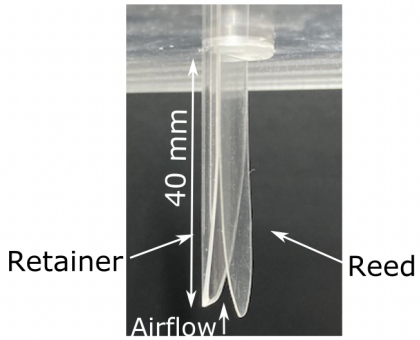

Figure 1: *Transparent reed-type vocal fold model.*

10.21437/Interspeech.2023-317

2.2. Experimental setups

The experimental apparatus and setups are depicted in Fig. 2. The vocal fold model was set in a chamber made of acrylic boards (268 × 134 × 280 mm^3), and a diaphragm made of rubber film of a large balloon in a lung model was pushed by a hand to exhale the air from the vocal folds. An air tube mimicking the trachea was omitted to clearly observe the mucus inside the vocal fold channel.

The artificial mucus was made of 76 g glycerin and 12 g sodium chloride dissolved in 1 L distilled water. To investigate the effects of viscosity on droplet production, we tested the mucus with doubled concentration. Before each measurement, 30 µL of mucus was put on the vocal fold surface by a micropipette. A green laser sheet (5W 532 nm CW Laser, CNI Co. Ltd., China) was set to illuminate the center of vocal folds and to visualize the droplets of mucus. A high-speed camera (Phantom T1340, Vision Research, NJ) was used to record the images at a frame rate of 3200 fps. At the same time, the sound produced by the vocal folds was recorded by a microphone (NA-28, Rion, Japan) at 300 mm from the vocal fold outlet with a sampling frequency of 48 kHz.

The experiments were conducted three times with the following procedures. First, the diaphragm of the chamber was pulled to inhale the air, and the diaphragm position was fixed. Second, the artificial mucus was put on the vocal fold surface by a micropipette. Third, the laser sheet was turned on, and the room lights were turned off. Finally, we started recording the high-speed camera and the microphone, and the diaphragm was rapidly pushed up. For the first and second measurements, we used the normal mucus, whereas the mucus with doubled concentration was used for the third measurement.

Duration of the reed oscillation was different in each recording in the range from 0.7 s to 2 s, and all data were recorded for 3 s including before and after the oscillation. The sound spectrograms were calculated using the Hann window with a window length of 2286 points and an overlap ratio of 80%.

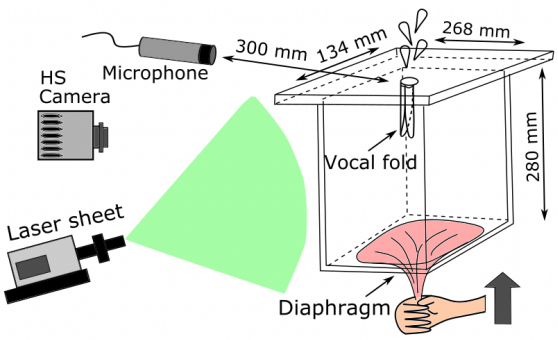

Figure 2: *Schematic of experimental setups.*

2.3. Droplet measurement

After the recordings of the high-speed camera, we measured the number and size of droplets produced by the reed vibrations. An example of recorded images is shown in Fig. 3. In the middle of the image, the mucus is in between the retainer and the vibrating reed. Above the reed and retainer, there is a 10 mm thick acrylic plate, and the droplets can be observed above the plate. To measure the size of droplets, we counted the number of white pixels within the region 14 × 81 mm^2 (325 × 2040 pixels) above the plate for each image. The images were firstly cropped for the measurement area from the original image, and the cropped images were binarized with a function *imbinarize* in MATLAB 2021b (MathWorks) using a threshold value of 0.4 for Otsu's threshold selection method [15]. Then, the areas of binarized droplet parts were counted for each image.

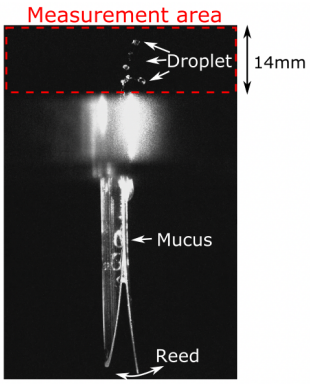

Figure 3: *Example of droplet image and measurement area.*

3. Results

The measured spectrogram, sound amplitude waveform, and area of droplets for the first measurement are shown in Fig. 4. The time t is set to 0 when the reed started vibrating with the complete closure. The average fundamental frequency f_0 was 83 Hz. When the reed started vibrating, f_0 slightly increased for initial 0.05 s, and then the oscillation became stable. The f_0 was gradually decreased until t = 0.6 s. After t = 0.6 s, the amplitude suddenly decreased by 11 dB, and the oscillation stopped at t = 0.7 s. In the initial stable oscillation, although some liquid films in the retainer (like Fig. 3 middle part) was vibrating, no droplet appeared from the outlet of the vocal fold. When the sound amplitude decreased at t = 0.3 s, small droplets started appearing. Then, at t = 0.33 s, larger droplets emerged and flew upwards. After the droplet production, the sound amplitude increased again at t = 0.34 s and reached a peak at t = 0.41 s. The sound amplitude rapidly decreased at t = 0.43 s, and vast number of droplets splashed at the vocal fold outlet at t = 0.44 s. These droplets remained in the measurement area from t = 0.44 to 0.48 s. Then the sound amplitude increased and decreased from t = 0.48 to 0.58. At t = 0.59 s, the droplets, which have similar in size to the first ones, were produced again, and the reed oscillation ceased after six cycles of small sound production.

The sound and droplet areas of the second measurement are plotted in Fig. 5. The initial f_0 was 77 Hz. At the beginning of the reed vibration, the first droplet was launched from a liquid film in the retainer at the third reed oscillation cycle, and the droplet went straight up towards the outlet. After that, the second droplet group flew at t = 0.14 s, and at this moment, the sound amplitudes were unstable. Then, the droplet production stopped at t = 0.2 s. The sound amplitudes became stable at around t = 0.4 s and gradually decreased by 6 dB from t = 1.3 to 1.7 s. The total duration of the sound production was approximately 1.7 s and much longer than that of the first measurement. With the decrease of the sound amplitudes, f_0 was also decreased to 59 Hz due to the flow rate decrease.

The third measurement was conducted using the mucus with increased concentrations of glycerin. Hence, the viscosity

of mucus was increased, and the mucus became stickier to the vocal fold walls. The initial f_o was 91 Hz. For the first time, small droplets started flying before the reed vibration initiation (t = −0.01 s). The droplets were formed by the initial airflow that produced the negative pressure pulling the reed to close. Then, at the fourth oscillation, the large droplets flew from the vocal fold outlet. The sound amplitude was gradually decreased by 6 dB from the beginning to t = 1.2 s. At t = 0.27 s, large droplets flew from the retainer, and the fundamental frequency and its harmonics started decreasing at the same time. Once the sound amplitude reached a minimum at t = 1.3 s, the droplets were generated again, and the sound amplitude increased. The large droplets flew at t = 1.4 and 1.6 s, and the sound amplitudes and the fundamental frequency were unstable around that period. The sound production continued until t = 1.9 s. The fundamental frequency at the last state was f_o = 62 Hz.

While the area of droplets was the largest in the first measurement (> 2000 pixels.), the number of droplet productions was the largest in the third measurement (> 20). The stable reed oscillation and sound generation were achieved only in the second measurement, and both fundamental frequency and sound amplitudes were unstable for the whole oscillation cycles of the first and third measurements.

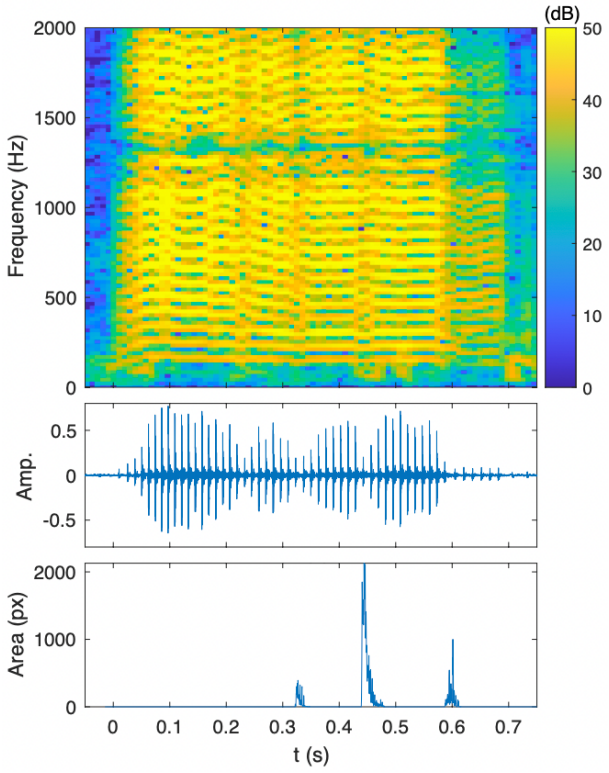

Figure 4: *Spectrogram, sound amplitude waveform, and pixel area of droplets for the first measurement.*

4. Discussion

Comparing the three trials in this study, the sound generation in the first trial was quite short and unstable. However, the size of droplets was the largest (~ 0.86 mm near the outlet) in the three trials. This is probably because the initial shapes and positions of mucus were different in the three trials. Since the mucus was set on the vocal folds by hand using a micropipette, it was difficult to make the exact same mucus conditions, and some

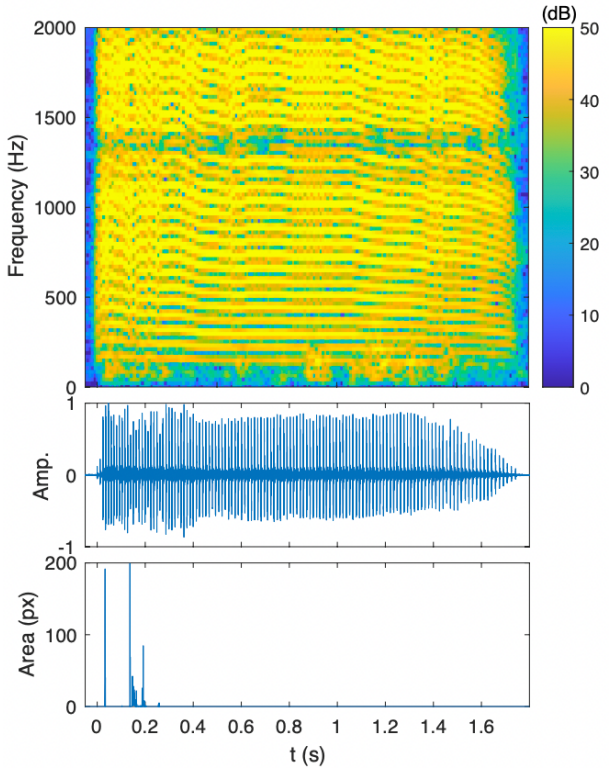

Figure 5: *Spectrogram, sound amplitude waveform, and pixel area of droplets for the second measurement.*

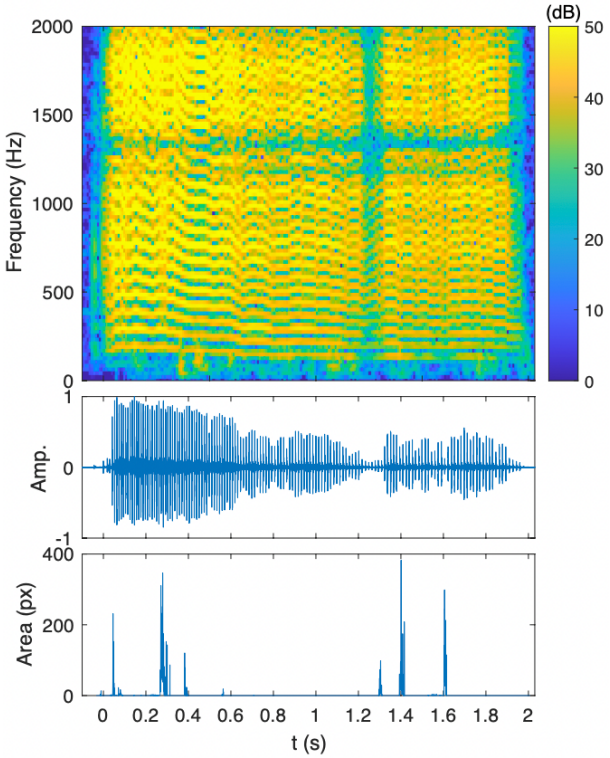

Figure 6: *Spectrogram, sound amplitude waveform, and pixel area of droplets for the third measurement. The concentration of mucus was increased from the first and second measurements.*

portions of mucus significantly disturbed the reed oscillation in the first trial. Therefore, large droplets were produced when the sound amplitude was decreased. In addition, all droplets were produced while the sound production was unstable in the other trials. The droplets flew only at the beginning of unstable oscillation for the second try, and the droplet launched when the f_0 or the amplitudes varied for the third try. These results indicate that the energy of pulsating airflow went to the droplet production and flow power to the reed oscillation became weak.

To observe the droplet production process, we extracted some pictures of mucus in the retainer around $t = 0.03$ of the second experiment in Fig. 7 [16]. At the beginning, liquid films of mucus waved at the middle of the retainer. After the closure of the reed at $t = 0.02375$ s (Fig. 7 (a)), the films splashed in the middle (white arrow) and formed a droplet above as shown in Fig. 7 (b). According to the numerical flow simulation for the reed-type artificial vocal fold [11], turbulent airflows are formed in the retainer when the reed is closed. A passage of the turbulent flow splashed the mucus films and produced the droplet in the retainer. Then, this droplet flew straight away from the vocal fold outlet as shown in Fig. 7 (c) and (d). This production mechanism is slightly different from the previously considered mechanisms that liquid bridges between the vocal folds produce the droplets [2]. Indeed, even though some liquid bridges between the reed and retainer were observed, we could not confirm that the liquid bridges produced droplets from the images. Instead, we observed that the turbulent airflow in between the vocal fold walls formed the droplets that flew to the vocal tract. Although the shape of the vocal fold model is different from the actual human vocal folds, this phenomenon can happen in a place where the turbulent airflow passes on the mucus films, like larynx ventricles and epiglottis.

It should be noted that most of the droplet areas observed in this study were larger than 10 pixels (*i.e.*, larger than 430 μm in diameter). The main aerosol droplet size of a human is less than 100 μm [17], and even if some parts of water were vaporized, the size was larger than the usual aerosol droplets that float in the air. However, there is a possibility that the flying droplets were further separated by the turbulent airflow and formed smaller droplets. In the next study, we have to further develop an experimental setup that enables us to observe the formation of smaller droplets.

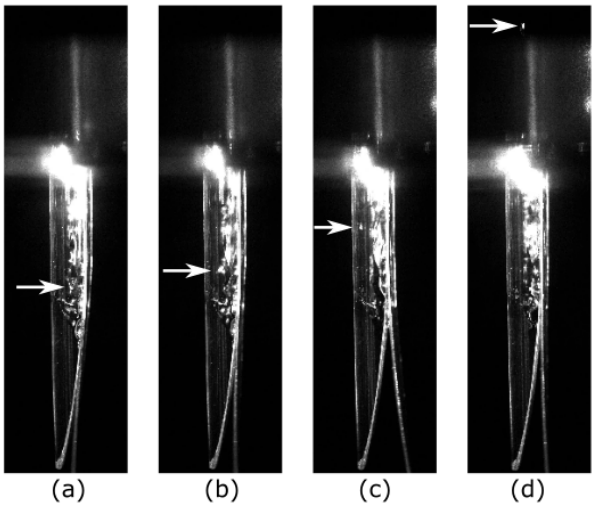

Figure 7: *Images of mucus in the retainer and produced droplet at t = 0.02375 s (a), 0.02434 s (b), 0.02656 s (c), 0.03219 s (d). (See multimedia file [16].)*

5. Conclusions

In this study, we measured droplet production in the reed-type artificial vocal folds using a high-speed camera to clarify the relationship between the vocal fold vibration and aerosol droplet production. The results showed that large droplets were produced when the sound amplitudes were decreased or the fundamental frequency became unstable. The sizes of the droplets were the maximum for the first experiment when the sound production was most unstable. Meanwhile, the number of droplet productions was the maximum for the third trial when the concentration of artificial mucus was increased. With the observation of high-speed imaging, we could observe that the drop of mucus splashed in the middle of the vocal folds and formed a droplet that flew from the vocal fold outlet. This was probably caused by the shear forces of turbulent airflow passing on the liquid films attached to the retainer wall. These results suggest that people should be cautious of aerosol production when producing voice with large and unstable flow rates. For future work, we need to further improve the experimental setups to observe the production of smaller droplets that can float in the air and more strongly affects the virus's airborne transmission.

6. Acknowledgements

This work was partially supported by JSPS KAKENHI, Grant Numbers JP21K02889, JP23K17195, and Sophia University Special Grant for Academic Research (Research in Priority Areas). This study was partly presented in Acoustical Society of Japan 2022 Fall Meeting, and we acknowledge insightful comments from audience in the conference.

7. References

[1] M. Z. Bazant, and J. W. Bush "A guideline to limit indoor airborne transmission of COVID-19," *Proceedings of the National Academy of Sciences*, 118(17), e2018995118, 2021.

[2] B. Patterson, and R. Wood "Is cough really necessary for TB transmission?" *Tuberculosis*, 117, 31-35, 2019.

[3] L. Morawska, G.R. Johnson, Z.D. Ristovski, M. Hargreaves, K. Mengersen, S. Corbett, C.Y.H. Chao, and D. Katoshevski, "Size distribution and sites of origin of droplets expelled from the human respiratory tract during expiratory activities," *Journal of Aerosol Science*, 40(3), 256-269, 2009.

[4] S. Asadi, A.S. Wexler, C.D. Cappa, S. Barreda, N.M. Bouvier, and W.D. Ristenpart "Aerosol emission and superemission during human speech increase with voice loudness," *Scientific Reports*, 9(1), 1-10, 2019.

[5] S. Asadi, A.S. Wexler, C.D. Cappa, S. Barreda, N.M. Bouvier, and W.D. Ristenpart "Effect of voicing and articulation manner on aerosol particle emission during human speech," *PLoS One*, 15(1), e0227699, 2020.

[6] S. Verma, M. Dhanak, and J. Frankenfield, "Visualizing the effectiveness of face masks in obstructing respiratory jets," *Physics of Fluids*, 32(6), 061708, 2020.

[7] K. Onishi, A. Iida, M. Yamakawa, and M. Tsubokura, "Numerical analysis of the efficiency of face masks for preventing droplet airborne infections," *Physics of Fluids*, 34(3), 033309, 2022.

[8] T. Arai "Vocal-tract models to visualize the airstream of human breath and droplets while producing speech," *in Proceedings of INTERSPEECH 2021 - 22nd Annual Conference of the International Speech Communication Association*, pp. 3021-3025, 2021.

[9] L. Fritzsche, R. Schwarze, F. Junghans, and K. Bauer, "Toward unraveling the mechanisms of aerosol generation during phonation," *Physics of Fluids*, 34(12), 121904, 2022.

[10] H. Dudley, and T.H. Tarnoczy “The speaking machine of Wolfgang von Kempelen,” *Journal of the Acoustical Society of America*, 22(2), 151–66, 1950.
[11] T. Yoshinaga, T. Arai, R. Inaam, H. Yokoyama, and A. Iida “A fully coupled fluid–structure–acoustic interaction simulation on reed-type artificial vocal fold,” *Applied Acoustics*, 184, 108339, 2021.
[12] T. Arai “Sliding three-tube model as a simple educational tool for vowel production,” *Acoustical Science and Technology*, 27(6), 384–388, 2006.
[13] T. Yoshinaga, T. Arai, H. Yokoyama, and A. Iida, “Effects of airflow in constricted vocal tracts on vowel production of the reed-type artificial vocal fold,” *Acoustical Science and Technology*, 43(5), 283-286, 2022.
[14] J. Sundberg, J. “The level of the ‘singing formant’ and the source spectra of professional bass singers,” *Speech Transmission Laboratory Quarterly Progress and Status Report*, 4, 21-39, 1970.
[15] N. Otsu, “A threshold selection method from gray-level histograms,” *IEEE transactions on systems, man, and cybernetics*, 9(1), 62-66, 1979.
[16] See the supplemental multimedia file for a movie of droplet production process.
[17] G. Bagheri, O. Schlenczek, L. Turco, B. Thiede, K. Stieger, J. M. Kosub, ... & E. Bodenschatz “Size, concentration, and origin of human exhaled particles and their dependence on human factors with implications on infection transmission.” *Journal of Aerosol Science*, 168, 106102, 2023.