11th International Conference on Electrical, Electronics and Computer Engineering (IcETRAN), Nis, 3-6 June 2024

Breath Sound Acquisition in Different Environments

Dejan Ćirić
Faculty of Electronic Engineering in Niš
University of Niš
Niš, Serbia
dejan.ciric@elfak.ni.ac.rs, ORCID: 0000-0003-4974-3131

Đurađ Bogunović
Faculty of Electronic Engineering in Niš
University of Niš
Niš, Serbia
djuradj@elfak.rs

Marko Janković
Faculty of Electronic Engineering in Niš
University of Niš
Niš, Serbia
marko.p.jankovic@elfak.rs, ORCID: 0009-0002-1109-3241

Miljan Miletić
Academy of Applied Preschool Teaching and Health Studies, section in Kruševac
Serbia
mmelektronik1@gmail.com

***Abstract*—Breathing is one of the vital functions of a human. This is why the monitoring of breathing represents an important tool for diagnosing a medical condition in a patient, that is, for assessing health status. Apart from medical diagnostics, breath monitoring can be used for other purposes such as managing stress, optimizing sports performance, security applications, or human-computer interaction. Breath detection employs diverse sensors and approaches. The initial step in this process involves acquiring breath sound. This paper explores the potential for acquiring breath sound using several types of microphones in different environments. To this end, four microphones of varying types (digital MEMS, electret, studio condenser, and smartphone microphone) are utilized. Acquisition is conducted in two distinct environments or scenarios: with the microphone in the air in close proximity to the nose and mouth, and with the microphone in a military gas mask. The recorded signals are subsequently analyzed in both the time and spectral domains.**

Keywords—breath sound acquisition, MEMS microphone, condenser microphone, sound source vicinity, audio analytics

I. Introduction

The breath pattern contains important information about a person's medical health and reflects the current condition of the human body during various activities, including physical exercises, or in specific situations such as wearing a military gas mask. Thus, the respiratory system's health, as well as the health of other organs like the heart can be indicated by breathing characteristics (both inhalation and exhalation).

Shortness of breath, known as dyspnea, is a prevalent indication in various acute and chronic medical situations. Immediate breathlessness can manifest during an asthma or heart attack, while persistent breathlessness often signals chronic obstructive pulmonary disease, low cardiovascular fitness and congestive heart failure [1]. Experiencing breathlessness during exertion also serves as an independent predictor of mortality and is a frequently utilized clinical measure to evaluate and track the progression of diseases. Establishing a systematic approach to identify breathlessness could alleviate the challenges associated with recognizing this symptom, potentially offering healthcare professionals early insights into patients' medical conditions well before any signs of disease progression become apparent in a clinical environment [1].

For the most precise estimations of breathing patterns, commonly used sensors include thermistors, respiratory gauge transducers, and acoustic sensors. However, they are sometimes referred to as intrusive due to the discomfort they may cause in daily use [1]. Viewed from a different perspective, the predominant method for consistently monitoring respiratory patterns involves using sensors in direct contact with the face of the body [2]. However, a minority of systems can monitor respiration without necessitating contact between the sensors and the body. The use of wearable or specialized non-contact devices for respiratory monitoring has not gained widespread adoption, rendering it impractical for comprehensive coverage across the entire population. Moreover, it is unsuitable for long-term health monitoring for all patients at risk.

The use of sound and signal processing in respiratory analysis is not a recent development. The analysis and classification of breath sounds through signal processing have been ongoing for several decades. Over time, various methods for analyzing breath sounds have been proposed. The recent surge in the performance and popularity of artificial intelligence (AI) has resulted in AI-based approaches for identifying respiratory events. In this context, commonly employed techniques with proven efficacy include the use of k-nearest neighbor (k-NN), naive Bayes, support vector machines (SVM), random forest, deep neural networks (DNN), and hidden Markov models (HMM) [3].

The primary motivation behind our work has been to develop a simpler and more widely applicable solution for acquiring breath data as well as identifying breat pattern components. Through our approach of exploring different options for acquiring breath sound, the goal is to investigate the differences of breat patterns captured in diverse scenarios. Looking from a

This work was supported by the Ministry of Science, Technological Development and Innovation of the Republic of Serbia [grant number 451-03-65/2024-03/200102].

general perspective, the aim is to achieve competitive results in breath acquisition and subsequently in breath pattern monitoring. The findings in this study hold significant practical utility for further analysis in similarly focused research endeavors.

We have collected audio samples of breathing in two conditions—forced and normal (relaxed) breathing—from both male and female volunteers of varying ages and breathing tendencies. The samples were recorded using four different types of microphones, including a digital MEMS microphone, as well as studio, electret, and smartphone microphones. The collected audio samples are analyzed, focusing on differences between breathing patterns.

The paper is organized as follows: methods for breath acquisition, monitoring and detection are first reviewed in Section II. Then, the methodology used in this paper for breath acquisition and analysis is described in Section III. The results and discussion thereof are provided in Section IV, while the final thoughts and conclusions are presented in Section V.

II. Breath Monitoring and Detection

A. *Review of Approaches for Non-Sound Breath Monitoring*

Different scenarios for respiratory monitoring present distinct challenges and limitations. In sporting scenarios, various technical challenges, such as those introduced by motion artifacts, need careful consideration to select an appropriate approach for signal acquisition and processing [4]. Conversely, medical scenarios pose their own challenges and come with specific requirements.

Respiratory activity sensing can be classified into three broad categories: a) built-in motion sensors (IMU)-based breath activity detection, b) chest displacement sensing for breath pattern estimation, and c) acoustic-based (audio-based) breath monitoring [5]. According to another criterion, a variety of sensors used for breath monitoring can be categorized into two groups: contact and contactless sensors [1].

The measurement of respiration rate has relied on devices like respiratory belts or impedance pneumography, which, however, tend to be both costly and impractical for non-clinical applications [5]. A commonly employed contact-based approach involves capturing chest wall movements induced by respiration using devices such as strain sensors embedded in straps or garments [4]. In the field of sports and exercise, there is a growing interest in a different method, which centers on extracting breath patterns (respiratory frequency) from cardiac signals recorded by devices commonly used by athletes and exercise practitioners [4].

In more recent times, research has indicated the viability of gauging respiration rate in non-clinical environments through the utilization of consumer devices. Thus, a wireless earphone equipped with the IMU sensors is employed to track breath rate, and the IMU sensors built-in commercial earbuds are applied to monitor respiration rate during physical activities [5]. In an approach presented in [2], the movement of the thorax following emission of ultrasonic waves by a microphone built into a smartphone is monitored. Due to the utilization of ultrasonic irradiation, it is necessary to position the smartphone in front of the chest, rendering it impractical for individuals resting on a bed with a dense covering. Image information has been utilized to ascertain respiratory rate by placing a finger on the integrated camera of a smartphone, and thoracic movement has been detected by a conventional smartphone positioned on an individual's chest [2].

B. *Sound-Based Breath Monitoring*

Previous research on sound-based or audio-based breath detection has concentrated on identifying and categorizing specific breath sounds by utilizing breath characteristics to differentiate, for example, between normal and abnormal breath sounds [3,6]. In real-world scenarios, employing audio sensing may demand significant resources and raise privacy concerns [5]. Within these audio-based methods, both contact and contactless approaches can be discerned. An example of the former is the estimation of respiratory rate using contact-based sensors for tracheal sounds [7]. In this context, breath sounds are captured by a tracheal microphone [3], representing an invasive method that may cause discomfort to the user.

Regarding contactless audio-based methods, they primarily involve estimating the respiratory pattern using recorded respiratory sounds captured by a microphone or microphones, such as those in smartphones used for acquiring nasal breath recordings [8]. These methods typically belong to a non-invasive ones having an advantage of an easy realisation, but also drawback that they are more sensitive to external noises. Besides, microphone sensors are considered to be less prone to motion artifacts than other sensors such as strain sensors [4].

A built-in microphone in a smartphone attached to the bed headboards is also employed for detecting respiratory sounds based on deep learning technology [2]. In this approach, the subtle sounds of breathing recorded by a smartphone are syncronized with respiratory information obtained by the polysomnography examination. The smartphones' microphones are situated around 50 cm away from the patient's nose. Another method for determining real-time detection of breathing phases involves placing a smartphone's built-in microphone near the mouth and the nose of a patient [9]. However, this approach, while avoiding patient discomfort, is not well-suited for prolonged and continuous monitoring over the long term.

In some studies, other types of microphones are applied including contactless wearable near-field microphones used for respiratory rate estimation from short audio segments obtained after physical exertion in healthy adults [1]. In such a model-driven approach, it is critical to determine if such a sensor is capable of providing the audio data with sufficient information to distinguish the breath sound patterns. Use of earphones as a microphone for breath sound detection has also been explored in [8]. This approach is anticipated to enhance performance in situations where breath sounds are subtle and challenging to discern. Implementing this method would necessitate the evaluation of each earphone's performance.

A single highly directional microphone affixed to an articulated microphone stand oriented toward the subject’s head placed at an approximate distance of 22 cm is used in [10] to capture high-quality audio. A cardioid pattern capacitor microphone with 32 dB of sensitivity is utilized for real-time detection and identification of respiratory movements, encompassing both mouth and nasal inspiration and expiration [3]. In the investigation outlined in [11], an acoustic sensor is employed in proximity to the mouth to monitor respiration while asleep. A Matlab application based on a microphone is developed in [3] for identifying respiratory movements - inspiration and expiration - in real-time.

In [4], a facemask (made of 3D-printed thermoplastic polyurethane) incorporates a condenser microphone to estimate respiratory frequency from breath sounds during walking and running. This estimation in the time domain involves determining the time elapsed between consecutive exhalation events extracted from breathing sounds at 30-second intervals. Microphones embedded in facemasks are also utilized for estimation of respiratory frequency in both indoor (office) and outdoor (public street, public bus, and subway) settings [12].

There are studies for breath monitoring and analysis where the sound generated by breathing is collected by multiple microphones. An example of such a scenario can be found in [13] where the respiratory sound is acquired by multiple microphones installed near patient's beds. On the other hand, employing microphone arrays and beamforming techniques allows for the concentration on the breath signal while minimizing interference from surrounding ambient noise.

One more example of using respiratory sounds is related to identifying the respiratory phases by the sounds acquired by a trained physician in performing lung auscultation by placing a phonendoscope Littmann 3200 at patient's jugulum [14].

III. Breath Sound Acquisition in This Study

Several sound sensors (microphones) were used here to acquire breath sounds in two environments (scenarios): the sensor placed in the air close to the nose and mouth of a subject (at an approximate distance of a couple of centimeters), and inside a standard military gas mask (also in the proximity of the nose and mouth, in the area between them), as depicted in Fig. 1. The sound in the first environment (referred to here as 'air') was recorded by four microphones: a digital MEMS microphone (Infineon Xensiv MEMS microphone connected to the Audiohub Nano board), a studio microphone (AKG 120 Perception connected to the Focusrite Scarlett audio interface), an electret microphone (connected to a custom-made microphone amplifier and the same Focusrite Scarlett interface), and the microphone of a smartphone (Samsung Galaxy A7-2018).

The sound acquisition in the second environment (referred to here as the 'gas mask') was conducted using only the digital MEMS microphone mentioned above. The microphone's position within the gas mask was selected to optimize the capture of respiratory activity while minimizing the influence of surrounding noise, as illustrated in Fig. 2.

Audio samples collected in the air have a duration of several tens of seconds and contain sounds of two types of breathing: forced (characterized by pronounced inhales and exhales) and normal (relaxed, i.e., natural) breathing. Audio samples collected in the gas mask last for a minute and contain only sounds of normal breathing. All audio samples were recorded with a sampling frequency of 44.1 kHz and stored in WAV format.

The collected audio signals are first subjected to analysis in the time domain to investigate their waveforms, followed by analysis in the spectral and spectrogram domains. Through these steps, the main characteristics of nose and mouth breathing, as well as inhalation and exhalation, are observed and discussed. Certain differences caused by variations in microphones, environments, and subjects are also identified and presented.

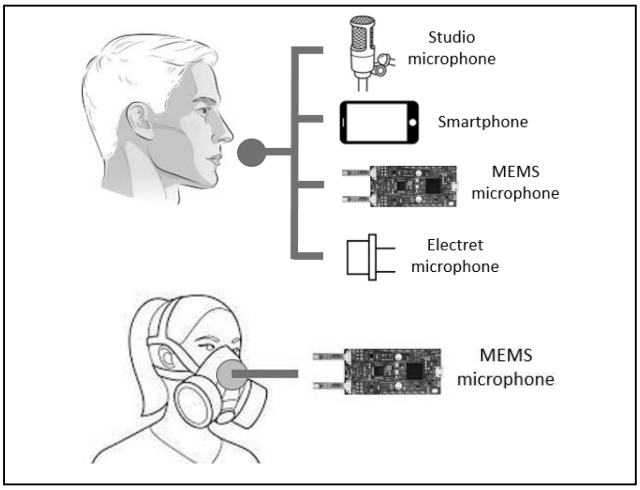

Fig. 1. Breath sound acquisition in air - in the vicinity of nose and mouth (top figure) and in military gas mask (bottom figure)

Fig. 2. MEMS microphone board atteched to the gas mask (left figure), where microphone is inserted in the mask interior (right figure)

IV. Brath Sound Characteristics

This section outlines key characteristics of breath sounds across three domains: time, spectral, and spectrogram. Initially, all recorded sounds undergo preprocessing to extract typical breathing patterns, including several inhalations and exhalations. Subsequently, these extracted patterns are visualized within the mentioned domains and analyzed to identify significant characteristics associated with specific breathing organs (nose or mouth), breathing phases (inhalation or exhalation), subjects, recording environments, and acoustic sensors.

Illustrations of typical patterns of breathing sounds in the time (waveform), spectral, and spectrogram domains are provided in Figs. 3, 4, and 5, respectively. These sounds were recorded in air using the studio microphone. In Fig. 3, the waveforms of inhalation and exhalation demonstrate distinct properties. Among them, the most notable difference between inhalation and exhalation lies in the amplitude values, with significantly larger values observed during the exhalation phase. Additionally, the duration and shape of the rise and decay portions typically differ between these two breathing phases. For instance, in several cases, the decay is prolonged during exhalation.

The spectra of individual inhalations and exhalations of the sound presented in Fig. 3 are depicted in Fig. 4. Notably, there is a clear mutual similarity between the spectra of two inhalations, as well as between two exhalations. When comparing the spectra of inhalation and exhalation, distinct differences emerge, particularly at lower frequencies, up to 500 Hz. In this frequency range, the shapes of the spectral curves differ, with significantly higher levels observed in exhalation spectra compared to inhalation spectra. However, the disparity between inhalation and exhalation diminishes notably at higher frequencies, especially above 1 kHz. It is noteworthy that characteristic peaks and dips are present in both inhalation and exhalation spectra. Moreover, in contrast to the trend observed at lower frequencies, the levels found in exhalation spectra are somewhat smaller than those in inhalation spectra at higher frequencies.

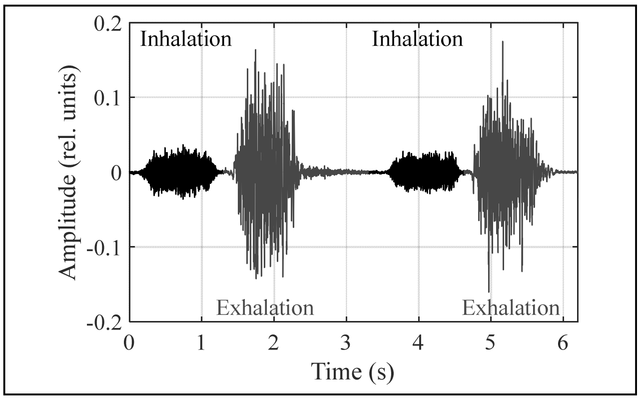

Fig. 3. Waveforms of inhalation and exhalation in breathing through mouth recorded in air by means of studio microphone

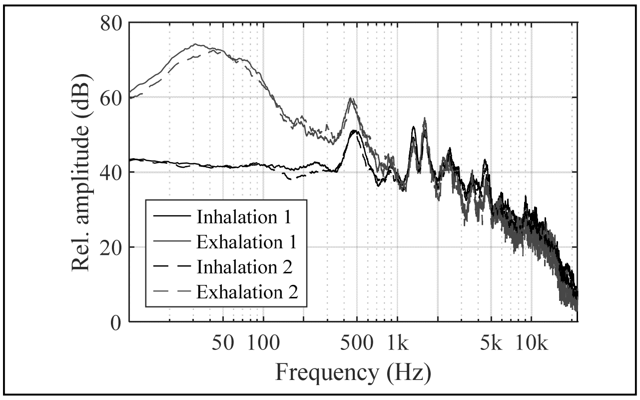

Fig. 4. Spectra of individual inhalations and exhalations in breathing through mouth recorded in air by means of studio microphone

The spectrograms as 3D plots, depicted in Fig. 5, illustrate the identified characteristics in both the time and frequency domains. Furthermore, specific features such as changes in the level of particular frequency components along the time axis can be observed within these spectrograms.

The comparison of spectra from a single inhalation and exhalation of a particular subject, extracted from sounds recorded in both environments (air and gas mask) using various microphones, is shown in Fig. 6. Notably, there are some consistent global trends evident separately for inhalation and exhalation across all breathing sounds recorded by different microphones in air. In this environment, the levels remain relatively constant at frequencies up to 1 kHz for inhalation, with a subsequent decrease in levels observed above this frequency. Somewhat different spectral shapes are observed in the breathing sounds recorded by the smartphone microphone, characterized by a visible roll-off from several hundred hertz towards lower frequencies. Another trend is characterized by significantly higher levels in exhalation below 500 Hz, or even 1 kHz, compared to inhalation during breathing in air. When comparing spectra recorded in air and through the mask, the most notable difference lies in the higher levels found in the mask, particularly during inhalation. This increase in levels is also observable during exhalation, albeit primarily at the lowest frequencies, mainly below 50 Hz.

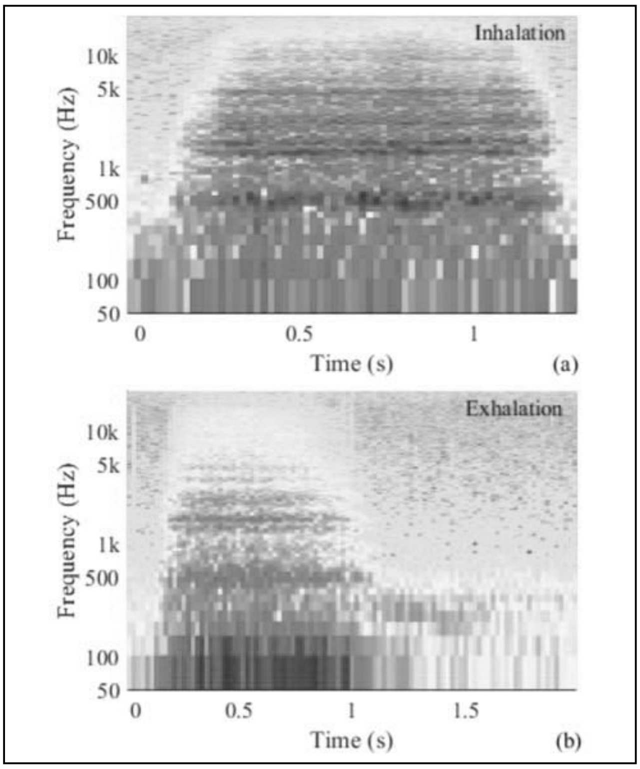

Fig. 5. Spectrogram of a single a) inhalation and b) exhalation in breathing through mouth recorded in air by means of studio microphone

The effects of both the acoustic sensor (microphone) and breathing phase (inhalation and exhalation) are evident in Fig. 7. In this regard, there is an influence from the microphone's frequency characteristics, as well as its physical attributes such as size, membrane position, and housing. The previously mentioned roll-off for the smartphone microphone is also notable here in both breathing phases, see Fig. 7.d). When comparing spectra for inhalation and exhalation, regardless of the microphone used, exhalation exhibits higher levels at lower frequencies, while the trend reverses at higher frequencies. The frequency at which this trend changes is in the range 1-3 kHz.

Each individual (subject) also contributes to the spectral characteristics of the breathing sound. An illustration of such contributions is provided in Fig. 8, which presents spectra of a single inhalation and exhalation for three subjects recorded in the air. The variations in levels among subjects at particular frequencies are significant, ranging as high as several tens of decibels. The largest differences are observed at low frequencies, up to several hundred hertz, while the smallest variances appear at higher frequencies for nasal breathing, see Fig. 8.b).

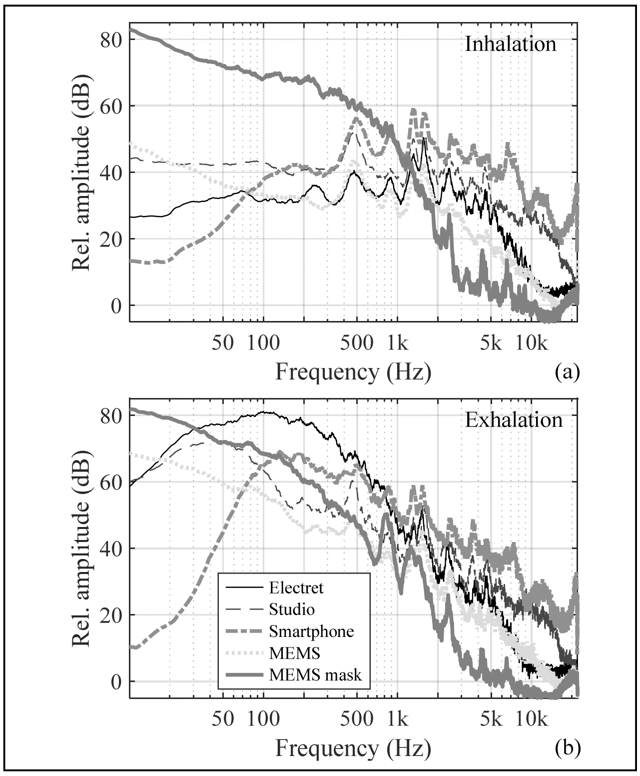

Fig. 6. Spectra of a single a) inhalation and b) exhalation through mouth recorded in air (by different microphones) and mask by MEMS microphone

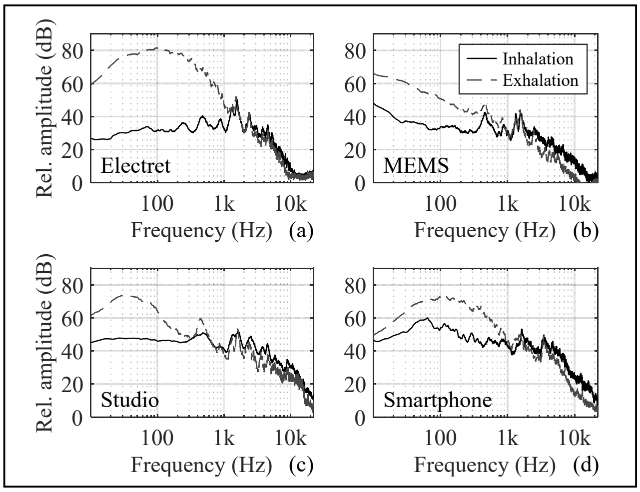

Fig. 7. Spectra of a single inhalation and exhalation through mouth of a single subject recorded in air by different microphones

All previous results for breathing in the air are derived from forced breathing, as its characteristics are more pronounced. Fig. 9 illustrates the distinctions between forced and normal breathing. In this specific case, the spectra for mouth breathing are very similar, whereas the disparities between forced and normal breathing are more noticeable during nasal breathing. Furthermore, in the latter scenario, the discrepancies between inhalation and exhalation are smaller during normal breathing compared to forced breathing. The described trends are highly subject-dependent; in some subjects, the differences between forced and normal breathing may be even more pronounced.

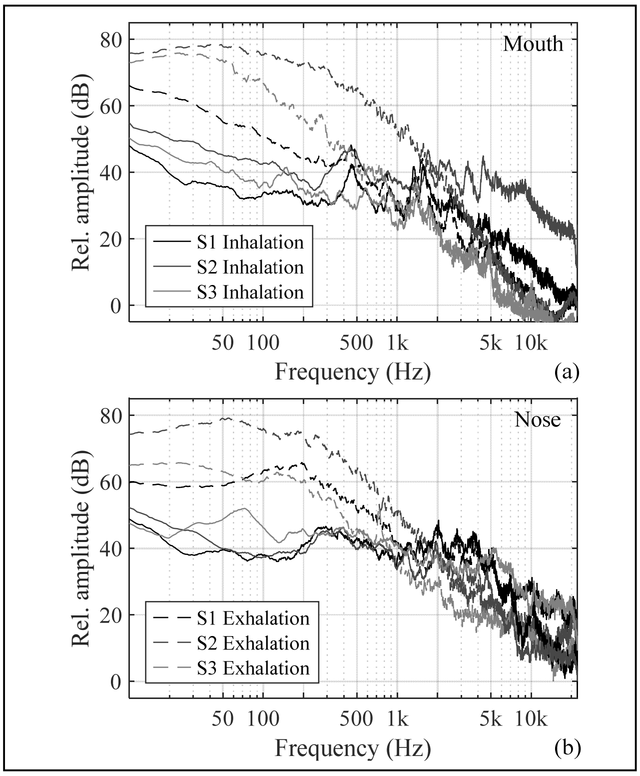

Fig. 8. Spectra of a single inhalation and exhalation through a) mouth and b) nose for three subjects (S1, S2 and S3) recorded in air by MEMS microphone

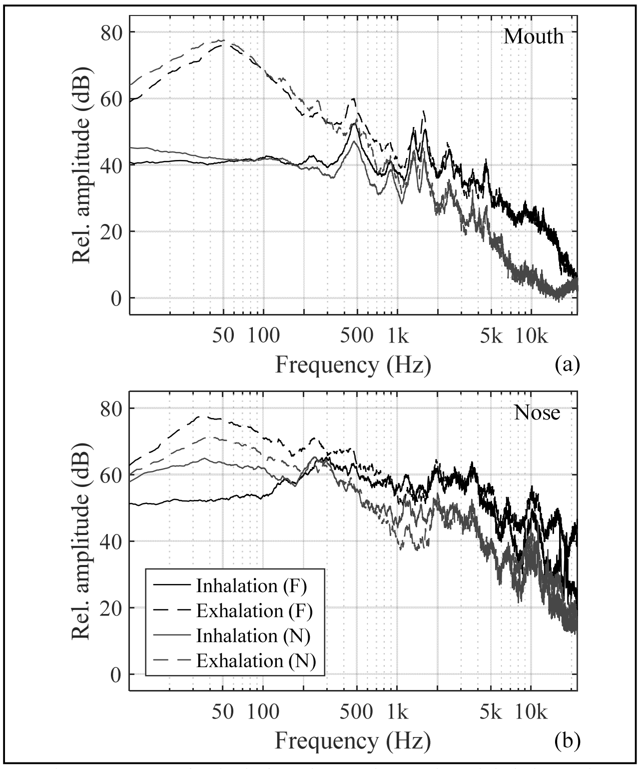

Fig. 9. Spectra of a single forced (F) and normal (N) inhalation and exhalation through a) mouth and b) nose for one subject (S1) recorded in air by studio microphone

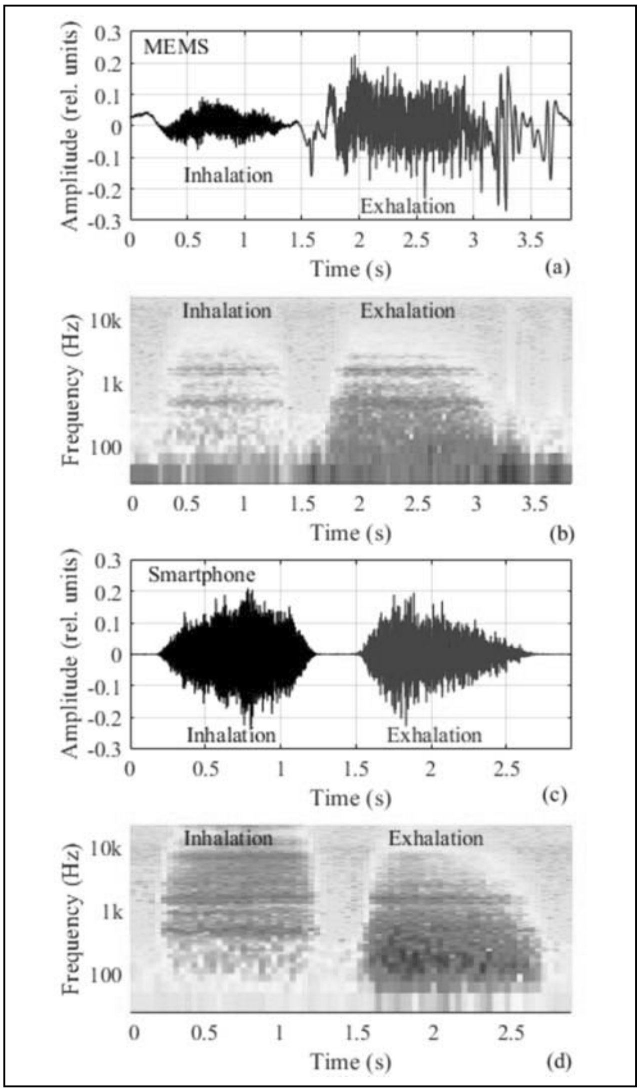

Fig. 10. Sound of forced breathing in air acquired by a) and b) MEMS microphone,and c) and d) smartphone microphone

Waveforms together with spectrograms of inhalation-exhalation pattern for a single subject recorded by two microphones are shown in Fig. 10. The presented waveforms for these two microphones are rather different considering elements sush as shape of inhalation-exhalation pattern, rise and decay. The shown spectrograms also differ, and some of the previously noticed trends are also visible – higher values for exhalation at lower frequencies and opposite situation at higher frequencies especially for the smartphone microphone, see Fig. 10.d).

V. Conclusions

This research analyzes various methods of acquiring breathing sounds, but also characterizing and visualizing distinctive properties of breathing patterns. Certain limitations of the current approach stem from a small sample size, which will be addressed in the near future by expanding the dataset of breath sounds. Employing multiple methods of breath acquisition, along with different breathing phases and modalities, has highlighted significant differences that environmental factors introduce into the process of breath acquisition and detection.

Significant differences in the characteristics of breathing phases (inhalation and exhalation) are identified across all three applied domains: time, frequency, and spectrogram. These differences encompass variations in breathing patterns in the time domain and the shapes of spectral characteristics in the frequency domain. Depending on the environment and acquisition sensor, breathing sounds exhibit specific features that enable clear distinctions between inhalation and exhalation, as well as between breathing through the mouth and nose. This study provides valuable insights into the utilization of different technologies and principles of breath acquisition and processing for further research. Despite the limited scope of the amassed data, it holds significant promise for informing the development of a systematic framework for automating the process of breath acquisition and its precise evaluation.

References

[1] A. Kumar, V. Mitra, C. Oliver, A. Ullal, M. Biddulph, and I. Mance, “Estimating respiratory rate from breath audio obtained through wearable microphones,” 43rd Annual International Conference of the IEEE Engineering in Medicine and Biology Society (virtual conference), 2021.

[2] K. Fukuyama, O. Sugiyama, K. Chin, S. S. Atou, S. Matsumoto, and M. M. Uto, “Identification of respiratory sounds collected from microphones embedded in mobile phones,” Advanced Biomedical Engineering, vol. 11, pp. 58–67, 2022.

[3] J. Castroa and P. Marti-Puigb, “Real-time identification of respiratory movements through a microphone,” Advances in Distributed Computing and Artificial Intelligence Journal, vol. 3, no. 3, pp. 1–11, 2014.

[4] C. Romano, A. Nicolò, L. Innocenti, M. Bravi, S. Miccinilli, S. Sterzi, M. Sacchetti, E. Schena, C. Massaroni, “Respiratory rate estimation during walking and running using breathing sounds recorded with a microphone,” Biosensors, vol. 13, no. 6, pp. 637–652, 2023.

[5] R. Y. Jin, M.M. Rahman, T. Ahmed, J. Kuang, A.J. Gao, “RRDetection: Respiration rate estimation using earbuds during physical activities,” 45Annual International Conference of the IEEE Engineering in Medicine & Biology Society, Sydney, Australia, 2023.

[6] S-H. Li, B-S. Lin, C-H. Tsai, C-T. Yang, and B-S. Lin, “Design of wearable breathing sound monitoring system for real-time wheeze detection,” Sensors, vol. 17, no. 1, pp. 171–175, 2017.

[7] G. Sierra, V. Telfort, B. Popov, M. Pelletier, P. Despault, R. Agarwal, and V. Lanzo, “Comparison of respiratory rate estimation based on tracheal sounds versus a capnograph,” IEEE Engineering in Medicine and Biology 27th Annual Conference, Shanghai, China, pp. 6145-6148, 2005.

[8] Y. Ren, C. Wang, J. Yang, and Y. Chen, "Fine-grained sleep monitoring: Hearing your breathing with smartphones," IEEE Conference on Computer Communications, Hong Kong, China, pp. 1194-1202, 2015.

[9] C-H. (Iris) Shih, N. Tomita, Y. X. Lukic, Á. H. Reguera, E. Fleisch, T. Kowatsch, “Breeze: Smartphone-based acoustic real-time detection of breathing phases for a gamified biofeedback breathing training,” Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies, vol. 31, no. 4, article no. 152, pp. 1–30, 2019.

[10] B. R. Snider, A. Kain, “Automatic classification of breathing sounds during sleep,” IEEE International Conference on Acoustics, Speech and Signal Processing, Vancouver, BC, Canada, pp. 699-703, 2013.

[11] R. Abbasi-Kesbi, A. Valipour, K. Imani, “Cardiorespiratory system monitoring using a developed acoustic sensor,” Healthcare Technology Letters, vol. 5, no. 1, pp. 7–12, 2018.

[12] C. Lim, J. Kim, J. Kim, B.-G. Kang, Y. Nam, “Breeze: Estimation of respiratory rate in various environments using microphones embedded in face masks,” Journal of Supercomputing, vol. 78, pp. 19228–19245, 2022.

[13] C. Doukas, T. Petsatodis, C. Boukis, and I. Maglogiannis, “Automated sleep breath disorders detection utilizing patient sound analysis,” Biomedical Signal Processing and Control, vol. 7, pp. 256–264, 2012.

[14] D. Skalicky, V. Koucky, D. Hadraba, M. Viteznik, M. Dub, F. Lopot, “Detection of respiratory phases in a breath sound and their subsequent utilization in a diagnosis,” Applied Sciences, vol. 11, pp. 6535–6546, 2021.