



HAL
open science

Respiratory Monitoring Based on Tracheal Sounds: Continuous Time-Frequency Processing of the Phonospirogram Combined with Phonocardiogram-Derived Respiration

Xinyue Lu, Christine Azevedo Coste, Marie-Cécile Nierat, Serge Renaux,
Thomas Similowski, David Guiraud

► **To cite this version:**

Xinyue Lu, Christine Azevedo Coste, Marie-Cécile Nierat, Serge Renaux, Thomas Similowski, et al.. Respiratory Monitoring Based on Tracheal Sounds: Continuous Time-Frequency Processing of the Phonospirogram Combined with Phonocardiogram-Derived Respiration. *Sensors*, 2021, 21 (1), pp.99. 10.3390/s21010099 . hal-03093720

HAL Id: hal-03093720

<https://inria.hal.science/hal-03093720v1>

Submitted on 27 May 2021

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.



Distributed under a Creative Commons Attribution 4.0 International License

Article

Respiratory Monitoring Based on Tracheal Sounds: Continuous Time-Frequency Processing of the Phonospirogram Combined with Phonocardiogram-Derived Respiration

Xinyue Lu ^{1,2}, Christine Azevedo Coste ^{3,*}, Marie-Cécile Nierat ⁴, Serge Renaux ^{2,5}, Thomas Similowski ^{4,6} and David Guiraud ^{3,5}

¹ Faculté des Sciences, University of Montpellier, F-34090 Montpellier, France; sindayue@gmail.com

² NeuroResp, F-34600 Les Aires, France; serge.renaux@neuroresp.com

³ INRIA, F-34090 Montpellier, France; david.guiraud@inria.fr

⁴ UMRS1158 Neurophysiologie Respiratoire Expérimentale et Clinique, INSERM, Sorbonne Université, F-75005 Paris, France; m-cecile.nierat@wanadoo.fr (M.-C.N.); thomas.similowski@upmc.fr (T.S.)

⁵ NEURINNOV, F-34090 Montpellier, France

⁶ AP-HP, Site Pitié-Salpêtrière, Service de Pneumologie, Médecine Intensive et Réanimation (Département R3S), Groupe Hospitalier Universitaire APHP-Sorbonne Université, F-75013 Paris, France

* Correspondence: christine.azevedo@inria.fr

Abstract: Patients with central respiratory paralysis can benefit from diaphragm pacing to restore respiratory function. However, it would be important to develop a continuous respiratory monitoring method to alert on apnea occurrence, in order to improve the efficiency and safety of the pacing system. In this study, we present a preliminary validation of an acoustic apnea detection method on healthy subjects data. Thirteen healthy participants performed one session of two 2-min recordings, including a voluntary respiratory pause. The recordings were post-processed by combining temporal and frequency detection domains, and a new method was proposed—Phonocardiogram-Derived Respiration (PDR). The detection results were compared to synchronized pneumotachograph, electrocardiogram (ECG), and abdominal strap (plethysmograph) signals. The proposed method reached an apnea detection rate of 92.3%, with 99.36% specificity, 85.27% sensitivity, and 91.49% accuracy. PDR method showed a good correlation of 0.77 with ECG-Derived Respiration (EDR). The comparison of R-R intervals and S-S intervals also indicated a good correlation of 0.89. The performance of this respiratory detection algorithm meets the minimal requirements to make it usable in a real situation. Noises from the participant by speaking or from the environment had little influence on the detection result, as well as body position. The high correlation between PDR and EDR indicates the feasibility of monitoring respiration with PDR.

Keywords: respiratory monitoring; breathing detection; tracheal sounds; Phonocardiogram-Derived Respiration (PDR)



Citation: Lu, X.; Azevedo, C.; Nierat, M.-C.; Renaux, S.; Similowski, T.; Guiraud, D. Respiratory Monitoring Based on Tracheal Sounds: Continuous Time-Frequency Processing of the Phonospirogram Combined with Phonocardiogram-Derived Respiration. *Sensors* **2021**, *21*, 99. <https://dx.doi.org/10.3390/s21010099>

Received: 29 November 2020

Accepted: 21 December 2020

Published: 25 December 2020

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2020 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

Impairment of ventilatory control (central respiratory paralysis) may lead to a dependence on external mechanical ventilation (usually with intermittent positive pressure via tracheotomy), even though the main inspiratory effector, the diaphragm, is intact. These may be due to the interruption in the transmission of the control commands (essentially post-traumatic upper cervical spinal cord injury) or abnormalities in the genesis of the control (central, congenital, or acquired alveolar hypoventilation) [1]. An improved treatment, diaphragm pacing, can restore a diaphragmatic contraction and thus relieve patients of the constraint of a mechanical ventilator. Compared to external mechanical ventilation, this technique provides more natural respiration [2,3], improves speeches and olfaction, and helps to reduce infection risks and health care costs [4].

In the context of ventilatory assistance, detecting breathing interruptions resulting from airway obstruction or ventilator disconnection is a fundamental safety issue. All ventilatory assistance devices include built-in flow and/or pressure sensors connected to alarms to alert caregivers of any flow interruption. Implanted phrenic nerve electric stimulators for diaphragm pacing are notable exceptions to this rule. Indeed, they include built-in electronic dysfunction alarms but are devoid of “efficacy” alarms. During diaphragm pacing, an apnea (absence of flow for at least 10 s [5]) due to airway obstruction (obstructive apnea, with respiratory effort) or diaphragmatic neurotransmission failure (central apnea, without respiratory effort) will not result in an immediate alarm if the patient is not equipped with an additional flow sensor [6].

Detecting breathing interruptions is key to diagnosing both central and obstructive apneas. Yet, the usual methods (pneumotachograph, thermistor, etc.) to assess breathing activities and detect apnea episodes rely on flow sensors connected to the airway, which is burdensome and limits their tolerability (imply to wear a mask or a mouthpiece). Notably, monitoring thoracic movements (respiratory effort, with plethysmograph) does not provide an adequate alternative of flow detection because this approach cannot detect obstructive apnoeic events and is sensible to patient’s movements.

The aim of our research was to improve the efficiency and safety of diaphragm pacing systems. We propose to use microphones for monitoring patient’s respiration via tracheal sounds (TS). The approach is non-invasive (no supplementary complexity due to implantation surgery) and miniaturized so that patients can carry the device easily without wearing a mask. It has long been known that ventilatory flow can be accessed transcutaneously by recording TS. Indeed, the turbulent airflow going through the airway causes pressure vibrations that are transmitted to the body surface through the tracheal wall and the surrounding tissues. Tracheal sounds can thus be recorded simply using a skin-taped microphone. They provide a signal that is representative of ventilatory airflow and has been described as “phonospirometry” [7]. A microphone affixed over the suprasternal notch was shown to provide the greatest signal amplitude-to-noise ratio [8]. At this location, cardiac sounds can also be recorded (phonocardiogram).

The present study proposes and seeks to validate firstly signal processing algorithms on healthy subjects. The proposed method is a robust real-time monitoring of respiratory activity through the percutaneous recording of TS. The two main contributions of the methods are: (i) TS embeds temporal and frequency information that, mixed together, enhanced the identification of respiratory phases, and (ii) a new processing of cardiac sound from the TS recording can further provide relevant information about respiratory states.

1.1. Apnea Detection

Studies have shown the feasibility of detecting respiration from TS mainly in sleep apnea. The detection algorithms can be divided into two categories:

1.1.1. Detection in the Temporal Domain

Corbishley et al. [8] proposed an algorithm that was based on temporal envelope detection with several user-dependent thresholds and could be embedded in a miniature and low-powered sensing device. Some time later, Yadollahi et al. [9] used the logarithm of TS variance combined with an oximeter to estimate the relative respiratory flow. Recently, Kalkbrenner et al. [10] processed short-term and long-term envelopes of TS combined with measurement of patient movements by IMUs to evaluate the respiratory amplitude changes; however, the processing is offline.

1.1.2. Detection in the Frequency Domain

Chuah et al. [11] proposed the detection of respiratory events using the peak picking technique in a power spectrum limited to specific frequency bands for both chest sounds and TS. Hult et al. [12] improved their own algorithm from Reference [13] using threshold detection in real-time in the frequency envelope of the 300–600 Hz power spectrum.

A similar method was used by Kalkbrenner et al. [14], who classified respiratory phases using threshold detection in the 10–600 Hz power spectrum. Kulkas et al. [15] proposed detecting apnea events based on the ratio of the power spectrum density of heart beat (0–50 Hz) to that of heart beat and respiration (0–600 Hz).

Other methods based on artificial neural networks have been proposed, such as by Lei et al. [16]. A complete review of TS processing methods is presented in Muthusamy et al. [17].

As a whole, few methods have been developed for real-time monitoring and even fewer for an ambulatory context. Indeed, most of the studies have been carried out in quiet and controlled acoustic environments with stable sources of noise, and limited participant movements, mainly during sleep. Furthermore, the studies did not incorporate multi-domain detection.

1.2. Phonocardiogram (PCG)-Derived Respiration

Electrocardiogram (ECG) corresponds to the electrical potential action of the cardiac muscle. Three principal waves appear together, called the QRS-complex, and are linked to the depolarization of the two ventricles.

During respiration, chest movements change the distances between the ECG electrodes and therefore affect the morphology of QRS-complexes. Based on these changes, respiration can be measured indirectly from the ECG signals: ECG-Derived Respiration (EDR).

M. Schmidt et al. [18] summarized several EDR methods. One widely used method is based on R-wave amplitude (R peaks) change analysis: R-peak amplitude follows the respiration cycles and can be used to detect respiration phases, including apnea.

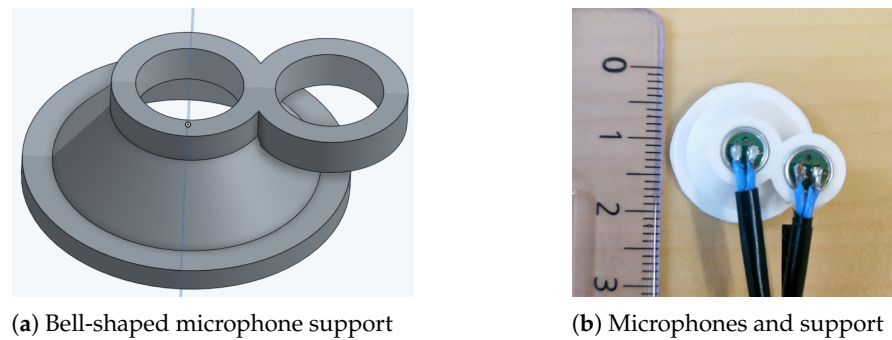
To further enhance the classification results while using only TS, we hypothesized that combining the time and frequency domain detections would provide more accurate results. We also made the assumption that cardiac sounds embed the same kind of information as classical EDR. Indeed, the phonocardiogram (PCG) is caused by the closure of the heart valves. There are two fundamental heart sounds, called S1 and S2. The period from S1 to S2 is systole, while the period from S2 to S1 is diastole. We call this signal, which has only been used for heart rate estimation [10,19], PCG-derived respiration (PDR). We show that detection can be improved by combining the detection results from the temporal and frequency detection domains and PDR. The remainder of this paper, thus, presents the combined algorithm and the new method to extract respiration information from cardiac sounds. We validated the algorithm through a trial and discuss the results.

2. Material and Method

2.1. Material

Two uni-directional microphones (Pro-signal, ABM-708-RC) were used to record sounds: one for TS and one for environmental noise. The recorded environmental noise was intended to be applied to further reduce noise in the recorded TS, but it was not used in the present study (details in discussion). The microphones were inserted into a bell-shaped support (Figure 1a, design inspired by Reference [8]). The two microphones were located next to each other as illustrated in Figure 1b. During manipulation, only the support was fixed on the neck at the suprasternal notch with two strips of adhesive tape (as shown in Figure 2).

Each microphone was connected to an inhouse-designed analog card, which filters and amplifies recorded sounds from 100 to 1200 Hz with a second-order Band-Pass Butterworth filter and an amplification of 230 times. The filtering narrows the useful band to that known to contain the respiratory and cardiac sounds [8]. Last, the pre-processed sounds were sampled at 10 kHz with a PowerLab 16/35 (ADInstruments, Dunedin, New Zealand).



(a) Bell-shaped microphone support

(b) Microphones and support

Figure 1. Illustration of the microphone setup.**Figure 2.** Position of the microphones: two microphones are inserted into the support that is fixed at suprasternal notch with two strips of adhesive tape.

In order to evaluate the detection results of the proposed method, a low-resistance pneumotachograph (Hans Rudolf, 3700 series, linearity range 0–160 L/min; Kansas City, MO, USA), an abdominal strap (plethysmograph, AD Instrument/Pneumotrace II, UFI, Morro Bay, CA, USA), and three Ag/AgCl ECG electrodes (amplified with a Bio Amp: FE132, AD Instruments, Castle Hill, Australia) were also synchronously acquired with PowerLab. The algorithm was developed with MATLAB (Natick, MA, USA).

2.2. Protocol

Thirteen healthy participants were enrolled (5 women and 8 men; age: 36 ± 10.37 years, weight: 66 ± 12.27 kg, height: 172 ± 7.84 cm). The study was conducted according to the principles of the Declaration of Helsinki and approved by the French Comité de Protection des Personnes Sud-Ouest & Outremer II, decision #2-18-12-MS1, 5 October 2018. The participants received detailed information about the study objective and methods and provided written informed consent.

Each underwent two recordings of 2 min each: one in sitting position and the other in lying position. The 2-min sessions consisted of four parts: (1) normal respiration for 60 s, (2) apnea (spontaneous breath holding) for the next 10 s, (3) normal respiration while speaking for the next 20 s, and (4) normal respiration with strong environmental noise (a video with city traffic noise played right next to the microphones [20]) to the end. An example of session recording is shown in Figure 3, and all the examples shown in the next sections were extracted from this session.

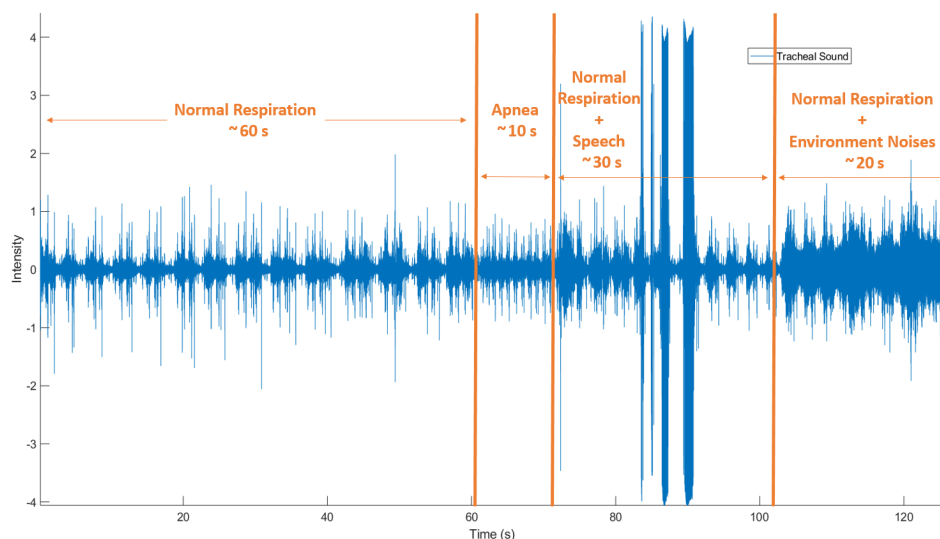


Figure 3. Recording procedure: 0–60 s normal respiration (NR); 60–70 s apnea (holding breathing spontaneously, AP); 70–100 s normal respiration + speech (RS); 100–120 s normal respiration + environmental noise (RN).

2.3. Algorithm

In our previous studies, we developed and showed the feasibility of one first respiration detection algorithm tested on healthy subjects [21] and on one patient under implanted phrenic nerve stimulation [22]. This first algorithm was based on patient-specific thresholds and tested with short recordings. On the contrary, the present work used non-patient-specific thresholds and also included the new PDR.

Recorded tracheal sounds ($ts(n)$) were processed in the temporal and frequency domains and also by the PDR method. The recordings were split into bins of $l_{seg} = 16,384$ sample (1.6384 s) duration. This is convenient for the Fast Fourier Transform (FFT) algorithm and enables the recording of at least one cardiac sound within each interval, assuming a heart-rate of minimum 38 beats per second (bps). The global algorithm is shown in Figure 4.

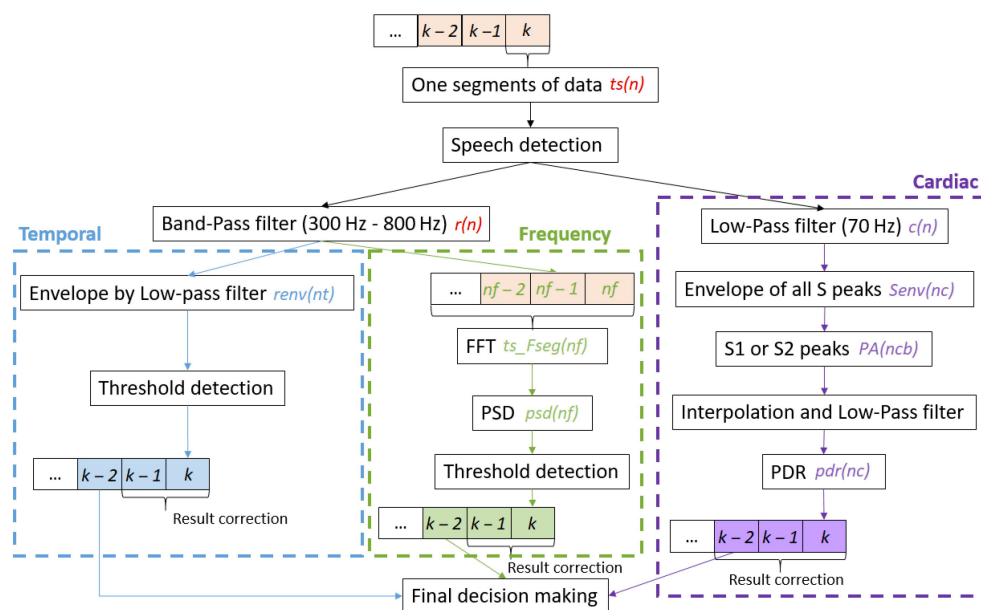


Figure 4. Diagram of the detection algorithm.

2.3.1. Speech Detection

Speech or snoring can be considered as the expirations with the strongest intensity. These sounds have a much higher intensity, and a simple detection based on a threshold on the Mean Absolute Value (MAV) was performed:

$$MAV(k) = \frac{\sum_{n=k \times l_{seg} + 1}^{(k+1) \times l_{seg}} |s(n)|}{l_{seg}},$$

where k is the index of the current bin. If this $MAV(k)$ exceeds 0.5, it indicates that speech/snoring has occurred during this bin.

2.3.2. Frequency Band Separation

Respiratory sounds are usually centered between 20 Hz and 2000 Hz, but this may differ depending on the acquisition system [23]. For our recording system, the respiratory frequency content was mostly centered between 100 Hz and 1500 Hz (Figure 5). Furthermore, assuming environmental noise at low and high frequencies just above 1200 Hz, we designed a steep band pass filter with an eighth-order Butterworth from 300 Hz to 800 Hz to extract the respiratory signal ($r(n)$). To extract the cardiac sound signal, we designed a steep low-pass filter that extracted its most significant part, an eighth-order Butterworth low-pass filter at 70 Hz to get the phonocardiogram ($c(n)$).

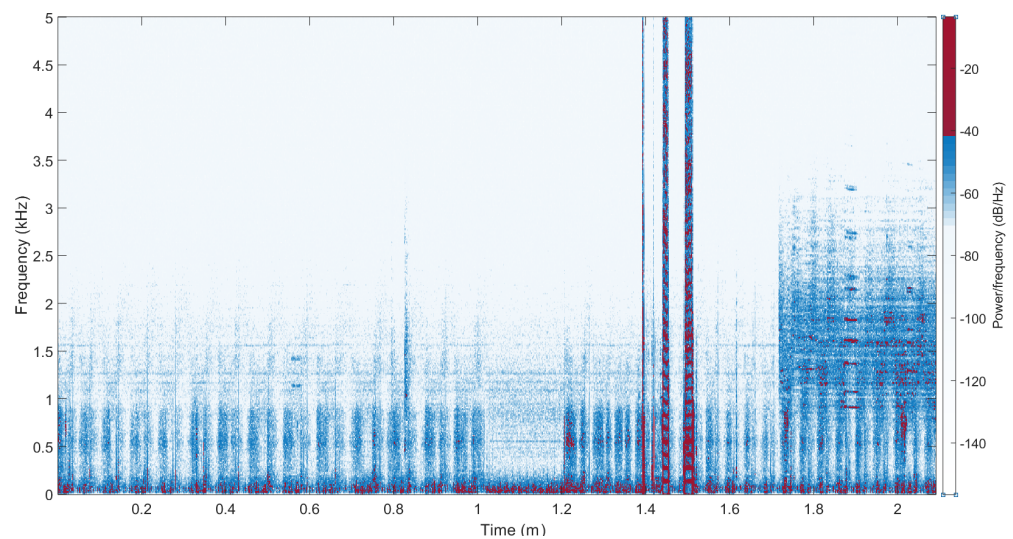


Figure 5. The spectrogram for one recording.

2.3.3. Temporal Envelope Detection

The temporal envelope ($renv(n)$) of the respiratory signal was obtained by applying a second-order low-pass Butterworth filter at 0.8 Hz on the rectified signal $|r(n)|$. Then, two thresholds were applied on $renv(n)$:

1. The minimum threshold:

$$th_{T_min} = 110\% \times renv(3_s_apnea),$$

where the $renv(3_s_apnea)$ is $renv(n)$ within any 3 s of apnea in the same recording;

2. The adaptive threshold:

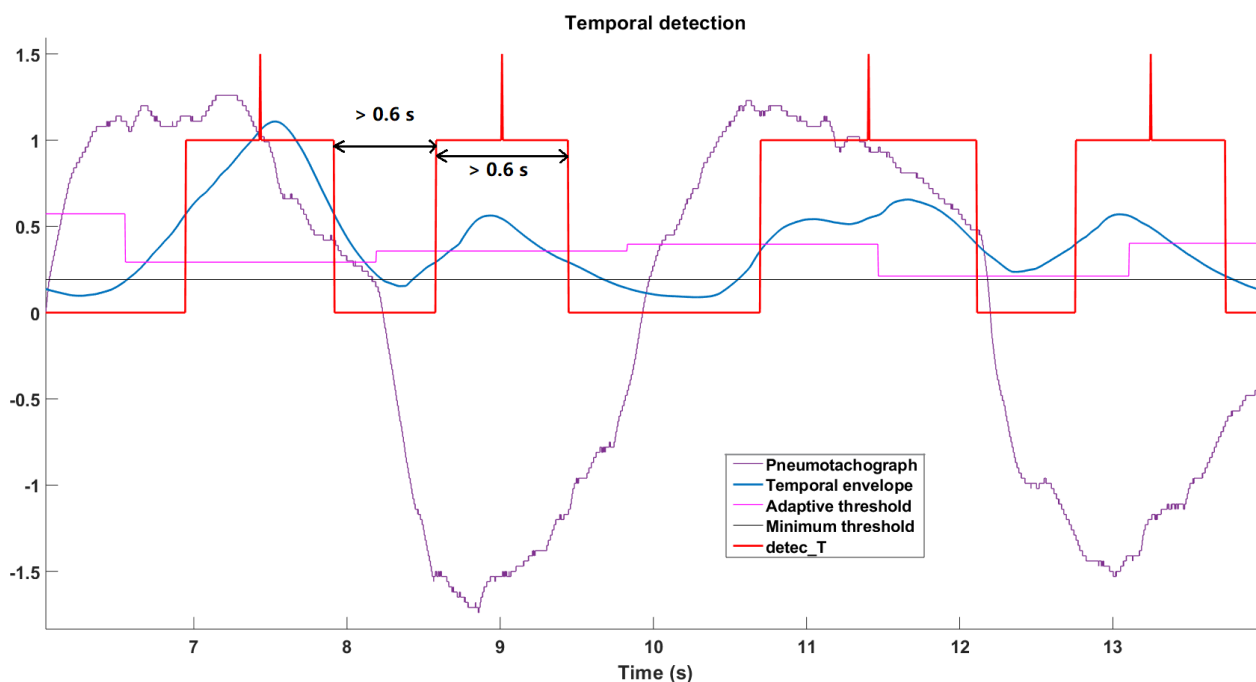
$$th_{T_adapt}(k) = 90\% \times \frac{\sum_{n=k \times l_{seg} + 1}^{(k+1) \times l_{seg}} renv(n)}{l_{seg}}.$$

The result of the temporal envelope detection ($detect_T(n)$) was fixed to three states, presented in Table 1.

Table 1. Temporal detection states.

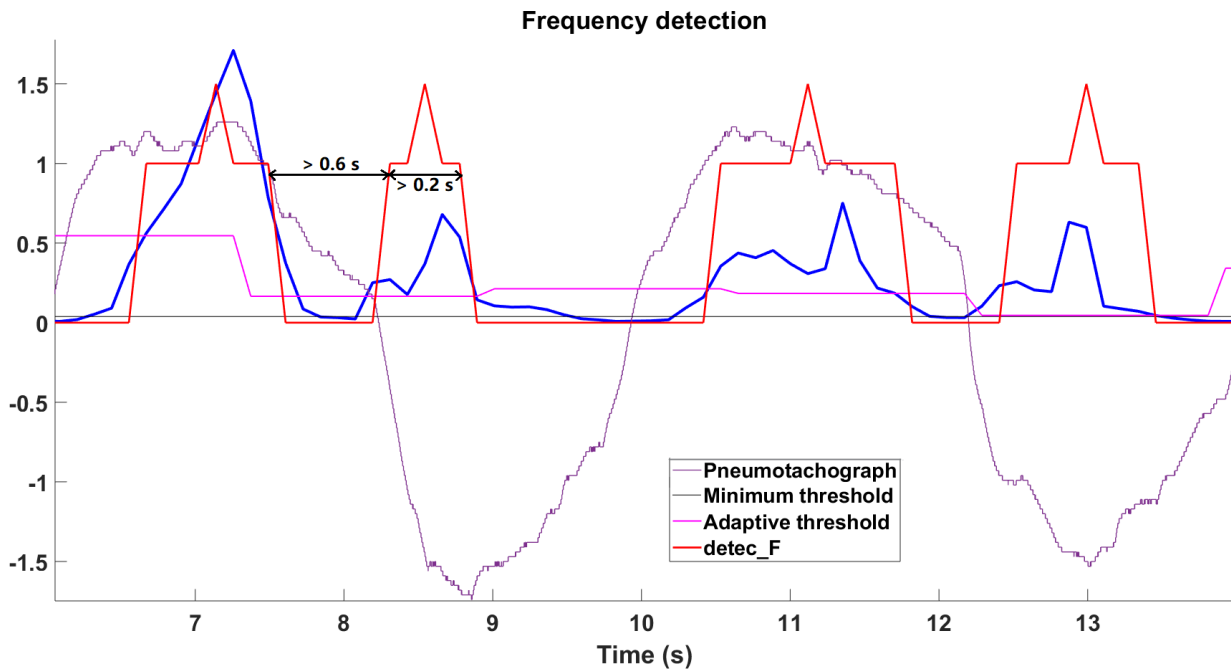
State of $detect_T(n)$	Explication	Conditions
0	Pause/apnea	$th_T_adapt(k) \leq th_T_min$
1	Inspiration/Expiration	$renv(n) > th_T_adapt(k)$
2	Speech/Snoring	$MAV(k) > 0.5$

If $detect_T = 1$ was obtained for bin k , an error correction method was applied on bins $k - 1$ and k : a detected respiratory event was eliminated if it lasted less than 0.6 s because its duration was not in the normal range for an inspiration/expiration; $detect_T(n)$ was forced to “0” if the interval from the previous detected event was less than 0.6 s. The middle-timing moment of a detected event was labeled ($detect_T(n) = 1.5$) to indicate the position of this event to serve for the final decision-making. A temporal detection example is shown in Figure 6a, where the pneumotachograph signal (in purple) is considered the reference, a positive signal indicates an inspiration, and a negative one indicates an expiration. We also plotted the temporal envelope (in blue), the adaptive threshold (in pink), and the minimum threshold (in gray). The final temporal detection result is in red with the peak “middle-timing moment”. The minimum interval and event duration are both 0.6 s.



(a) Temporal envelope detection. The minimum interval and event duration are both 0.6 s.

Figure 6. Cont.



(b) Frequency detection. The minimum interval is 0.6 s, and the minimum event duration is 0.2 s.

Figure 6. Temporal and frequency detection result of the same period from the same recording. A positive (resp. negative) pneumotachograph reference signal (purple) indicates inspiration (resp. expiration). The adaptive (pink) and minimum thresholds (gray) are marked in the figure. The envelope signal (temporal or PSD) is in blue, and the final detection result is in red with the peak “middle-timing moment”.

2.3.4. Frequency Detection

In the frequency domain, one bin k of tracheal sounds $ts(n)$ was divided into segments of 2048 samples (15 segments per bin). An FFT was processed on these segments with a Hanning window and 50% overlap ($TS_Fseg(nf)$). According to our recordings and frequency analysis, the useful frequency content was indeed between 400 Hz and 700 Hz: the PSD of this band was calculated ($psd(nf)$) as the frequency envelope. The detection part was similar to that described for the temporal envelope detection. Two thresholds were applied on $psd(nf)$:

1. The minimum threshold:

$$th_F_min = 110\% \times psd(3_s_apnea),$$

where the $psd(3_s_apnea)$ is $psd(nf)$ within any 3 s of apnea;

2. The adaptive threshold:

$$th_F_adapt(k) = 90\% \times \frac{\sum_{nf=k \times nb_seg+1}^{(k+1) \times nb_seg} psd(nf)}{nb_seg},$$

where nb_seg is the number of segments per bin, which equals 15.

Like $detec_T(n)$, the frequency detection signal $detec_F(nf)$ also had three states, as presented in Table 2, and the middle-timing moment of detected events was labeled as $detec_F(nf) = 1.5$. For the error correction, the interval of each respiratory event should be longer than 0.6 s, and the minimum duration of each respiratory event is 0.2 s (illustrated in Figure 6b).

Table 2. Frequency detection stats.

State of $detect_F(nf)$	Explication	Conditions
0	Pause/apnea	$th_F_adapt(k) \leq th_F_min$
1	Inspiration/Expiration	$psd(nf) > th_F_adapt(k)$
2	Speech/Snoring	$MAV(k) > 0.5$

2.3.5. PDR Detection

For the extracted cardiac signal, the MAV of the cardiac sound was computed every 0.06-s window (600 samples) with 0.03 s (50%) overlapping to obtain an envelope of S1 and S2 ($Senv(nc)$) [24].

Inspired by the method of Kalkbrenner et al. [10], the detection of S1 or S2 was more robust in that the time intervals between these events could be bounded, as illustrated in Figure 7. Time boundaries were defined as follows: where pa_tmin and pa_tmax are borders of this area in the time axis, based on the previous beat-to-beat interval marked with a black line (in Figure 7) and defined with the following conditions:

$$pa_tmin(ncb) = 80\% \times b_b(ncb - 1),$$

$$pa_tmax(ncb) = 120\% \times b_b(ncb - 1).$$

$b_b(ncb - 1)$ is the previous beat-to-beat time interval. Moreover, the lower boundary is limited to 0.4 s (150 beats per minute (bpm)) and the higher to 1.2 s (50 bpm). The local maxima in the predicted area $PA(ncb)$ were considered the cardiac peak if its amplitude was higher than the adaptive threshold:

$$pa_adap = 2 \times mean(Senv(nc)).$$

At the same time, the pa_adap should be greater than the minimum intensity threshold presented with a black dotted line in Figure 7 and defined as:

$$pa_min = 30\% \times mean(Senv(first_10_seconds)),$$

where $Senv(first_10_seconds)$ is the cardiac envelopogram for the first 10 s of recording. If there was no matching peak in this area, $PA(ncb)$ was moved forward two samples to verify another local maxima.

A first-order interpolation between each detected peak (green dotted line in Figure 7) was applied, and the signal was upsampled at the same frequency of $Senv(nc)$. Then the signal was filtered at 0.25 Hz by a second-order low-pass Butterworth filter. The filtered signal represents the upper envelope of S1 (or S2). Our hypothesis was that its variation would be due to respiratory effort (changes in thoracic volume): an increase in the signal would indicate an inspiration, while a decrease indicates an expiration. To reflect the correlation between the reference signals and the variation in the envelopogram, we further assumed that its derivative was representative of the respiration phases as it was linked to the variation in the S1/S2 peak envelopogram. We called this new curve PDR: phonocardiogram-derived respiration ($pdr(nc)$), indicating an inspiration when it is positive and an expiration when it is negative.

Unlike the two previous detection domains, the PDR signal ($pdr(nc)$) was able to distinguish inspiration and expiration. Thus, the cardiac (PDR) detection signal $detec_C(nc)$ had four states as presented in Table 3, and the middle-timing moment of detected events were labeled as $detec_C(nc) = 1.5$. If there were continuous inspirations/inspirations, they were merged because usually inspiration and expiration alternate. The respiratory effort signal showed few respiratory pauses (about 0.2 s) or long duration respiratory events, although some detected event durations could last more than 3 s. Thus, the error correction was applied on three bins: k , $k - 1$, and $k - 2$. This made a processing delay of about 3.2 s

(2 bins), which is acceptable according to the definition of apnea (10 s of absence of air flow). An example of cardiac detection is shown in Figure 8. The minimum respiratory event interval becomes 0.1 s, and minimum respiratory event duration fixes as 0.7 s.

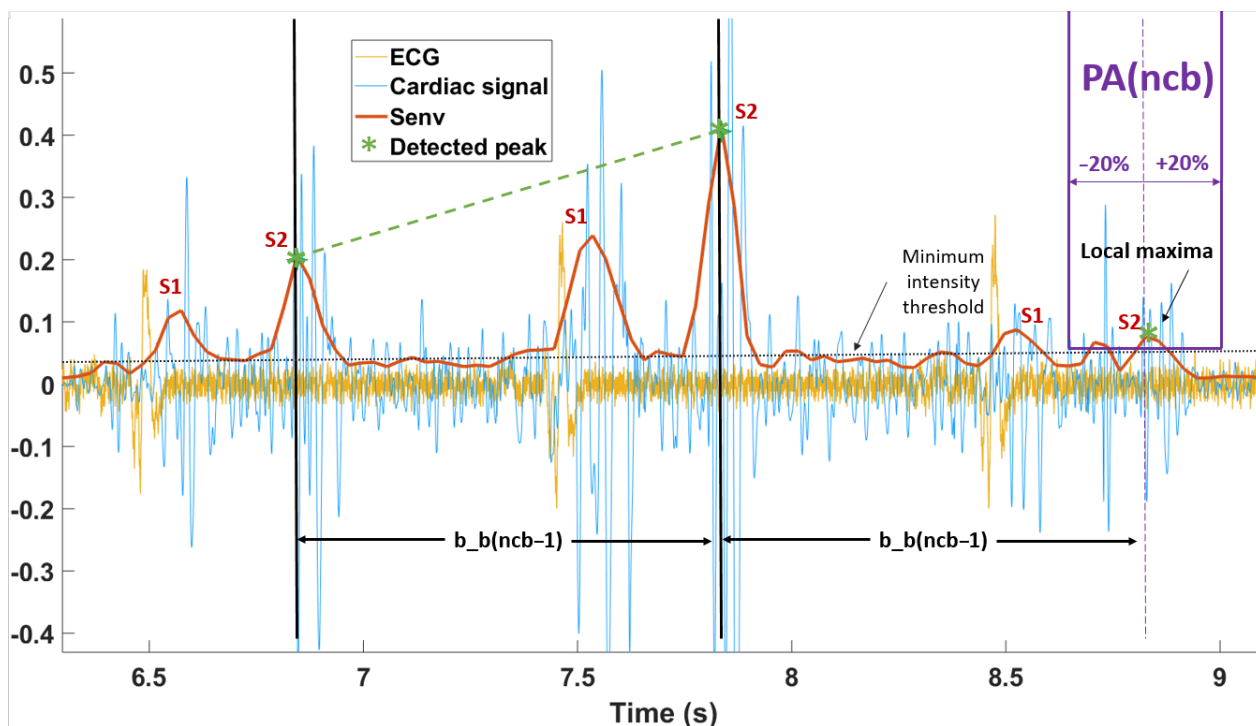


Figure 7. Cardiac S peak detection. There is a small delay between the electrocardiogram (ECG) (yellow) and the extracted cardiac signal (blue). Predicting area $PA(ncb)$ (purple) defined by $\pm 20\%$ of last heart beat-to-beat interval $b_b(ncb - 1)$ and the adaptive threshold. The cardiac envelopogram is in red, and the detected cardiac peak (S1/S2) is marked by a green star.

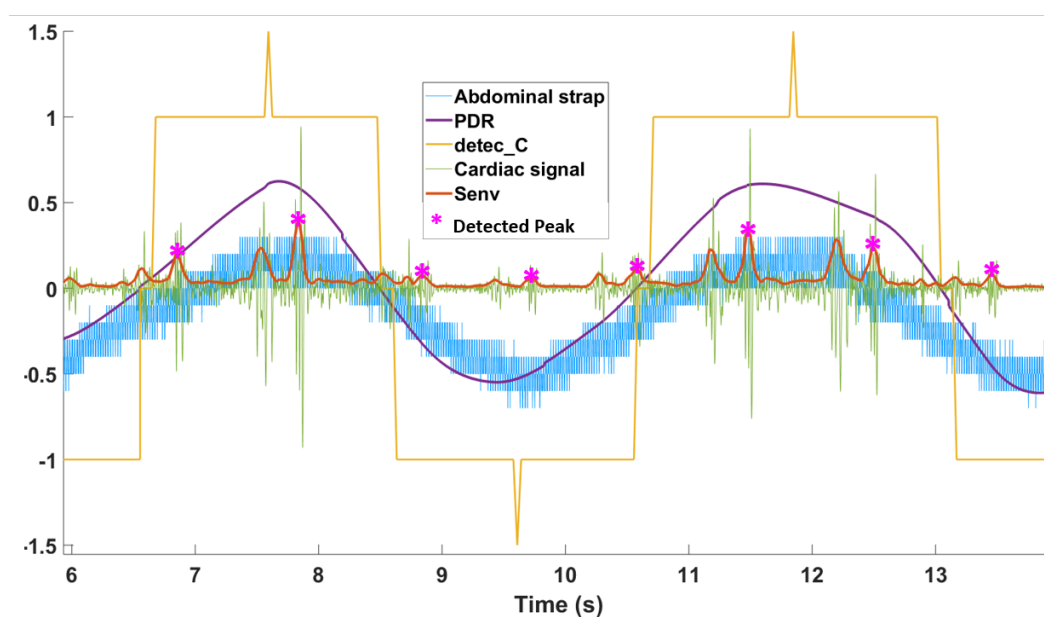


Figure 8. PDR. The extracted cardiac signal is shown in green, and the processed envelopogram is in red. Only one cardiac peak (S1/S2) detected (pink star) for each cardiac cycle. Positive (Negative) of the PDR detection result (yellow) indicates the inspiration (expiration) as reference signal.

Table 3. Cardiac detection (phonocardiogram (PCG)-derived respiration (PDR)) stats.

State of $detect_C(n)$	Explication	Conditions
−1	expiration	$pdr(nc) < 0$
1	Inspiration	$pdr(nc) > 0$
0	Pause/apnea	$eventduration < 0.7\text{ s}$ or $intervall < 0.1\text{ s}$
2	Speech/Snoring	$MAV(k) > 0.5$

2.3.6. Decision-Making

The temporal and frequency envelope detected the presence of airflow (apnea or not), whereas PDR also assesses respiratory efforts to distinguish between central and obstructive apnea. The final detection was the fusion of the three detection domains. To detect apnea alone or the complete set of respiration phases, two decision-making methods are defined:

Apnea Detection

The parameter Q is used to evaluate the respiratory status and was calculated by the weighted average as follows:

$$Q = 1.5 \times \frac{\sum_{nt=1}^{NT} detect_T(nt)}{NT} + 1.5 \times \frac{\sum_{nf=1}^{NF} detect_F(nf)}{NF} + 0.5 \times \frac{\sum_{nc=1}^{NC} |detect_C(nc)|}{NC}$$

For the bin $k - 2$, the final decision ($detec(k - 2)$) is made as:

$$detec(k - 2) = \begin{cases} 1 \rightarrow \text{respiration if } Q \geq 1.5 \\ 0 \rightarrow \text{pauses if } Q < 1.5 \\ 2 \text{ if speech/snoring} \\ -1 \rightarrow \text{apnea if pauses} > 10\text{s} \end{cases}$$

An example is shown in Figure 9. In theory, the PDR is used to evaluate respiratory effort. However, when the signal-to-noise ratio is low because of weak respiratory sounds, PDR detection may have a better result than those obtained in the temporal and frequency domains. Nevertheless, most of the time, temporal and frequency domain detections were more sensitive, so the PDR detection results are less weighted.

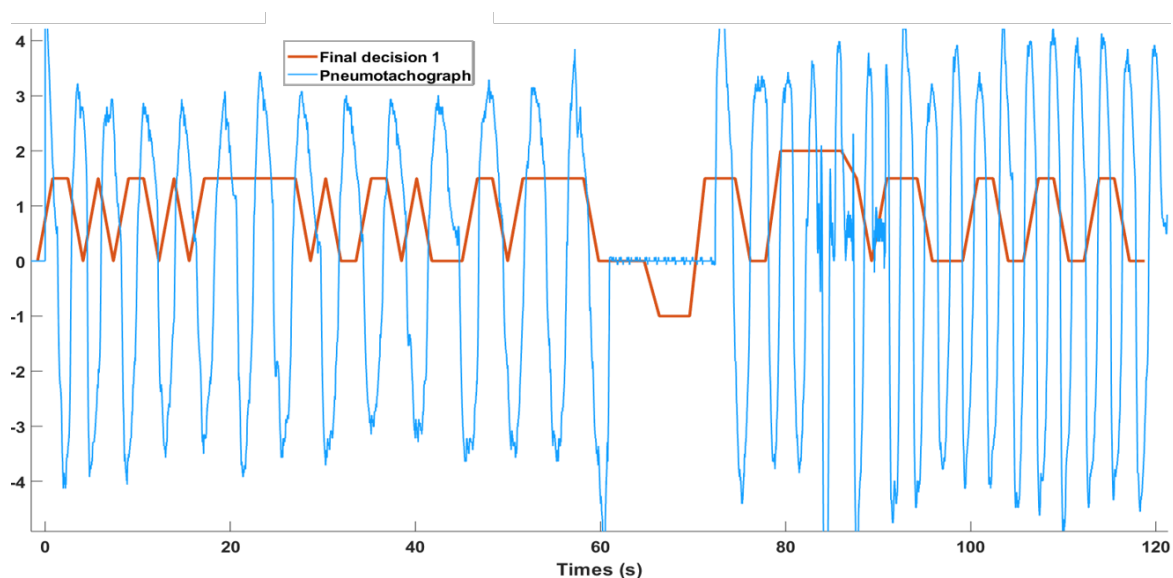


Figure 9. Final decision-making for apnea detection. Detection stat: 2—Speech; 1.5—Respiration; 0—Pauses; −1—apnea. The reference signal: pneumotachograph is in blue, and the apnea detection signal is in red.

Breathing Activities Assessment

The second algorithm detects the presence of each respiratory event. A respiratory event was considered to be detected if it was detected in at least two of the three detection domains, and the intervals between them were less than 1 s estimated with their middle-timing peak distances. An example is shown in Figure 10.

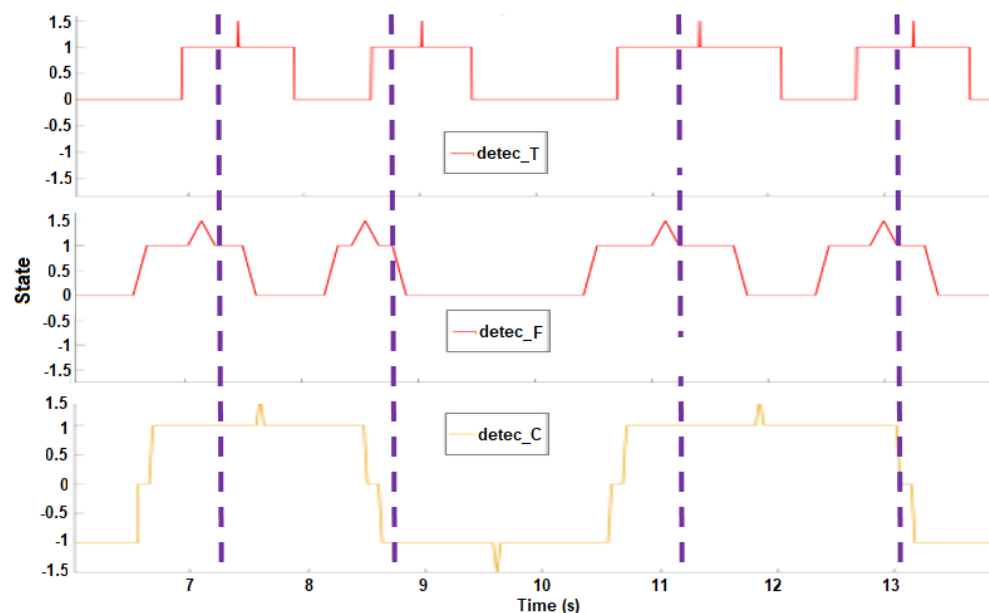


Figure 10. Final decision-making for breathing activities assessment. (**Top**): results from temporal detection. (**Middle**): results from frequency detection. (**Bottom**): results from cardiac (PDR) detection. The final detected events are marked in purple dashed lines.

Statistics

Several means comparing statistical tests were applied to show the influence of speech, noise, and participant position. The compared data were: (1) specificity, sensitivity, and accuracy (SSA) during normal respiration+apnea (NR+AP) versus normal respiration + speech + normal respiration + environmental noise (RS+RN) (Figure 3), (2) SSA from sitting versus lying, (3) SSA of each detection domain versus the final combined results, (4) correlation between PDR and EDR, and (5) correlation between S-S and R-R intervals; all compared data were paired. Based on the Shapiro-Wilk test, if the data were normally distributed, the paired Student test was applied to compare means; otherwise, the Wilcoxon test was applied to compare medians. If the p -value was ≥ 0.05 , we concluded that the difference between the two paired samples was not significantly different.

3. Results

3.1. Evaluation Method

The proposed algorithm was applied on a total of 26 recordings of 2 min from 13 healthy participants, 13 in sitting and 13 in lying positions. All speech episodes were detected successfully in all the recordings. According to the two final decision-making methods, the respiratory detection results were evaluated on: (1) the ratio of apnea detection, and (2) the performance of breathing activities assessment with specificity, sensitivity, and accuracy, which are defined as follows:

$$\text{Specificity} = \frac{TN}{TN + FP}$$

$$\text{Sensitivity} = \frac{TP}{TP + FN}$$

$$\text{Accuracy} = \frac{TP + TN}{TP + TN + FP + FN}$$

where TN , FP , TP , and FN are, respectively, the number of true negatives, false positives (Figure 11a), true positives, and false negatives (Figure 11b).

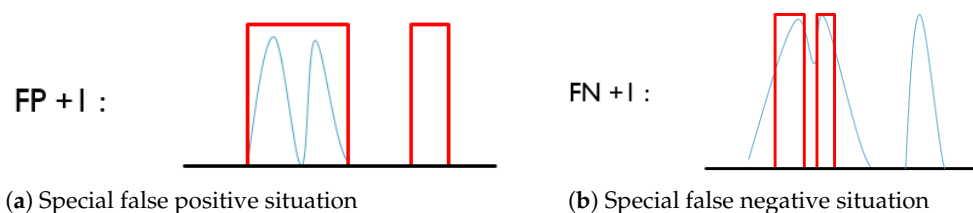


Figure 11. Example of breathing activity counting. Respiratory envelope (in blue), detection signal (in red). (a) If two continuous breathing activities are recognized as a single event, or the pause/apnea is detected as one breathing activity, a false positive detection will be reported; (b) if one breathing activity is recognized as two events, or one breathing activity is not detected, a false negative detection will be reported.

In our context, the detection system must not detect breaths that do not occur during apnea, i.e., FP . Given the definition of apnea, which is an absence of respiratory flow for at least 10 s, and the observation that the respiratory rate is on average 15 cycles per minute, this leads to:

- the specificity of respiratory detection (breathing/apnea) from TS must be $>82.9\%$.
- the accuracy of respiratory detection (breathing/apnea) from TS must be $>90.5\%$.

The FN are less important for patient safety. Given that the sensitivity is inversely proportional to FN , we fixed a minimum of 80% corresponding to $<25\%$ of FN compared to TP .

3.2. Apnea Detection

The rate of successful apnea detection was 92.3%. Among the 26 recordings that included a single apnea phase, two were missed and detected as long pauses because of unexpected environmental noise. In seven recordings, one or two more apneas were reported because the recorded respiratory sounds were too weak.

3.3. Breathing Activities Assessment

This method detects all breathing phases, including apnea. If only apnea detection is looked at, the success rate was 88.5%, slightly less than the previous method (92.3%).

Looking at all breathing activities, for all 26 recordings, the average of these detection results (Figure 12) is 99.4% for specificity (min 82.9%), 85.3% for sensitivity (min min 80%), and 91.5% for accuracy (min 90.5%), which meets the minimum performance required for the system (described in Section 3.1). In more detail, all 26 recordings meet the required specificity, 19 out of 26 meet the required sensitivity, and 18 out of 26 meet the required accuracy.

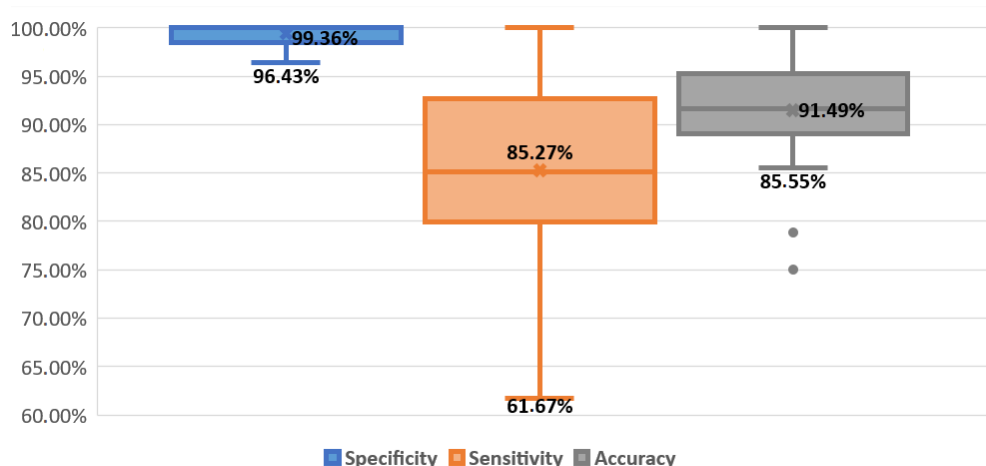


Figure 12. Box plot for final respiratory detection result. The line and the cross inside the box correspond to the median and the mean value, respectively. The box-limits represent the first and the third quartiles, the points highlight outliers, and the upper and lower bars represent the maximum and minimum values, respectively. Specificity: blue box. Sensitivity: orange box. Accuracy: gray box.

3.3.1. Influence of Speech and Noise

The influence of speech and noise on the detection results is presented in Figure 13. Comparing the results of the NR + AP versus RS + RN phases, only the specificity in the frequency domain and cardiac detection domain were better during the RS + RN phase ($p < 0.05$, Table 4); no other parameters showed significant differences. This shows that noisy environments did not alter the detection results.

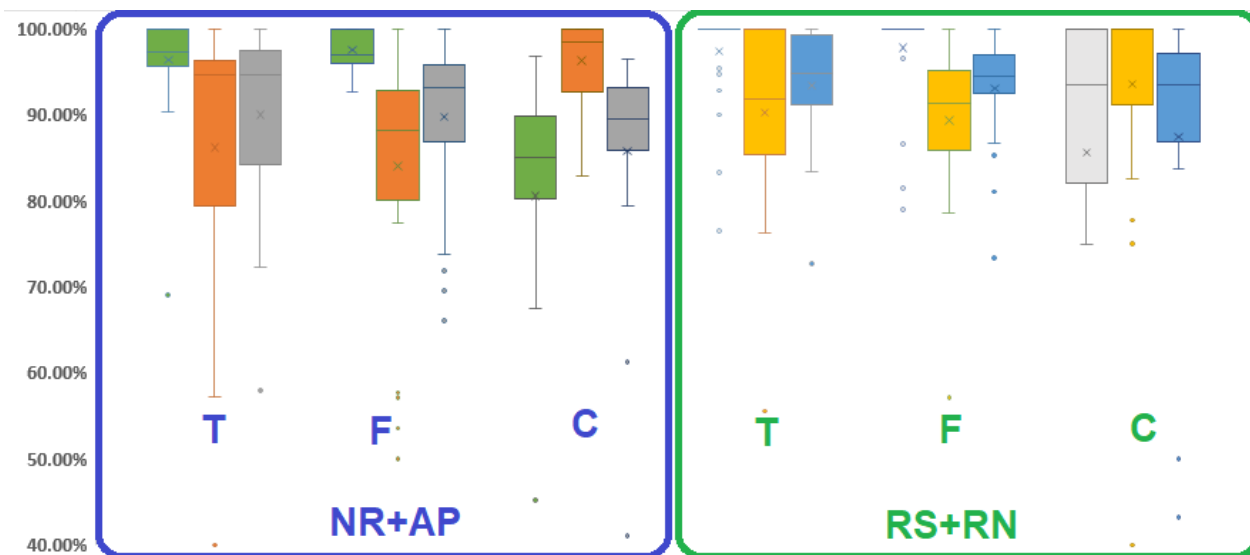


Figure 13. Influence of speech and noise on respiratory detection. Box plot (as box plot of Figure 12) for specificity (green and light gray boxes), sensitivity (orange and yellow boxes) and accuracy (dark gray and blue boxes) of temporal (T), frequency (F), and cardiac (C) detection domains, separated in NR+AP phases (in blue square) and RS + RN phases (in green square).

Table 4. Median values comparing for specificity, sensitivity and accuracy of each detection domain—NR + AP vs. RS + RN. * p -value < 0.05, difference between the two paired samples was significantly different.

Detection Domain	Median Comparison	p -Value
Temporal	Specificity	$p = 0.0966$
	sensitivity	$p = 0.8829$
	Accuracy	$p = 0.3638$
Frequency	Specificity	$p = 0.0088$ *
	sensitivity	$p = 0.2032$
	Accuracy	$p = 0.1219$
Cardiac (PDR)	Specificity	$p = 0.016$ *
	sensitivity	$p = 0.7369$
	Accuracy	$p = 0.0851$

3.3.2. Influence of Position

Further examination of the detection results in the different participants positions were compared (Figure 14) The specificity, sensitivity, and accuracy showed no significant differences between sitting and lying positions (Table 5).

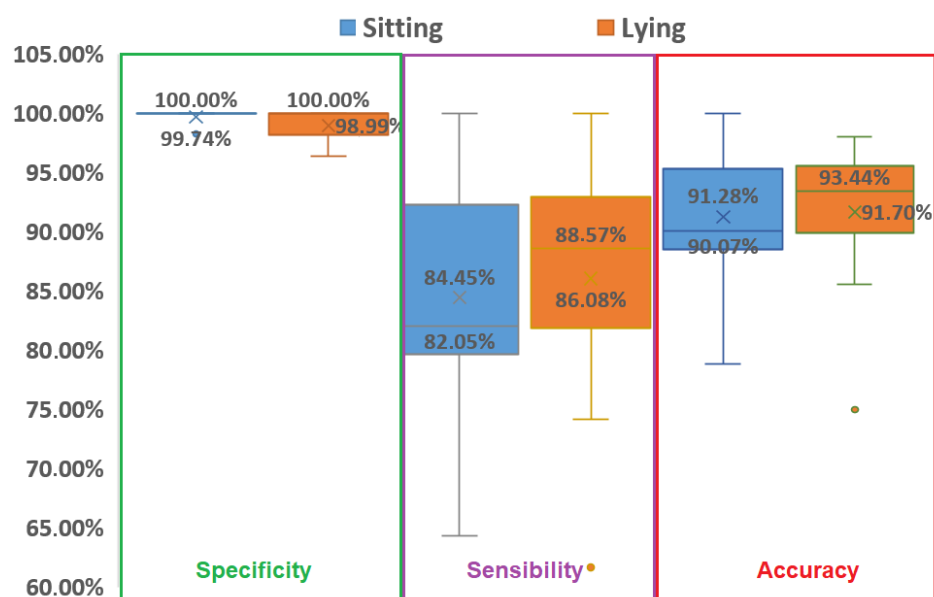


Figure 14. Box plot (like box plot of Figure 12) for specificity (green rectangle), sensibility (purple rectangle), and accuracy (red rectangle) in sitting (orange boxes) and lying (blue boxes) positions.

Table 5. Median values comparing for specificity, sensitivity, and accuracy—sitting vs. lying.

Median Comparison	p -Value
Specificity	$p = 0.2079$
sensitivity	$p = 0.7354$
Accuracy	$p = 1$

3.3.3. Advantage of Combining Several Detection Domains

The SSA of each detection domain and the global results (combined results from the 3 domains) are shown in Figure 15. Specificity was significantly improved (reduced false positives) with the combined detection approach compared to single domain detection (Table 6). Sensitivity was significantly reduced (increased false negatives) with the com-

bin detection approach compared to PDR detection. In all other comparisons, there was no significant difference.

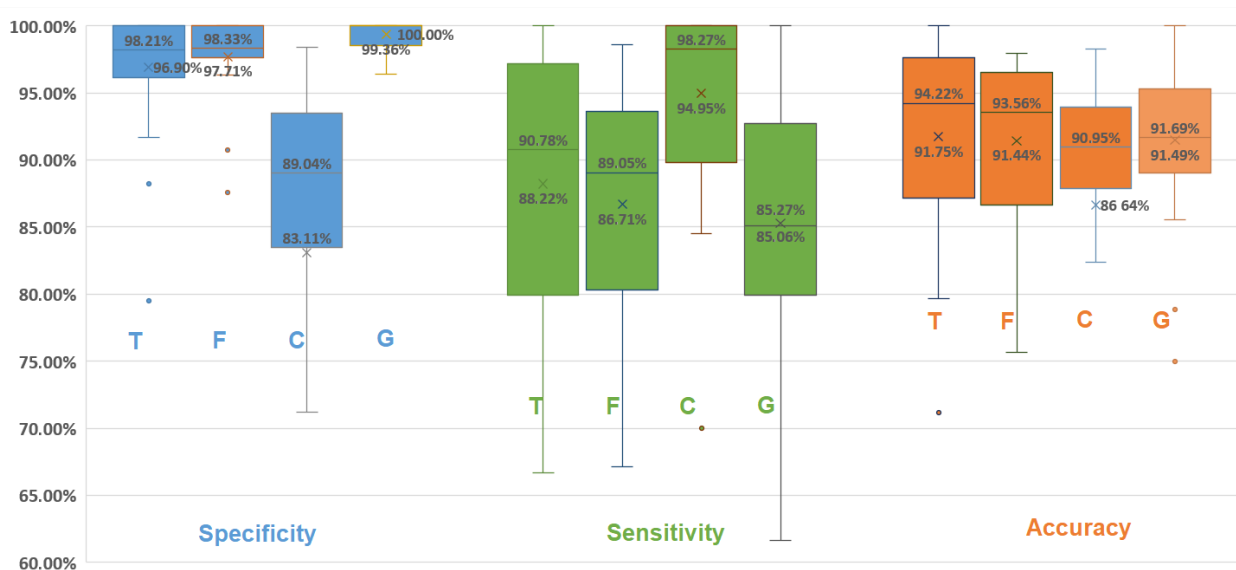


Figure 15. Box plot (as box plot of Figure 12) for specificity (blue), sensitivity (green), and accuracy (orange) of temporal (T), frequency (F), and cardiac (C) detection domains, as well as global result (G).

Table 6. Median values comparing for specificity, sensitivity, and accuracy of each detection domain—single detection domain vs. combining detection domain. * *p*-value < 0.05, difference between the two paired samples was significantly different.

Median Comparison	Detection Domain	<i>p</i> -Value
Specificity	Temporal	<i>p</i> = 0.00889 *
	Frequency	<i>p</i> = 0.01489 *
	Cardiac	<i>p</i> = 9.896 × 10 ⁻⁶ *
Sensitivity	Temporal	<i>p</i> = 0.09933
	Frequency	<i>p</i> = 0.6348
	Cardiac	<i>p</i> = 0.001801 *
Accuracy	Temporal	<i>p</i> = 0.8417
	Frequency	<i>p</i> = 0.8028
	Cardiac	<i>p</i> = 0.2079

3.4. Cardiac Sound Processing

As the PDR method is a new approach, we validated S1/S2 detection against classical R-peak detection from ECG recordings: R-peaks were detected through threshold detection and manually verified. In general, there should be one S1/S2 complex for each cardiac cycle corresponding to one R-peak. Both the reliability of the S-S interval (versus R-R interval) and PDR versus EDR (envelope of R-peak amplitude) were assessed.

3.4.1. Correlation between PDR and EDR

The obtained PDR was compared with EDR because they both resulted from thoracic volume changes due to respiration.

For the example shown in Figure 16, the amplitude variations of EDR versus PDR were obviously not correlated and thus PDR cannot be used for assessing the quantitative volume of respiration. However, both signals seemed to oscillate in phase: we thus compared the respiration cycles in the PDR vs. EDR signals. To do so, local maxima of each curve ($iPDR(k)$, $iEDR(k)$) were detected and their timing compared (Figure 17).

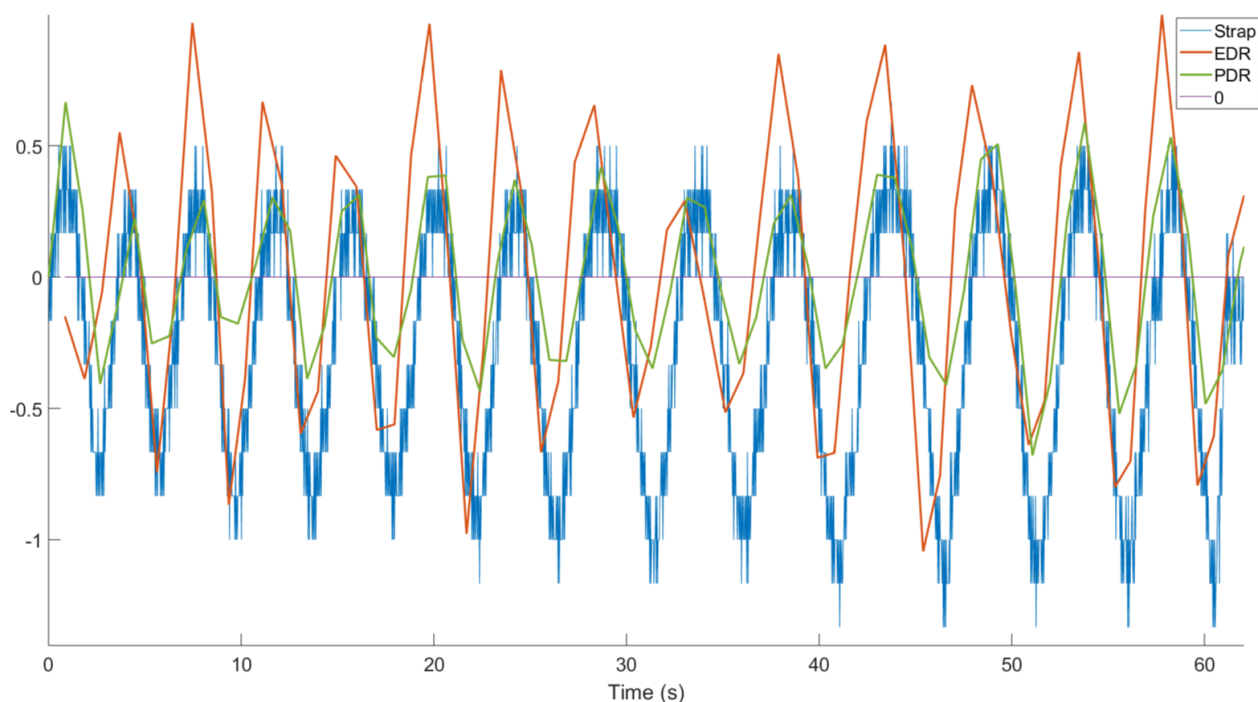


Figure 16. PDR (green) processed from tracheal sounds and ECG-Derived Respiration (EDR) (red) processed from ECG and the thoracic strap signal (blue). Their positive (negative) slopes indicate inspiration (expiration).

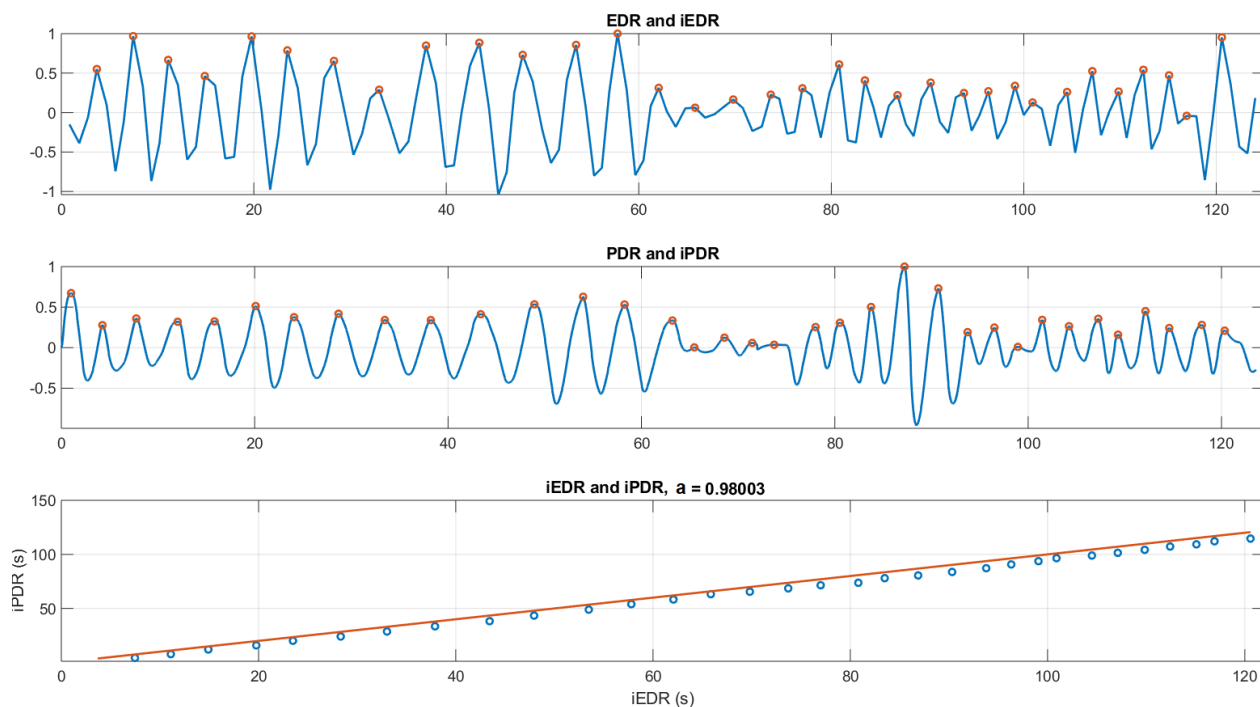


Figure 17. (Top): the EDR signal and its local maxima (peaks) $iEDR$. (Middle): the PDR signal and its local maxima $iPDR$. (Bottom): the linear regression fit line between $iEDR$ and $iPDR$ ($a = 0.98$).

A perfect match between the two should lead to a linear equation, such as

$$iPDR(k) = a \times iEDR(k),$$

where a in theory is equal to 1. Indeed, when $iPDR(k)$ and $iEDR(k)$ are not matched (missed, shifted, or additional peaks), it results in a variation ($a \neq 1$, (Figure 18)). Thus, the

slope of the first-order linear regression fit line (a) could be a good indication to evaluate the correlation between EDR and PDR. Essentially, the closer slope (a) is to 1, the more PDR and EDR are correlated.

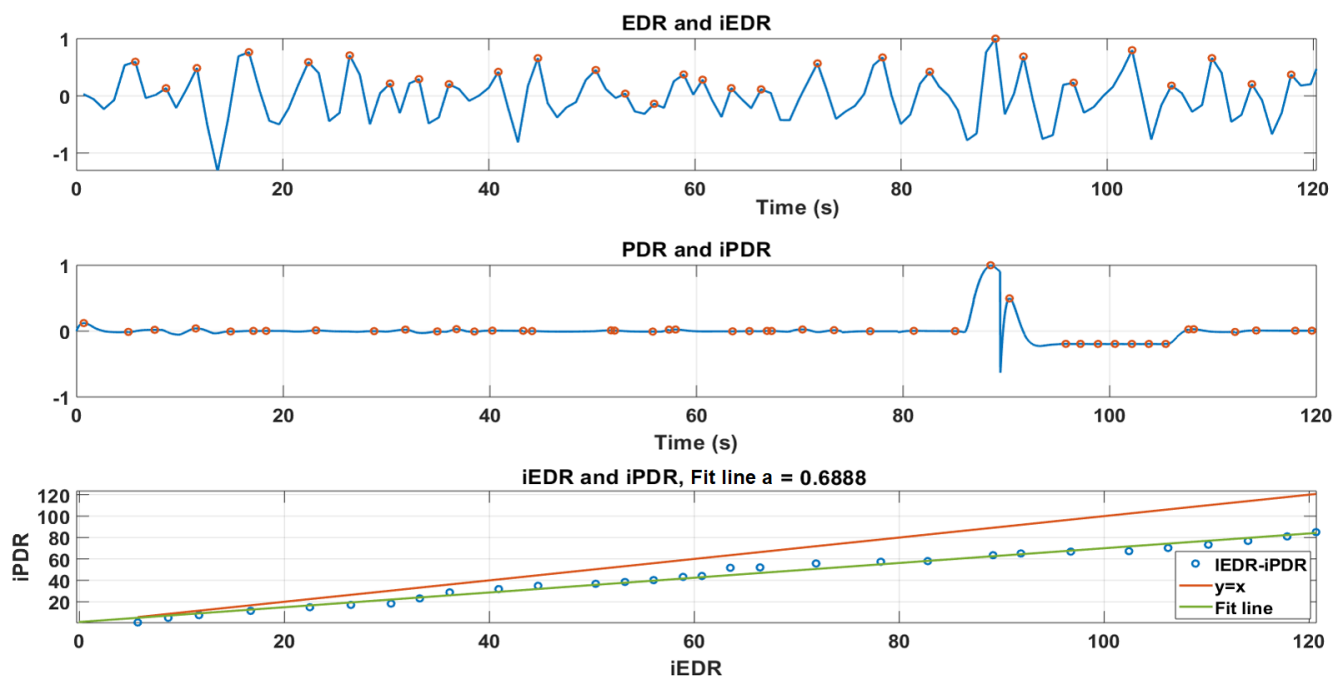


Figure 18. (Top): the EDR signal and its local maxima (peaks) $iEDR$. (Middle): the PDR signal and its local maxima $iPDR$. (Bottom): the theory fit line ($y = x$) in red and the linear regression fit line between $iEDR$ and $iPDR$ ($a = 0.68$).

The result of all recordings is illustrated in Figure 19, separated into the NR + AP phase, RS + RN phase (as described in Figure 3), and global results. The global mean slope was 0.77, which shows a good correlation between EDR and PDR. The Wilcoxon test ($p = 0.0067$) showed that the results during NR + AP were better than the ones during RS + RN because the cardiac peaks were difficult to detect during speech.

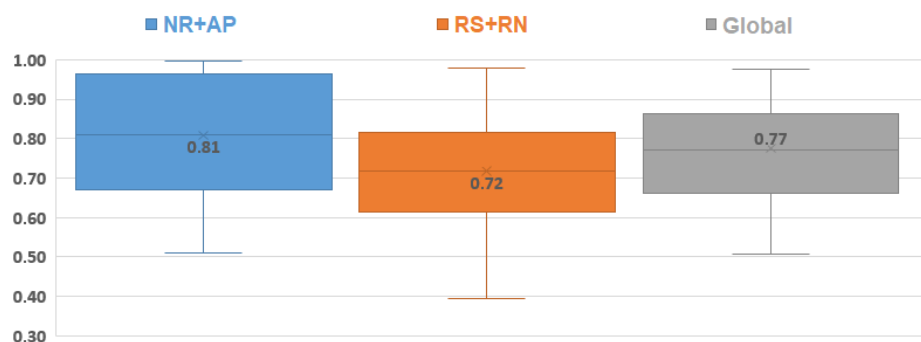


Figure 19. Box plot (like box plot of Figure 12) for the evaluation of the results of PDR according to EDR, separated into NR+AP (blue), RS+SN (orange), and global results (gray).

The PDR results were also evaluated in sitting and lying positions (Figure 20), and the Student test showed no significant difference in the means ($p = 0.5359$).

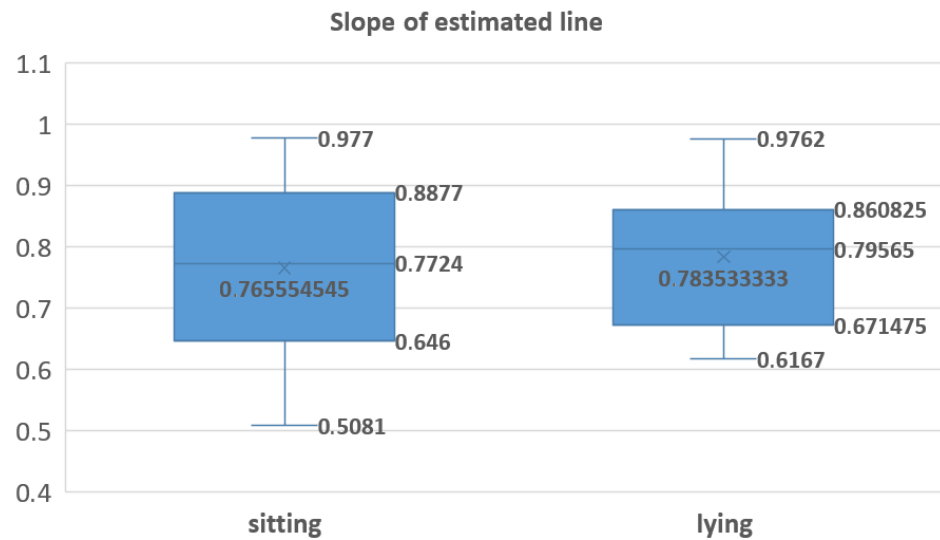


Figure 20. Box plot (like box plot of Figure 12) for the evaluation of the results of PDR according to EDR, separated in sitting (left) and lying (right) positions.

3.4.2. Heart Beat-to-Beat Interval

The heart beat-to-beat intervals derived from cardiac sounds were compared to the heart rate variation (HRV) (R-R interval) calculated from ECG. For each recording, the result was also separated into NR + AP and RS + RN. The ideal relationship should be a straight line $y = \text{offset}$: the initial offset is due only to the delay between R-peak and S sounds. However, the small oscillations are explained by the detection of either S1 or S2. The NR + AP phase showed a good fit, but the introduction of noise led to strong S complex mislocations.

Thus, plotting S-S versus R-R intervals showed mainly four groups of points corresponding to R-R vs. S1-S1 (or S2-S2), R-R vs. S1-S2, R-R vs. S2-S1, and outliers when the S1-S2 complex was missed. In the latter case, all the S-S intervals greater than two times the R-R interval were considered outliers and thus removed: a ratio R_{error} was calculated as:

$$R_{error} = 1 - \frac{N_{abnormal}}{N_{total}},$$

where $N_{abnormal}$ is the number of outliers, and N_{total} is the total number of points. This ratio was used to weight the final assessment. Once outliers were removed, a first fit line slope (β) was estimated using the solution of: $\underline{SS} = \beta \underline{RR}$. Based on this line, two other shifted lines defined the group separation in considering the averaged S1-S2 interval (0.29 s) and minimum S2-S1 interval (0.5 s). Thus, two boundary lines and one central line are obtained:

$$\begin{aligned} y1 &= \beta x + 0.5, \\ y2 &= \beta x, \\ y3 &= \beta x - 0.29. \end{aligned}$$

Thus, all points were then separated into three groups (by looking for their nearest line). As in the PDR and EDR assessment, the slope of the linear regression fit line was used to evaluate the S peak detection performance, three different fit lines could be calculated separately for these three groups of points, as in the example shown in Figure 21.

The assessment was based on how close the slope is to 1 (the best theory slope) as follows:

$$a_x = 1 - |a'_x - 1|,$$

where a'_x is the slope of the fit line for group x , and a_x is the new slope of group x to evaluate the correlation. Thus, an average slope was weighted as:

$$a = R_{error} * \frac{n_1 * a_1 + n_2 * a_2 + n_3 * a_3}{n_1 + n_2 + n_3},$$

where n_x presents the number of points in group x . For example, in Figure 21, the averaged slopes (a) for NR + AP, RS + RN, and the whole recording were 0.84, 0.73, and 0.79, respectively. The results for the 26 recordings are presented in Figure 22. The slope of group 1 in each part was very close to 1, which means a high correlation between the S-S and R-R interval lengths. The medians of the results from RS + RN and NR + AP were significantly different ($p = 0.00004$). The cardiac peaks were more difficult to detect from TS during speech or with environmental noise. Group 2 and group 3 showed even poorer correlations, but, despite their relatively small numbers, the global mean still reached 0.89. Thus, the S-S interval obtained from the proposed cardiac peak detection method is a reliable way to estimate heart rate or as an alternative method to present the heart rate variation (HRV).

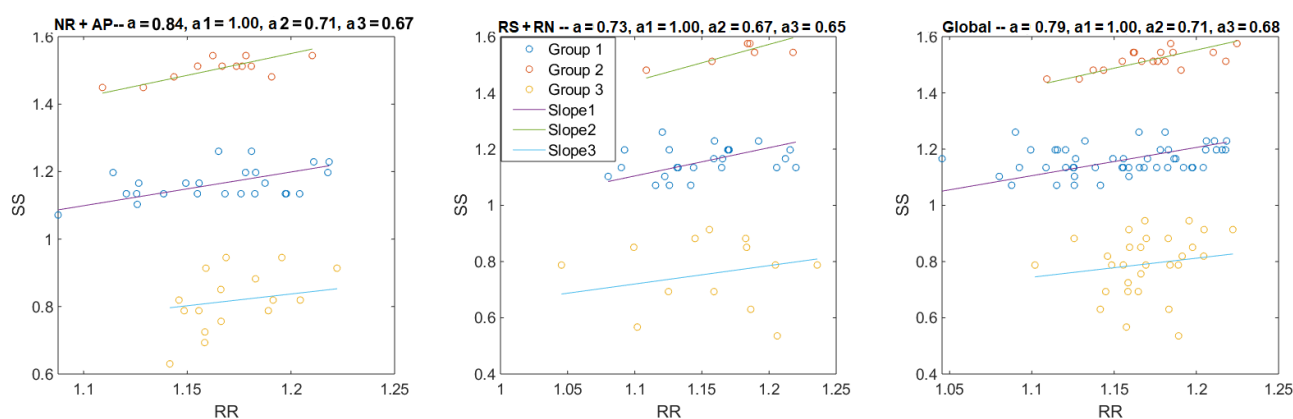


Figure 21. The left, middle, and right figures show the linear regressions for NR + AP, RS + RN, and the whole recording. The blue points are from the group of R-R vs. S1-S1 or S2-S2 (group 1, at the middle group of each figure), the orange points are from the group of R-R vs. S2-S1 (group 2, at the top of each figure), and the yellow points are from the group of R-R vs. S1-S2 (group 3, at the bottom of each figure). The slopes of each fit line are presented the in figure title (a_1 , a_2 , and a_3 are slopes for the group 1, group 2, and group 3, respectively).

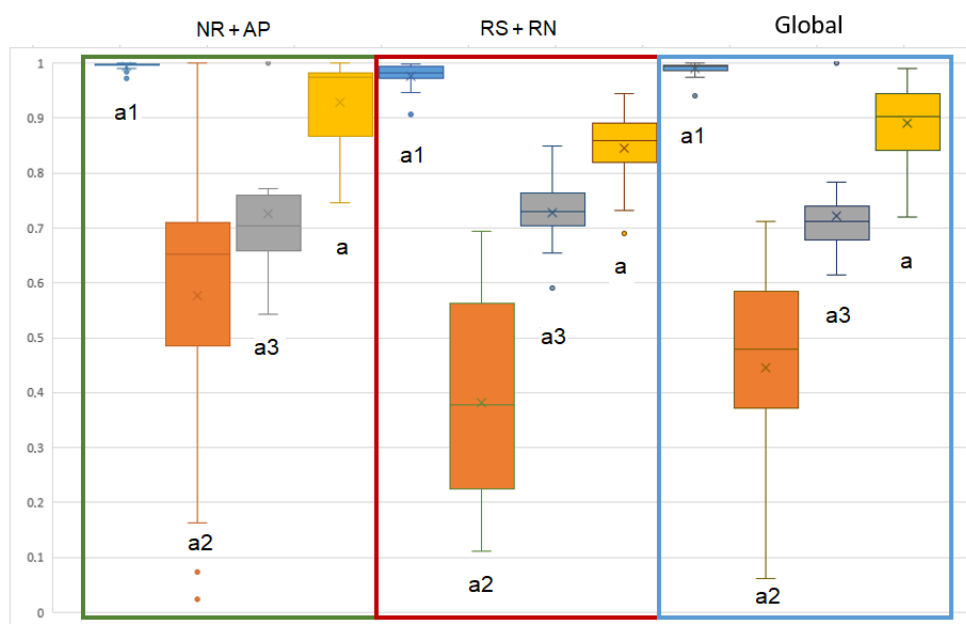


Figure 22. Box plot (like the box plot of Figure 12) for the distribution of fit line slopes for each group in each part and the global signal (in different rectangle).

4. Discussion

In this paper, we present an original method for continuous respiratory monitoring based on TS. The first aim was to detect apnea, and the second aim was to further classify breathing activities. Previous methods used a detection solution based on either the temporal features of sound or its frequency content. On the contrary, we combined classifications from the two domains with new signal processing based on cardiac sounds. We thus increased the performances of the classification. This preliminary method was proved to be feasible on 13 healthy subjects with 26 apnea events. For a better evaluation of the proposed algorithm, the next step is to validate on the main target population—patients with implanted phrenic nerve stimulation. Otherwise, we can underline the following outcomes.

Figure 15 and Table 6 show greater sensitivity with PDR than the final combined results. This is because the final decision-making method had, in fact, missed some events due to delays and different durations in the results of each detection domain (as shown in Figure 10), even though they were well detected in other domains. As a result, false negatives were added and the sensitivity was reduced. On the other hand, some false positives caused by noise could be eliminated in other detection domains (e.g., the temporal envelope could eliminate brief noises and the frequency domain might eliminate long duration noises with different frequency content than respiration). The number of false positives was reduced, and, therefore, the specificity was increased. Even though the sensitivity was ultimately reduced, it was still above the minimum requirement for a detection system (>80%), so the results remain acceptable. Furthermore, in this study, high specificity was more important than high sensitivity or high accuracy because false positives would mean that apnea had been detected as respiration. In other words, patients would stop breathing, but the detection system would not deliver an alarm.

This study used two microphones to record TS and environmental noise separately. The ideal would have been to use recorded noise to reduce the environmental noise in the recorded TS by using the Wiener adaptive filter [25]. But, in the obtained recordings, the noisy environment was simulated by playing specific videos, in which the frequency band was higher than the respiratory frequency band (Figure 5) so that it could be easily filtered. A better way would be to record data in a real noisy environment (e.g., on street) to see if the design with noise reduction by two microphones would improve the detection result.

To our knowledge, the proposed PDR method had never been studied before. In this study, however, we showed the feasibility of evaluating respiratory effort with PDR by comparing the result with EDR. This means that, using only a microphone, it is possible to continuously monitor cardiac activity, detect the apnea episodes, and eventually discriminate between the type of apnea based on the PDR (apnea with respiratory effort: obstructive apnea; otherwise: central apnea). In our study, PDR also allows to detect correctly respiration phases and apnea, in particular, when the respiration noise is very weak: in this case, time and frequency domains detection may fail, while PDR succeeds. It should, nevertheless, be noted that the EDR method is still not widely applied in clinic routine, and it is not known to be a good indicator of respiratory effort. More comparisons with a plethysmograph (PSG) are thus needed. In addition, the proposed algorithm has not yet been tested in individuals with real apnea. Thus, the next step will be to record patients' TS synchronized with PSG to test whether the same algorithm still works correctly. Furthermore, cardiac peak detection may be improved by combining wavelet decomposition [26] (better detection of peak position) and the envelopegram by absolute value [24] (better detection of peak amplitude).

Monitoring TS could be a useful way beyond the very limited niche of patients with diaphragm pacing or sleep apnea. Indeed, it would provide a non-invasive way to approximate inspiratory flow that would be useful in all patients requiring respiratory monitoring in acute situations (e.g., as a safety measure during the administration of morphine for acute pain) and in chronic situations (e.g., home mechanical ventilation).

5. Conclusions

The method we presented is able to continuously monitor healthy subject's respiration (apnea detection, breathing activities assessment, respiratory effort evaluation) and cardiac activity (heart beat detection) using only a microphone recording tracheal sounds (TS). The next step will be to validate the approach on patients with implanted phrenic nerve stimulation. Recorded tracheal sounds were processed in temporal envelope detection, frequency envelope (PSD) detection, and PDR (cardiac) detection, and then combined for the results. The proposed algorithm was tested on 26 recordings from 13 healthy participants, in sitting and lying positions, and the apnea detection rate was 92.3%, with 99.36% specificity, 85.27% sensibility, and 91.49% accuracy, which meets the minimum requirements for the system. Indeed, noise, like speech or from the environment, may influence the detection result, but within an acceptable range. The different positions could lead to different recorded sound intensities but did not influence the detection result. Compared to using single detection domain, combining the three detection domains significantly improved the algorithm's specificity, which is the most important parameter in apnea detection. But, it reduced its sensitivity (within an acceptable range) due to the decision-making method.

The new proposed method, phonocardiogram-derived respiration (PDR), showed a good correlation with EDR, which demonstrates the feasibility of using PDR to monitor respiration. At the same time, the comparison of R-R intervals and S-S intervals also indicated a good correlation; thus, it is also possible to monitor heart activity from PCG extracted from TS. More comparisons with PSG are, nevertheless, needed with patients for further confirmation of the method's performance.

Author Contributions: Conceptualization, C.A.C., D.G.; Methodology, X.L., T.S., C.A.C., D.G.; Software, X.L.; Validation, X.L.; Formal Analysis, X.L.; Investigation, T.S., M.-C.N., T.S.; Resources, S.R.; Data Curation, X.L.; Writing—Original Draft Preparation, X.L.; Writing—Review and Editing, X.L., C.A.C., M.-C.N., S.R., T.S., D.G.; Visualization, X.L.; Supervision, C.A.C., D.G., S.R., T.S.; Project Administration, C.A.C.; Funding Acquisition, S.R., C.A.C. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by [ANRT (Association Nationale de la Recherche et de la Technologie, France)] grant number [2016/1600] and [regional funding Occitanie] grant number [18007760].

Ethical Statement: All subjects gave their informed consent for inclusion before they participated in the study. This study was submitted to the CPP (Committee for the Protection of Persons, France) and has been validated on 5 October 2018.

Data Availability Statement: Data sharing is not applicable to this article.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Similowski, T.; Derenne, J.P. Stimulation phrénique implantée. *Méd. Théor.* **2001**, *7*, 457–469.
2. Weese-Mayer, D.E.; Silvestri, J.M.; Kenny, A.S.; Ilbawi, M.N.; Hauptman, S.A.; Lipton, J.W.; Talonen, P.P.; Garcia, H.G.; Watt, J.W.; Exner, G.; et al. Diaphragm pacing with a quadripolar phrenic nerve electrode: An international study. *PACE Pacing Clin. Electrophysiol.* **1996**, *19*, 1311–1319. [[CrossRef](#)]
3. Glenn, W.W.; Holcomb, W.G.; Gee, J.B.; Rath, R. Central hypoventilation; long-term ventilatory assistance by radiofrequency electrophrenic respiration. *Ann. Surg.* **1970**, *172*, 755–773. [[CrossRef](#)] [[PubMed](#)]
4. Le Pimpec-Barthes, F.; Gonzalez-Bermejo, J.; Hubsch, J.P.; Duguet, A.; Morélot-Panzini, C.; Riquet, M.; Similowski, T. Intrathoracic phrenic pacing: A 10-year experience in France. *J. Thorac. Cardiovasc. Surg.* **2011**, *142*, 378–383. [[CrossRef](#)] [[PubMed](#)]
5. Iber, C.; Ancoli-Israel, S.; Chesson, A.L.; Quan, S.F. *The AASM Manual for the Scoring of Sleep and Associated Events: Rules, Terminology and Technical Specifications*; American Academy of Sleep Medicine Westchester: Darien, IL, USA, 2007.
6. DiMarco, A.F. Diaphragm Pacing. *Clin. Chest Med.* **2018**, *39*, 459–471. [[CrossRef](#)]
7. Que, C.I.; Kolmaga, C.; Durand, L.G.; Kelly, S.M.; Macklem, P.T. Phonospirometry for noninvasive measurement of ventilation: Methodology and preliminary results. *J. Appl. Physiol.* **2002**, *93*, 1515–1526. [[CrossRef](#)] [[PubMed](#)]
8. Corbishley, P.; Rodríguez-Villegas, E. Breathing detection: Towards a miniaturized, wearable, battery-operated monitoring system. *IEEE Trans. Biomed. Eng.* **2008**, *55*, 196–204. [[CrossRef](#)]

9. Yadollahi, A.; Giannouli, E.; Moussavi, Z. Sleep apnea monitoring and diagnosis based on pulse oximetry and tracheal sound signals. *Med. Biol. Eng. Comput.* **2010**, *48*, 1087–1097. [[CrossRef](#)] [[PubMed](#)]
10. Kalkbrenner, C.; Eichenlaub, M.; Rüdiger, S.; Kropf-Sanchen, C.; Rottbauer, W.; Brucher, R. Apnea and heart rate detection from tracheal body sounds for the diagnosis of sleep-related breathing disorders. *Med Biol. Eng. Comput.* **2017**, *56*, 671–681. [[CrossRef](#)] [[PubMed](#)]
11. Chuah, J.S.; Moussavi, Z. Automated detection of respiratory phases by acoustical means. In Proceedings of the 20th Annual International Conference of the IEEE Engineering in Medicine and Biology Society. Vol.20 Biomedical Engineering Towards the Year 2000 and Beyond (Cat. No.98CH36286), Hong Kong, China, 1 November 2001.
12. Hult, P.; Fjallbrant, T.; Wranne, B.; Engdahl, O.; Ask, P. An improved bioacoustic method for monitoring of respiration. *Technol. Health Care* **2004**, *12*, 323–332. [[CrossRef](#)] [[PubMed](#)]
13. Hult, P.; Fjallbrant, T.; Dahle, S.; Danielsson, P.; Ask, P. A method for respiration monitoring by the use of a bioacoustic signal. In Proceedings of the First International Conference on Advances in Medical Signal and Information Processing, Bristol, UK, 4–6 September 2000; IET: London, UK, 2000; Volume 2000, pp. 22–25. [[CrossRef](#)]
14. Kalkbrenner, C.; Stark, P.; Kouemou, G.; Algorri, M.E.; Brucher, R. Sleep monitoring using body sounds and motion tracking. In Proceedings of the 2014 36th Annual International Conference of the IEEE Engineering in Medicine and Biology Society, Chicago, IL, USA, 26–30 August 2014; Volume 6000, pp. 6941–6944. [[CrossRef](#)]
15. Kulkas, A.; Huupponen, E.; Virkkala, J.; Tenhunen, M.; Saastamoinen, A.; Rauhala, E.; Himanen, S.L. New tracheal sound feature for apnoea analysis. *Med. Biol. Eng. Comput.* **2009**, *47*, 405–412. [[CrossRef](#)] [[PubMed](#)]
16. Lei, B.; Rahman, S.A.; Song, I. Content-based classification of breath sound with enhanced features. *Neurocomputing* **2014**, *141*, 139–147. [[CrossRef](#)]
17. Muthusamy, P.D.; Sundaraj, K.; Abd Manap, N. Computerized acoustical techniques for respiratory flow-sound analysis: A systematic review. *Artif. Intell. Rev.* **2019**, *53*, 501–3574. [[CrossRef](#)]
18. Schmidt, M.; Schumann, A.; Müller, J.; Bär, K.J.; Rose, G. ECG derived respiration: comparison of time-domain approaches and application to altered breathing patterns of patients with schizophrenia. *Physiol. Meas.* **2017**, *38*, 601–615. [[CrossRef](#)] [[PubMed](#)]
19. Martin, A.; Voix, J. In-Ear Audio Wearable: Measurement of Heart and Breathing Rates for Health and Safety Monitoring. *IEEE Trans. Biomed. Eng.* **2018**, *65*, 1256–1263. [[CrossRef](#)] [[PubMed](#)]
20. White Noise Soundscapes. City Traffic White Noise | 10 Hours | For Sleeping, Studying or to Block Out Noise. Available online: <https://www.youtube.com/watch?v=8s5H76F3SIs> (accessed on 25 December 2020).
21. Lu, X.; Guiraud, D.; Renaux, S.; Similowski, T.; Azevedo, C. Breathing detection from tracheal sounds in both temporal and frequency domains in the context of phrenic nerve stimulation. In Proceedings of the 2019 41st Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC), Berlin, Germany, 23–27 July 2019; pp. 5473–5476. [[CrossRef](#)]
22. Lu, X.; Guiraud, D.; Renaux, S.; Similowski, T.; Azevedo Coste, C.; Azevedo, C. Monitoring phrenic nerve stimulation-induced breathing via tracheal sounds. In Proceedings of the 2019 XXIII Annual Conference of the International Functional Electrical Stimulation Society (IFESS), Toronto, ON, Canada, 24–28 June 2019.
23. Penzel, T.; Sabil, A.K. The use of tracheal sounds for the diagnosis of sleep apnoea. *Breathe* **2017**, *13*, e37–e45. [[CrossRef](#)] [[PubMed](#)]
24. Liang, H.; Lukkarinen, S.; Hartimo, I. Heart sound segmentation algorithm based on heart sound envelopogram. In Proceedings of the Computers in Cardiology, Lund, Sweden, 7–10 September 1997; Volume 24, pp. 105–108. [[CrossRef](#)]
25. Plapous, C.; Marro, C.; Scalart, P. Improved signal-to-noise ratio estimation for speech enhancement. *IEEE Trans. Audio Speech Lang. Process.* **2006**, *14*, 2098–2108. [[CrossRef](#)]
26. Ozkan, I.; Member, S.; Yilmaz, A.; Celebi, G. Improved Segmentation with Dynamic Threshold Adjustment for Phonocardiography Recordings. In Proceedings of the 2019 41st Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC), Berlin, Germany, 23–27 July 2019; pp. 6681–6684.