A Single-Channel Acoustic Method For Portable Diagnosis of Sleep Apnea

by

Hisham Alshaer, MD

A thesis submitted in conformity with the requirements for the degree of Doctor of Philosophy
Graduate Department of IBBME
University of Toronto

Copyright © 2013 by Hisham Alshaer, MD
Abstract

A Single-Channel Acoustic Method For Portable Diagnosis of Sleep Apnea

Hisham Alshaer, MD
Doctor of Philosophy
Graduate Department of IBBME
University of Toronto
2013

Background:
Sleep apnea is a very common disease with serious health consequences. It affects approximately 85% of adults, most of whom remain undiagnosed, which puts them at risk of car accidents, because of sleepiness, and hypertension, heart failure, and stroke. The objective of this thesis was to develop a single-channel device for portable monitoring of sleep apnea that uses breath sounds collected via a microphone distal to the nose and mouth.

Methods:
Project 1: Frequency characterization was used to identify the basic components of the respiratory cycle, i.e., inspiration and expiration.
Project 2: Inspiratory sounds were used to determine upper airway narrowing by means of Linear Predictive Coding (LPC), which was validated against objective measures of upper airway resistance.
Project 3: An algorithm to calculate the frequency of apneas and hypopneas per hour (apnea-hypopnea index or AHI) was developed and validated against polysomnography in 50 subjects during sleep.
Project 4: Finally, a self contained device was then developed and evaluated in 49 subjects in the home setting, 11 of whom used the device on 2 different nights.
Results:

Project 1: Inspiratory and expiratory phases had characteristically different spectra that allowed them to be distinguished with up to 97% correct classification.

Project 2: LPC coefficients were found to be modulated by the ensuing of upper airway narrowing.

Project 3: AHI determined by acoustic analysis showed up to 94% correlation with PSG and up to 90% diagnostics accuracy.

Project 4: In the home unattended setting, the overall rating for ease-of-use was excellent and the success rate of independent use was 94%. The portable device showed excellent performance with very high signal-to-noise ratio of 31.7 dB. The intra-subject 2-night AHI scores were reproducible in 9 out of 11 (82%) home subjects, which is well within reported inter-night variability.

Conclusion:

Acoustic analysis of breath sounds is a powerful tool for characterization of respiratory patterns including upper airway narrowing and accurate identification of apneas and hypopneas as compared with the current gold standard. This technology can be packaged in a compact device that can be used independently and reliably in the home environment.
Dedication

To my mother and my father, Naderah and Ismael and my wife Inas.

Acknowledgements

Words are not enough to express my gratitude to those who have been supportive of me throughout this journey of research.

I am especially grateful to my supervisors Geoff Fernie and Douglas Bradley for giving me this valuable opportunity to work on this important project. I value the trust they showed, the advice that they gave, and the resources they offered.

I extend my thanks to my colleagues and friends in the Technology Team and the Sleep Team. Almost everyone has left a touch of kindness or an offer of help when it was needed. I would like to thank those who shared advice on data analysis especially Kaveh Momen, Cesar Marquez, Tilak Duta, Yue Li, Azadeh Yadollahi, and others. Thanks to the sleep laboratory technologists who helped in subject recruitment, especially Fiona Rankin, Ruth Rutherford, Wen-Hou Tseng, and all the other technologists who helped with that. Many thanks to the industrial designers Adam Shobchak and Steve Pong, and the engineer Alex Levchenko for their valuable role in creating the hardware.

I would like to acknowledge my committee members, Pascal van Lieshout and Willy Wong, who were always ready to answer my questions and address my concerns. I am also grateful to my examiners who dedicated time to read and evaluate my work on top of their busy schedule, Alex Mihailidis and Rangaraj Rangayyan.

I acknowledge the generous support of various federal, provincial, and private agencies that funded this work since its early stages. Many thanks to NSERC for granting me a 3-year PhD scholarship. The Ontario Ministry of Research and Innovation has funded this project several times via the Technology Transfer Toronto program. MaRS Innovation has been supportive with funds and experience. Our group also received generous grants
from the Ontario Centre of Excellence and the Ontario Brain Institute. I am grateful also to private sector contributions, namely Johnson and Johnson Inc.

Behind the scenes, there have always been those who cared and showered me with their unconditional love and support. Inas was always there for me, in easy times and in hard times. I thank her for her understanding and patience during all those years and for facing all the hardship with strength and dedication. My parents, Naderah and Ismael, receive my deepest gratitude and love for the many years of support during my undergraduate and masters studies that provided the foundation for this work.
Contents

1 Background ......................................................... 1

1.1 Introduction .................................................... 1

1.2 Pathophysiology of Sleep Apnea ............................... 2
  1.2.1 Pathophysiology of OSA ..................................... 2
  1.2.2 Pathophysiology of CSA ..................................... 5

1.3 Epidemiology of Sleep Apnea ................................... 7
  1.3.1 Prevalence of OSA and CSA ................................. 7

1.4 The Undiagnosed Portion of Sleep Apnea ...................... 9
  1.4.1 The Impact of Sleep Apnea .................................. 9
      Sleep Apnea and Cardiovascular Disease ....................... 9
      Sleep Apnea and Motor Vehicle Accidents ...................... 10

1.5 Diagnosis of Sleep Apnea ....................................... 12
  1.5.1 Overnight Polysomnography ................................ 12
      Quantification of the Severity of Sleep Apnea ............... 13
  1.5.2 Sleep Apnea Syndrome ...................................... 14
  1.5.3 Sleep Studies in Ontario and Around the World .......... 15
  1.5.4 Portable Monitoring of Sleep Apnea ....................... 16
      Evaluation of Currently Existing Portable Monitors .......... 17

1.6 Treatment of Sleep Apnea ...................................... 18
  1.6.1 Treatment of OSA ........................................... 18
Stage 1: Analysis of Breath Sounds in the Test Group and Feature Extraction ........................................ 41
Stage 2: Breathing Phase Tracking Algorithm Development and Validation ........................................ 44

3.5 Discussion ......................................................................................................................... 48

4 Identification of Breathing Cycle Phases, Part II ........................................................................... 53

4.1 Abstract ................................................................................................................................. 54
4.2 Introduction .............................................................................................................................. 54
4.3 Methods .................................................................................................................................. 56
  4.3.1 Polysomnography and Sleep Staging .................................................................................. 56
  4.3.2 Data Acquisition ............................................................................................................... 56
  4.3.3 Breathing Acoustics Analysis ............................................................................................ 56
  4.3.4 Statistical analysis ............................................................................................................. 58
4.4 Results ...................................................................................................................................... 58
4.5 Discussion ................................................................................................................................. 61

5 Detection of UA Narrowing ........................................................................................................ 67

5.1 Abstract ...................................................................................................................................... 68
5.2 Introduction .................................................................................................................................. 68
5.3 Methodology ............................................................................................................................... 70
  5.3.1 Experimentation and Data Acquisition ............................................................................... 70
    Subjects ...................................................................................................................................... 70
    Breath Sound Recordings ......................................................................................................... 70
    Lower Body Positive Pressure Application ............................................................................. 70
    Measurement of Upper Airway Resistance (R_{UA}) ............................................................... 71
5.4 Data Analysis ............................................................................................................................. 72
    Breath Sounds Segmentation and Annotation ......................................................................... 72
6 Preprocessing of Breathing Acoustic Data

6.1 Abstract ......................................................... 80
6.2 Introduction .................................................... 81
6.3 Data Acquisition ............................................... 83
6.4 Proposed Algorithm ............................................. 84
  6.4.1 Signal Envelope Creation .................................. 84
  6.4.2 Segmentation and Normalization .......................... 87
6.5 Results and Discussion ......................................... 89
6.6 Conclusion ..................................................... 94

7 Detection of Apneas and Hypopneas

7.1 Abstract ......................................................... 96
7.2 Introduction .................................................... 97
7.3 Methods ........................................................ 99
  7.3.1 Subjects ...................................................... 99
  7.3.2 Acquisition of Breath Sounds ............................ 99
  7.3.3 Polysomnography ........................................... 99
  7.3.4 Development of the Automated Algorithm .............. 101
    Transformation of the Raw Acoustic Signals ............... 101
    Scanning for Respiratory Events ............................ 102
    Apnea Tests ................................................... 104
    Hypopnea Tests ............................................... 106
    Algorithm Implementation .................................. 108
7.3.5 Algorithm Evaluation and Statistical Analysis 108
7.4 Results 111
    7.4.1 Subjects’ Characteristics 111
    7.4.2 Comparison Between AASM and TV50 PSG Scoring 111
    7.4.3 Agreement Between AHI-a and AHI-p 111
    7.4.4 Diagnosis of Sleep-Disordered Breathing 114
    7.4.5 Effect of Sleep Time on AHI 115
7.5 Discussion 117

8 Implementation of Portable Monitoring in The Home 123
    8.1 Abstract 124
    8.2 Introduction 125
    8.3 Materials and Methods 128
        8.3.1 Development of the Diagnostic System 128
            Self-Contained Acoustic Device With an Embedded Data-Capturing Module 128
            Central Data Analysis 130
        8.3.2 System Evaluation in the Home Environment 130
            Evaluation of Functionality and Signal Quality in the Home Environment 131
            Evaluation of Usability in the Home Environment 132
            Evaluation of Diagnostic Capability and Reproducibility in the Home Environment 133
        8.3.3 Agreement with PSG 133
    8.4 Results 134
        8.4.1 Functionality and Signal Quality in the Home Environment 134
        8.4.2 Usability in the Home Environment 134
        8.4.3 Diagnostic Capability and Reproducibility in the Home Environment 136
<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>8.4.4 Agreement with PSG</td>
<td>136</td>
</tr>
<tr>
<td>8.5 Discussion</td>
<td>138</td>
</tr>
<tr>
<td>8.6 Conclusion</td>
<td>141</td>
</tr>
<tr>
<td>9 Concluding Remarks</td>
<td>142</td>
</tr>
<tr>
<td>9.1 Limitations</td>
<td>144</td>
</tr>
<tr>
<td>9.2 Future Directions</td>
<td>145</td>
</tr>
<tr>
<td>Appendices</td>
<td>147</td>
</tr>
<tr>
<td>A Model-Based Analysis of Bland-Altman Limits of Agreement</td>
<td>148</td>
</tr>
<tr>
<td>A.1 Statistical Model</td>
<td>148</td>
</tr>
<tr>
<td>A.2 ANOVA Table</td>
<td>149</td>
</tr>
<tr>
<td>A.3 Interrater Reliability</td>
<td>150</td>
</tr>
<tr>
<td>Bibliography</td>
<td>151</td>
</tr>
</tbody>
</table>
Acronyms

AASM  American Academy of Sleep Medicine

ADC  Analog to Digital Convertor

ADVENT-HF  Effect of Adaptive Servo Ventilation (ASV) on Survival and Hospital Admissions in Heart Failure

AHI  Apnea Hypopnea Index

ANOVA  Analysis of Variance

APAP  Autotitrating Positive Airway Pressure

ASDA  American Sleep Disorders Association

ASV  Adaptive Servo Ventilation

BE  Breathing Envelope

BMI  Body Mass Index

BP  Blood Pressure

BR  Bands Ratio

CANPAP  CANadian Positive Airway Pressure trial for patients with congestive heart failure and central sleep apnea

COPD  Chronic Obstructive Pulmonary Diseases

CPAP  Continuous Positive Airway Pressure

CSA  Central Sleep Apnea

dB  Decible
DMA  Direct Access Memory

ECG  Electrocardiogram

EE  Effort Envelope

EEG  Electroencephalogram

EMGsm  Submandibular Electromyography

EOG  Electrooculogram

FE  Falling Edge

FEF  Falling Edge Factor

FFT  Fast Fourier Transform

GPIO  General Purpose Input Output

$H_{CT}$  High Clustering Tendency

HF  Heart Failure

HR  Heart Rate

$L_{CT}$  Low Clustering Tendency

LED  Light Emitting Diode

LPC  Linear Prediction Coding

LV  Left Ventricular

MCU  Microcontroller

MEMS  Micro-Electrical Mechanical Systems

NPV  Negative Predictive Value
OSA  Obstructive Sleep Apnea

PE  Potential Event

PF  Peak Frequency

PPV  Positive Predictive Value

PSG  Polysomnography

\( R_{UA} \)  Upper Airway Resistance

RAM  Random Access Memory

REM  Rapid Eye Movement Sleep

RIP  Respiratory Inductance Plethysmography

\( \text{SaO}_2 \)  Oxyhemoglobin Saturation

SD  Secure Digital

SDB  Sleep Disordered Breathing

SERVE-HF  Treatment of Sleep Disordered Breathing by Adaptive Servo-Ventilation in Heart Failure Patients

SNA  Sympathetic Nervous System Activity

SNR  Signal to Noise Ratio
Chapter 1

Background

1.1 Introduction

Sleep apnea syndrome is a breathing disorder characterized by repetitive cessations of breathing during sleep for 10-40 second long intervals. The frequency of these events ranges from 5 to 100 times/hour depending on the severity of the case. This causes episodes of oxygen deprivation and provokes arousals and sleep fragmentation. As a result, patients suffer from poor sleep quality, daytime sleepiness, and poor cognitive performance. Repetitive apneas and intermittent hypoxia also elicit sympathetic nervous system activation, oxidative stress, and elaboration of inflammatory mediators that cause repetitive surges in blood pressure at night and increase the risk of developing daytime hypertension, atherosclerosis, heart failure, and stroke independently of other risk factors [1, 2, 3, 4, 5].

Sleep apnea syndrome is classified according to the underlying mechanism into 2 types, obstructive sleep apnea (OSA) and central sleep apnea (CSA). OSA is the most common type of sleep apnea. As the name implies, it results from collapse of the upper airway (UA) either partially (hypopnea) or totally (apnea) during sleep, as illustrated in Figure 1.1. These events alternate with episodes of hyperventilation (i.e., hyperpnea)
Chapter 1. Background

during which loud snoring occurs [6].

1.2 Pathophysiology of Sleep Apnea

1.2.1 Pathophysiology of OSA

The UA is a muscular and compliant structure susceptible to collapse during inspiration because of negative airway pressure generated during inspiration and the end of expiration due to maximum withdrawal of neural input to the UA dilator muscles [7]. UA patency is maintained by tonic and phasic neural activation of pharyngeal dilator muscles. This is particularly important during inspiration, when intermittent inspiration-linked phasic activation of pharyngeal dilator muscles elicits the heightened muscular activity necessary to maintain airway patency [7]. This neural activation of dilator muscles is attenuated during sleep, thus predisposing to airway collapse.

Narrowing of the UA contributes to the pathogenesis of OSA, thus, conditions that lead to UA narrowing can predispose to OSA. In obese patients, for example, increased adipose tissue in the neck predisposes to UA narrowing [8]. Other factors such as tonsillar hypertrophy or craniofacial skeletal abnormalities can also predispose the UA to narrowing and thus result in development of OSA even in patients of normal body weight [8].

In conditions that cause reduced UA lumenal diameter, such as obesity, there is increased UA resistance. Poiseuille’s law necessitates greater inspiratory negative pressure generation so as to maintain airflow in the presence of a narrow UA [7]. During attenuated pharyngeal dilator muscle activity during sleep, the heightened negative inspiratory pressure potentiated by the Bernoulli effect generates further airway narrowing and airway resistance, hence resulting in a vicious cycle leading to eventual complete airway collapse [9, 7]. Another more common scenario is the collapse of the UA at the end of expiration. During this phase of respiration, the airway collapsibility might be the result
of the fall of neuromuscular activity and further influenced by mechanical factors, including a reversal in airflow direction and declining lung volume throughout the expiratory phase [10].

The predominant site of UA collapse during sleep in patients with OSA resides within the pharyngeal segment between the posterior aspect of the hard palate and the glottic inlet. This comprises the nasopharynx, the oropharynx, and the hypopharynx [11]. Studies using the nasopharyngoscope strongly suggest that regardless of the origin of the collapse, eventually the entire pharyngeal segment may collapse [11]. The specific sites of narrowing or closure are influenced by the underlying neuromuscular tone and UA muscle synchrony [8]. Additionally, the stage of sleep plays an important role in UA collapsibility since obstructive events are generally most prominent during rapid-eye-movement (REM) sleep due to more pronounced hypotonia of the UA muscles, than in other stages of sleep [8].

Despite UA obstruction and reduced or absent airflow, the respiratory centre continues to generate respiratory drive to the intercostal muscles and diaphragm so that rib cage and abdominal movements continue against the obstructed UA lumen. This in turn results in paradoxical movements of the rib cage and abdomen, i.e., when one moves inwards the other moves outwards and vice versa. In PSG traces, this phenomenon manifests as an out-of-phase rib cage and abdominal signals, as illustrated in Figure 1.1.

Obstructive apneas elicit a series of mechanical, chemical, hemodynamic, and neural responses that have adverse consequences, particularly on the cardiovascular system, as illustrated in Figure 1.2.1. The futile inspiratory effort against an obstructed lumen results in decreased intrathoracic pressure, which in turn enhances venous return to the right side of the heart and thus a shift of the interventricular septum to the left resulting in impedance of left ventricular filling and reduced stroke volume [4]. The ensuing hypopxia and hypercapnea (increased blood CO$_2$ levels) stimulates the chemoreceptors, which results in vasoconstriction via eliciting sympathetic nervous system activity. Arousal from
Figure 1.1: Polysomnographic tracing of obstructive apneas. Apneas manifest in near flat line in the ‘Sum’ trace, which represent the net movement of the rib cage and abdomen. Rib cage and abdomen traces show a typical out of phase (paradoxical) movement during apneas, while the oxygen saturation drops during and after the apnea. The apnea is terminated by an arousal as shown by increased activity in EEG, EOG, and EMG. EEG=electroencephalography; EOG=electrooculography; EMGsm=submandibular electromyography; SaO$_2$=Oxyhemoglobin.

Sleep takes places at the termination of obstructive apneas, which causes a further burst of sympathetic outflow. The arousal helps to restore pharyngeal dilator muscular tone temporarily, which facilitates UA opening and restoration of airflow until sleep ensues and muscular tone is reduced again. These cycles of apnea and arousal expose the heart and circulation to high amplitude oscillations in central sympathetic nerve traffic, blood pressure, and heart rate [4]. The repetitive arousals result in sleep fragmentation and poor sleep quality, which manifests in daytime sleepiness, fatigue and poor quality of life.

Physiologically, the presence and severity of sleep apnea is most commonly defined by the frequency of apneas and hypopneas per hour of sleep, which is termed the apnea-hypopnea index (AHI), further described in section 1.5.1.
Figure 1.2: Effects of OSA on the cardiovascular system. Obstructive apnea generates negative intrathoracic pressure (Pit) leading to increased left ventricular (LV) transmural pressure. Systemic blood pressure (BP) increases secondary to increased sympathetic nervous system activity (SNA) due to hypoxia and arousals from sleep. Elevated BP and heart rate (HR) increases myocardial O$_2$ demand in the face of a reduced myocardial O$_2$ supply. These conditions predispose a patient acutely to cardiac ischemia and arrhythmias, and chronically could contribute to heart failure. PaO$_2$: O$_2$ Partial Pressure, PaCO$_2$: CO$_2$ Partial Pressure. Bradley and Floras [4]

1.2.2 Pathophysiology of CSA

---

1Reproduced with permission. For this figure, promotional and commercial use of the material in print, digital or mobile device format is prohibited without the permission from the publisher Lippincott Williams Wilkins. Please contact journalpermissions@lww.com for further information.
CSA is significantly less common than OSA in the general population as mentioned above in Section 1.3.1. Therefore, it will be discussed briefly in this section given that this thesis does not aim at distinguishing CSA from the much more common OSA. In contrast to OSA, UA obstruction is not a key factor in CSA. Rather, CSA is a consequence of unstable central respiratory control. Central apnea is triggered when partial pressure of CO$_2$ (PaCO$_2$) is reduced (hypocapnea) to below the apnea threshold during sleep due to preceding hyperventilation. The low PCO$_2$ signals the respiratory centre in the brain to reduce or cease ventilation. Subsequently, the ventilatory drive continues to decrease until breathing becomes very shallow or absent resulting in a hypopnea or an apnea respectively. As a result, CO$_2$ levels build up, which in turn stimulates respiratory drive until hyperventilation ensues again. Hyperventilation results in another episode of hypocapnea and the same cascade of events is triggered again in the form of a pathophysiological vicious cycle. The continuation of that vicious cycle results in a cyclical rhythmic pattern of breathing in which central apneas alternate with periods of hyperventilation that have a characteristic waxing-waning pattern of tidal volume [12] as illustrated in Figure 1.3. This type of breathing is also called Cheyne-Stokes respiration. In contrast to OSA, rib cage and abdominal movements remain in-phase during central hypopnea and are absent in central apnea, because airflow decreases as a consequence of reduced respiratory drive and not UA obstruction.

Although the pathophysiology of central apnea differs from that of obstructive apnea and is not accompanied by the generation of exaggerated negative intrathoracic pressure, it shares several detrimental consequences on the cardiovascular system with OSA [13]. Central apneas increase sympathetic nervous system activity due to hypoxia, which in turn results in elevated blood pressure and heart rate and increased myocardial O$_2$ demand in the face of reduced supply, similar to OSA as explained in Section 1.2.1.
Chapter 1. Background

1.3 Epidemiology of Sleep Apnea

1.3.1 Prevalence of OSA and CSA

Cohort studies have been conducted to estimate the prevalence of OSA in the population. Table 1.1 lists the findings of 3 different cohort studies that have used in-laboratory PSG to examine the occurrence of OSA in the general population. According to the Wisconsin Study, the prevalence of moderate and severe OSA (AHI $\geq 15$) was 9% and 4% in men and women, respectively [14].
Table 1.1: Prevalence of OSA from 3 studies with similar design

<table>
<thead>
<tr>
<th>Study Location</th>
<th>n</th>
<th>Age Range (years)</th>
<th>Estimated Prevalence of AHI ≥ 5 events/hour (% [95% CI])</th>
<th>Estimated Prevalence of AHI ≥ 15 events/hour (% [95% CI])</th>
</tr>
</thead>
<tbody>
<tr>
<td>Wisconsin</td>
<td>626</td>
<td>30–60</td>
<td>24 (19–28) 9 (6–12)</td>
<td>9 (6–11) 4 (2–7)</td>
</tr>
<tr>
<td>Pennsylvania</td>
<td>1,741</td>
<td>20–99</td>
<td>17 (15–20) Not given</td>
<td>7 (6–9) 2 (2–3)</td>
</tr>
<tr>
<td>Spain</td>
<td>400</td>
<td>30–70</td>
<td>26 (20–32) 28 (20–35)</td>
<td>14 (10–18) 7 (3–11)</td>
</tr>
</tbody>
</table>


In contrast to OSA, CSA is far less common in the general population [12]. Bixler et al. have found that the overall prevalence of CSA is 0.4% in men, but 7.8% have central events even if their central AHI is less than the threshold for having the disorder [15]. Therefore, the great majority of subjects referred to sleep laboratories and found to have sleep apnea have predominantly OSA even if they have some central events [12].

The prevalence of CSA in certain patient populations, such as those with heart failure and stroke, is significantly higher than in the general population. In heart failure patients, for example, 15% to 40% of patients were found to have CSA [12, 13]. The main risk factors for CSA are male sex, hypocapnia, atrial fibrillation, and increasing age, but not obesity [13]. Similar to heart failure, the prevalence of CSA is relatively higher in stroke patients than the general population. In-laboratory PSG at the Toronto Rehabilitation Institute revealed that 76% of stroke patients have sleep apnea; 64% with OSA and 12% with CSA [16].

The incidence (i.e., the occurrence of new cases over a given time interval) of sleep apnea is not well known [14] because of paucity of prospective epidemiological studies that have examined this issue. In addition, when a subject is diagnosed with sleep apnea in a sleep disorders clinic or in a cross-sectional epidemiological study, it is not possible
to predict the time he/she started to have the disease. Consequently, it has been proven difficult to determine the incidence of sleep apnea in the general population.

Survival of patients with sleep apnea has been studied in the context of the outcome such as cardiovascular diseases and vehicle accidents, which will be explained below.

1.4 The Undiagnosed Portion of Sleep Apnea

In the Wisconsin Sleep Cohort study, 4,925 middle aged (30-60 years) men and women were examined with in-laboratory PSG [17]. Of those subjects found to have moderate to severe sleep apnea, 93% of women and 82% of men had not been diagnosed prior to participating in the study [17]. This study did not differentiate between OSA and CSA, and thus one would assume these proportions of undiagnosed OSA and CSA are similar, but that the great majority have OSA. If these figures are extrapolated to the Canadian population, then approximately 1,500,000 individuals in Canada have undiagnosed OSA.

1.4.1 The Impact of Sleep Apnea

Sleep Apnea and Cardiovascular Disease

OSA increases the risk of developing hypertension [2], heart failure [5, 3, 4] and stroke by approximately 3 fold compared to subjects without OSA. Since both heart failure (HF) and stroke are major causes of death and disability that consume a large proportion of health care resources, and since OSA is a very common disorder, it is now recognized as a major public health problem. In a Canadian study, it was reported that patients with OSA are more frequently treated for cardiovascular diseases within the 5 years prior to the diagnosis of OSA than subjects without OSA [18]. Costs for health care among these patients with OSA were twice as high as for subjects without OSA [18]. Untreated OSA in patients with heart failure was found to be associated with increased mortality rates independently of confounding factors [19].
In an observational study, Marin et al. [20] studied 1651 healthy individuals (controls), non-apneic snorers, patients with untreated mild and severe OSA, and OSA patients treated with continuous positive airway pressure (CPAP), as shown in Figure 1.4. Participants were followed up over a mean of 10 years. End points were either fatal cardiovascular events (death from myocardial infarction or stroke) or nonfatal cardiovascular events (nonfatal myocardial infarction, nonfatal stroke, coronary artery bypass surgery, and coronary angioplasty). The risks of fatal and nonfatal cardiovascular events were approximately 3 times higher in patients with untreated severe OSA compared to healthy subjects. On the other hand, risk for events in patients with CPAP-treated OSA was similar to that of non-apneic snorers and those with mild OSA (Figure 1.4), which suggests but does not prove that treating OSA reduces cardiovascular event rates. It is noteworthy, however, that this study was not a randomized controlled study and treatment was provided according mainly to physician judgement and patient preference, which is a limitation of that study. Another randomized controlled clinical trial has shown contradictory results i.e. CPAP treatment of OSA has no significant effect on cardiovascular event rate [21]. Therefore, the evidence of CPAP on cardiovascular events in OSA is deemed as being inconclusive.

Similar to OSA, CSA has a detrimental effect on HF patients; CSA is an independent risk factor for death or cardiac transplantation [13]. Unlike OSA, however, CSA is thought to be a result of HF, not a cause, due to chronic hyperventilation and hypocapnea [13]. Thus, impaired cardiac function in heart failure explains the relatively high incidence of CSA in those HF patients.

**Sleep Apnea and Motor Vehicle Accidents**

Sleep apnea results in poor sleep quality due to sleep fragmentation caused by arousals. Therefore, patients often suffer from excessive daytime sleepiness, which predisposes them to motor vehicle accidents. In a study that included 913 adults, Young et al. found that
men and women with an AHI > 15 were significantly more likely to have multiple accidents in 5 years compared with those with no disease (AHI < 5) with an odds ratio of 7.3 [22]. In a Swiss study, Horstmann et al. showed that patients with moderate to severe OSA syndrome have up to 15-fold higher risk of motor vehicle accidents compared to controls [23]. In patients with severe sleep apnea (AHI > 34/h, n=78), the motor vehicle accident rate was 13.0 per million km as compared to 1.1 in patients with milder OSA (AHI 10-34/h, n=78), and 0.78 in a control group (AHI < 10), respectively. The investigators of this study used a strictly anonymous questionnaire to assess the occurrence of accidents as opposed to Young et al. who used official traffic databases, which allowed them to account for accidents that might not have been reported. Risks of motor vehicle accidents associated with OSA from various studies are presented in Figure 1.5.

Sleep apnea-related motor vehicle accidents exert a huge financial burden on the economy. Sassani et al. reported that OSA related vehicle accidents cost $15.9 billion and resulted in 1,400 deaths in the year 2000 in the U.S [24]. According to their analysis, treating all drivers with OSA would cost $3.18 billion, which would save $11.1 billion in collision costs, and 980 lives annually.
Figure 1.5: A meta-analysis Forest Plot of risk of collisions in drivers with OSA reported by Sassani et al [24]. The odds ratios of 6 published studies (on the left column) and the corresponding odds ratio and the 95% confidence interval (on the right column) are presented.

1.5 Diagnosis of Sleep Apnea

1.5.1 Overnight Polysomnography

The current gold standard method of choice for diagnosing sleep apnea is overnight PSG [25]. The patient has to sleep in a sleep laboratory attached to many monitoring electrodes under the supervision of a technician. PSG involves the measurement of several physiologic recordings including the electroencephalogram (EEG), electrooculogram (EOG), electrocardiogram (ECG), submental and leg electromyograms (EMG), body position, pulse oximetry, measurements of airflow, and measurements of thoracic and abdominal respiratory effort [25], as illustrated in Figure 1.6. PSG is used to record those physiologic variables in order to diagnose a wide spectrum of respiratory and non-respiratory disorders of sleep [26].

Respiratory movements of the thorax and abdomen are the most important signals
for diagnosis of sleep apnea. Respiratory inductance plethysmography (RIP) is the gold standard for detecting thoracoabdominal movement and respiratory activities including apneas and hypopneas and the discrimination between OSA and CSA [26]. EEG records global neural activity from electrodes placed on the patient’s scalp. It is performed to monitor whether the patient is awake or asleep, and if asleep, what sleep state he/she is in, in addition to recording arousals from sleep that may or may not be associated with respiratory events [26]. Eye movement (EOG) and submental EMG are crucial for distinguishing wakefulness and rapid eye movement (REM) sleep from other sleep stages [26]. ECG is used to detect cardiac rhythm and identification of nocturnal arrhythmias with 2 to 3 electrodes on the chest. Pulse oximetry is used to detect reductions in blood oxygen saturation as a result of apneas and hypopneas. Data output of PSG channels is used to identify individual apneas and hypopneas as illustrated in Figures 1.1 and 1.3 in Section 1.1. Scoring criteria for apneas and hypopneas are explained in details in Section 7.3.3.

Quantification of the Severity of Sleep Apnea

AHI is the most commonly used variable to determine the presence and physiological severity of sleep apnea. The classification of the severity of sleep apnea based on AHI (as practiced in the Sleep Laboratory of Toronto Rehabilitation Institute) is as follows:

<table>
<thead>
<tr>
<th>Diagnosis</th>
<th>AHI</th>
<th>Severity</th>
</tr>
</thead>
<tbody>
<tr>
<td>Normal</td>
<td>&lt; 10</td>
<td>N/A</td>
</tr>
<tr>
<td>Sleep Apnea</td>
<td>10 − 20</td>
<td>Mild</td>
</tr>
<tr>
<td></td>
<td>20 − 30</td>
<td>Moderate</td>
</tr>
<tr>
<td></td>
<td>≥ 30</td>
<td>Severe</td>
</tr>
</tbody>
</table>
1.5.2 Sleep Apnea Syndrome

Patient evaluation is performed based on laboratory results in addition to signs and symptoms. Therefore, sleep apnea syndrome is defined as an AHI above the normal level as found by overnight PSG plus one of the following: excessive daytime sleepiness that is not explained by other factors, or two of the following symptoms: choking or gasping during sleep, recurrent awakening from sleep, unrefreshing sleep, daytime fatigue, and/or impaired concentration [27]. The AHI threshold for sleep apnea is defined as 5 apneas and hypopneas per hour according to the American Association of Sleep Medicine criteria or 10 according to others as explained in detail in Section 7.3.3. Accordingly, AHI is one, but not the only, factor in establishing the diagnosis. Both AHI and clinical presentation
need to be examined for the diagnosis and treatment of sleep apnea.

1.5.3 Sleep Studies in Ontario and Around the World

In Canada and Ontario, approximately 370 and 776 PSGs per 100,000 individuals, respectively, are performed annually [28]. The Ontario Ministry of Health has indicated that in Ontario the number of PSG studies performed in 2006 was approximately 115,000 at a cost of approximately $58 million. Figure 1.7 shows the number of sleep studies performed in Ontario from the year 1999 to 2007. Corresponding rates internationally are 427/100,000 in the United States, 42.5/100,000 in the United Kingdom, 177/100,000 in Belgium, and 282/100,000 in Australia [28]. Based on these numbers, the Ontario Medical Advisory Secretariat stated that “The rate of sleep studies performed in Ontario is thus very high in relation to other provinces in Canada, as well as other countries.” They also stated that “There may be a 10-fold rise in the rate of sleep tests in the next few years.”

![Figure 1.7: Total annual sleep tests in Ontario. (Obtained from direct communication between the Ontario Ministry of Health and the Toronto Rehabilitation Institute.)](image)

Although attended PSG is a comprehensive and reliable means for diagnosing sleep apnea, it is expensive and access to it is limited, resulting in long waiting lists [26] reaching up to 2 years in some places in Canada. This drove many researchers to find strategies to eliminate the shortcomings of PSG. One strategy is using unattended monitoring systems
outside the laboratory, i.e., portable home monitoring.

1.5.4 Portable Monitoring of Sleep Apnea

Due to the large number of undiagnosed subjects and the huge economic burden of diagnosing and treating sleep apnea syndrome, there have been numerous attempts to develop more accessible and economical monitors for diagnosing sleep apnea. An important aspect in this development is making this monitor usable in patients’ homes in order to save them spending a night in a sleep laboratory, in other words making it portable. Such a portable diagnostic device could be very widely applicable due to the large number of individuals with undiagnosed sleep apnea.

Several attempts took place to create reliable monitors by research and commercial entities. The greatest majority of those devices consisted of a subset of PSG electrodes. The American Sleep Disorders Association (ASDA, now known as American Academy of Sleep Medicine) classified the different monitors used for sleep studies into 4 categories, conventionally known as level (or type) I to IV, based on the number of channels they evaluate and record [29, 30] as presented in Table 1.2. Level I is full attended PSG. The numbers of channels decreases progressively in levels II and III, down to as low as 1 channel in level IV.

<table>
<thead>
<tr>
<th>Type or Level</th>
<th>Portability</th>
<th>Indicative N_channels</th>
<th>Indicative signals</th>
</tr>
</thead>
<tbody>
<tr>
<td>I</td>
<td>Facility-based</td>
<td>14-16</td>
<td>EEG, EOG, EMG, ECG/HR, airflow, effort, SaO₂</td>
</tr>
<tr>
<td>II</td>
<td>Portable</td>
<td>7-13</td>
<td>(may have EEG), HR, EOG, chin EMG, ECG/HR, airflow, effort, SaO₂</td>
</tr>
<tr>
<td>III</td>
<td>Portable</td>
<td>4-6</td>
<td>Airflow and/or effort, ECG/HR, SaO₂</td>
</tr>
<tr>
<td>IV</td>
<td>Portable</td>
<td>1-3</td>
<td>Airflow, HR, peripheral arterial tonometry, SaO₂</td>
</tr>
</tbody>
</table>

Table 1.2: Classification of sleep apnea monitors (Trikalinos and Lau [29])
Evaluation of Currently Existing Portable Monitors

In their extensive systematic review that covered 50 studies, Flemons et al concluded that there is no convincing evidence that any of the available portable monitors (types II, III, or IV) could be used in un-attended settings and still provide reliable signals [31]. In agreement with Flemons et al., in a later review, Ahmed et al. concluded that there is a lack of evidence to support the validity of available portable monitors to date for use within the general population [32]. This demonstrates the present weakness in the area of portable sleep apnea monitoring. I attribute the weakness of the current technology to the dilemma of developers to balance between the portability of the device and ease of use on the one hand versus the sophistication, accuracy, and reliability of the device on the other hand. The challenge in developing a portable sleep apnea monitor is creating a simple device with a minimal number of channels (sensors) such that a lay person is able to use it in the home without professional help, yet is reliable and accurate in making a diagnosis. However, devices that employ fewer channels have been found to be less accurate than devices with more monitoring channels [33]. However, as more channels are added to a portable sleep apnea monitor to improve accuracy, the failure rates increase to as high as 33% [34] due to the inability of the patient to properly connect the electrodes and accidental detachment of electrodes during sleep. This paradox is illustrated in Figure 1.8. I believe that a successful portable monitor for sleep apnea should overcome this challenge by providing sufficient information to make an accurate diagnosis, yet with a structurally simple design to make it usable in unattended settings.
1.6 Treatment of Sleep Apnea

1.6.1 Treatment of OSA

The objective of treatment of OSA is relieving airway obstruction in order to restore airflow at night and improve sleep continuity and alleviation of excessive daytime sleepiness and other symptoms of sleep apnea. The first line of treatment is patient education on sleep apnea. Patients should be educated on the sleep apnea pathophysiology, impact of weight loss, sleep position, avoiding alcohol, modification of risk factors, and effect of medications on sleep apnea [35]. Patients with supine related OSA are advised to avoid sleeping in the supine position using a method that keeps the patient in a non-supine position [35]. Since obesity is a known risk factor of OSA, weight loss is recommended for overweight patients and can lead to reduced severity, improved AHI, and in some cases, resolution of OSA. If these measures fail or are not sufficient for treating OSA, then more aggressive approaches are used, as follows:

Continuous Positive Airway Pressure (CPAP) is currently the most common treatment for OSA. CPAP provides air under constant pressure throughout the respiratory cycle and transmitted through a hose to a tight-fitting nasal, oronasal, or oral mask. It functions by distending the UA and allowing uninterrupted airflow. Treatment of OSA by CPAP has been shown to alleviate daytime hypsomolence, and improve the quality of life and neurocognitive function [36, 37]. An observational study suggests
that it may also reduce traffic accident rates [38].

Depending upon the patient’s clinical situation, positive airway pressure can be applied in other modes. For example, the autotitrating CPAP mode (AutoPAP or APAP) can automatically adjust the level of CPAP to the minimum required to maintain UA patency [35]. So, for example, it will increase pressure when UA resistance increases such as when moving from the lateral to the supine position and when going from non-REM to REM sleep to overcome the higher UA resistance in these two conditions. Another mode, the bilevel positive airway pressure (BPAP) devices are usually employed in patients with coexisting OSA and hypoventilation who require ventilation support such that inspiratory pressure exceeds expiratory pressure [35]. Although CPAP has been shown to be an effective treatment for OSA, poor compliance is the main problem in long-term care of patients with OSA.

**Oral Appliances** are devices consisting of articulated upper and lower bite plates worn by the patient during sleep. By adjusting the lower bite plate to displace the mandible forward, they move the tongue anteriorly thus increasing the UA size. Oral appliances are typically fitted by a dentist specializing in sleep apnea and most require adjustment of the distance of protrusion. They are usually reserved for non-obese patients with mild to moderate OSA who cannot tolerate CPAP; however, they are generally less effective than CPAP and have some side effects, such as difficulty chewing, dry mouth, occlusive changes, jaw discomfort, and temporo-mandibular joint pain [39].

**Surgery** for removing excess pharyngeal soft tissue is another modality for treating OSA [39]. This option is frequently undertaken when there is a definable anatomical cause of UA obstruction, such as tonsillar or adenoid hypertrophy, most often in children. In adults, there are a variety of surgical options involving the resection of various combinations of the uvula, soft palate and peripharyngeal tissue including uvulopharyngopalatoplasty, either by surgery, laser, or radiofrequency ablation, as well as septoplasty, rhinoplasty, and partial glossectomy [39]. Also, repositioning of muscle attachments such
as genioglossus advancement, maxillary, or maxillomandibular advancement have been employed.

### 1.6.2 Treatment of CSA

Although CSA can occur in the absence of HF or other conditions, in which case it is termed idiopathic CSA, CSA is most often a manifestation of advanced heart failure; therefore, the first line of CSA treatment is to optimize medical therapy for heart failure [12]. If CSA persists, then specific therapy for CSA can be tried. Acetazolamide, a drug initially used to treat and prevent acute mountain sickness, has been shown to attenuate CSA in HF patients with HF [40, 41]. However, its effect on clinical outcome has not been tested. Its main mechanism of action is induction of metabolic acidosis that causes constant drive to breathe thus preventing central apneas. The use of supplemental oxygen also attenuates CSA in HF patients by reducing hypoxic respiratory drive and thus damping hyperventilation between apneas [40]. However, its effect on long-term clinical outcome has not been tested.

Initially, small, short-term, randomized trials of CPAP were shown to attenuate CSA and improve cardiovascular function in HF [13]. A larger, long-term, multi-center trial to test the effects of CPAP on the combined rate of mortality and cardiac transplantation in HF patients with CSA, the CANadian Positive Airway Pressure trial for patients with congestive heart failure and central sleep apnea or CANPAP, found that although CPAP attenuated CSA, improved nocturnal oxygenation and left ventricular ejection fraction and 6-minute walking distance, and reduced sympathetic nervous system activity, it had no effect on heart-transplant free survival [42].

Adaptive Servo Ventilation (ASV) is another form of positive airway pressure designed specifically to treat CSA in patients with heart failure. When it detects central apnea, it generates inspiratory pressure to support ventilation and abort the apnea. When it detects airflow generation by the patient, it shuts off. Although ASV is more effective
than CPAP in alleviating CSA [43], it has not been shown to improve long-term outcome in patients with heart failure. Currently 2 international multi-centre trials to test the effect of ASV on morbidity and mortality in heart failure patients, ADVENT-HF [44] and SERVE-HF [45] are undergoing.

1.7 Summary

Sleep apnea is a major public health problem that affects a significant portion of the adult population. Untreated sleep apnea appears to increase the risk of cardiovascular mortality and morbidity. The most frequent manifestation of sleep apnea is excessive daytime sleepiness that causes deterioration in quality of life and cognitive performance, and increase in motor vehicle accident rates. Treatment of sleep apnea reduces daytime sleepiness, improves quality of life and neurocognitive function, and may reduce accident rates.

Despite this evidence, majority of subjects with sleep apnea remain undiagnosed and untreated. This is attributed partially to the high cost of and limited access to attended PSG. On the other hand, home-based (portable) devices lack accuracy or ease-of-use required for implementation in unattended settings. The goal of this work, therefore, is to investigate breath sounds as a tool for accurate diagnosis of sleep apnea that can be implemented in portable settings.
Chapter 2

Review of Acoustic Analysis of Breath Sounds

Acoustic analysis of respiratory sounds has gained an increasing role in the diagnosis of respiratory disorders because of advances in techniques for sound measurement and signal analysis [46]. Respiratory sound analysis has been used to identify pathological respiratory sounds, such as wheezes and crackles [47, 48, 49, 50, 51, 52, 53]. Sound recordings have also been used during sleep to diagnose snoring and locate the site from which it arises [54, 55, 56]. In these studies the microphone used to capture the acoustic signal was located either on the upper lip [54, 55], on the forehead [57], on the trachea [58, 59], or suspended above the patient [56]. Breath sound analysis, and snoring in particular, has also been used to distinguish patients with OSA from simple snorers [60, 61, 62, 63].

This review is a summary of published literature on techniques used to detect sleep apnea by means of breath sounds analysis.
2.1 Necessary Definitions

The term breath sounds’ in this review refers to all types of sounds produced while breathing during sleep, including normal breathing and snoring, captured using a microphone over the nose, mouth, or trachea. Since snoring sounds have so far been the most studied type of breath sound, it is worthwhile providing a definition for snoring. Existing definitions such as that of American Sleep Disorders Association (1990) define primary snoring as “loud UA breathing sounds in sleep, without episodes of apnea or hypoventilation [63].” Some researchers, including those in the field of acoustic analysis, have introduced a definition of snoring that characterizes its production mechanism [63, 64]. However, those definitions considered only snoring that involves tissue vibration with a definable fundamental frequency [64] or pitch [63]. The author of this thesis did not find a comprehensive definition of snoring that fully encompasses its physical characteristics in the literature.

The author proposes a modified version of the definition of Salmi et al. [64] of snoring: snoring is a harsh, discontinuous respiratory sound occurring usually during sleep with or without periodic components (definable fundamental frequency) that is induced by vibrations in the walls of the oropharynx or by passage of air through a constricted airway without vibration. It is typically an inspiratory sound but a small expiratory component can appear, especially in patients with OSA.

2.2 Acoustic Techniques for Diagnosing Sleep Apnea

2.2.1 Time Domain-Based Techniques

Cavusoglu et al. found that snoring episodes have different temporal regularity in OSA patients compared with non-apneic patients [60]. Snores recorded from OSA patients had larger variability in duration, separation between episodes, and power. This was
not linked to the severity of the disease nor to the underlying mechanism, although
abnormalities of breathing in OSA were the catalyst for investigating snoring regularity
in that study.

Finnish researchers [58, 59] have examined the amplitude variations of tracheal sounds
recorded from a microphone located on the suprasternal notch. In their work, they
reduced the resolution of the recording by taking only maximum and minimum values of
successive 20-second intervals. Subsequently, an amplitude modulation or envelop of the
breathing sounds was formed. This approach facilitated the visual distinction between
segments of regular breathing, snoring, and apneas [58]. The next step was automating
the process by running non-linear filters on the same time domain representation, which
basically detected local extreme values and their ratios in a sliding window [59]. The ratio
of the local range to higher extreme values was the most reliable feature for distinguishing
the 3 types of breathing. This method relied on deflections of the long-term envelop
disregarding individual apneas. Therefore, it cannot be used to provide an accurate
count of apneas as required in clinical practice, although it has potential for use as a
screening method. Another study that used tracheal sounds depended on the drop of
sound energy recorded over the trachea and its duration [65], which was the only study
in which the number of apneas and hypopneas was counted. The investigators, however,
required the use of an extra signal, oxygen saturation, in order to support their algorithm
when tracheal sounds were insufficient to make a decision. This presents a limitation in
the translation of this work into a simple portable device.

2.2.2 Frequency-Domain-Based Techniques

Several researchers have examined the frequency characteristics of snoring sounds, at-
ttempting to obtain information about the site and level of obstruction and the degree
of narrowing of the UA in OSA. Canadian researchers pioneered this field by their first
attempt to distinguish patients with and without OSA using spectral characteristics of
snores [66]. They showed that post-apneic snores consisted mainly of broad-band noise with relatively more power at higher frequencies, a finding noted repeatedly by other researchers [61, 62, 67]. The broad-band noise pattern was also observed by Fiz et al. in the majority of OSA patients as opposed to identifiable harmonics that prevail in simple snorers [68]. However, this study reported a negative correlation between the severity of the disease and peaking frequencies, which contradicts the majority of other studies.

A group of otolaryngologists [61] found that snoring that takes place in simple snorers has different harmonic and frequency distribution than that in patients with OSA. They observed rhythmical repetitions in spectral analysis and harmonics of the Fast Fourier Transform (FFT) in the former type, which indicates a vibrating structure such as the uvula or the velum. On the other hand, the spectral analysis of snoring sounds in patients with a high AHI revealed a high-frequency, non-harmonic noise pattern, and absence of rhythmic repetitions, as presented in Figure 2.1. The latter was explained as being due to the sound of air passing through a constricted airway at the base of the tongue. One should notice, however, that snores included in this study were quite selective. Only the first snore after a group of apneas was selected as an apneic snore, which does not necessarily reflect all types of breath sounds in between apneas. Thus, the conclusion about the mechanism of this sound production could only be valid in the light of assumptions drawn from previous studies about the site and mechanism of obstruction.

A class of features derived from the frequency spectrum of sounds produced in the UA are called formant frequencies. Formants are maximal energy frequencies that could be identified by peaks in the linear prediction coding (LPC) spectrum [62]. They represent resonant frequencies of the UA, which can be influenced by the shape and size of the UA [62]. It has been shown that the 1st formant (F1) is associated with the degree of constriction in the pharynx; F2 is related to the degree of advancement of the tongue relative to its neutral position; and F3 is correlated with the degree of lip-rounding,
Figure 2.1: Left: an example of a spectrogram (top) and amplitude spectrum (bottom) of a snore in a patient after an obstructive apnea. Right: an example of a spectrogram (top) and amplitude spectrum (bottom) of a snore that is not associated with an obstructive apnea. Source: Michael et al. [61].

(although lip-rounding affects all formants in principle.) Therefore, it is plausible that formant frequencies of apneic snorers and the benign snorers are different. Ng et al. examined the first 3 formant frequencies of snoring sounds recorded from subjects with and without sleep apnea [62]. They found that only F1, with a threshold of 470 Hz, was significantly higher in apneic snoring and can best classify apneic snores from benign (non-apneic) ones as illustrated in Figure 2.2. This study included snores only with sharp rising and falling peaks, which suggests that all the analyzed snores were of a periodic nature, which does not necessarily represent all types of snores that accompany sleep apnea as explained earlier. Subsequently, they expanded their test on natural snores collected from 40 subjects [67]. Their results persistently suggested that higher values
of F1 can better identify apneic snorers from non-apneic ones. Using receiver operating characteristic (ROC) curves, the cut-off frequency of F1 was 513 Hz for males, 327 Hz for females, and 497 Hz for both combined. Formant frequencies were also found in sound recordings taken from microphones on the trachea [69]. However, the author of this thesis believes that the spectral peaks found in this recording do not represent UA resonance, which does not make them valid formants, because at that location the sound will not have passed through the UA acoustic filter that shapes and determines the formant structure. Therefore, works using tracheal recordings will not be covered in this review. However, tracheal sounds captured over the neck can still be useful as explained below.

**Figure 2.2:** Notched box plots of apneic and benign (non-apneic) snores for males (M), females (F), and for both males and females combined (C), under first formant frequency (F1) analysis, with marked threshold values. Only F1, with a threshold of 470 Hz, was found to be significantly higher in apneic snoring and can best classify apneic snores from benign (non-apneic) for males and females combined. Source: Ng et al. [62].

Another frequency domain variable is pitch, which is the fundamental frequency of
a signal if present. Autocorrelation with centre clipping was used in order to detect the pitch of snores. Abeyratne et al., showed that individual snores have in fact variable pitch, which they called ‘pitch jump’ and proposed the intra-snore pitch jump probability as a feature to detect OSA patients. This resulted in 86-100% sensitivity but a relatively low specificity at 50-80% [63].

The Welch spectrum has also been used as a frequency domain feature. With 50% overlap of 10 minute segments, researchers found that segments with airflow limitation, as confirmed by nasal pressure transducer, contained significantly higher frequencies in all bins above 100 Hz up to 5512 Hz than segments of normal breathing and those containing full apneas [70]. It was supposed that during flow limitation, the UA constriction would cause airflow to interact more with the vocal tract resulting in high frequency components of sound. A drawback of this technique is that the segments’ durations chosen for analysis in that study were noticeably long, which makes it prone to contamination by transient events. For example, sharp transient noise peaks, such as those due to cough, movement etc, will translate to wide band noise in the frequency representation.

2.3 Acoustic Techniques for Detecting Dynamics of the UA During Sleep

Since obstructive apneas are directly related to narrowing of the UA during sleep, several attempts have been made to infer the occurrence of UA narrowing from acoustic analysis. Earlier studies looking into that used linear predictive coding and the resultant formant frequencies to retrieve information about the caliber of the UA [71, 72]. In recent years,

---

1Pitch can also be considered a time-domain feature since it is derived from the interval (period) between successive waves in the time-domain representation of the signal. We refer to it as a frequency-domain feature, since we present it as the frequency of successive tissue collisions in snoring expressed in Hz.
Ng et al. have utilized a transmission line model of the UA and generated synthetic snores based on the source filter theory that is well established in the speech processing field [67]. They calculated the peak frequency (PF) from the Welch power spectrum and the first formant (F1) from the discrete all pole modeling algorithm, which they argued was more reliable than the linear prediction algorithm for their study [62]. Using that model, they demonstrated that narrowing of the pharyngeal airway with mouth opening can increase the value of F1 and occasionally PF.

Abeyratne et al. claimed that higher order statistics are needed for modeling the UA response to the source excitation sequences, in this case snoring or normal breathing [73]. They proposed a mixed-phase model of the UA and demonstrated that the total airway response differed in post-apneic snores in comparison with other snores and with normal breathing. This study was an exploratory work in which the algorithm was introduced without validation by distinguishing between patients with and without OSA.

### 2.4 Summary of Acoustic Analysis of Breath Sounds

Breath sounds are rich in features that could provide information about breathing patterns, flow limitation, site of obstruction, caliber of the airway, and presence or absence of sleep apnea. Several attempts have been made to develop an alternative diagnostic method for sleep apnea using acoustic analysis of breath sounds. The vast majority of studies that examined breath sounds have focused on snoring sounds. It stands to reason that many researchers have taken this direction, because snoring is a result of UA narrowing during sleep, which in the extreme case is a cause of OSA. Most of these studies examined snoring sounds as a tool to distinguish between snoring with or without OSA. This was achieved using time domain features, such as temporal distribution, and frequency-domain features, such as spectral peaks.

In summary, breath sound recording is a potentially valuable tool that could be used
in a simple portable device for diagnosing sleep apnea.

## 2.5 Goals and Objectives of This PhD Thesis

Despite several promising steps towards the development of a portable monitor for diagnosis of sleep apnea, the currently available technology suffers from considerable limitations. The purpose of this dissertation is the development of a portable device for the diagnosis of sleep apnea that addresses the hardware and software limitations of currently available technologies.

### 2.5.1 Addressing Hardware and Device Montage Limitations

After examining the literature on portable devices (Section 1.5.4), it can be clearly seen that the current technology suffers many limitations including the lack of reliability in unattended settings, the high failure rates of complex devices due to inappropriate use and electrode detachment, and the poor accuracy of simple devices. My goal, therefore, was to develop a method for diagnosing sleep apnea in portable settings that combines simplicity and accuracy (Figure 1.8). The strategy for achieving this goal was to use the minimum number of physical channels yet derive multiple physiological measures from the acoustic analysis that can be used in combination to increase the reliability in diagnosing sleep apnea. These measures include the dynamics of the UA and the occurrence of apneas and hypopneas as will be discussed in individual chapters of this thesis. This strategy is illustrated in Figure 2.3. Breath sounds were the signal of choice for this work because they offer the single richest signal with information on breathing patterns. The techniques for analyzing breath sounds were developed to accurately identify apneas and hypopneas regardless of the presence or absence of snoring, which was not sufficiently addressed in previous studies as (Section in 2.4).
2.5.2 Addressing Software and Signal Processing Limitations

The vast majority of studies that analyzed breath sounds to detect sleep apnea examined snoring as the sole input, and were designed merely to detect the presence or absence of OSA. However, in medical practice, knowledge of the disease severity in terms of AHI is essential. Based on this information, in combination with the clinical assessment, clinicians can decide on the treatment strategy and follow up their patients to assess their response to it. In addition, snoring does not necessarily occur in CSA. For this reason, studies that used snoring analysis were aimed at detecting OSA, and relevant issues such narrowing of the UA and its site. In certain patient groups, such as heart failure and stroke patients, the prevalence of CSA is much higher than in the general population. Therefore, an absence of snoring in CSA will render snore-driven techniques less useful in detecting apneas and hypopneas in such patients. Therefore, another goal of this thesis is to translate the acoustic analysis of breath sounds into a clinically useful tool in such a way that the severity of the disease can be predicted in terms of the AHI, regardless respiratory events. It should be noted, however, that the distinction between OSA and CSA is not one of the objectives of this thesis.

2.5.3 Thesis Content Overview

This PhD thesis will describe the steps undertaken to develop a portable home monitor including processing of breath sounds, its in-lab validation, and field implementation. The first step was to detect the basic constituents of breath sounds, i.e., breathing phases (inspiration and expiration), which are utilized in the subsequent steps. Next, I investigated the ability of acoustic analysis of breath sounds to detect pathological changes that occur during OSA due to UA narrowing. The subsequent studies were performed using data collected from clinical subjects. In the first, a method was introduced to preprocess breath sounds to make them suitable for subsequent pattern recognition. Subsequently,
Figure 2.3: Comparison between the approach of standard sleep apnea monitors and the approach proposed for this project.

A rule-based pattern recognition algorithm was developed in order to detect apneas and hypopneas, which was validated in 50 subjects against PSG. In the last study presented in this thesis, a portable device that implements the acoustic analysis of breath sounds was introduced and examined in the patients’ home environment.
Chapter 3

Identification of Breathing Cycle

Phases, Part I

The content of this chapter was published as a peer reviewed journal article: Hisham Alshaer, Geoff R. Fernie, T. Douglas Bradley, Monitoring of Breathing Phases Using a Bioacoustic Method in Healthy Awake Subjects, J Clin Monit Comput. 2011 Oct;25(5):285-94. The paper has been reproduced according Springer’s self-archiving policy.¹

Except for the formatting and some organizational and stylistic improvements, the content of this chapter is almost identical to the aforementioned publication. Contribution of authors is as follows: Alshaer did the feature extraction of respiratory phases, developed the algorithm, and wrote the bulk of the paper. Fernie supervised the technical aspects of this work and edited the mathematical content of the paper. Bradley provided clinical and laboratory facilities and supervised the protocol and medical aspects of this work and edited medical content of this paper.

In order to fully utilize the richness of the breath sounds in evaluating the underlying

¹Springer allows author-created version of his/her article on his/her institutional repository. Springer’s self-archiving policy can be found at this link: http://www.springer.com/open+access/authors+rights?SGWID=0-176704-12-683201-0

33
conditions, breath sounds should be separated into their basic components that correspond to various breathing activities. Different breathing phases are associated with different physiological and pathophysiological phenomena. In this stage of the work, features that distinguish inspiratory and expiratory sounds are examined. The work was performed in two setups; one on awake subjects (Part I) and the other on sleeping subjects (Part II), which are presented in this and the following chapter respectively. This part of the work establishes the utility of breath sounds in tracking breathing phases, which was then used in sleeping subjects. I implemented frequency-domain characterization of the signal, which has been an established method in biomedical signal analysis [74]. In this chapter, I also introduced an algorithm for real-time tracking of breathing phases. Identification of breathing phases achieved in this step was used in performing subsequent data analysis of breath sounds, such as detecting UA narrowing, presented in Chapter 5.

3.1 Abstract

Objective: To test the ability of a microphone recording system, located distal to the respiratory outflow tract, to track the timing of the inspiratory and expiratory phases of breathing in awake healthy subjects.

Methods: Fifteen subjects participated. Breath sounds were recorded using a microphone embedded in a face frame in a fixed location in relation to the nostrils and mouth, while simultaneously recording respiratory movements by respiratory inductance plethysmography (RIP). Subjects were studied while supine and were instructed to breathe normally for 2 minutes: through their noses only (nasal breathing), during the first minute, and through their mouths only (oral breathing) during the second minute. Five subjects (test group) were chosen randomly to extract features from their acoustic data. Ten breaths (5 nasal and 5 oral breaths) from each subject were studied. Inspiratory and
expiratory segments of breath sounds were determined and extracted from the acoustic data by comparing it to the RIP trace. Subsequently, the frequency spectrum of each phase was then determined. Spectral variables derived from the 5 test subjects were applied prospectively to detect breathing phases in the remaining 10 subjects (validation group.)

**Results:** Test group data showed that the mean of all inspiratory spectra peaked between 30 Hz and 270 Hz, flattened between 300 Hz and 1100 Hz, and peaked again with a center frequency of 1400 Hz. The expiratory spectra peaked between 30 and 180 Hz and its power dropped off exponentially after that. Accordingly, the bands ratio (BR) of magnitude values of the frequency spectrum between 500 - 2500 Hz to those between 0 - 500 Hz was chosen as a feature to distinguish between breathing phases. BR for the mean inspiratory spectrum was 2.27 and for the mean expiratory spectrum was 0.15. The route of breathing did not affect the BR ratio within the same phase. When this BR was applied to 436 breathing phases in the validation group, 424 (97%) were correctly identified (Kappa = 0.96, p<0.001) indicating strong agreement between the acoustic method and the RIP.

**Conclusion:** Frequency spectra of breathing sounds recorded from a face frame, reliably identified the inspiratory and expiratory phases of breathing. This technique may have various applications for respiratory monitoring and analysis.

**Key Words:** Acoustic, Respiratory Monitoring, Breathing Phases, Frequency Spectrum

### 3.2 Introduction

Several clinical conditions require close monitoring of respiratory activity including respiratory failure, respiratory tract infections, as well as respiratory depression associated with anesthesia and sedatives [75, 76, 77]. The recovery phase after anesthesia is a period
of increased risk of respiratory depression that requires continuous monitoring of vital circulatory and respiratory functions. The diagnosis of OSA [6] also requires continuous respiratory monitoring.

Besides monitoring the respiratory rate, the timing of the inspiratory and expiratory phases of breathing is essential in several situations. For example, echocardiographic investigations of transvalvular and venous flows are respiratory phase-related and thus require knowledge of the breathing phases [78, 79]. When performing computed tomography and magnetic resonance imaging, it is also necessary to gate image acquisition to a particular phase of the respiratory cycle to reduce image blurring caused by respiratory motion artefact [80]. In respiratory physiology, many experiments require monitoring of breathing phases, such as measuring end tidal PCO$_2$ or changing the concentration of inspired gas content [81, 82].

Clearly, the development of improved non-invasive, easy to apply, and reliable methods for monitoring of breathing would enhance the ability to detect breathing abnormalities in all of the above situations. The currently available respiratory monitoring techniques generally involve detection of chest wall movement, sensing of airflow, or determination of arterial blood gas composition [75], which have been well established in the clinical practice and research fields. However, each one of these methods requires the attachment of a dedicated physical transducer on the patient’s body such as chest bands, a nasal cannula, or an oximeter. Therefore, a technique that could be integrated or combined with existing medical equipment will reduce the number of physical channels and ameliorate the burden on the patient and caregiver. For example, if a face mask that is already in use for oxygen delivery can also be used for respiratory monitoring, this will not increase the net number of physical transducers or electrodes on the patient’s body during the medical procedure.

The aim of this study was to test the ability of a microphone recording system located distal to the respiratory outflow tract embedded in a face mask to monitor breathing in
awake healthy volunteers. I hypothesized that analyzing the frequency components of breath sounds will reliably identify individual breathing cycles as well as their inspiratory and expiratory components.

3.3 Methods

3.3.1 Subjects

I recruited, by advertisement, healthy subjects at least 18 years of age with no history of respiratory or cardiovascular disease, and who were not taking prescribed medications: 15 subjects, 6 men and 9 women, participated. Subjects were divided randomly into 2 groups with 5 subjects in one group (test group) and 10 in the other (validation group). The data from the test group were used to examine acoustic characteristics of breathing phases and extract distinguishing features. Acoustic features were then incorporated into an algorithm that was applied to breath sounds from the validation group to determine how accurate they were at distinguishing breathing phases. The study protocol was approved by the Toronto Rehabilitation Institute Research Ethics Committee, and all subjects provided written consent prior to participation.

3.3.2 Breath Sound Recording

Breath sounds were recorded using a unidirectional, electret condenser microphone (Knowles Acoustics, Model MB6052USZ-2). The microphone’s unidirectional pattern reduces pickup of sounds from the sides and rear, improving isolation of the sound source. The microphone was embedded in a respiratory mask that was modified by cutting away material so that only a structural frame remained to keep the microphone in a fixed location in relation to the nostrils and mouth approximately 3 cm in front of the subject’s face, as shown in Figure 3.1(a). Digitized sound data were transferred to a computer using a USB preamplifier and audio interface (M-Audio, Model Fast Track Pro USB) with a
sampling rate (Fs) of 22,050 Hz and resolution of 16 bits. The external audio interface was preferred over the regular built-in audio adapters because of its better signal-to-noise (SNR) ratio, which is 60 dB at 1 kHz. Sound recordings were passed through a 4th order band-stop digital filter with a centre frequency of 60 Hz to suppress power line interference.

3.3.3 Respiratory Inductance Plethysmography

Respiratory inductance plethysmography (RIP) (Respitrace™; Ambulatory Monitoring Inc., White Plains, NY) was used to monitor respiratory motion and timing of the breathing phases of the subjects. RIP is composed of two flexible sinusoidal wires each embedded in individual stretchy fabric bands, one wrapped around the chest and the other around the abdomen (Figure 3.1(a)). The inductance of each band changes upon rib cage and abdominal displacement and generates a voltage signal proportional to its inductance. When calibrated against a spirometer, the sum of the electrical signals arising from the ribcage and abdomen provide an estimate of tidal volume. The signals are digitized at 150 Hz and stored in a computer. The electrical sum of the ribcage and abdominal signals is displayed and provides the total thoracoabdominal displacement, which reflects changes of tidal volume during respiration. RIP and breath sounds signals were captured simultaneously to assess the timing of breath sounds against the RIP waveform. In contrast to other types of respiratory monitoring devices, such as pneumotachography, RIP has the advantage of being applied on the chest wall, not on the face, so that it does not interfere with the capture of breath sounds.

3.3.4 Study Protocol

Subjects were studied in the supine position and were instructed to breathe normally for two minutes at their regular breathing rate with the microphone holding frame placed on the subjects’ faces. They were instructed to breathe through their noses only (nasal
Figure 3.1: Illustration of the experimental setup and signal outputs. As shown in (a), the microphone is attached to a frame in front of the subject’s face; RIP bands are placed around the subject’s chest and abdomen to measure thoracoabdominal motion. A 25-second long recording of breath sounds, and simultaneously summed thoracoabdominal RIP signals from a representative subject are shown in (b). Dashed vertical lines separate inspiration and expiration of the second breath from the left.
breathing), during the first minute, and through their mouths only (oral breathing) during the second minute. All the breaths during these periods were included in the analysis.

**Stage 1: Analysis of Breath Sounds in the Test Group and Feature Extraction**

The aim of stage 1 was to determine the spectral characteristics of the inspiratory and expiratory components of breath sounds. Five subjects (3 women and 2 men) were chosen randomly from the 15 subjects to study the frequency characteristics of the acoustic signals during the inspiratory and expiratory phases of the respiratory cycle.

Inspiratory and expiratory segments of breath sounds were determined and extracted from the acoustic data by comparing it to the rising edge (inspiration) and falling edge (expiration) of the RIP trace, as illustrated in Figure 3.1(b). The first 10 complete breaths of each subject’s recording were analyzed, half of which were during nasal breathing and the other half during oral breathing. Data from the inspiratory and expiratory phases of 50 breaths from all of the 5 subjects were analyzed.

Subsequently, the frequency spectrum of each phase was calculated separately using Welch’s method, which is the average of a 2048-point ($\approx$ 93 ms) Fast Fourier Transform (FFT) of sliding Hanning windows with 50% overlap. FFT arrays were normalized in amplitude, by dividing FFT(f) by the maximum frequency bin, in order to compare the relative changes in spectrum among resultant spectral arrays. Variables derived from nasal breathing were statistically compared with oral counterparts using the Mann-Whitney U Test in order to test the effect of the route of airflow before proceeding to the phase classification.

**Stage 2: Breathing Phase Tracking Algorithm Development and Validation**

A computerized algorithm was generated using features derived from spectra from the 5 test subjects (as explained in the results section). This algorithm was then applied
prospectively to the 10 subjects in the validation group. The algorithm was tested for its ability to determine breathing phases from acoustic data independently from other inputs. All breaths of the validation group were included in this stage. A total of 218 breaths (i.e., 436 phases) were recorded and included in the analysis with an average of 21.8±8.2 breaths per subject. Data analysis was performed and the algorithm was developed using Matlab R2007b software package (Mathworks, Natick, Massachusetts). The Fisher Exact Test was used to test the relation between breathing phases determined by the algorithm and by RIP, and Kappa was also calculated to find the degree of agreement between the 2 methods.

3.4 Results

The average respiratory rate of all the 15 subjects was 10.7 ± 3.7 breaths/minute, which lies well within in the normal range indicating that the face mask did not provoke tachypnea.

Stage 1: Analysis of Breath Sounds in the Test Group and Feature Extraction

Data obtained from the test group yielded 100 arrays of FFT spectra, 50 inspiratory and 50 expiratory acoustic spectra, normalized in amplitude. The average spectrum of all normalized arrays belonging to each phase with the corresponding standard deviation are presented in Figure 3.2. Figure 3.2(a) demonstrates that the frequency spectra of the 2 phases have different energy distributions. The mean inspiratory spectrum peaked between 30 Hz and 270 Hz. The spectrum exhibited flatness between 300 Hz and 1100 Hz before the next major peak with a center frequency of 1400 Hz. The expiratory spectrum, on the other hand, peaked between 30 and 180 Hz as shown in Figure 3.2(b). Its power dropped off exponentially until 500 Hz after which it flattened at low power.

The signal power above 500 Hz was consistently higher in inspiration than expiration.
Accordingly, the ratio of magnitude values of the frequency spectrum between 500-2500 Hz\(^2\) to values between 0-500 Hz was used to distinguish the two phases of the breathing cycle. This ratio is presented in Equation 3.1 as the frequency Bands Ratio (BR).

\(^2\)Although in this study the derived feature was limited to 2,500 Hz, other parts of the thesis, such as Chapter 5, utilize features that require higher frequencies, which justifies sampling at higher frequencies for purpose of the entire thesis.
Chapter 3. Identification of Breathing Cycle Phases, Part I

\[
BR = \frac{\sum_{f=2500\,Hz}^{f=500\,Hz} |FFT(f)|}{\sum_{f=2500\,Hz}^{f=0\,Hz} |FFT(f)|}
\]

Equation 3.1

The numerator of Equation 3.1 represents the sum of FFT magnitude bins between 500 - 2500 Hz, and the denominator represents the sum of FFT magnitude bins below 500 Hz. BR was calculated for each of the six curves shown in Figure 3.2 which include the mean spectrum curve ± standard deviations curves for both inspiration and expiration. These results are presented in Table 3.1.

<table>
<thead>
<tr>
<th>Table 3.1: BR calculated for inspiratory and expiratory spectra</th>
</tr>
</thead>
<tbody>
<tr>
<td>Inspiration</td>
</tr>
<tr>
<td>Mean inspiration spectrum</td>
</tr>
<tr>
<td>Mean inspiration spectrum + Std</td>
</tr>
<tr>
<td>Mean inspiration spectrum − Std</td>
</tr>
</tbody>
</table>

Table 3.1 shows that the mean BR for inspiration was 15.1 times higher than mean BR for expiration. Comparing the two extremes, ‘BR for mean inspiration − Std’, and ‘BR for mean expiration + Std’, BR for inspiration was 10.2 times that for expiration. This suggests that the frequency-based variable BR can distinguish the phases of the breathing cycle.

In order to test if the proposed variable BR is affected by the type of breathing (nasal vs. oral) each phase was further sub-classified according to the type of breathing: inspiratory nasal, inspiratory oral, expiratory nasal, and expiratory oral. The frequency characteristics of the 4 categories are displayed in the Figure 3.3. BR was compared within the same phase between nasal and oral breathing using the Mann-Whitney U Test. The results show that there was no effect of the route of breathing on the BR variable within the same phase (P= 0.464 for inspiration and 0.255 for expiration). Accordingly, both nasal and oral breaths were treated similarly in the next stage.
Stage 2: Breathing Phase Tracking Algorithm Development and Validation

In this stage, the BR parameter found in stage 1 that distinguished inspiration from expiration was utilized to track the breathing phases in the subjects in the validation group. This was achieved by developing an algorithm that depends on past readings of acoustic data to predict the current phase as illustrated in Figure 3.4. The advantage of using past values rather than post-processed statistics is that such technique can be adopted for real-time implementation. In this algorithm, the acoustic data stream is segmented into 200 ms segments. The BR of each segment is calculated by applying the Welch method as described in stage 1. Subsequently, the mean BR of the past 7 segments (i.e., 1.4 seconds) or the mean of all the past BRs, whichever was greater, was calculated. The time window of 1.4 seconds was chosen to optimize time and frequency
resolution in consideration of breathing physiology. The normal breathing rate is 12 to 20 breaths/minute [83]. At the highest breathing frequency (20 breaths/minute), each breathing phase lasts for about 1.5 seconds. Thus averaging the BR values of the past 1.4 seconds should encompass 1 full phase of rapid breathing. On the other hand, a window width of 200 ms is small enough to provide sufficient time resolution of breath sounds. Each newly found BR was then compared with the past BR average. If the currently examined BR was greater than twice the past 1.4 seconds BR average then it was labeled as inspiration. Likewise, a segment was labeled as expiration if the corresponding BR is 50% below the average of the past segments. Although initial results showed there is a 10-15 fold difference in the BR between the breathing phases, a lower threshold was considered in the classification approach. This is because the moving average window incorporates transitional BR points between inspiration and expiration that dilute the BR average of a pure breathing phase.

Figure 3.5(a) displays all BR values calculated from the acoustic data with the corresponding breathing cycle phases determined by RIP for comparison. Visual examination shows a correlation between the BR waveform and its RIP counterpart. Each rising edge of RIP trace in Figure 3.5(a) corresponds to inspiration and vice versa. The BR trace shows a quasi-plateau corresponding to each rising edge of RIP and a trough corresponding to each falling edge of RIP. Although BR showed some ripple within the same phase, such as that at time 5-10 seconds in Figure 3.5(a), that ripple was smoothed out by moving window averaging. The BR trace also shows a sharp transition between each successive plateau and trough that corresponds to the inter-phase transition of breathing. In summary, BR is consistently higher for inspiration than for expiration for each 200 ms segment.

The algorithm was tested prospectively on the breath sounds of 10 subjects in the validation group. Out of 436 breathing phases, 424 were correctly identified. The algorithm was equivocal in 9 phases; that is, it changed the classification transiently in the
Figure 3.4: Flowchart of the algorithm used for estimating the phase of the breathing cycle from acoustic data

middle of the phase, and misclassified 3 phases. In order to test the agreement between RIP and the acoustic method, Kappa was calculated and the Fisher Exact Test was performed. For both tests, equivocal and misclassified phases were grouped with incorrect determinations. Kappa was 0.96, indicating strong agreement between the 2 methods. The Fisher Exact Test results were also highly significant (P < 0.001) indicating a strong association between RIP and the acoustic method for classification of breathing phases.
Figure 3.5: (a) The BR variable calculated from 200 ms segments of acoustic breathing data is presented in comparison with the RIP signal. Each point on the BR waveform represents BR value for a 200 ms segment of acoustic breathing data. Amplitudes were normalized to present the 2 waveforms on the same scale. (b) Breathing phases found using the proposed algorithm are presented by the dashed line. The positive values represent inspiration; negative values represent expiration; zero values represent undetermined activity
3.5 Discussion

Using acoustic spectra of breath sounds, the inspiratory and expiratory phases of respiration were reliably identified, owing to their remarkably characteristic and distinct spectra. The inspiratory spectrum was concentrated in two main regions, the first below 300 Hz, and the second between 1200 and 1800 Hz (Figure 3.2(a)), with a relatively uniform distribution in between. On the other hand, the expiratory spectrum was confined to the lower region below 500 Hz (Figure 3.2(b)). The difference in spectral characteristics can be attributed to the physiological events involved in these different phases and to interaction of airflow with the transducer. In inspiration, air travels down the respiratory tract creating air turbulence around the nose and mouth. In expiration the resultant spectrum is due to two main events; sound traveling up the respiratory tract and the effect of the stream of expired air striking the microphone diaphragm. My findings suggest that a simple comparison between the upper and lower frequency content of breath sounds can be used to classify the corresponding segment into inspiratory and expiratory phases, respectively. I found that the ratio of the upper to the lower frequency band above and below 500 Hz (i.e., BR) reliably identified the inspiratory and expiratory phases of the breathing cycle, respectively. These results are in congruence with our study done on sleeping subjects in whom sleep was confirmed by EEG [84]. Similar to the analysis used in this work, differences in BRs provided a reliable means of distinguishing the inspiratory from expiratory phases of normal breaths in sleeping subjects.

RIP was used to validate the proposed method because it does not interfere with airflow through the nose and mouth as would be the case with spirometry and pneumotachography. The RIP waveform rises during inspiration and falls during expiration, with the increase and decrease in lung volume, respectively. RIP is a well validated, widely used technique for measuring tidal volume and monitoring breathing under both clinical and experimental conditions [85, 86, 87, 88].

The strategy for developing a respiratory monitor was based on examining the BR
waveform, which is derived from the short-time Fourier transform of a time series of breath sounds. One possible scenario for dealing with this waveform would have been calculating its histogram and determining the averages of the upper and lower percentiles. Also, clustering techniques could have been used to segment the data into 2 clusters. My approach has the advantage over those approaches in that it can be used for real-time monitoring of respiration. In this approach, BRs are examined in sequence and each BR is compared with the preceding BR. The preceding BRs are subject to a moving average window the length of a breathing phase (approximately 1.4 seconds). The empirical threshold of 2, chosen for the validation phase, yielded excellent results with 97.4% of all breathing phases classified correctly.

Respiratory monitoring has an important place in various clinical applications such as during sleep, anesthesia, medical imaging, and in severely ill patients with cardio-respiratory diseases [75]. Besides breathing rate, the identification of the inspiratory and expiratory phases of the breathing cycle could be of importance in various clinical and research situations. Such conditions include synchronizing image acquisition with respiration [80], delivery of radiotherapy [89], and detection of sleep apnea. For example, obstruction of the UA could be present in one of the phases only, selectively altering the acoustic characteristics of that phase. Breathing sounds are cyclic by nature and each cycle is non-stationary within itself because of alterations of inspiration and expiration. Thus, treating inspiratory and expiratory phases of the cycles as one might obscure some important physiologic and diagnostic information.

Acoustic analysis of respiratory sounds has gained an increasing role in the diagnosis of respiratory disorders because of advances in techniques for sound measurement and signal analysis [90]. Respiratory sound analysis, for example, has been used to identify pathological respiratory sounds, such as wheezes, crackles, and particularly, snoring [47, 48, 49, 50, 91, 51]. Various techniques have been used to identify snoring for the purpose of diagnosing OSA [68, 92, 63, 93]. Monitoring breath sounds is another appli-
cation of acoustics in the respiratory field. A group of researchers utilized breath sounds for monitoring breathing rate to detect postoperative respiratory depression [94, 95]. In their configuration the sensor responded to the jet of air emitted at expiration, whereas inspiratory air-flow was not monitored. Although the presented configuration shares the sensitivity of the microphone to the expired air pressure, which causes the expiratory sounds to be relatively larger in amplitude, it does not depend on the amplitude variations, but rather on the frequency bands ratios. This provides the advantage of being more immune to noise and gain variations that result from alignment changes or from using different mask structures. Another advantage of my technique is that it utilizes the acoustic features of each of the 2 phases, i.e. a decrease in BR denotes inspiration and an increase denotes expiration. This approach allows detection of the start and end of each phase independently of each other. The ability to delineate the timing of individual phases increases its reliability and provides the potential for a wider range of applications, such as inhaled drug delivery or modifying of inspired gas content that requires defining the onset of inspiration.

Recently, there has been more interest in trying to characterize the different phases of respiration acoustically [90, 96, 97]. However, these studies used a contact microphone placed on the chest [90] or over the trachea [96, 97]. Since the UA acts as an acoustic filter that modifies the spectrum of sound traveling through it, the characteristics of sound captured distal to the mouth are different from those captured from contact microphones proximal to the mouth through intervening tissues [98]. Different analyses are required to characterize respiratory phases at these different locations. The advantage of the presented technique is that the microphone positioning permits seamless embedding into a standard face mask that is used during other clinical procedures, such as oxygen and drug delivery. Potentially, this could be clinically applicable for monitoring breathing during such interventions. Additionally, that configuration saves the caregiver having to attach a separate transducer for respiratory monitoring. This also gives an advantage over
standard respiratory monitoring techniques such as inductance plethysmography, which require the attachment of a dedicated thoracic band. The deployment of a microphone as a sensing element allows combining the described technique with other acoustic analysis. For example, we and others have described the utility of acoustic analysis of snoring and non-snoring breath sounds for the diagnosis of sleep apnea [99, 100, 63, 68]. The phase identification can be added to such analysis algorithms using the same microphone signal output without adding an extra channel.

To date, I have investigated respiratory monitoring in healthy subjects, and demonstrated its potential. However, its usefulness for respiratory monitoring in disease states would have to be the subject of future studies. A common example would be chronic obstructive pulmonary diseases (COPD) in which expiration is prolonged due to expiratory airflow limitation that may be accompanied by wheezing, whereas inspiration is usually normal. Therefore, I don’t expect any difference in inspiratory acoustic properties. Similarly, this technique will still be sensitive to the expiratory air pressure, but filtering of accompanying wheezes might be needed to ensure its accuracy in identifying breathing phases.

The main drawback of the presented method is the lag between the occurrence of a change in the respiratory phase, and its detection by the computer algorithm. Since this method depends on the FFT of acoustic data, there must be a reasonable length of data upon which to apply the FFT. In this study, FFT was applied on 200 ms segments. This means that there is at least a 200 ms lag between capturing the beginning of the window and making the decision on what phase it is. In addition to the window length, the FFT calculations require approximately 60 ms of processing time using a current PC. Thus the total delay in this case would be approximately 260 ms. This delay could be acceptable in certain clinical applications in which approximate timing of inspiration and expiration is required, such as breathing under sedation. However, in other applications, like gating image acquisition to the breathing cycle, this delay might not be acceptable and a faster
technique should be implemented. The calculation time could be shortened by several
techniques such as reducing the sampling rate, shortening the segmentation window, and
reducing the number of FFT points. Repeating the same calculations for an application
where Fs is 10 kHz, the segmentation interval is 100 ms, and FFT length is 1024, the
resultant delay will decrease by 100.7 ms. This lag could be further reduced by deploying
techniques to predict the current breathing activity based on past BR values before the
FFT calculation of the current segment is completed. Future studies will be required to
address those performance issues.

To our knowledge, this study is the first to classify the inspiratory and expiratory
phases of respiration using acoustic data gathered by a microphone situated in front of
the mouth and nostrils in awake healthy subjects. It employs relative changes in spectral
characteristics, and thus it is not susceptible to variations in overall signal amplitude that
result from inter-individual variations. This work suggests the potential for using more
phase-dependant bio-acoustic data to assist in the clinical interpretation of breathing
data. It could be applied for both real-time (assuming the development of more rapid
signal processing), and offline applications. Future work could be directed towards clas-
sification of breathing phases during abnormal types of breathing such as snoring, apnea,
hyperventilation, hypoventilation and wheezing, and towards optimizing the speed of
data processing to provide real-time output.
Chapter 4

Identification of Breathing Cycle
Phases, Part II

The contents of chapter were published in a special issue of the “International Journal of Healthcare Technology and Management” as a journal article: Hisham Alshaer, Geoff R. Fernie, T. Douglas Bradley, “Phase tracking of the breathing cycle in sleeping subjects by frequency analysis of acoustic data”, 2010, Volume 11, Number 3/2010, pages 163-175. Except for formatting and a few organizational modifications, this chapter is almost identical to the aforementioned paper. Inderscience Publishers allow authors to incorporate the article content in other works by the author.

This work establishes the utility of breath sounds in tracking breathing phases as presented in chapter 3 but in sleeping subjects. The findings of this paper confirm that the frequency bands of inspiration and expiration can be used to distinguish breathing phases. In this part of the work, I studied a narrower frequency range than in the previous one, which was deemed sufficiently discriminatory.

1The Copyright, author rights, responsibilities and entitlements of Inderscience Publishers can be accessed at this link: http://www.inderscience.com/info/inauthors/author_copyright.php#entitlement
4.1 Abstract

The aim of this study was to test the ability of breath sound recordings from a non-contact microphone to detect breathing phases (inspiration and expiration) in sleeping subjects. Sleep stages and breathing sounds were digitally recorded from 10 subjects during sleep for a full night. Ten normal breaths were selected from different parts of the night from each subject, and frequency spectra of the inspiratory and expiratory phases were calculated. The ratio of frequency magnitude bins between 400 - 1000 Hz to frequency bins between 10 - 400 Hz was calculated for inspiration (Ri) and expiration (Re) separately for each breath. Ri was significantly higher than Re (p < 0.001). Retrospectively, the ratio of Ri to Re was used to distinguish inspiration and expiration. A ratio of 1.5 was found a suitable threshold to differentiate between respiratory phases in an acoustic signal with a correct decision in 90% of cases.

**Keywords:** breathing phases, breath sounds, microphone, respiratory monitoring, frequency spectrum

4.2 Introduction

Acoustic analysis of respiratory sounds has gained an increasing role in the diagnosis of respiratory disorders because of advances in techniques for sound measurement and signal analysis [90]. Respiratory sound analysis has been used to identify pathological respiratory sounds, such as wheezes and crackles [47, 48, 49, 50, 51, 52, 53]. Sound recordings have also been used during sleep to diagnose snoring and locate the site from which it arises [54, 55, 56]. In these studies, the microphone used to capture the acoustic signal was located either on the upper lip [54, 55], on the forehead [57], or suspended above the patient [56]. Sound recording during sleep has also been used to diagnose (OSA), a condition characterized by repeated complete (apnea) or partial (hypopnea) cessation of airflow due to complete or partial collapse of the pharynx, respectively.
These events alternate with episodes of hyperventilation (i.e., hyperpnea) during which loud snoring occurs [6]. Recurrent apnoeas and hypopnoeas lead to intermittent hypoxia that provokes arousals and sleep fragmentation that cause restless sleep, and excessive daytime sleepiness. Repetitive apnoeas and intermittent hypoxia also elicit sympathetic nervous system activation, oxidative stress and elaboration of inflammatory mediators that cause repetitive surges in blood pressure at night and increase the risk of developing daytime hypertension, atherosclerosis, heart failure, and stroke independently of other risk factors [1, 2, 3, 4, 5].

A normal breathing cycle is composed of two main phases, inspiration and expiration. The acoustic characteristics of inspiration and expiration could vary significantly within a subject. For example, narrowing of the UA that takes place in OSA could be present in one of the phases only, selectively altering the acoustic characteristics of that phase. Therefore, treating the two phases similarly may lead to inaccurate conclusions about the dynamics of the airway. Therefore, the cyclic nature of breathing sounds necessitates dividing them into phases.

Several algorithms have been developed for detecting different types of breathing sounds. Duckitt et al. have used Hidden Markov models for separating snoring from normal breathing and other types of noise [101]. Abeyratne et al. have used a combination of time and frequency domain features for categorizing breathing sounds into breathing, silence, voiced/unvoiced snoring sounds, and speech [63]. Those algorithms didn’t take into consideration the cyclic nature of breathing, and treated each breath or breathing event as a single entity without distinguishing its phases. The aim of this work, therefore, was to use breath sound recordings from a microphone in front of the face to classify inspiratory and expiratory phases of regular breathing during sleep. I hypothesized that numerical comparative analysis of the frequency spectra of breath sounds could differentiate between inspiration and expiration.
4.3 Methods

4.3.1 Polysomnography and Sleep Staging

Data were collected from 10 consecutive men and women at least 18 years of age referred for overnight PSG. During PSG, sleep stages were recorded during the course of the night using standard PSG techniques that included EEG, EOG, and submental EMG [102]. The corresponding sleep stage for the selected breath samples was determined from the PSG recording.

4.3.2 Data Acquisition

During PSG, breath sounds were recorded by a cardioid condenser microphone (Audiotechnica condenser microphone, Model PRO 35x). The microphone’s cardioid polar pattern reduces pickup of sounds from the sides and rear, improving isolation of the sound source. The microphone was embedded in the centre of a loose fitting full face mask frame. The mask provided a structural frame to keep the microphone in a fixed location 3 cm in front of the subject’s face, as depicted in Figure 3.1.

Digitized breath sounds were transferred to a computer using a USB preamplifier and audio interface (M-Audio, Model MobilePre USB) with a sampling rate of 22050 Hz and resolution of 16 bits. The external audio interface was preferred to the regular built-in audio adapters because of its better SNR ratio which is 91 dB (typical, A-weighted).

4.3.3 Breathing Acoustics Analysis

Full night breath sound recordings were displayed on a computer screen. The recordings were visually scanned to identify periods of regular breathing. After visual scanning, the recordings were played back for auditory analysis.

Since my objective was to identify phases of normal regular breathing, sequences of normal breaths that did not have signs of obstructive breathing such as snoring, inter-
ruptions, or other irregularities such as tachypnea (rapid breathing), or hyperventilation (deep breathing) were then included in the subsequent frequency analysis. This process included 3 segments of the night selected randomly. If a portion of the recording fulfilled the aforementioned inclusion criteria, then 3 to 4 consecutive breaths were selected from that portion. A total of 10 breaths was selected from each subject. The investigator was blind to the sleep stage while selecting the analyzed breaths except that sampling started after the onset of sleep. The real time stamp of each breath was registered in order to identify, retrospectively, the sleep stage in which it took place. Subsequently, the investigator listened to these breaths again to divide each breath into its inspiratory, expiratory, and interbreath phases. Each phase was labelled manually.

The frequency spectrum of each phase was calculated separately using Welch’s method i.e., the average of 2048-point FFT of sliding Hanning windows with 50% overlap. The FFT window length is equivalent to 92.87 ms, which is enough to capture frequencies as low as 10.7 Hz. Yet, that time window is short enough to provide sufficient time resolution of breathing activity, since the normal breathing rate is approximately 14 to 20 breaths per minute. Therefore, in rapid breathing, where the shortest window for good time resolution is required, one cycle lasts approximately 3 seconds with breathing phases lasting approximately 1.5 seconds each. The window length I selected should, therefore, provide sufficient time resolution for all breathing rates within the normal range.

The resultant frequency spectrum was displayed on a computer screen for visual analysis. Frequency spectra of the interbreath pauses were also calculated and incorporated in the analysis to control for the effect of ambient noise. Careful visual examination of the spectra revealed that during inspiration, the amplitude of signals above 400 Hz was consistently higher than during expiration. Following from this observation I formulated the hypothesis that the ratio of the spectral magnitude between 400 - 1000 Hz to those between 10 - 400 Hz will be higher in inspiration than in expiration. To test this hypothesis, we calculated this ratio for each breathing cycle using the following equation:
This ratio is conventionally named parameter R, where the numerator represents the sum of FFT magnitude bins that lie between 400 - 1000 Hz, and the denominator represents the sum of FFT magnitude bins that lie between 10 - 400 Hz. Bins below 10 Hz were excluded to avoid DC contamination. To ensure the repeatability of the results, R was calculated for 3-4 successive breaths in the included sequence and for a total of 3 randomly selected sequences from different parts of the night. A total of 100 breaths were collected from the 10 subjects—10 breaths from each subject. The R value was calculated for inspiration (conventionally named Ri) and expiration (conventionally named Re) for each breath using Equation 4.1.

4.3.4 Statistical analysis

Data are expressed as mean±SD unless stated otherwise. Ri and Re were compared by Wilcoxon signed rank test. This test compares 2 related variables drawn from non-normally distributed populations. One-sample sign test was performed to test whether a data set with a non-normal distribution, in this case Ri/Re described below, is greater than that of the specified threshold.

4.4 Results

The age, sex, and body mass index of the 10 subjects are shown in Table 4.1. A total of 100 breaths were sampled from these subjects with a mean number of 10 breaths per subject. Seventy percent of the breaths analyzed were from non-rapid-eye movement sleep (NREM), 18% from rapid eye movement sleep (REM), and 12% from brief awakenings that occurred spontaneously after sleep onset, according to objective PSG criteria.
Table 4.1: Characteristics of subjects

<table>
<thead>
<tr>
<th>Subject</th>
<th>Age(years)</th>
<th>Sex</th>
<th>Body Mass Index</th>
</tr>
</thead>
<tbody>
<tr>
<td>Subject 1</td>
<td>51</td>
<td>F</td>
<td>39.1</td>
</tr>
<tr>
<td>Subject 2</td>
<td>43</td>
<td>M</td>
<td>25.6</td>
</tr>
<tr>
<td>Subject 3</td>
<td>49</td>
<td>M</td>
<td>23.7</td>
</tr>
<tr>
<td>Subject 4</td>
<td>27</td>
<td>M</td>
<td>36.8</td>
</tr>
<tr>
<td>Subject 5</td>
<td>64</td>
<td>M</td>
<td>26.3</td>
</tr>
<tr>
<td>Subject 6</td>
<td>60</td>
<td>M</td>
<td>33.0</td>
</tr>
<tr>
<td>Subject 7</td>
<td>68</td>
<td>F</td>
<td>28.5</td>
</tr>
<tr>
<td>Subject 8</td>
<td>31</td>
<td>M</td>
<td>30.3</td>
</tr>
<tr>
<td>Subject 9</td>
<td>48</td>
<td>F</td>
<td>31.6</td>
</tr>
<tr>
<td>Subject 10</td>
<td>56</td>
<td>M</td>
<td>26.7</td>
</tr>
</tbody>
</table>

Inspiration and expiration were characterized by distinct frequency spectra, as depicted in Figure 4.1. For inspiration, there was a sharp narrow band of harmonics usually below 200Hz. The spectrum exhibited a valley between 200Hz and 400Hz and a peak again after 400Hz as shown in Figure 4.1(b). Another variation of the inspiratory spectrum was the same initial narrow band followed by a relatively smooth spectrum without the 400 Hz drop. The expiratory spectrum, on the other hand, formed a wider band that spanned frequencies up to 500Hz and whose power dropped off rapidly above this frequency as shown in Figure 4.1(c). The spectrum of the interbreath pause did not show a consistent pattern.

The mean value of Ri (0.66±0.46) was significantly greater than the mean value of Re (0.18±0.17), (P<0.001) such that 95% of breaths had a Ri greater than Re. Since small differences between Ri and Re might be attributed to randomness, two thresholds of 50% and 100% difference between Ri and Re were tested, corresponding to ratios of
1.5 and 2 consecutively. To this end, the ratio Ri/Re was calculated for each breath. Determining this ratio also allowed Ri and Re to be treated as dependant pairs just as in the case in actual breathing because they take place in close proximity under very similar physiological conditions. These ratios were then tested to assess whether they were greater than the aforementioned thresholds. The Ri/Re drawn from 100 breaths was significantly greater than both 1.5 (p<0.001) and 2 (p<0.001). Accordingly, the observation that Ri is greater than Re of the same breath by 50% and 100% is not attributable to chance.

The overall reliability of this method to detect the correct order of breathing phases was tested retrospectively for each of the chosen thresholds. Out of 100 breaths, 90 had Ri 1.5 times greater than Re and 73 had Ri 2 times greater than Re.

Brief spontaneous awakenings after sleep onset did not significantly alter the Ri/Re
from those taking place during EEG confirmed sleep \(p=0.958\). Therefore, they were included in the aforementioned analysis since such brief awakenings normally occur during the sleeping period.

### 4.5 Discussion

In the current work, the ratio of the higher frequency (400-1000 Hz) to the lower frequency band (10-400 Hz) derived from breath sound recordings of sleeping subjects was very effective in distinguishing inspiration from expiration. Inspiration was characterized by a narrow band below 200 Hz, a trough towards 400 Hz and a wider but shorter peak above 400 Hz (Figure 4.1(b)), or a smooth frequency distribution after the decline of the initial narrow peak.Expiration, on the other hand, was characterized by a wider peak that had a relatively sharp increase from 10 to 50 Hz and a smooth drop from 50 to 400 Hz. Frequency content above 400 Hz was relatively sparse (Figure 4.1(c)).

Respiratory monitoring has important clinical applications such as during sleep, anaesthesia, medical imaging, and after severe injuries [75]. Several researchers have demonstrated the importance of bioacoustics in monitoring and diagnosing abnormal breathing. Respiratory monitoring has particular importance in sleep medicine because of its potential to detect repetitive apneas in patients with sleep apnea. There has been growing interest in using breath sounds for the diagnosis of OSA. Given the relatively high prevalence of the OSA syndrome (i.e., repetitive obstructive apneas accompanied by symptoms of excessive daytime sleepiness, disturbed sleep and/or neurocognitive impairment), which affects approximately 4% of men and 2% of women [6], sleep apnea constitutes a major public health burden. Yet, it was reported that approximately 87% of individuals with this conditions remain undiagnosed in the United States [17]. This may be partially attributable to the lack of access to, and high cost of full-night attended PSG, which is currently the standard means of diagnosing OSA [26]. As a result, there is
a need for alternative methods to diagnose OSA, in a more accessible and cost-effective manner such as acoustic analysis of breath sounds in the ambulatory setting [63, 68, 92]. The majority of previous studies in the field of acoustics have focused on snoring sounds to diagnose OSA using various signal processing methods such as time domain [92, 60] and spectral analysis [66, 68, 92], pitch characteristics [63], and neural networks [93]. A few other studies have investigated temporal variations in tracheal sounds to differentiate between normal and obstructive breathing [58, 59]. The strategy of all those studies was to compile large numbers of data segments from various disease conditions and compare the predominant features of each. This approach does not identify individual apneas and hypopneas, the frequency of which is required to make a precise diagnosis of OSA. Detecting individual apnoeas and hypopoeas requires tracking breathing breath-by-breath, which in turn requires consideration of the phasic nature of breathing, as we have done herein.

Breathing sounds are cyclic by nature and each cycle is non-stationary within itself because of alterations of inspiration and expiration. Thus, treating inspiratory and expiratory phases of the cycles as one, as in the aforementioned studies, might wash out some important physiologic and diagnostic information. Recently, there has been more interest in trying to distinguish, from acoustic data, the different phases of the breathing cycle. Guler et al. have proposed a classification scheme that takes into consideration the cyclic nature of respiratory sounds [90]. Hult et al. [96, 97] have used time and frequency parameters to distinguish inspiratory and expiratory phases. These studies used a contact microphone placed on the chest [90] or over the trachea [96, 97]. Since the UA acts as an acoustic filter that modifies the spectrum of sound travelling through it, the characteristics of sound captured distal to the mouth are different from that captured proximal to the mouth, and through intervening tissues, such as from the trachea via the surface of the neck, or from the lungs via the surface of the chest [103, 98]. Therefore, a different algorithm may be required to characterize the respiratory cycle when sound
is captured directly distal to the mouth, which is the level at which breath sounds are
normally heard and interpreted by humans. The location of the microphone distal to the
airway tract provides more information about the influence of the UA, the site of the
obstruction in OSA, on breath sounds as compared to lower locations such as over the
trachea or chest wall [103, 98], which is the reason for my choice for this location. Besides
the location of the transducer, those studies have used data collected from subjects who
were fully awake. These findings might or might not be applicable for the detection of
disorders of breathing that occur exclusively during sleep, such as OSA. Therefore, there
is a need to develop a simple algorithm for tracking the phases of the breathing cycle by
sound recordings in sleeping subjects.

In the present study, the frequency spectrum of each respiratory phase was utilized in
order to distinguish inspiration from expiration. Frequencies below 10 Hz were excluded
in order to avoid the effect of low frequency and any intruding DC component. Fre-
quencies above 1 kHz were not deemed to influence the results. The frequency spectrum
pattern of each phase was consistent. A cut-off point of 400 Hz was chosen to distinguish
between inspiration and expiration because visual examination of frequency spectra of
breath sounds revealed that during inspiration, the amplitude of signals above 400 Hz was
consistently higher than during expiration. Following from this observation, I formulated
the hypothesis that the ratio (R) of frequency magnitude values between 400 – 1000 Hz
to values between 10 – 400 Hz will be higher in inspiration than in expiration. I avoided
comparing the mean Ri and Re because this approach does not consider that these two
variables are paired and are drawn from the same breathing cycle. Therefore the ratio
of Ri/Re for each breathing cycle was calculated in order to examine the intra-breath
relationship between Ri and Re. Ri/Re was found to be significantly greater than one.
In other words, for each individual breath Ri is significantly higher than Re. In 90% of
breaths Ri was at least 1.5 times greater than Re, and in 73% of the breaths Ri was at
least 2 times greater than Re. Accordingly, 1.5 is a suitable threshold to differentiate
between inspiration and expiration from breath sound recordings using the experimental setup described herein with a correct decision in 90% of cases. Therefore, this appears to be a promising technique to reliably distinguish breathing phases from continuous breath sound recordings during sleep.

Normal breaths from randomly selected parts of the overnight recording were included, without knowledge of the sleep stage. Abnormal breathing patterns, such as snoring, were excluded from the analysis because the scope of this work covers normal breaths only, whose features could be useful as reference points for developing similar methods for analyzing abnormal breathing. The subsequent identification, from the PSG recording, of the sleep-wake stage from which selected breaths were derived revealed that 70% of analyzed breaths took place in NREM sleep, 18% in REM sleep, and 12% during spontaneous awakenings from sleep. Such awakenings are a normal occurrence: normal humans wake up spontaneously approximately 30-40 times per night according to PSG criteria, yet most such awakenings are not perceived by the subject [104]. Therefore, brief awakenings after the onset of sleep are considered a normal occurrence on PSG recordings, and part of normal sleep structure. Therefore, the latter breaths, happening during brief awakenings, were included in the analysis. The statistical analysis showed that the frequency characteristics for those breaths did not differ from those derived from sleep during the rest of the night.

The variations between inspiratory and expiratory spectra are attributable to a combination of factors such as the acoustic characteristics of the airway and to the physical effect of airflow. The frequency spectrum of expiration is the result of both expiratory sound and direct airflow sensed by the microphone diaphragm. The effect of expiratory direct air contact with the transducer has been previously studied using a different transducer to detect airflow. For example, Vegfors et al. (1994) used a fibre-optic probe that changes refractivity due to condensation of water vapour in expired air [105]. The ability of the microphone’s diaphragm, used in this study, to sense expiratory airflow provides
the ability to monitor breathing without physically touching the subject as opposed to the fibre-optic probe. The inspiratory signal, on the other hand, is due to characteristic sound produced by turbulence of air travelling down the airway only and not to direct airflow.

The presented method for tracking breathing phases has some limitations. At this stage in its development, we have only used it to identify the inspiratory and expiratory phases of normal breathing, but not of breathing accompanied by snoring. Snoring is a common complaint among middle aged men and women: 33% of men, and 19% of women in this age group report having loud snoring [106]. Snoring results from the vibration of pharyngeal tissue with or without collision [107]. This vibration introduces a set of harmonics into the frequency spectrum that relate to the number of vibrations or collision per second, which constitutes the fundamental frequency of snoring [107]. Although airflow is partially interrupted by vibrations of the pharyngeal tissue during snoring, inspiratory and expiratory airflow continue, which suggests that the distinct frequency patterns of inspiration and expiration will persist on the background of the snoring-related harmonics. Accordingly, phase tracking in such conditions might be achieved by extracting the non-harmonic frequency structure and calculating the bands ratios as described in this paper. This could be the subject of future work to distinguish inspiration from expiration during snoring. Although snoring is the most frequently encountered sound aside from normal breathing during sleep, there are other contaminants of acoustic data such as somnolent speech, cough, and extra-respiratory sounds such as Duvet noise. To automate and make this technique generally applicable will require the detection and exclusion of such contaminants as a step prior to analysis of the underlying breath sounds. Techniques described in preliminary work of others that attempted to classify respiratory and non-respiratory sounds during sleep, could be employed for this purpose [101].

To our knowledge, this study is the first to classify breathing phases in sleeping subjects using a microphone distal to the UA. It provides an objective numerical method for
distinguishing each phase by simple comparison of their frequency spectra. This method employs relative changes in spectral characteristics, and thus it is not susceptible to variations in overall signal amplitude that result from inter-individual variations. This work suggests the potential for using bioacoustic analysis to assist in the clinical interpretation of breath sound recordings for purposes of respiratory monitoring. In future works, the calculations proposed by this study could be integrated in a more comprehensive algorithm for automating the classification of breath sounds. It could be applied for both real-time and offline applications. In either case, phase monitoring could be accomplished by tracking fluctuations of the R variables. The suggested sensor configuration provides the ability of integrating the transducer with a medical mask that provides a means for non-intrusive monitoring of breathing activities. In the future, there is scope to expand application of this technique to include classification of breathing phases during abnormal types of breathing such as snoring, apnoeas, hypopnoeas, and post-apnoeic hyperventilation. More sophisticated techniques of pattern recognition might also be employed to increase the robustness of characterizing the frequency spectrum of each phase.
Chapter 5

Detection of UA Narrowing

This chapter is based on the peer reviewed IEEE proceeding: Hisham Alshaer, Martha Garcia, M. Hossein Radfar, Geoff R. Fernie, and T. Douglas Bradley, Detection of upper airway narrowing via classification of LPC coefficients: Implications for OSA diagnosis. In IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), pages 681–684 © 2011. The material presented in this chapter is similar to the published paper except for the formatting and some organizational and stylistic improvements, which conforms with the University of Toronto thesis style. The IEEE does not require individuals working on a thesis to obtain a formal reuse license. ¹

UA narrowing and collapse play a major role the pathophysiology of OSA. I hypothesized that breath sounds are modulated by UA caliber and thus, acoustic analysis of breath sounds can detect UA narrowing. Some studies have shown that snoring analysis can be used for this purpose. However, this is subject to deficiencies of pure snoring-

¹Reprinted with permission from the senior author Bradley D. as per the IEEE reuse policy. In reference to IEEE copyrighted material used in this thesis, the IEEE does not endorse any of University of Toronto products or services. Internal or personal use of this material is permitted. If interested in reprinting/republishing IEEE copyrighted material for advertising or promotional purposes or for creating new collective works for resale or redistribution, please go to http://www.ieee.org/publications_standards/publications/rights/rights_link.html
Chapter 5. Detection of UA Narrowing

Driven techniques discussed earlier in Section 2.4. In the work presented herein, we were the first to show that non-snoring sounds carry the acoustic signature of UA narrowing. I used the technique described in Chapter 4 and 3 in order to separate inspiratory sounds that were then used for the analysis described in this paper.

5.1 Abstract

The similarities between unvoiced speech sounds and turbulent breath sounds were used to detect change in sound characteristics caused by narrowing of the UA, similar to that occurring in OSA. In 18 awake subjects, UA resistance ($R_{UA}$), an index of UA narrowing, was measured simultaneously with breath sound recordings. Linear Prediction Coding was applied on turbulent inspiratory sounds drawn from low and high $R_{UA}$ conditions and K-means was used to cluster the resulting coefficients. The resulting 2 clusters were tested for agreement with the underlying $R_{UA}$ status. Distinct clusters were formed when $R_{UA}$ increased relatively high but not in cases with lower rise in $R_{UA}$ ($P < 0.01$ for all indicators.) This is the first work to show the utility of LPC in breath sounds analysis confirmed by an objective indicator of UA narrowing.

5.2 Introduction

OSA is a breathing disorder characterized by repetitive cessations of breathing from 5 to 100 times/hour during sleep, each lasting 10-60 seconds, due to narrowing and collapse of the UA [14]. This causes episodes of oxygen deprivation and provokes arousals from sleep and consequent sleep fragmentation. As a result, patients suffer from poor sleep quality, daytime sleepiness, and impaired cognitive performance [27]. OSA increases the risk of developing hypertension, heart failure and stroke by 3 to 4 fold compared to subjects without OSA [13, 108]. It is a common disease affecting approximately 7% of adults. Nevertheless, the majority of patients with OSA remain undiagnosed; it was shown in an
epidemiological study that of those subjects found to have moderate to severe OSA, 93% of women and 82% of men had not been previously diagnosed clinically\[17\]. The current method of choice for diagnosing OSA is overnight PSG in which the patients have to sleep in a laboratory attached to many monitoring electrodes under the supervision of a technician. PSG is expensive and access to it is limited, resulting in long waiting lists. For this reason, several attempts have been made to devise new methods to diagnose OSA using simple techniques that patients can use independently at home such as the acoustic analysis of respiratory sounds [66, 61]. A strikingly common aspect of almost all available studies on the acoustic analysis for the diagnosis of OSA is the focus on snoring sounds as a basis for differentiating people with and without OSA. Although snoring is a hallmark of OSA, it might not necessarily take place with each apnea and hypopnea. Accordingly, the disease severity might be underestimated if some apneas are missed due to absence of snoring. Therefore, comprehensive acoustic analysis of breath sounds should take into consideration both their snoring and non-snoring components. The latter result from turbulence created during the passage of air into and out of the lung through the UA. The degree and character of air turbulence should be influenced by changes in UA caliber and airflow rate. We pioneered this field by our previous research showing that UA narrowing in OSA is at least partially a consequence of fluid shift from the lower body into the neck. For example, fluid displacement from legs due to application of lower body positive pressure (LBPP) via inflatable trousers increases neck circumference, narrows the UA [109] and increases UA resistance ($R_{UA}$) [110, 111], presumably due to accumulation of fluid around the UA. The narrowing of the UA without tissue vibration is analogous to generation of unvoiced fricative sounds in speech production. This notion suggests that the quality of breath sounds will vary according to the degree of narrowing similar to the case of unvoiced frication. A major challenge in this work, however, is detecting objectively the occurrence of UA narrowing. The goal of this work is to detect variations in pure turbulent breath sounds qualities in relation to changes in a quantitative index
5.3 Methodology

5.3.1 Experimentation and Data Acquisition

Subjects

Data were collected from 18 subjects recruited by advertisement (4 women, 14 men, age 55.6 ± 10.2 years, body mass index (BMI) 32.2 ± 8.7, the frequency of apneas and hypopneas per hour of sleep (apnea hypopnea index or AHI) 36.73 ± 20.80. No specific inclusion or exclusion criteria were imposed. The study protocol was approved by the research ethics board, and all subjects provided written informed consent.

Breath Sound Recordings

Breath sounds were recorded using a cardioid condenser microphone (MX185, Shure®) in front of the subject’s nose and embedded in a full face mask that was strapped to the head as shown in Figure 5.1. Digitized sound data were transferred to a computer using a USB preamplifier and audio interface (M-Audio, Model Fast Track Pro USB) with a sampling rate (Fs) of 22,050 Hz and resolution of 16 bits. Acquired sound was bandpass-filtered at 20-10,000 Hz.

Lower Body Positive Pressure Application

The experimental protocol has 2 arms (parts), a control arm and an LBPP arm, separated by a washout period. A pair of deflated medical anti-shock trousers (MAST III-AT; David Clark, Inc.) was wrapped around both legs from the ankles to the upper thighs of supine awake subjects. For the control arm, trousers were left deflated, and for the LBPP arm, trousers were inflated to 40 mmHg to force fluid out of the legs. The subjects were then
Figure 5.1: Setup of the face mask, microphone, pharyngeal catheters, and the pneumotachometer with a sample waveform of breath sounds

Measurement of Upper Airway Resistance (R_{UA})

Airflow was measured using a pneumotachometer attached to the outlet of the mask. Two pressure catheters were inserted from the nose to the nasopharynx and hypopharynx to measure nasopharyngeal and hypopharyngeal pressures respectively as shown in Figure 5.1. Transpharyngeal pressure was measured as the difference between nasopharyngeal and hypopharyngeal pressure. As an index of UA narrowing, we measured R_{UA} by dividing the transpharyngeal pressure by airflow, given by R_{UA} = \Delta P/F and is expressed in cm.H_{2}O/Litre/second, where P is pressure and F is flow. R_{UA} was calculated at the lowest value of airflow every 30 seconds. Breath sound recordings were synchronized with

crossed over to the opposite arm. The duration of each arm was 20 minutes separated by 15 minutes washout period. The first five minutes of each arm was a baseline (BL) period, which was used as a reference for the subsequent changes in R_{UA} and breath sounds. Breath sounds and R_{UA} values from the same arm were compared to each other to avoid any possible effect of the change of microphone position during the cross-over.
the pressure and airflow signals in order to correlate sound characteristics with $R_{UA}$. In this study, I focused solely on the relationship between $R_{UA}$ and breath sounds. EEG, EOG, and $EMG_{sm}$ were recorded to monitor sleep-wake status to ensure that subjects remain awake throughout the experiment. This study was approved by Research Ethics Board of the Toronto Rehabilitation Institute.

5.4 Data Analysis

Breath Sounds Segmentation and Annotation

Breath sounds included in this analysis are turbulent inspiratory sounds only. Expiratory sounds were excluded to avoid the effect of expired airflow on the microphone. One of the experimenters listened to the breath sounds to exclude snoring and wheezing because they involve tissue vibration and thus are voiced in nature. Two sets of sounds were collected from each experimental arm: one set from the BL and another set at the point at which peak $R_{UA}$ occurred in each of the control and LBPP arms. Therefore, 36 data sets were retrieved from the 18 subjects. Each data subset is referred herein to as M. Each subset of inspiratory sounds was annotated according to the $R_{UA}$ value that accompanied that subset of sounds. Depending on the length of the breathing cycles, 2 to 5 inspirations were selected within each epoch for each $R_{UA}$ value (in shorter cycles more were included and vice versa). Breath sounds were displayed and examined in LabVIEW™ (version 9.0 2009).

Linear Predictive Coding (LPC) Implementation

LPC has been accepted as one of the best modeling techniques for speech signals, in particular unvoiced speech sounds, in order to capture the shape of the vocal tract [112, 113]. The LPC model of unvoiced speech sounds assumes a random noise generator as an excitation source. Turbulent breath sounds share this feature with unvoiced speech
Chapter 5. Detection of UA Narrowing

sounds because both are generated as a result of the passage of air through the UA, whether fully patent or narrowed, but without the occurrence of tissue vibration such as snoring. LPC models the vocal tract, or the UA in this context, as an all-pole filter given by Equation 5.1:

\[ H(z) = \frac{1}{1 - \sum_{i=1}^{p} a_i z^{-i}} \]  

(5.1)

with an LPC order \( p = 6 \) where \( a_{1-p} \) are the LPC coefficients. Figure 5.2 demonstrates the similarity between LPC implementation in speech and breath sounds. LPC was applied in the following steps:

1) Because breath sounds vary in amplitude due their cyclic nature, they were normalized in amplitude to remove the effect of gain in the LPC model. The signal’s envelope was found by calculating a moving average of the signal’s variance using a 1,100 point (50 ms) window and then normalizing to that envelope as described previously [114].

2) Pre-emphasis was applied to compensate for the inherent spectral tilt similar to application in speech [113].

3) In order to apply LPC on segments of equal duration, normalized breath sounds were segmented with a Hamming window of \( \sim 250 \) ms duration with a frame shift of 200 ms. This way, an average of \( 272 \pm 82 \) vectors of LPC coefficients from the 36 experimental arms (M) were obtained, which comprises the training data sets.

Pattern Recognition and Clustering of Breath Sounds

The next step is to classify training data into a number of clusters. To detect the presence of distinct clusters in each of the 36 M vectors, each derived from an experimental arm, the following steps were followed (as also illustrated in Figure 5.3):
Figure 5.2: Proposed analogy of LPC modeling of unvoiced speech sounds and turbulent breath sounds

1. **Feature selection:** The 6th order LPC coefficients were selected as a feature for the classifier.

2. **Clustering algorithm:** I implemented the $K$-means algorithm on $M_{1-36}$ with a total of $272 \pm 82$ LPC vectors in each $M$. Number of clusters was forced into 2 based on the knowledge of the 2 underlying conditions i.e. BL and peak $R_{UA}$.

3. **Finding clustering tendency:** To measure the ability of $K$-means to separate LPC vectors in $M$, based on the underlying $R_{UA}$ status, BL and peak $R_{UA}$, the ratio of LPC vectors that belong to each of the resulting 2 clusters is calculated as:

   \[ T = \sum_{i=1}^{n} \left( \frac{x_i}{S_i} \right) \in C_j \]  
   \hspace{1cm} (5.2)

   which is the sum of the number of LPC vectors $x_i$ in each inspiratory sound segment $S$, where $n$ is the total number of vectors in $M$ and $C_j$ is each of the resulting clusters ($j=1,2$). If that sum showed that 75% or more of sound segments originating from BL aggregate in a distinct and different cluster from those originating from Peak $R_{UA}$, then each of the 2 clusters is said to be the correct cluster ($C_{corr}$) for that $R_{UA}$ state and that arm is said to have high clustering tendency. On the other hand, if the result is below 75% or if BL and Peak $R_{UA}$ sounds do not aggregate in distinct clusters, then this case is said to have low clustering tendency.
4. **Calculating the overall classification accuracy:** Accuracy of this method in differentiating between supposedly different sounds was achieved by calculating the weighted sum of the percentages of LPC vectors $x_i$ in each segment $S$ that were classified in $C_{corr}$ as in Equation 5.3:

$$ A = \sum_{l=1}^{m} w_l \cdot \sum_{i=1}^{n} \left( \frac{x_i}{S_l} \right) \in C_{corr} $$  \hspace{1cm} (5.3)

where weight $w_l$ is equal to the number of frames in each inspiration divided by the total number of frames in a single arm. The rest of the variables are similar to Equation 5.2. All acoustic processing techniques were implemented in MATLAB\textsuperscript{TM} (version 7.9.0 R2009b)

**Figure 5.3:** Flow chart of the data clustering and analysis algorithm

5.4.1 **Relation between Sound Properties and $R_{UA}$**

From the aforementioned calculations, inferences were made on the relation between $R_{UA}$ values of BL and Peak $R_{UA}$ on one hand and clustering tendency on the other. The relations were statistically tested using the Wilcoxon rank sum test or t-test depending on the data distribution type.
5.5 Results and Discussion

Out of 36 experimental arms, 27 showed high clustering tendency (H\textsubscript{CT} group) and 9 showed low clustering tendency (L\textsubscript{CT} group). The characteristic of those groups are as displayed in table 5.1 and Figure 5.4. In the H\textsubscript{CT} group, the peak R\textsubscript{UA} was 14.9±10.2 units, which was significantly higher than that in L\textsubscript{CT}, 8.0±3.8 (p=0.0041). Similarly, the difference between BL and peak R\textsubscript{UA} (ΔR\textsubscript{UA}) in H\textsubscript{CT} group was 11.0±9.4, which is significantly higher than ΔR\textsubscript{UA} in the L\textsubscript{CT} group, 5.7±3.0 (p=0.0089). These results show that the increase in R\textsubscript{UA} causes a change in breath sound qualities that can be detected with LPC. The overall accuracy of breath sounds classification was 84.7±7.9% vs. 58.6±5.7% in H\textsubscript{CT} and L\textsubscript{CT} respectively (P<0.0001). All of those parameters show clearly that LPC coefficients of turbulent breath sounds vary when a rise of R\textsubscript{UA} takes place above a certain level, but do not when the rise is to a lower degree or absent. Since R\textsubscript{UA} is an indicator of UA narrowing, I suggest that this method can be used to detect UA narrowing.

Table 5.1: Summary of R\textsubscript{UA} values according to the clustering tendency

<table>
<thead>
<tr>
<th>R\textsubscript{UA} Status</th>
<th>H\textsubscript{CT}</th>
<th>L\textsubscript{CT}</th>
</tr>
</thead>
<tbody>
<tr>
<td>BL R\textsubscript{UA}</td>
<td>Average=3.9±1.9 Median = 3.6</td>
<td>Average=2.2±1.3 Median = 2.1</td>
</tr>
<tr>
<td>Peak R\textsubscript{UA}</td>
<td>Average=14.9±10.2 Median = 11.7</td>
<td>Average=8±3.8 Median = 7.6</td>
</tr>
<tr>
<td>Δ R\textsubscript{UA}</td>
<td>Average=11.0±9.4 Median = 8.3</td>
<td>Average=5.7±3.0 Median = 6.3</td>
</tr>
<tr>
<td>A</td>
<td>Average=84.7±7.9 Median = 84.5</td>
<td>Average=58.6±5.7 Median = 58.8</td>
</tr>
</tbody>
</table>

A: overall classification accuracy (%); BL: baseline; Δ: change

We previously showed that fluid displacement by LBPP narrows the UA and increases R\textsubscript{UA}, which is a potential mechanism for OSA [109, 115, 116]. Progressive narrowing and collapse of the UA is the cause of obstructive apneas. Therefore, detection of the patency status of the UA is a valuable key in diagnosing this disease. To our knowledge, this is the
Figure 5.4: Box plots of $R_{UA}$ values in the $H_{CT}$ and $L_{CT}$ groups. The horizontal lines represent the median.

First study to show the ability of LPC to characterize turbulent breath sounds confirmed by an objective measure of airway narrowing. Snoring sounds have been repeatedly studied for distinguishing people with and without sleep apnea. However, the ability to distinguish totally normal breath sounds, in this study represented by the BL conditions, from those resulting from partial narrowing, in this study represented by peak $R_{UA}$, was not investigated earlier. Snoring in breath sound is analogous to voiced speech. Although snoring is an important sign of OSA, it does not comprise the whole spectrum of breath sounds. For example, snoring might be absent if the cause of apnea originates from the central nervous system, also called central apnea. Also, temporary arousal from sleep could result in increased UA muscle tonus and temporary disappearance of snoring.
In such cases, tissue vibration disappears and breath sounds revert to unvoiced sounds resulting from air passage through the airway. Therefore, a comprehensive acoustic diagnosis should take in consideration such sounds and be able to distinguish those that pass through normal vs. narrowed UA. Future works should focus on aspects needed for clinical implementation of these promising findings. For example, the specific LPC feature that were used by $K$-means to distinguish the different types of breath sounds need to be investigated in each group. Such features can be extracted from the centroids of clusters that were found by $K$-means. Moreover, the relation between those features and the $R_{UA}$ values could be investigated, which could be further developed to non-invasive estimation of airway narrowing. This, in turn, will help create more accurate tools to diagnose sleep apnea that can be used at home, increase access to diagnosis, and ameliorate the burden of this disease.
Chapter 6

Preprocessing of Breathing Acoustic Data

The content of this paper was accepted as peer-reviewed IEEE conference proceedings: Hisham Alshaer, Geoff R. Fernie, Ervin Sejdić, T. Douglas Bradley, “Adaptive Segmentation and Normalization of Breathing Acoustic Data of Subjects with Obstructive Sleep Apnea” in the proceedings of IEEE Toronto International Conference on Science and Technology for Humanity (TIC-STH), © 2009. The material presented in this chapter is almost identical to the published paper except for the formatting and some organizational and stylistic improvements, which conforms with the University of Toronto thesis style. IEEE does not require individuals working on a thesis to obtain a formal reuse license\(^1\).

Contribution of authors: Alshaer developed the iterative and adaptive method de-

\(^1\)Reprinted with permission from the senior author Bradley TD as per the IEEE reuse policy. In reference to IEEE copyrighted material, which is used with permission in this thesis, the IEEE does not endorse any of University of Toronto products or services. Internal or personal use of this material is permitted. If interested in reprinting/republishing IEEE copyrighted material for advertising or promotional purposes or for creating new collective works for resale or redistribution, please go to http://www.ieee.org/publications_standards/publications/rights/rights_link.html
scribed in this paper and wrote most of the manuscript. Fernie reviewed the mathematical content of this paper and led the project. Sejdič helped in editing the mathematical content and equations. Bradley reviewed the integrity of the manuscript and revised the medical content.

In this chapter, I present an algorithm for preprocessing breathing acoustic data in preparation of respiratory events detection in order to calculated the AHI and establish a diagnosis. Breath sounds are very dynamic in nature and, therefore, the baseline changes frequently vary throughout the course of sleep. Since apneas and hypopneas are deviations from baseline, it is essential to find the true baseline in order to identify those events. I proposed an adaptive algorithm for finding transitions in breath sounds levels and normalizing individual segments for establishing a homogenous baseline. Although several methods were developed for similar purposes in other fields of signal processing such as for EEG and for music signal processing, this method was developed especially for breath sounds and for the diagnosis of sleep apnea. The proposed method uses the signal’s envelope (envelogram) as a tool to track trends in the level of activity as demonstrated by Rangayyan [117].

6.1 Abstract

Breath sounds in patients with OSA are very dynamic and variable signals due to their versatile nature. In this paper, an adaptive segmentation algorithm for these sounds is presented. The algorithm divides the breath sounds into segments with similar amplitude levels. As the first step, the proposed scheme creates an envelope of the signal characterizing its long term amplitude variations. Then, $K$-means clustering is iteratively applied to detect borders between different segments in the envelope, which will then be used to segment and normalize the original signal.

Keywords- sleep apnea, acoustic analysis, breath sounds, adaptive segmentation, $K$-
means clustering

6.2 Introduction

Acoustic analysis of respiratory sounds has gained an increasing role in the diagnosis of respiratory disorders because of advances in techniques for sound measurement and signal analysis [90]. For example, respiratory sound analysis has been used to identify pathological respiratory sounds, such as wheezes and crackles [47, 48, 49, 50, 51, 52, 53]. In addition, sound recording during sleep has also been used to diagnose OSA [63, 118], a condition characterized by repeated cessation of airflow due to complete (apnea) or partial (hypopnea) collapse of the pharynx. For simplicity, both types of events will be referred to as apnea in this chapter. These events alternate with episodes of hyperventilation (i.e., hyperpnea) during which loud snoring occurs [6]. Recurrent apneas lead to intermittent hypoxia that provokes arousal from sleep and sleep fragmentation, thus, causing restless sleep and excessive daytime sleepiness. Repetitive apneas and intermittent hypoxia also elicit repetitive surges in blood pressure at night and increase the risk of developing daytime hypertension, atherosclerosis, heart failure, and stroke, independently of other risk factors [1, 2, 3, 4, 5]. An individual apnea is usually 10-30 seconds in duration and can take place as frequently as 100 times per hour of sleep.

In standard respiratory practice, the most common signals used to monitor and diagnose OSA are nasal airflow and thoracic and abdominal movement. These signals are quasi-stationary over periods longer than three minutes and show minimal change in overall amplitude over extended time duration. In such signals, individual apneas can be identified as a gradual decrease in amplitude of the waveform (decrescendo pattern) followed by a complete cessation in the case of apnea, or a reduction in the signal amplitude in the case of hypopnea. On the other hand, breath sounds in patients with OSA are very dynamic with wide variations in amplitude due to alternating periods of loud snoring,
mild snoring, snorting (intermittent snoring), normal breathing, hyperventilation, and silence during apneas. Apneas manifest as a reduction in amplitude of the sound waveform with a pattern similar to that described above. However, during the ventilatory phase between apneas, the amplitude of breath sounds tends to vary widely over time during sleep due to the varying nature of breath sounds. Apneas that occur in the middle of low level sounds will, therefore, be difficult to detect by an examiner. Therefore, in order for an examiner to detect the apneas in a breath sound waveform, the waveform should be normalized in amplitude. This normalization could be performed manually by detecting and normalizing segments that have stable ventilatory levels. However, such a process is time consuming, subject to examiner error, and requires medical personnel familiar with OSA and breath sounds. Another approach is fixed segmentation in which the waveform is sliced into short segments, presumably quasi-stationary and with a fixed length. Then each segment is normalized independently. This method results in segmentation of data regardless of the degree of stationarity. Thus, this method can yield false positives, i.e., more segments than actually exist. Furthermore, whenever a segment is created and normalized, a distortion can be introduced at the border with other segments, potentially resulting in some information loss. Therefore, the best method is an adaptive segmentation algorithm that segments the data only when the characteristics of the signal change significantly.

When developing an automatic normalization algorithm, two important challenges specific to the breath sounds of patients with OSA are encountered. First, this segmentation process should not detect apnea as an individual segment. In such a case, the waveform within the apnea region will be normalized according to its own level, which will equalize it with the adjacent signals and make it impossible to detect. The second challenge is the presence of transient breath sounds that could interfere with the segmentation. Those considerations should be dealt with when normalizing the breath sounds signal of patients with OSA. The aim of this work, therefore, is to develop a
segmentation algorithm that automatically segments and normalizes breath sounds data of subjects with OSA.

6.3 Data Acquisition

Data were collected from patients referred for an overnight PSG. Breath sounds were recorded by a cardioid condenser microphone (Audi-Technica condenser microphone, Model PRO 35x). The microphone’s cardioid polar pattern reduces pickup of sounds from the sides and rear, improving isolation of the sound source. The microphone was embedded in the center of a loose-fitting plastic frame held in place with a headstrap in a fixed location, approximately 3 cm in front of the subject’s face. Digitized sound data were transferred to a computer using a USB preamplifier and audio interface (M-Audio, Model MobilePre USB) with a sampling rate of 22050 Hz and a resolution of 16 bits. The external audio interface was preferred over the regular built-in audio adapters because of its superior SNR ratio which is 91 dB. Figure 6.1 shows an 80 second sound recording with a representative apnea. The study protocol was approved by the local research ethics board.

![Figure 6.1](image)

**Figure 6.1:** A representative apnea, approximately 20 seconds in duration, is shown as an interruption of the breath sounds waveform. AU=arbitrary units.
6.4 Proposed Algorithm

The proposed algorithm consists of two major parts that are described in the subsequent sections.

6.4.1 Signal Envelope Creation

The first step of the proposed algorithm is the creation of so-called low-pass envelopes. These signal envelopes are created to detect overall changes in the amplitude of the acquired signal, $x(n)$, of length $N$, where $0 \leq n \leq N - 1$. Envelopes are calculated in such a way that sharp transitions in the signal’s levels are maintained because they represent the borders of varying segments.

As a first step, the signal is divided into $K$ non-overlapping segments, where the number of segments can be calculated as:

$$K = \left\lfloor \frac{N}{L} \right\rfloor$$  \hspace{1cm} (6.1)

with $L$ being the desired segment length and the symbol $\left\lfloor \right\rfloor$ representing the greatest integer function. In this paper, $L = 11025$ points (or 500 ms according to the sampling rate). This interval is chosen in order to preserve the fine details of individual breaths, such as inspiratory and expiratory phases. The shortest breathing phase is 1.5 seconds in rapid normal breathing (20 breaths/minute); thus, the bin size (500 ms) provides sufficient resolution to capture breathing details. Accordingly, the sampling rate of the new envelope ($E_{FS}$) becomes 2 Hz. On the other hand, extending this interval might result in the merging of apnea borders and thus, a false representation of the apnea’s duration. Additionally, transient high amplitude outliers produced by coughing and snorting (transient load snoring) will merge with the surrounding signals, thus making it more difficult to remove them in the next steps. It should be noted that $N$ is not necessarily an integer multiple of $K$, making it necessary to omit some of the data points.
Therefore, the data are trimmed from the end of the signal. The trimmed version of the signal is denoted by $x_{tr}(m)$ where $0 \leq m \leq M - 1$ and $M = KL \leq N$. As a second step, we form a vector $e \in \mathbb{R}^K$ of squared values whose points are assigned as follows:

$$e(k) = \sum_{j=kL}^{(k+1)L-1} x_{tr}^2(j) \quad \text{for } 0 \leq k \leq K - 1 \quad (6.2)$$

The resulting signal is a train of peaks, each representing a breathing phase, which is interrupted by apneas as illustrated in a 3-minute recording in Figure 6.2. Wide variations in overall signal levels take place over periods longer than 15 minutes as will be illustrated in later sections. Outliers can result in high amplitude spikes in the breathing envelope. These outliers can affect subsequent statistics and therefore should be removed. In order to remove these outliers, $e(k)$ is segmented into short overlapping intervals, $s(p)$, of length $P$ with each segment representing a pattern of breathing. Mathematically,

$$s(p) = e(p + rq) \quad (6.3)$$

Figure 6.2: $e(k)$ waveform showing alternating periods of ventilation and apneas, where the peaks represent breaths where $0 \leq p \leq P - 1$, $q$ is the amount of the overlap, $0 \leq r \leq R - 1$ and $R = \left\lfloor K/q \right\rfloor$. In this case, the presence of apneas and various breathing patterns should be considered. In patients with severe sleep apnea (i.e., worst case scenario), breathing is present only
50% of the time and is interrupted by apneas that are approximately 30 seconds in duration. Thus, approximately every 60 seconds, an alternating pattern of apnea and ventilation occurs during sleep and this constitutes the basic unit of segmentation. In order to incorporate multiple patterns in one window, we chose a segmentation window (P) of 180 seconds, i.e., 3 times the aforementioned basic unit of 60 seconds. This window slides in 30 second intervals (q). The windows length in samples is given by: window length in seconds x $E_{FS}$. Accordingly, P=180x2=360 samples and q=30x2=60 samples, given that each second of $e(k)$ is represented by 2 samples (from $E_{FS} =2$ derived in Section 6.4.1). Next, for each segment the standard deviation is calculated:

$$\sigma_s = \sqrt{\frac{1}{P} \sum_{p=0}^{P-1} (s(p) - \mu_s)^2}$$  (6.4)

where $\mu_s$ is the mean value of the segment. Using the standard deviation, we form a segment-based indicator sequence, $I_s(p)$, based on the following rule:

$$I_s(p) = \begin{cases} 
1 & \text{if } s(p) > 4\sigma_s \\
0 & \text{otherwise}
\end{cases}$$

Using this segment-based sequence, we form a reassigned version of $e(k)$ using the following rule:

$$e_r(p + rq) = \begin{cases} 
4\sigma_s & \text{if } I_s(p) = 1 \\
e(p + rq) & \text{otherwise}
\end{cases}$$

It should be pointed out that, in the case of consecutive points that indicate the presence of outliers, the duration of these consecutive points should not exceed 5% of the length of the segment. Otherwise, the strong amplitude deviations are not considered outliers, but could still contain physiologically relevant information. The next step is to trace the overall changes in the waveform’s level by interpolating its maxima. To find the local maxima points, we compare each value of $e_r(k)$ against its neighbouring values.
If this value is greater than both of its neighbours, then it is a local maximum. Using the detected local maxima, we interpolate points between these maxima using piecewise cubic Hermite interpolation [119]. The resultant curve connects individual breath peaks and represents short-term changes in signal level and is conventionally called the Short-term Envelope (SE) as shown in Figure 6.3. Next, we created another envelope that traces long term changes that bypass local variations. This envelope is created by implementing the same method described for finding the SE to get the long-term envelope, by interpolating the maxima of the SE. The resultant envelope is normalized in amplitude and is referred to as the level envelope (LE) as shown in Figure 6.3(D). Normalization in this work is achieved by dividing any time series by the value at the 95\textsuperscript{th} percentile rather than the maximum in order to avoid dividing by local outliers that could result at the borders of segments. Only the LE will be used in the next step.

### 6.4.2 Segmentation and Normalization

The second part of the algorithm essentially determines the number of segments having variable amplitude in the LE, if any. The first step is to evaluate the standard deviation of the LE:

\[
\sigma_{LE} = \sqrt{\frac{1}{K} \sum_{k=0}^{K-1} (LE(k) - \mu_{LE})^2}
\]  

(6.5)

where \(\mu_{LE}\) is the mean value of the LE\((k)\). This is based on the fact that an obstructive apnea is the reduction of breathing effort to below 50\% of the baseline volume. Thus, the change in overall signal level of the ventilatory components should be sufficiently less variable than 50\% for apneas to be easily distinguishable. Therefore, a threshold of \(\sigma_{LE}=0.10\) for an acceptably smooth LE was selected. Accordingly, if \(\sigma_{LE}>0.10\), the LE amplitude has strong variations and is subject to normalization, as in the signals in Figures 6.3, 6.5, and 6.6.

The main challenge in this work is the lack of knowledge about the number of clusters
Figure 6.3: Illustration of different stages of the algorithm on an artificial set of data simulating breathing interrupted by apneas. A. $e(k)$ waveform with the SE, LE envelopes. B. Dashed line shows the location of the border between different segments found in the 1st iteration and resultant $e(k)$ after normalization based on that border. C. The location of all borders after the 2nd iteration and $e(k)$ after the 2nd normalization. D. magnified sample of A.

in a given breathing recording and the length of each cluster. A cluster, in this case, is a segment with a stable/similar ventilatory level. Therefore, in order to find clusters with different amplitude levels, the $K$-means clustering technique [120] was implemented to segment the LE into two initial clusters. Once the initial clustering is done, the amplitudes of both clusters are normalized (by the 95th percentile). The normalized
segments of LE are joined again to form a new LE. $\sigma_{LE}$ is then tested and this process is repeated iteratively until $\sigma_{LE}$ is less than 0.10. This process is illustrated in Figure 6.4. Therefore, in each iteration, a new segment is found and normalized until all the segments of LE reach an equal level. Clustering LE into only 2 clusters in each iteration overcomes the absence of knowledge about the number of clusters. Moreover, implementing $K$-means overcomes the challenge of locating the points at which the signals change in level and determining the length of each segment. Subsequently, the points at which the LE was segmented using this method are considered the borders between segments with variable levels in the original acoustic data, which are then normalized independently.

6.5 Results and Discussion

To test the adaptive segmentation algorithm, a simulated breathing waveform was created. Figure 6.3(A) shows the simulated $e(k)$ 15 minutes in duration, with a breathing
rate of 15 breaths/minute and 22 apneas interrupting breathing. Each apnea is 20 seconds in duration. The signal has three different levels. The algorithm needed two iterations to segment and to normalize the signal. As a result, all ventilatory and apneic components of each segment reached the same level as the ventilatory and apneic components of all other segments, so that all apneas became equally visible. It is worthwhile mentioning that the method used to create the envelope of e(k) guarantees the preservation of sharp transitions in a signal’s level, such as that occurring at five minutes in Figure 6.3(A). Other methods such as low-pass filtering might result in losing this information. Figure 6.5 shows the performance of this algorithm on a 30-minute recording from a patient with sleep apnea. Figure 6.5(A) represents e(k) of the raw acoustic data. This signal has variable levels and several outliers. The apneas can be spotted in the high level segment of the signal (between 12-22 minutes); however, apneas in low level segments, such as between 10-12 minutes, are obscured. Figure 6.5(B) shows the signal after removing the outliers. Some high amplitude breath sounds were also truncated in this process; however, this does not affect the performance of the algorithm and does not obscure apneas. The locations of the borders suggested by the segmentation algorithm are also shown as dashed lines. Figure 6.5(C) shows the final result after the normalization of each segment separately by its 95th percentile. Apneas appear as valleys in waveforms 30 seconds in duration or more. They are approximately equal in amplitude and depth and are distributed all through the recording. This forms a pattern that better facilitates the identification of apneas compared to the original shown in Figure 6.5(A).

The ultimate purpose of studying breath sounds is not only to enable the detection of individual apneas but also to distinguish subjects who have sleep apnea from those who do not. Therefore, although the design of this algorithm has considered apneic breathing, it is important for such a segmentation algorithm to function universally, without corrupting data of subjects who do not have the disease. Figure 6.6 presents the segmentation and normalization of e(k) taken from a subject who does not have sleep
Figure 6.5: Segmentation and normalization of a 30 minute recordings from a patient with severe sleep apnea. A. \( e(k) \) waveform with variable levels and several outliers. B. \( e(k) \) after removing the outliers. Dashed lines represent the location at which the signal is segmented as found by the algorithm. C. \( e(k) \) after normalization.
apnea. By comparing Figure 6.5(C) and Figure 6.6(C), the subject with sleep apnea can be distinguished from the subject who does not by the presence of characteristic troughs 30 seconds in duration. This shows that the algorithm did not create artifacts in the normal breathing recording that could be mistaken for apneas.
Figure 6.6: Segmentation and normalization of a 30 minute recording from a patient without sleep apnea from the same stages as those included in Figure 6.5
6.6 Conclusion

This chapter presented a $K$-means based algorithm for adaptive segmentation of breath sounds of patients with sleep apnea. The proposed scheme was based on long term variations in signal levels. I have shown that the proposed algorithm is capable of segmenting waveforms of breath sounds with an unknown number of segments and of an unknown duration.
Chapter 7

Detection of Apneas and Hypopneas

The material of this chapter is based on the paper: Hisham Alshaer, Geoff R Fernie, Ellen Maki, T Douglas Bradley, “Validation of an Automated Algorithm for Detecting Apneas and Hypopneas by Acoustic Analysis of Breath Sounds”, Sleep Medicine (2013). The chapter is almost identical to the published paper except for formatting and some organizational and stylistic improvements, which conform with the University of Toronto thesis style. The publisher, Elsevier, allows the authors to use their papers in subsequent compilations of the their works.

In this study, I show that applying a sophisticated computer algorithm to breath sounds acquired from a microphone embedded in a face frame during sleep provides a reliable means for detecting apneas and hypopneas and quantifying the AHI compared to PSG without requiring an additional input such as oximetry. Authors’ contributions: Alshaer developed the device and algorithm for acoustic analysis of breath sounds, recruited subjects, acquired and analyzed the results, drafted and revised the manuscript. Dr Fernie helped in designing the study, provided technical advice on mathematical and engineering aspects of the project and revised the manuscript. Dr Maki developed sta-

---

1For more information please refer to Elsevier’s policy on ‘Authors’ Rights and Responsibilities’ at this link: [http://www.elsevier.com/authors/author-rights-and-responsibilities](http://www.elsevier.com/authors/author-rights-and-responsibilities)
Prior to this stage of the project, I have tried a variety of methods to detect apneas and hypopneas in breath sound signals. For example, a set of representative templates of apneas and hypopneas were selected. These were then used as references for event detection by means of autocorrelation-based template matching. The template matching technique was discarded because of the wide variations of hypopnea patterns. This required adding more templates, which was impractical. Eventually, the author decided to examine each potential event against a set of rules that is derived from the fundamental definition of apneas and hypopneas and from their pathophysiological properties in order to overcome the shortcoming of template matching. Techniques and algorithms described in previous chapters were carried over and integrated in this part of the work.

Sleep apnea was referred to as ‘sleep disordered breathing’ in the original work. In this chapter the term ‘sleep apnea’ will be used instead for consistency with the rest of this thesis.

7.1 Abstract

Background: Sleep apnea is common and associated with increased cardiovascular disease risk. However, most patients remain undiagnosed due to lack of access to sleep laboratories. I therefore tested the validity of a single-channel monitoring setup that captures and analyzes breath sounds to detect sleep apnea.

Methods: Breath sounds were recorded from 50 patients undergoing simultaneous PSG. Using custom-designed automatic software, breath sounds were subjected to a set of pattern recognition rules to identify apneas and hypopneas from which the acoustic apnea-hypopnea index (AHI-a) was calculated. Apneas and hypopneas from PSG were
scored by three technicians blinded to AHI-a and the PSG-derived AHI scored by the other two technicians according to two criteria; one relying solely on the drop of the respiratory signal by $>90\%$ for an apnea and by $50$-$90\%$ for a hypopnea for $\geq 10$ seconds (TV50 criteria), and another that also required a desaturation or an arousal for a hypopnea (American Association of Sleep Medicine, AASM criteria). PSG AHI (AHI-p) was calculated for each technician according to both criteria.

**Results:** There was no significant difference between AHI-p scores according to TV50 and AASM criteria. AHI-a correlated strongly with AHI-p according to both TV50 ($R=94\%$) and AASM criteria ($R=93\%$). Bland-Altman analysis revealed that $98\%$ and $92\%$ of AHI-a fell within the limits of agreement for AHI-p according to TV50 and AASM criteria, respectively. Based on a diagnostic cut-off of AHI-p $\geq 10$ for sleep apnea, overall accuracy of AHI-a reached $88\%$ while negative predictive value reached $100\%$.

**Conclusion:** Acoustic analysis of breath sounds is a reliable method for quantifying AHI and diagnosing sleep apnea compared to simultaneous PSG.

### 7.2 Introduction

Sleep-disordered breathing (SDB) is associated with poor sleep and hypersomnolence that causes daytime fatigue and increases the risk of motor vehicle accidents [17]. OSA, which is the most common type of SDB, also increases the risk of developing hypertension, HF, and stroke [4, 5], and of death from cardiovascular diseases [20]. Patients with untreated SDB consume twice as many healthcare resources for treatment of cardiorespiratory diseases as subjects without SDB [18]. On the other hand, treating SDB alleviates hypersomnolence, and lowers blood pressure and improves cardiovascular function in patients with hypertension or HF [121, 122, 123, 124, 125]. Therefore, widespread diagnosis and treatment of SDB could have a significant beneficial medical and public health impact [126]. Unfortunately, it has been estimated that up to $85\%$ of people with
SDB remain undiagnosed due to the lack of awareness of the disease and lack of accessibility to a sleep laboratory [17]. Therefore, there is an increasing demand for developing reliable yet simple instruments to diagnose SDB that are more accessible and less costly than PSG.

Several attempts have been made towards creating portable monitors for SDB that are less expensive and more available than PSG and can be used in the patients’ home. Most of these devices reproduced a subset of PSG channels in a more compact form such as nasal flow, oximetry, and thoracoabdominal effort [127]. This approach, although resulting in relatively less expensive montages than PSG, still requires a combination of channels in order to achieve an acceptable accuracy [128]. It is well-known, however, that the more the channels that are added to a portable monitor, the more difficult it is to use and the higher the failure rates in unattended settings [127].

Breath sounds have recently emerged as a rich source of data on respiratory patterns. Several groups have shown that acoustic analysis of breath sounds can be used to identify pathological respiratory sounds such as wheezing [129] and crackles [47], as well as identification of snoring site [55, 56]. In the quest for a reliable and simple home monitor for SDB, acoustic analysis of breath sounds has also been used to distinguish normal breath sounds and simple snoring from those resulting from SDB [130, 60, 63, 131, 132]. Although such techniques could have utility in screening for SDB, in medical practice, knowledge of disease severity in terms of frequency of apneas and hypopneas per hour (AHI) is usually taken into account when recommending treatment. Hence, to improve accuracy and reliability of acoustic analysis of breath sounds for diagnosing SDB, respiratory sound analysis should be able to identify individual apneas and hypopneas. The objective of this study, therefore, was to develop and test the accuracy of acoustic analysis of overnight breath sound recordings to detect the presence and quantify the severity of SDB.
7.3 Methods

7.3.1 Subjects
Fifty consecutive subjects at least 18 years of age were studied. Subjects were referred for PSGs because of a history suggestive of SDB including at least 2 of the following symptoms: a history of loud habitual snoring, restless sleep, morning headaches or excessive daytime sleepiness. No exclusion criteria were imposed.

7.3.2 Acquisition of Breath Sounds
Breath sounds were recorded by a unidirectional condenser microphone embedded in the centre of a loose fitting face frame, which kept the microphone in a fixed location approximately 3 cm in front of the subject’s face as shown in Figure 7.1. Digitized sound data were transferred to a computer using a USB preamplifier and audio interface (M-Audio, Model MobilePre USB) with a sampling rate of 22050 Hz and resolution of 16 bits.

7.3.3 Polysomnography
Subjects underwent overnight PSG using standard techniques and scoring criteria for sleep stages and arousals from sleep [133, 102]. Thoracoabdominal movements and tidal volume were measured by respiratory inductance RIP [27]. Airflow was measured by nasal pressure cannulae [27] and arterial oxyhemoglobin saturation ($\mathrm{SaO}_2$) by oximetry.

Apneas and hypopneas were scored according to two different criteria. The first was the American Academy of Sleep Medicine (AASM) criteria which defines an apnea as a drop in the respiratory signal, in this study the electronic sum of thoracoabdominal movement, by $\geq 90\%$ lasting $\geq 10$ seconds [134], and a hypopnea as an event that satisfies either of the following two conditions: a drop of the respiratory signal by $\geq 30\%$ lasting $\geq 10$ seconds and accompanied by either a $\geq 4\%$ desaturation or terminated by an arousal,
Chapter 7. Detection of Apneas and Hypopneas

Figure 7.1: Illustration of the face frame and location of the microphone

or a drop of the respiratory signal by $\geq 50\%$ lasting $\geq 10$ seconds and accompanied by either a $\geq 3\%$ desaturation or terminated by an arousal [134]. For the second criteria, apneas were similarly defined, but hypopneas were defined as a 50% to 90% reduction in thoracoabdominal sum lasting $\geq 10$ seconds, regardless of any desaturation or arousal as described previously [135], herein referred to as TV50. This was done because the implemented acoustic recording setup does not include oximetry. The AHI was quantified as the number of apneas and hypopneas per hour of sleep. The protocol was approved by the Research Ethics Board of Toronto Rehabilitation Institute and subjects provided written consent prior to participation.
7.3.4 Development of the Automated Algorithm

My approach for detecting apneas and hypopneas in this work is to scan breath sound waveforms for apnea and hypopnea-specific features. The features were derived from the basic definitions of apneas and hypopneas and their pathophysiological properties. The algorithm was developed to detect respiratory events based on the way a sleep technician would manually identify them in other traces such as nasal airflow or thoracoabdominal effort, i.e., by finding a baseline and the characteristics of signal reductions from the baseline. For this purpose, raw breath sound waveforms are pre-processed to obtain a more uniform version, which is then subjected to a set of mathematical rules each to examine a certain feature as described hereafter.

Transformation of the Raw Acoustic Signals

The aim of this step was to convert the raw acoustic signals into waveforms proportional to breath sound amplitude with a uniform baseline. To do this, I used the technique of adaptive segmentation and normalization whose mathematical and physiological bases were described previously [99] and are mentioned briefly in this section. Initially, the envelope of breath sounds was formed by the summation of absolute values of the raw sound signal samples within 400 ms long moving windows (L) overlapping by 75%\(^2\). The resulting envelope models individual breathing cycles and is referred to as breathing envelope (BE) as presented in Figure 7.2. Transient outliers in BE, such as coughs and transient loud snorting were removed. BE models all the remaining breath sounds including inspiration, expiration, and regular snoring.

Subsequently, a second envelope that traces the longer term variations was formed by interpolating the maxima of BE to create another envelope that is equal in length to BE. This latter is referred to as effort envelope or EE, as illustrated in Figure 7.2b.

\(^2\)The values used in this work are slightly modified from the ones used earlier in [99]
EE was normalized to establish a uniform baseline using the adaptive segmentation and normalization method I described earlier [99]. The new sampling frequency of EE is given by:

$$ Fs = \frac{1}{L \times (1 - \text{Overlap Ratio})} \quad (7.1) $$

**Scanning for Respiratory Events**

The normalized and outlier-free representation of respiratory activities, EE, is expected to contain the characteristic patterns of apneas and hypopneas. Scanning for potential respiratory events (PE) starts by detecting all valleys in the EE that drop below an empirical threshold of 0.4 of the standard deviation of EE as in equation 7.2:

$$ PE \in \left\{ EE < 0.4 \times \sqrt{\frac{\sum_{i=1}^{n} (EE_i - EE)^2}{n - 1}} \right\} \quad (7.2) $$

where $n$ is the number of elements in EE, which equals the duration of EE in seconds $\times$ Fs (found in equation 7.1). This threshold was selected to detect signal locations that have sufficient depth that might correspond to a drop in respiration and at the same time exclude negligible troughs that could be attributed to breath-to-breath variation (Figure 7.2b). Each PE segment is then extracted starting from the peak preceding to the peak following its minimum, as illustrated in Figure 7.3. Each PE segment is normalized to unity to facilitate subsequent testing.

Since individual apneas and hypopneas differ in their nature, their manifestations in breath sounds are also different. This gives rise to the main challenge in detecting respiratory events, which is to distinguish true events from a random non-pathological variation in ventilation such as due to snoring episodes. Accordingly, the design of a pattern recognition system should incorporate features of respiratory events that are flexible enough to encompass their different forms, yet sufficiently specific to exclude non-pathological
Figure 7.2: A three-minute segment of (a): raw acoustic signal waveform and (b): the corresponding breath envelope (BE) and effort envelope (EE) and 2 instances of potential apneas (PE). AU: arbitrary units
variation in ventilation and breath sounds. Therefore, my strategy was to develop a set of rules based on the basic definition and pathophysiology of apneas and hypopneas. These rules take in consideration the nature of pre and post apneic ventilation. Two sets of rules have been created, one for apneas and another for hypopneas, which are outlined separately below.

**Apnea Tests**

The hallmark that differentiates an apnea from a hypopnea is the complete absence of breathing and breath sounds during an apnea due to complete collapse of the upper airway in OSA or complete cessation of respiratory drive in CSA (Figure 7.3a). On the other hand, ventilation and breath sounds continue to occur during a hypopnea. A PE is classified as an apnea if the following tests are positive:

1. **Flatness Test** is performed to check the absence of breath sound signal data corresponding to an equivalent breathing cessation interval. This test is positive if the PE has a flat segment between the side edges (ventilation) characterized zeros or near-zero (<0.01) amplitude in the normalized PE (Figure 7.3a).

2. **Width Test**: The the length of the flat segment in seconds is calculated using equation 7.3:

   \[
   \text{Flat Segment Length (s)} = \frac{\text{Number of PE samples} < 0.01}{F_s} \tag{7.3}
   \]

   The test is positive if the Flat Segment Length was \( \geq 10 \) seconds by the definition of apnea.

3. **Depth Test**: In an apnea, the flat segment should lie between periods of ventilation that have an amplitude close to baseline level. This is tested by calculating the amplitude of the higher of the two apnea borders. The depth test is positive if the amplitude of either of the borders is \( \geq 0.9 \), which is physically interpreted as reduction in ventilation by greater than 90%.
Figure 7.3: Representative samples of (a): an apnea and (b): a hypopnea, with labelling of key features used to identify each from breath sound envelopes
Chapter 7. Detection of Apneas and Hypopneas

PE is classified as an apnea if the three apnea tests are positive. Otherwise, hypopnea tests are executed.

**Hypopnea Tests**

Hypopneas are characterized by partial collapse of the upper airway and reduction in airflow by 50% but above zero. Breathing and breath sounds continue during the course of the event. The following tests were developed to classify PE as a hypopnea:

**1-Falling Edge Test** is based on the assumption that a hypopnea evolves as a gradual reduction in net airflow as a result of gradual collapse of the upper airway in the obstructive type [136], or gradual decrease in respiratory drive in the central type [9] (Figure 7.3b). This reduction, however, does not always manifest in an ideal smooth negative slope because of the variable nature of breath sounds on a breath-to-breath basis. Therefore, the non-linearity of the drop in breath sounds amplitude prior to the hypopnea should be taken in consideration in this test. First, the falling edge (FE) of the PE is extracted as the segment starting from the first point of the PE until its minimum point. Then, the discrete derivative of FE (ΔFE) is calculated, which is the difference between each point and the preceding one. If FE is decreasing at all points, then the derivative will consist of negative values only. Positive elements in the ΔFE array arise when transient peaks are present. The sum of only positive elements in ΔFE is found and extracted from the total sum of ΔFE. The result is divided by the difference between the maximum and minimum points in FE to normalize it by the amplitude of the FE. The absolute value of the result is called the *falling edge factor* (FEF) and can be obtained using equation 7.4 that summarizes the aforementioned steps:

\[
FEF = \left| \frac{\sum \Delta FE - \sum (\Delta FE > 0)}{FE_{max} - FE_{min}} \right| \quad (7.4)
\]

FEF will always have a value from 0 to 1. An FEF of 1 means that amplitudes of successive breaths decrease gradually (without transient increase) from baseline to the
minimum level in a hypopnea. On the other hand, if a falling edge contains transient peaks, the FE derivative will contain positive values that will decrease the numerator of Equation 7.4. Thus, FEF will be less than 1 depending on the number of peaks that interrupt the gradual decreasing pattern and their relative amplitude. The falling edge test is considered positive if FEF is 0.7 or greater.

2- **Width Test**, which is a measure of the PE duration. The width test is performed by measuring the time interval (duration) between the 2 borders of the PE at the lower quarter of the PE given by the equation:

$$\text{PE duration} = \frac{||PE_{lq}||}{F_s}$$

(7.5)

where $||PE_{lq}||$ denotes the number of elements in the lower quarter of PE (Figure 7.3a). The width test is positive if it is greater than 10 seconds.

3- **Depth Test**, which is similar to the one used to evaluate an apnea and calculated similarly as the difference between the maximum and minimum values of the PE. The maxima appear at the start and end points of PE. The starting peak represents the level of the pre-apneic breathing. The end peak, on the other hand, corresponds to the post-apneic hyperventilation, which is usually higher in amplitude. Therefore, it stands to reason to expect that the end peak is higher than the start peak. Accordingly, I define a threshold for the starting peak of 0.5 and 0.8 for the end peak. This threshold is lower than that of the set for the apnea, in which total cessation of breathing takes place. The combination of those rules (falling edge, width, and depth tests) was designed to encompass the specific physiological characteristics of hypopneas, yet flexible enough to detect different forms that result from the dynamic nature of breath sounds.

These steps are illustrated in the algorithm flowchart in Figure 7.4. The acoustically-determined AHI (AHI-a) was calculated as the number of apneas and hypopneas per
hour of recording time, which corresponded to time-in-bed.

**Algorithm Implementation**

A subset of PSGs from 5 subjects with mild, moderate, and severe SDB was used to test the functionality of the aforementioned pattern recognition rules. Once the algorithm’s functionality was verified with this subset, it was applied on all the 50 subjects and the AHIs generated were compared with AHIs scored from PSGs as explained below.

### 7.3.5 Algorithm Evaluation and Statistical Analysis

The AHI obtained from acoustic recordings (AHI-a) was compared with that obtained from PSG (AHI-p) as follows. Since the AHI-p is obtained by a technician visually scoring the PSG recordings, one would anticipate variability in scoring between technicians for the same PSG. This is mostly due to subjectivity in evaluating borderline events. As a consequence, if there is a difference between AHI-a and AHI-p, some of this difference might be attributable to the inherent variability in the AHI-p. To determine the degree of inter-rater variability in the scoring of the AHI, three research technicians scored the AHI of each of the 50 patients, blinded to the original clinically scored AHI and to the AHIs scored by the other technicians and to the AHI-a. Therefore, the biomedical engineer did not have previous knowledge of the AHIs scored by these 3 technicians. Since the AHI-p scores of the three technicians represent the reference standard, we assessed the degree of agreement amongst the three technicians prior to comparison with the AHI-a. The inter-rater reliability among the three technicians and its 95% confidence interval were calculated using a two-factor mixed effects ANOVA model as described in the online statistical supplement. The agreement between the two methods was assessed by Pearson correlation and Bland-Altman limits of agreement. For those analyses, the AHI was evaluated according to the time-in-bed period rather than sleep time to simulate home recordings of breath sounds where sleep stages are not recorded. Correlation coefficients
with all the three scorers were calculated using pairwise differences in Pearson correlation and using bootstrap (n=2000) to obtain the 95% confidence interval (CI).

To assess the ability of acoustic analysis to distinguish between the presence or absence of SDB, sensitivity, specificity, and accuracy were calculated using a cut-off AHI of \( \geq 10 \). To assess the reliability in ruling out SDB, negative predictive values, and negative likelihood ratios were also calculated. These were first calculated according to time-in-bed for both AHI-a and AHI-p, and then, according to time-in-bed for AHI-a and sleep time for AHI-p.
Figure 7.4: Flowchart of the algorithm for detection of apneas and hypopneas from breath sound recordings
7.4 Results

7.4.1 Subjects’ Characteristics

Fifty subjects, 36 men and 14 women, whose characteristics are presented in Table 7.1 were recruited.

Table 7.1: Characteristics of the subjects.

<table>
<thead>
<tr>
<th>Subject Characteristics</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>BMI, kg/m²</td>
<td>30.1±6.4</td>
</tr>
<tr>
<td>Age, yr</td>
<td>53.5±13.5</td>
</tr>
<tr>
<td>Sleep efficiency, %</td>
<td>81.1±14.1</td>
</tr>
<tr>
<td>Sleep time, hr</td>
<td>5.3 ± 1.4</td>
</tr>
<tr>
<td>Mean SaO₂ during sleep, %</td>
<td>94.8 ± 1.8</td>
</tr>
<tr>
<td>Lowest SaO₂ during sleep, %</td>
<td>83.6 ± 10.3</td>
</tr>
</tbody>
</table>

Data are expressed as mean ± standard deviation.

7.4.2 Comparison Between AASM and TV50 PSG Scoring

The average of all AHI-p was 14.9 and 16.6 according to AASM and TV50, respectively, whereas the average for AHI-a was 17.1. Inter-rater reliability between the scores of the three technicians was 0.95 (95% CI of 0.67-0.98) for TV50 and 0.95 (95% CI of 0.63-0.98) for AASM criteria. AASM scores were slightly, but not significantly, lower than TV50 (P=0.59), and there was strong agreement between the TV50 and AASM scores (R = 0.99, Figure 7.5).

7.4.3 Agreement Between AHI-a and AHI-p

There was a strong correlation between AHI-a and AHI-p with a mean R of 0.94 and a 95% CI of 0.87-0.97 according to TV50 criteria, and mean R of 0.93 and 95% CI of
Figure 7.5: Comparison between the polysomnographic (PSG) scores according to American Association of Sleep Medicine (AASM) and TV50 showing the line of identity 0.85-0.96 according to AASM criteria. Figure 7.6 displays the distribution of the AHI-p scored by each of the three technicians and the relationship between the AHI-a and the mean AHI-p for TV50 (plot 7.6a) and AASM (plot 7.6b).

The Bland-Altman limits of agreement were calculated to assess agreement between the AHI-a and the AHI-p of each of the three technicians and the mean of all three. Forty nine (98%) of AHI-a fell within the limits of agreement of the mean AHI-p for TV50 as shown in Figure 7.7. Similarly, 96%, 96%, and 98% of AHI-a scores fell within the limits of agreement of AHI-p scored by technician 1, 2, and 3 respectively. The proportion of AHI-a scores that fell within the limits of agreement of PSG-p according to AASM was 92%, 94%, 92%, and 92% in comparison with technicians 1, 2, 3, and their mean scores, respectively. The Bland-Altman bias and limits of agreements were 1.4±10.3 AHI points for mean score according TV50 criteria and 3.1±10.7 according to the AASM criteria.
Figure 7.6: Distribution of the apnea-hypopnea index by acoustic analysis (AHI-a) and AHI by polysomnography (AHI-p) for the 3 technicians and against the average of the 3 AHI-p. (a): scored according TV50, (b): according to AASM
According to our criteria, a diagnosis of SDB is made if the AHI is $\geq 10$, whereas SDB is ruled out if the AHI is $<10$. In comparing the diagnosis of SDB based on AHI-a to that based on the three AHI-p, a decision rule for combining the diagnoses from the 3 technicians must be obtained. I considered two ways of achieving this. First, I considered a diagnosis based on the average of the three technicians, such that if the mean score was $\geq 10$, then SDB is considered to be present. Second, I considered a diagnosis based on the agreement of AHI-a with at least one technician. In this case, if AHI-a is $\geq 10$ and at least one of the three AHI-p is $\geq 10$, then the AHI-a diagnosis of SDB is considered to be a true positive, whereas a false positive ensues if AHI-a is $\geq 10$ and all three AHI-p are $<10$. The same concept applies to true negative and false negative values. The rationale behind investigating this approach is that the agreement of the acoustic analysis with
one technician indicates that the first lies within the range of inherent variability among
different human scorers, which could indeed result in fluctuations of scores around the
nominal cut-off of AHI 10 among the technicians themselves.

The comparisons of diagnostic accuracy of the AHI-a compared to either the mean of
the three AHI-p values, or compared to the AHI-p scored by one or more technicians using
TV50 or AASM criteria for hypopnea are presented in Tables 7.2 and 7.3. Considering
that the agreement with at least one technician incorporates the range of the three scores
for the same subject, it factors in the inter-rater variability around the nominal cut-off
point. When comparing agreement with at least one of the three technicians, validity
measures were 100%, 73%, and 88% for sensitivity, specificity, and accuracy, respectively,
according to TV50. When comparing against the mean AHI-p those decreased to 95%,
69%, and 84% (Table 7.2). These values were slightly lower when comparing AHI-a
against AHI-p according to the AASM criteria (Table 7.3).

### 7.4.5 Effect of Sleep Time on AHI

When employing PSG for diagnosis of SDB, the AHI is calculated by dividing the number
of apneas and hypopneas by total sleep time. However, because the aim of this study
is towards developing a device to be used in the home setting, which would not record
sleep, this might affect the extrapolation of AHI-a values calculated with time-in-bed as
the denominator to AHI-p values with total sleep time as the denominator. To examine
the effect of this potential limitation, I compared AHI-a based on time-in-bed to AHI-p
according to TV50, but based on total sleep time rather than recording time. Validity
measures revealed improvement over AHI-p based on recording time, with an overall
accuracy up to 90% as displayed in Table 7.4. The average sleep efficiency in the current
study was 80.1%. Seven subjects (14%) had sleep efficiency below 66%. In all of them the
diagnostic decision based on AHI-a agreed with AHI-p according to sleep time. Therefore,
none of the patients with OSA was missed as a result of dividing by the time in bed.
Furthermore, out of those 7 subjects with low sleep efficiency, 4 had the same disease severity category as determined by both measures and only 3 (6%) dropped from a higher disease severity to a lower one according to AHI-a.

**Table 7.2:** Diagnostic agreement according to the TV50 scoring criteria.

<table>
<thead>
<tr>
<th></th>
<th>According to the closest PSG-p</th>
<th>According to the mean PSG-p</th>
</tr>
</thead>
<tbody>
<tr>
<td>sensitivity</td>
<td>100%</td>
<td>95%</td>
</tr>
<tr>
<td>specificity</td>
<td>73%</td>
<td>69%</td>
</tr>
<tr>
<td>Accuracy</td>
<td>88%</td>
<td>84%</td>
</tr>
<tr>
<td>PPV</td>
<td>0.82</td>
<td>0.81</td>
</tr>
<tr>
<td>NPV</td>
<td>1</td>
<td>0.90</td>
</tr>
</tbody>
</table>

PPV, positive predictive value; NPV, negative predictive value

**Table 7.3:** Diagnostic agreement according to the AASM scoring criteria.

<table>
<thead>
<tr>
<th></th>
<th>According to the closest PSG-p</th>
<th>According to the mean PSG-p</th>
</tr>
</thead>
<tbody>
<tr>
<td>sensitivity</td>
<td>100%</td>
<td>96%</td>
</tr>
<tr>
<td>specificity</td>
<td>70%</td>
<td>64%</td>
</tr>
<tr>
<td>Accuracy</td>
<td>86%</td>
<td>81%</td>
</tr>
<tr>
<td>PPV</td>
<td>0.79</td>
<td>0.75</td>
</tr>
<tr>
<td>NPV</td>
<td>1</td>
<td>0.94</td>
</tr>
</tbody>
</table>

PPV, positive predictive value; NPV, negative predictive value
Table 7.4: Diagnostic agreement between AHI-a based on time-in-bed and AHI-p based on total sleep time using TV50.

<table>
<thead>
<tr>
<th></th>
<th>According to the closest PSG-p</th>
<th>According to the mean PSG-p</th>
</tr>
</thead>
<tbody>
<tr>
<td>sensitivity</td>
<td>97%</td>
<td>93%</td>
</tr>
<tr>
<td>specificity</td>
<td>79%</td>
<td>72%</td>
</tr>
<tr>
<td>Accuracy</td>
<td>90%</td>
<td>85%</td>
</tr>
<tr>
<td>PPV</td>
<td>0.88</td>
<td>0.84</td>
</tr>
<tr>
<td>NPV</td>
<td>0.94</td>
<td>0.88</td>
</tr>
</tbody>
</table>

PPV, positive predictive value; NPV, negative predictive value

7.5 Discussion

The most important finding of this study was that acoustically determined AHI, using a novel algorithm, agreed remarkably well with PSG determined AHI. To maximize rigor of the comparisons, I took into account variations in scoring of AHI by having three technicians score the AHI-p blindly according to two different criteria. The overall accuracy for diagnosis of SDB reached 90% with 94% correlation across the spectrum of AHIs, and 98% of AHI-a falling within Bland-Altman limits of agreement with AHI-p.

SDB is common and affects approximately 9% of men and 4% of women in the United States [6]. Nevertheless, approximately 85% remain undiagnosed [17]. Low diagnostic rates can be attributed partly to lack of accessibility to, and high cost of PSG, the current standard for SDB diagnosis. Extrapolating to the adult US population, this corresponds to more than 15 million people with undiagnosed SDB. Since SDB increases the risk of developing cardiovascular diseases [2, 3, 4, 5] this constitutes a major public health problem. Accordingly, to increase accessibility to SDB diagnosis, several portable devices, simpler and cheaper than PSG, have been developed to diagnose SDB at home.

The presented setup consists of a single channel and would therefore be classified as a Level I monitoring device. Level I devices have the advantage of being easier to apply...
than multiple-channel devices (Level II-IV). This is a desirable feature in a portable
monitor because it is more compatible with use in the home environment by untrained
users and reduces failure rates due to electrode detachment. The Portable Monitoring
Task Force of the AASM, for example, recommended developing easy to use portable
sleep apnea monitors that would reduce time required to attach sensors, and simplify
data collection and transfer [137]. On the other hand, currently available single channel
devices lack accuracy in determining the AHI compared to PSG and usually require
additional input to achieve sound accuracy [133, 128]. The present findings indicate
that the herein single-channel microphone input and acoustic analysis algorithm have
accuracy approaching that of PSG for determining the AHI. This approach, therefore,
combines both the clinical reliability and structural simplicity desirable for use in the
unattended home setting.

Breath sounds are rich in information on breathing patterns. For example, alter-
ations in non-snoring components of breath sounds can detect UA narrowing as observed
in OSA [100, 138]. Others have shown utility of breath sound analysis to infer the site of
UA obstruction [56] and estimate airflow [139]. Researchers aiming at diagnosing SDB
from breath sounds have focussed on snoring signal as the main input. For example, sev-
eral studies have implemented acoustic analysis of snore sounds to distinguish between
between simple snorers and snoring in SDB using time domain and frequency domain
features [62, 60, 63, 68, 140, 132]. The purpose of most snore-only driven techniques was
to screen for probability of SDB to determine diagnosis, or to provide a binary decision on
whether the subject has SDB or not. Some studies [141, 142, 143] went a step further by
estimating the correlation between snoring acoustic features and AHI. Although earlier
studies showed high sensitivity, specificity was noticeably low [141]. In more recent stud-
ies, ‘inter-event silence’ was the feature that resulted in the best performance for which
researchers reported good correlation with AHI (R² =0.62) using a floating microphone
[142]. A better performance in detecting the AHI from analysis of breath sounds from
a microphone located over the trachea was reported when combined with an oximetry signal [144]. In the present study, a microphone located in front of the mouth and nose was used. This configuration helps capture all types of breath sounds including normal inspirations and expirations and snoring that are all incorporated in the analysis. I was able to detect individual apneas and hypopneas on an event-by-event basis by transforming breath sounds into a waveform proportional to ventilatory depth. After establishing its baseline, the amplitude modulation envelope is then examined for individual apneas and hypopneas using a set of rules based on their definition and pathophysiological properties. These rules examine the deviation from baseline, the duration of the potential event, and the trend of the ventilation surrounding it. This approach resulted in excellent performance in distinguishing true events from noisy variations in breath sounds, without any additional input channel. The accurate identification of respiratory events by tracking of breath sound envelopes in this study demonstrates that the intensity of breath sounds, including snoring-related and non-snoring-related events, is modulated by the ventilatory depth. Even when breathing continues, such as during hypopneas, the overall intensity of breath sounds, including snoring, is typically reduced when flow limitation is present, and rises again when airflow increases at hypopnea termination.

To rigorously evaluate the agreement between acoustically calculated AHI and PSG, the latter was scored using two different criteria. The TV50 criteria take in consideration the basic physiological definition of apneas and hypopneas that depends solely on a decrease in respiratory volume or effort. Since the method examined in this study does not record SaO₂, the TV50 criteria may be more comparable with the analysis of variations in amplitude of breath sounds that reflect variations in ventilation than the AASM criteria, which requires additional criteria of oxygen desaturation or arousal for defining hypopneas. Thus, one might anticipate that fewer respiratory events would be detected by AASM than by either the TV50 criteria or the criteria used for breath sound analyses described herein. In this study, although AASM derived AHI scores were slightly lower
than those derived from the TV50 as anticipated (average of 14.9 vs. 16.6), there was no significant difference between mean AASM and TV50 derived AHI-p. Additionally, there was a very strong correlation between AHI-p scored according to TV50 and AASM criteria ($R = 0.99$). This suggests that those two criteria are very comparable for determining the AHI. Therefore, breath sound analysis alone, as described herein, appears to be sufficient to determine the AHI and to detect apneas and hypopneas without an $O_2$ desaturation or arousal criterion for hypopnea. Therefore, I will focus on comparisons of AHI-a and AHI-p determined by the TV50 criteria in the following discussion.

Because there is inter-rater variability in a technician-scored AHI-p [137], such variability should be taken into account when determining whether a subject is positive or negative for SDB according to a nominal AHI cut-off of $\geq 10$. In this study, for example, seven subjects (14%) were classified differently as to the presence or absence of SDB according to different AHI scores by the 3 technicians that fluctuated above or below the cut-off AHI of 10. This consideration therefore affects sensitivity and specificity calculations. The accuracy of detecting SDB by AHI-a using the cut-off criterion of 10 was 84% compared to the mean AHI-p of the three technicians. However, when considering agreement with one or more technicians, to account for the full range of inter-rater variability, accuracy increased to 88%.

The high sensitivity of this method could be attributed to the slight but systematic overscorining of respiratory events in the lower range (AHI $< 15$) as illustrated in Figure 7.6. It is clinically safer to over-score than to under-score borderline cases to avoid missing the diagnosis of SDB that may need treatment. Importantly, false positive cases were close to the cut-off AHI point of 10. This limitation can be addressed by defining a zone of uncertainty between the AHI-a of 10 to 18 where false positives lie. Treatment of SDB is ordinarily prescribed for the presence of a SDB syndrome based on an AHI and the symptoms of SDB determined by a clinical evaluation. Therefore, as would be the case for a borderline AHI-p, the clinical significance of an AHI-a in this zone of uncertainty
for a given patient would require a clinical evaluation to assess for symptoms of an SDB syndrome. In the presence of such symptoms, a trial of SDB therapy would be justified, but in the absence of such symptoms, treatment of the borderline AHI-a would not be mandated [21, 145, 146, 147]. The tendency to overscore the AHI from breath sound analysis compared to AHI-p in the lower range did not compromise the ability to discard negative cases as revealed by the NPV of 100% and negative likelihood ratio of zero. These data indicate that an AHI-a <10 reliably rules out the presence of SDB.

During PSG, the AHI is obtained by dividing the number of apneas and hypopneas by total sleep time. However, this requires recording of sleep from multiple electrodes. Since one objective of a portable monitor is to permit its self-application by untrained individuals, the great majority of portable monitors do not record sleep [128]. Consequently, portable monitors generally quantify the AHI using recording time as an estimate of sleep time as the denominator. Since recording time will include periods of wakefulness as well as sleep, it will almost always be longer than sleep time. Therefore, the AHI derived from recording time will usually be somewhat lower than those derived from total sleep time. To evaluate the effect of this limitation in this study, I calculated the validity measure of the recording time-derived AHI-a, versus the sleep time-derived AHI-p. This resulted in 97% sensitivity, 79% specificity, and an overall accuracy of up to 90% for diagnosing SDB with a cut-off AHI of ≥ 10. The validity measures according to sleep time-derived AHI-p were indeed better than that according to recording time-derived AHI-p. This is because sleep time is usually shorter than bed time, and therefore, having a smaller denominator for AHI-p results in higher values for AHI-p. Since the current iteration of the algorithm slightly over scores events in the lower range around the nominal AHI cut-off of ≥ 10, it agreed better with the sleep time-derived AHI-p in that region. Although wrist actigraphy does not record sleep, lack of actigraphic movement has been shown to correlate reasonably well with sleep time [148]. Therefore, its use in conjunction with the presented breath sounds analysis might provide a reasonable estimate of sleep time.
to calculate a more accurate AHI than that derived from time in bed.

Because the presented method does not measure sleep, it cannot detect non-respiratory disorders of sleep. Another limitation is that the method does not measure SaO$_2$. Although this limitation did not reduce the accuracy of acoustically determined AHI, it cannot provide an estimate of the degree of hypoxia associated with SDB. The current algorithm is designed to detect all types of apneas and hypopneas and cannot distinguish obstructive from central events. Although this does not impose a limitation for diagnosing SDB in the general population, in whom the vast majority of SDB is OSA [6], it may present a problem in patient populations who have higher prevalence of central events such as patients with heart failure [149]. These limitations are subject for future research. Finally, although the detection algorithm was designed based on pathophysiological characteristics of apneas and hypopneas, future studies will further examine the robustness of this approach in a larger set of subjects and in more versatile environments.

The AHI derived from the presented device and computer algorithm achieved excellent agreement with the AHI derived from PSG, with high sensitivity and specificity for detecting SDB. Because this method uses only one channel, it has the potential to be used as a portable system that should be easy to use at home by an untrained individual. Future work will be directed at validating the device in the home setting by comparing the AHI derived from home recordings to that obtained from in-laboratory PSG performed on the same subject.
Chapter 8

Implementation of Portable Monitoring in The Home

The material in this chapter is in press as: Hisham Alshaer, Alexander I. Levchenko, T. Douglas Bradley, Wen-Hou Tseng, Geoff R. Fernie, “Development of an Acoustic Portable Sleep Apnea Diagnostic System Using an Embedded Data Capturing Module” in Journal of Clinical Monitoring and Computing, 2013. The paper has been reproduced according Springer’s self-archiving policy.¹

After the successful validation of the acoustic analysis for detecting apneas and hypopneas in the laboratory environment, in this part of the project, the system moved from the laboratory to the home. This migration required the design of a special unit that untrained users can take home and use independently. In this chapter, I examine the concept of home monitoring and diagnosis of sleep apnea using a self contained unit that has no external wires. I evaluated the reliability of this method in terms of hardware functionality, the quality of captured data, and reproducibility of results in the home en-

¹Springer allows author-created version of his/her article on his/her institutional repository. Springer’s self-archiving policy can be found at this link: http://www.springer.com/open+access/authors+rights?SGWID=0-176704-12-683201-0

123
8.1 Abstract

**Introduction**: Sleep apnea is a very common disease with serious health consequences, yet is very under-diagnosed, partially because of the high cost and limited accessibility of in-laboratory PSG. The purpose of this work is to introduce a newly developed portable system for the diagnosis of sleep apnea at home that is both reliable and easy to use.

**Methods**: The system includes personal devices for recording breath sounds and airflow during sleep and diagnostic algorithms to process the recorded data. The data capturing device consists of a wearable face frame with an embedded electronic module featuring a unidirectional microphone, a differential microphone preamplifier, a microcontroller with an onboard differential analogue to digital converter (ADC), and a microSD memory card. The device provides continuous data capturing for 8 hours. Upon completion of the recording session, the memory card is returned to a location for acoustic analysis. Forty nine subjects were recruited to use the device independently at home, after which each subject answered a usability questionnaire. Random data samples were selected to measure the SNR as a gauge of hardware functionality. A subset of 11 subjects used the device on 2 different nights and their results were compared to examine diagnostic reproducibility. Independently of those, the system’s performance was evaluated against PSG in the laboratory environment in 32 subjects.

**Results**: The overall success rate of applying the device in un-attended settings was 94% and the overall rating for ease-of-use was excellent. Signal examination showed excellent capturing of breath sounds with an average SNR of 31.7 dB. Nine of the 11 (82%) subjects had equivalent results on both nights, which is consistent with reported inter-
night variability. The system showed 96% correlation with simultaneously performed in-laboratory PSG.

**Conclusion:** The results of this study suggest excellent usability and performance of this system and provide a strong rationale to further improve it and test its robustness in a larger study.

**Keywords:** home diagnosis, usability, portable system, embedded module

### 8.2 Introduction

Sleep apnea is a breathing disorder characterized by repetitive cessations of breathing during sleep for intervals of 10-90 seconds in length. The frequency of these events ranges from 5 to 100 times/hour depending on the severity of the case. OSA, the most common type sleep apnea, results from collapse of the upper airway either partially (hypopnea) or totally (apnea) during sleep. CSA, on the other hand, results from intermittent loss of central respiratory drive, which can also be partial or complete. These events alternate with episodes of hyperventilation during which loud snoring may occur. These events result in episodes of oxygen deprivation, and arousals from sleep that cause sleep fragmentation. Subsequently, patients suffer from poor sleep quality, daytime sleepiness, and poor cognitive performance. Repetitive apneas and intermittent hypoxia also elicit sympathetic nervous system activation that cause repetitive surges in blood pressure at night and increase the risk of developing daytime hypertension and atherosclerosis independently of other risk factors [1, 2, 3]. Patients with sleep apnea are at three to four fold greater risk of developing heart failure and stroke than subjects without sleep apnea [4, 5]. In the US, it has been estimated that sleep apnea-related motor-vehicle collisions due to sleepiness in the year 2000 caused 1,400 deaths and cost $15.9 billion [24]. Sleep apnea also imposes a significant financial burden on the health system; Canadian data show that patients with untreated sleep apnea consume twice as much health care re-
sources for treatment of cardio-vascular diseases than subjects without sleep apnea [18]. On the other hand, we have demonstrated that treating sleep apnea lowers blood pressure in hypertensive patients, improves cardiovascular function in patients with heart failure and hastens recovery from stroke [121, 122, 123, 150]. Therefore, diagnosing and treating such patients could have a very substantial beneficial medical and public health impact [126].

Sleep apnea poses a significant challenge to anesthesiologists in surgical populations. It has been shown that sleep apnea is more prevalent in surgical patients yet, most have not been diagnosed, which put them at higher incidence of complications and death. This includes difficult intubation, postoperative complications, increased intensive care unit admissions, and greater duration of hospital stay [151, 152, 153]. Therefore, identification of sleep apnea patients in the first place would be of great value for anesthesiologists towards the prevention of peri-operative complications.

Despite the high prevalence of sleep apnea, which affects approximately 7% of adults [6], the majority of patients (87%) remain undiagnosed [17], corresponding to approximately 15,000,000 patients in Canada and the US alone. This is partly attributable to the lack of accessibility to expensive overnight monitoring in a sleep laboratory that is currently required for diagnosis [126]. Presently, the “gold standard” for diagnosing sleep apnea is PSG. PSG requires patients to be monitored overnight in a sleep laboratory using multi-channel recordings while they are connected to sophisticated equipment with a technician in attendance [96]. Because of the expense and specialized training required to perform such testing, access to PSG is very limited in most jurisdictions.

As a result, there is a need for alternative methods to diagnose sleep apnea at home, in a more accessible and cost-effective manner. The challenge in developing a portable monitoring device is finding the right balance between simplicity on the one hand and accuracy on the other. Portable sleep apnea diagnostic devices have been divided into 4 categories on the basis of the number of transducers attached—Type IV devices are those
with 1 or 2 channels increasing up to Type I, which is full in-lab PSG [30] as illustrated in Figure 1.8. The currently available portable single channel devices suffer from low accuracy while multi-channel devices suffer from high failure rates due to data loss in unattended home settings. As more channels are added to improve accuracy, failure rates increase to as high as 33% [34]. Failure of the study in home settings is due to the patient’s difficulty connecting the electrodes to the body and the inadvertent detachment of these electrodes during sleep. On the other hand, devices that employ fewer channels have been found to be less accurate than devices with more monitoring channels [154].

Extensive reviews have concluded that there is no convincing evidence that any of the available portable monitors could be used in unattended settings and still provide reliable signals, which reduces their validity for use within the general population [127, 32].

In recent years, analysis of breath sounds recorded by a microphone has been an emerging tool for the diagnosis of sleep apnea. Breath sounds are rich in information about breathing patterns, differentiating habitual snoring from snoring in sleep apnea [61], the site from which snoring arises [155], and patency of the upper airway [100]. I have shown that respiratory airflow has a characteristic acoustic signal signature [84]. I have shown that acoustic analysis of breath sounds can accurately identify respiratory events, apneas and hypopneas, as compared with an overnight simultaneous PSG. My next step towards a fully portable system for the diagnosis of sleep apnea was to develop a portable personal unit that patients can use at home for data collection of breath sounds during sleep in unattended settings. Achieving this objective required the design of a unit that is easy to use and apply by a lay person, yet sophisticated enough to obtain un-interrupted capture of high quality breath sound signals required for the bioacoustic analysis and identification of apneas and hypopneas. The purpose of this paper, therefore, is to describe the technical aspects of an innovative single channel portable device and system for inexpensive and reliable diagnosis of sleep apnea in the home. The initial evaluation of the system’s functionality, usability, and diagnostic capability in the home
environment will also be presented.

8.3 Materials and Methods

8.3.1 Development of the Diagnostic System

At the Toronto Rehabilitation Institute, we have developed a system for respiratory monitoring and diagnosis of sleep disorders where data capture is done at home and data processing is centralized. The system consists of 2 components. First, a self-contained, breath sound capturing module, with removable data storage media, embedded in a custom made face frame with no external wires to an external unit or power supply—hereafter referred to as the ‘acoustic device’. This makes it self-administrable and easy to use at home. Upon completion of a recording session the memory card is returned to a diagnostic office for data extraction and analysis. Second, acoustic signal processing algorithms for data analysis and diagnosis hosted on a server. The 2 components are further explained below.

Self-Contained Acoustic Device With an Embedded Data-Capturing Module

The acoustic device consists of a data-capturing module embedded in an open lightweight face frame to hold the microphone in the optimum location. The block diagram of the data-capturing module is shown in Figure 8.1. The module includes an ATXmega128A3 microcontroller (MCU), a TS472 differential microphone preamplifier, a precision voltage reference source, a MCP1640D step-up DC-DC converter with bypass option, a microSD memory card. The ATXmega MCU was selected for its high performance 12-bit ADC, featuring a differential mode to utilize the advantages of the TS472, 4-channel direct memory access (DMA) controller, and the ability to operate down to 1.6 V that makes it particularly suitable for battery powered applications.

On power-on/reset the MCU initializes peripherals, puts the TS472 in a standby
Figure 8.1: Block diagram of the embedded module for sleep apnea diagnosis. ADC: analog to digital converter; DC: direct current; DMA: direct access memory; RAM: random access memory; GPIO: general purpose input output.

mode, activates the bypass feature of the MCP1640D and enters power saving mode. While the converter is in bypass mode, the MCU is essentially powered directly from batteries as the voltage range is sufficient for operation in sleep mode thus minimizing overall power consumption of the device. As the device is switched on, the MCU enables the DC-DC converter, initializes the microSD memory card, configures the ADC and DMA channels and starts continuous sampling, at sampling frequency of 16,000 Hz, with the DMA controller configured in a double buffer mode. In this mode one of the memory buffers is directly loaded with the samples from ADC, while the previously stored buffer is being recorded on a memory card. The device is powered by 2 alkaline AAA batteries that provide 8 hours of continuous data sampling sufficient for a typical overnight recording.

A unidirectional microphone with noise cancelling properties was selected in order to minimize capture of external noise. Noise cancelling properties of the microphone are achieved by a highly directional dual-port design. The microphone in the mask is specifically oriented towards the mouth and nostrils to capture desired sounds and airflow signals and to filter out ambient noise. This feature helped to achieve very low baseline noise levels and thus a good contrast between breath sounds and no activity status. This in turn enhanced capture of breathing interruptions. The microphone is placed in the
centre of a specially designed funnel-shaped directional element, as illustrated in Figure 8.2, that helps to capture not only audible breath sounds but also airflow from both the mouth and nostrils.

Light emitting diodes (LEDs) provide visual feedback to indicate the status of the device. The device does not have any external wires, since the transducer and electronics are attached to the same frame, which simplifies the application of the device by untrained persons. Users are required to power the device by pressing the power button, wait for the LED feedback signal, and wear the device by attaching a pair of head straps. Patients have the option of terminating data collection upon completion of the session or otherwise leave it until it turns off automatically.

Central Data Analysis

Upon completion of each session, data on the microSD card were uploaded to a server. Apneas and hypopneas were detected from breath sounds using algorithms developed and validated in our laboratory earlier, which have shown excellent agreement (R=95%) with PSG [99, 156]. Briefly, the amplitude modulation envelopes of breath sound waveforms were adaptively normalized to a uniform baseline, which was then scanned for valleys below a certain threshold. Each valley is then examined against a set of rules including its depth, width, and patterns of the falling and rising edges, to be classified as an apnea or hypopnea, or to be discarded. The frequency of apneas and hypopneas per hour of recording time, or the AHI, was determined for each overnight session.

8.3.2 System Evaluation in the Home Environment

The acoustic device was evaluated on 49 subjects. Each subject was given brief instructions on operating the device and was asked to wear it during his/her usual sleep time in his home.
Figure 8.2: Hardware prototype of the acoustic device showing the data capturing module embedded in an open face frame. The microphone is held in a position that allows maximum capture of sound and nasal and oral airflow.

Evaluation of Functionality and Signal Quality in the Home Environment

Signal quality is a reflection of the functionality of the embedded module. To assess the quality of the breath sound signals captured in the home environment, respiratory signal to background noise ratio ($SNR_{\text{breath}}$) was calculated. For this purpose, data were classified manually into breathing and non-breathing components, similar to the method implemented by Duckitt et al. [101]. Here, 5-minute segments (L) were extracted from the first, middle, and last parts of randomly selected 10 home overnight recordings. A total of 150 minutes were extracted. An experienced operator listened to each segment and
manually annotated each sound unit into one of 5 classes: inspiration, expiration, snoring, not-audible, other-noise. The first and last 25% of each sound unit were discarded in order to avoid boundary contamination from the adjacent sounds. The first 3 classes (inspiration, expiration, and snoring) originate from the physiological breathing events. These were combined in each L segment and treated as the true breath sound signals. The ‘in-audible’ class included low amplitude un-intelligible portions such as brief pauses between breaths and during apneas (Figure 8.3), which consist mainly of baseline noise injected into the module in the absence of a true acoustic source, and thus were treated as the background noise. The ‘other-noise’ class included incidental non-respiratory sounds, such as noise from bedsheets, somnolent speech, and other external sounds, and was discarded because it does not belong to the classes required in the analysis. To calculate the $SNR_{breath}$, first the normalized energy of breath sounds in L segments was calculated as:

$$E_{breath} = \frac{1}{n} \sum_{i=1}^{n} S_{breath}^2(i)$$  (8.1)

where $breath \in$ classes: {inspiration, expiration, snoring}, $S(i)$ are data samples, and $n$ is the number of all data samples of breath sounds in L. The normalized energy of background noise ($E_{noise}$) was calculated similarly. The $SNR_{breath}$ was calculated in decibels in each L, in which $E_{noise}$ was not zero, as:

$$SNR_{breath} = 10 \log_{10} \left( \frac{E_{breath}}{E_{noise}} \right)$$  (8.2)

The system’s overall $SNR_{breath}$ was then calculated as the average SNR of all included L segments.

**Evaluation of Usability in the Home Environment**

Participants were given brief instructions on using the device and were asked to take the device home to use it independently. Subsequent to the overnight sessions, each
participant answered 5 questions (Table 8.1) about his/her experience, including the ease-of-use and comfort of the device. Answers were given qualitatively as one of four options: very poor, poor, good, or excellent, each of which was quantified to a numerical value 1 to 4, respectively. In the case of inability to operate the device or sleep with it on, a rating of 0 was given. The median of all ratings was calculated for each question.

**Evaluation of Diagnostic Capability and Reproducibility in the Home Environment**

One important limitation of PSG that needs to be addressed is the difficulty to perform multiple sessions because of the high cost, limited accessibility, and patient’s discomfort. On the other hand, it has been shown that approximately one third of patients have remarkably variable AHI between 2 nights, which might result in improper diagnosis if based on the results of one night only. Therefore, a potential application of a portable sleep apnea diagnostic device is performing multiple night recordings and inter-night comparisons. To examine the feasibility and reproducibility of multiple-night based diagnosis, a subset of eleven volunteers, who did not undergo a sleep study for at least 1 year, agreed to use the acoustic device for 2 nights. The AHI for each night was calculated using acoustic analysis of breath sounds (as described in Section 8.3.1.)

**8.3.3 Agreement with PSG**

To evaluate the accuracy of AHI derived from data captured by the embedded module, the portable device was worn by 32 consecutive patients referred to the Toronto Rehabilitation Institute Sleep Laboratory for suspicion of sleep apnea. None of the data from these 32 subjects data were used in developing the algorithm described in Section 8.3.1. Subjects underwent overnight PSG using standard techniques and scoring criteria [133, 102]. Thoracoabdominal movements and tidal volume were measured by RIP [27]. Airflow was measured by nasal pressure cannulae [27] and arterial oxyhemoglobin
saturation ($SaO_2$) by oximetry. A sleep laboratory technician scored PSG manually to find apneas and hypopneas. Apneas were scored as a drop in sum of thoracoabdominal movement by $\geq 90\%$ lasting $\geq 10$ seconds [134]. Hypopneas were defined as a 50% to 90% reduction in thoracoabdominal sum lasting $\geq 10$ seconds as described previously [135].

Breath sounds were recorded simultaneously with PSG. At the end of the study, data were transferred from the microSD memory card to a computer for acoustic analysis using the aforementioned procedure described in Section 8.3.1. The AHI derived from PSG (AHP-p) and that from acoustic analysis (AHI-a) were quantified as the number of apneas and hypopneas per hour of recording. AHI-a was compared against AHP-p using Pearson correlation coefficient and diagnostic parameters were calculated.

The study protocol was approved by the Research Ethics Board of the Toronto Rehabilitation Institute and subjects provided written consent prior to enrollment.

8.4 Results

8.4.1 Functionality and Signal Quality in the Home Environment

The overall breath sounds’ $SNR_{breath}$ in home sessions was 31.7 dB. Figure 8.3 displays representative samples of breath sounds captured by the embedded module and the low amplitude background noise in the inter-breath segment.

8.4.2 Usability in the Home Environment

Feedback from the 49 subjects was used to evaluated ease-of-use and usability in the home environment. All the subjects (100%) indicated being able to fully administer the acoustic device at home in the absence of a trained person. Upon examination of the
Figure 8.3: A: single breathing cycle showing inspiration (time ≈ 0.5 to 2 seconds) and expiration (time ≈ 2-3.5 seconds) separated by the vertical dashed line. The inter-breath interval is marked to show a sample of ambient background noise and its relative amplitude. B: magnified waveform of a part of the inspiration revealing a snoring episode characterized by quasi-periodicity, resulting from tissue vibration. C: magnified waveform of a part of expiration showing a turbulent non-periodic pattern.

Data, 3 subjects (6%) were found to have faulty data resulting from improper powering of the device, which resulted in failure of data collection. Therefore, the overall success rate in using the device was 94%. Answers to the usability questionnaire are displayed in Table 8.1. The overall subjective rating of the ease-of-use related questions was ‘excellent’ and for the comfort related questions was ‘good’.
Table 8.1: Usability questionnaire and the median rating of each question for the 49 participants

<table>
<thead>
<tr>
<th>Question</th>
<th>Mean Rating</th>
<th>Qualitative Equivalent</th>
</tr>
</thead>
<tbody>
<tr>
<td>How easy could you operate the device?</td>
<td>4</td>
<td>Excellent</td>
</tr>
<tr>
<td>How easy you wore the device?</td>
<td>4</td>
<td>Excellent</td>
</tr>
<tr>
<td>How comfortable the mask was on your face after wearing it?</td>
<td>3</td>
<td>Good</td>
</tr>
<tr>
<td>How quickly you fell asleep comparing with other nights?</td>
<td>3</td>
<td>Good</td>
</tr>
<tr>
<td>How comfortable was your sleep comparing with usual nights?</td>
<td>3</td>
<td>Good</td>
</tr>
</tbody>
</table>

8.4.3 Diagnostic Capability and Reproducibility in the Home Environment

AHI resulting from acoustic analysis of data of the 11 subjects who used the device for 2 nights are displayed in Figure 8.4. Their AHI ranged from \(\sim 0\) to 80 apneas and hypopneas/hour. Nine out of 11 patients (82%) showed equivalent results on both nights, i.e., difference in AHI of less than 10 between the 2 nights.

8.4.4 Agreement with PSG

I recruited 32 consecutive subjects who underwent full PSG while wearing the acoustic device simultaneously. AHI-a derived from data captured by the acoustic device showed 96% correlation (\(R^2=0.93\)) with AHI-p obtained from manual scoring of PSG data as illustrated in Figure 8.5. Using a diagnostic cut-off of AHI \(\geq 10\), sensitivity was 100%, specificity was 85%, NPV was 100%, PPV was 80%, and the overall accuracy was 91%. Using a diagnostic cut-off of AHI \(\geq 15\), sensitivity was 89%, specificity was 96%, NPV
Figure 8.4: Comparison between the AHI obtained from the first and second nights in the home (sorted by severity.) Panel (A) shows subjects whose AHI varied by $<10$ on the 2 nights, while panel (B) shows subjects whose AHI varied $\geq 10$ on the 2 nights.

was 96%, PPV was 89%, and the overall accuracy was 94%.
8.5 Discussion

This work presents the development and initial evaluation of an integrated system for accessible diagnosis of sleep apnea at home. The core of this system is a self-contained device with a small embedded module to capture breath sounds. The acoustic device was well accepted by users as shown by the very high success rates and complete subject compliance. It also demonstrated excellent performance as revealed by the high SNR and reproducibility of AHI in the home environment supported with remarkable agreement with PSG.

The device captured high quality breath sound signals including snoring sounds (characterized by a semi-periodic waveform) and other inspiratory and expiratory sounds (characterized by turbulent non-periodic waveforms.) The use of a unidirectional microphone, the low-noise electronic design, and the funneling structure in front of subjects’ mouths were all deployed to enhance capture of breath sounds and improve signal quality, thus achieving good SNR. SNR of breath sounds to ambient background noise was
31.7 dB. Other researchers in the field recommended a minimum SNR of 10.5 [157] and 8 dB [158] to effectively distinguish breath sounds and snoring from ambient noise. Thus, the SNR in uncontrolled home environments with this design was clearly sufficient to perform the desired acoustic analysis. Beside the low noise properties of the embedded module, the low power design is also an essential factor for obtaining a continuous 8-hour overnight recording required in this monitoring application.

An essential, but quite overlooked aspect of a home sleep apnea monitor is usability. The Portable Monitoring Task Force of the American Academy of Sleep Medicine, for example, recommended developing easy to use portable sleep apnea monitors that would reduce time required to attach sensors, and simplify data collection and transfer [137]. The acoustic device was designed to be single-channel and self-contained without external wires in order to make it easy to use by an untrained person in unattended settings, which is an important feature of a portable sleep apnea monitor. In the current study, 100% of the subjects could administer the device independently without the support of a trained assistant at home, with 94% success rate as revealed by objective examination of the data, which demonstrates excellent usability of the device. It is noteworthy, however, that comfort ratings were lower, i.e., ‘good’ for comfort versus ‘excellent’ for ease-of use (Table 8.1), which was attributed mostly to the experience of wearing the acoustic device while sleeping for the first time. These findings, in addition to other qualitative feedback, can be used to improve the ergonomics of future iterations of the device, such as reducing the size of the embedded module.

The AHIs obtained from home sessions in 11 subjects were used to assess performance and reproducibility in the home environment, all of whom did not have a recent PSG. PSG studies have previously shown that the majority of patients have reproducible AHI as measured by 2-night difference in AHI <10 [159, 160]. Nevertheless, there may very well be true inter-night variability, denoted by difference in AHI exceeding 10, in approximately 18% [160] to 32% [159] of patients between the 2 nights. In the current study, 9
out of 11 subjects demonstrated the equivalence of AHI on both nights (2-night difference in AHI under 10). In 2 out of 11 (18.2%) subjects, greater variability in the AHI of $\geq 10$ was observed, which lies well within, indeed at the lower end of the previously reported proportions. This is a sound indicator of the reproducibility of the results in the home settings and that captured signals reflected the underlying respiratory condition in the subjects who used the device while sleeping at home.

To ensure the validity of the AHI results obtained in the home environment, the system’s performance was assessed against PSG, the current gold standard for diagnosing sleep apnea. The system showed excellent agreement (96%) with PSG in terms of AHI concordance in 32 subjects who wore the acoustic device simultaneously with PSG. Diagnostic accuracy was 91% and 94% for an AHI cut-off $\geq 10$ and 15, respectively. This is inline with our previously reported agreement with PSG (95%) in our work describing the development the apnea and hypopnea detection algorithm (Chapter 7 and [156]). In the current study, however, the AHI was obtained from data collected using the newly designed acoustic device and embedded module that has no external wires. This step of validation of the acoustic device in a controlled environment demonstrates its accuracy and supports the validity of the AHI results obtained in the home studies.

Accordingly, I believe that this system has the potential to make diagnosis of sleep apnea more accessible, particularly in remote areas where sleep laboratories are not located. There is also the possibility to make multiple recordings in subjects’ homes without the need for multiple visits to a sleep laboratory, so that, for example, severity of sleep apnea could be monitored over time in response to interventions such as weight loss. This device makes it feasible to perform multiple studies and subsequently to explore the causes of variability in AHI and lead to new treatment options that recreate the circumstances that are associated with the lowest AHI scores. In the proposed model, data collection is performed independently from the data analysis. This setup gives the advantage of developing and improving the diagnostic algorithms centrally without the
need to re-design the data acquisition module and disrupt the peripheral data collection. This setup helps to maintain a simple structure of the portable unit which will facilitate its use by untrained individuals and potentially reduce cost per test.

Future improvements of the embedded module might include migration to one of the rapidly growing ARM Cortex-M MCU families offering more functional and cost effective platform for embedded applications with their 16 bit onboard ADC, SD host controller, and high RAM-to-flash ratio. Digital Microelectromechanical systems (MEMS) microphones with their small form factor, high noise immunity, and low power consumption also offer a viable improvement choice.

8.6 Conclusion

I have described a novel sleep apnea diagnostic system using a portable battery operated acoustic data acquisition device as well as its functionality and applicability to diagnose sleep apnea in the home environment. Its excellent usability by lay people makes it suitable for use at home in unattended settings. The evaluation results point to the feasibility of making multiple home recordings from the same subject with capture of high quality data. These data provide a strong rationale to perform home studies in a larger population and compare them with PSG. This new portable home sleep apnea monitor has the potential to improve accessibility and reduce the costs of diagnosing sleep apnea.
Chapter 9

Concluding Remarks

The present PhD thesis project has shown that acoustic analysis of breath sounds is a powerful tool for obtaining information about respiratory patterns including UA narrowing and sleep apnea. I have shown that an automated algorithm for identification of apneas and hypopneas from breath sounds captured during sleep yielded an AHI that is accurate and highly correlated with the current gold standard, PSG. I have also shown that this technology can be packaged in a compact device that is highly portable and can be used independently and reliably in the home environment.

Obtaining the proper biophysiological signal is the first step in ensuring accurate analysis of the condition of interest. In this project, breath sounds were captured directly from the most distal end of the respiratory tract. This setup not only helped in improving the SNR ratio, but also in obtaining profound insights into breathing patterns.

LPC analysis, for example, is a classical technique for obtaining resonant frequencies in speech processing. Narrowing of the UA similar to that which takes place in OSA has an impact on airflow turbulence that resulted in differences in the LPC spectrum. Transducer location helped in utilizing the physical sensitivity to expired airflow and thus provide a tool for monitoring the short-term breathing activities and also the longer term variations in ventilatory depth such as apneas and hypopneas. Therefore, both temporal
and spectral features of breath sounds were shown to be modulated by the respiratory
patterns and UA status and were used effectively to monitor breathing patterns.

The challenge that this project has tackled was creating a portable device for diag-
nosing sleep apnea at home that is more accessible and less expensive than PSG, yet
accurate and reliable. This is a major issue that physicians and engineers in this field
have been facing as expressed repeatedly in reviews of sleep apnea monitors. This issue
has been addressed throughout the chapters of this work. Eventually, the single chan-
el input device presented herein satisfied the simple setup required for ease-of-use in
the home. At the same time, the multi-faceted data analysis provided comprehensive
information about breathing patterns and the UA status.

An essential question during the design of this single channel device is why breath
sounds were chosen over other physiological signals? Breath sounds were chosen because
they are the single signal with the greatest range of information among all other phys-
iological signals used for monitoring breathing and therefore the optimum choice for a
single channel device. Although other single channel devices, such as pulse oximetry,
can be useful as a screening tool, it fails to detect all respiratory events, because oxygen
desaturation does not necessarily occur with all apneas and hypopneas. This manifests
in a lower mean AHI according to AASM criteria as compared to TV50 criteria, as de-
scribed in Chapter 7. Additionally, pulse oximetry does not take into account the UA
status; therefore, it does not have the potential to distinguish between events that occur
due to UA obstruction, as in OSA, from those due to lack of respiratory drive from the
brain and which are not associated with UA narrowing, as in CSA. Since snoring is a
major symptom of sleep apnea, knowledge of the presence of snoring and its severity is
an advantage during clinical evaluation and this can be obtained from breath sounds,
but not from other channels. The combination of those features gives this single channel
device an advantage over pulse oximetry for detecting snoring and sleep apnea. Similar
comparisons can be drawn against other single channel setups such as nasal air flow and
On the other hand, multiple channel devices (Level I, II, and III) provide more data streams than a standard single channel device. However, portability and usability of such devices is an issue because of high failure rates in the unattended setting due to inability of patients to attach electrodes and their detachment during sleep. The single channel device proposed in this work has been shown to be easy to use and has yielded very positive user feedback as discussed in Chapter 8. The subjective feedback was substantiated by data indicating reproducibility in unattended settings. I therefore conclude that this single channel device and acoustic analysis provide a satisfactory balance between diagnostic accuracy for sleep apnea and ease of use in unattended settings.

9.1 Limitations

The most obvious limitation of the present work is the lack of information on sleep status. This limitation did not have an impact on the diagnostic outcome based on an AHI cut-off $\geq 10$, as discussed in 7.4.5. However, this limitation might have more significant consequences in certain scenarios, such as in patients with both insomnia and sleep apnea. In such patients, the total sleep time might be much shorter than the recording time. Thus, dividing the number of apneas and hypopneas by the total recording time might result in a significantly lower AHI than would have been calculated if the sleep time was the denominator and thus the severity of sleep apnea might be underestimated or even missed. Although this situation was not encountered during the course of this project, it should be accounted for during clinical implementation. The simplest and most straightforward way to identify that group is via self-reporting of insomnia and the time it takes the patient to fall asleep (sleep latency), using a subjective questionnaire similar to that presented in Chapter 8. Until a suitable technology for monitoring sleep status in this context is developed, patients with self-reported insomnia...
might not be good candidates for using this device and may still require full PSG to obtain a diagnosis.

Another significant limitation of this device is that it cannot detect sleep disorders other than sleep apnea, such as narcolepsy or periodic leg movement that would still require PSG for diagnosis. Since body position is not detected, it is not possible to distinguish supine-related sleep apnea. Additionally, our device cannot be used to titrate CPAP for the treatment of sleep apnea.

### 9.2 Future Directions

In this work, I have shown that some breath sounds are modulated by UA caliber (Chapter 5). In addition to data that show that snoring characteristics are also influenced by UA narrowing, we suggest the possibility that use of breath sound analysis could be used to distinguish respiratory events that are associated with UA narrowing and collapse, i.e., OSA, from those that are not usually associated with UA narrowing and collapse, i.e., CSA \(^1\). Future research will examine these findings in a number of subjects who have OSA and CSA in order to design an algorithm that uses those features to reliably distinguish those two types of the disease. Future research will, therefore, examine the ability of breath sound analysis to distinguish OSA from CSA, that would assist in managing sleep apnea in patients with stroke or heart failure in whom CSA is relatively common. This will open the door for evaluating the performance of this device in specific populations. Indeed, a study including only stroke patients is underway.

I have shown the performance of the proposed device in the laboratory environment and reported some results from recordings in the home. Next, the performance of the device at home will be evaluated against in laboratory attended PSG. For this purpose,

\(^1\)An international patent with this content has been filed: “Breathing Disorders Identification, Characterization and Diagnosis Methods, Devices and Systems”, File number: PCT/CA2011/000555
subjects will undergo both in-laboratory PSG and at home monitoring in a larger sample size. For further scrutiny, reproducibility studies at home and inter-night variability will be also be expanded, beyond what I have already shown. These validation studies will pave the way towards clinical implementation.
Appendices
Appendix A

Model-Based Analysis of Bland-Altmanman Limits of Agreement

A.1 Statistical Model

The statistical model used to describe and analyse the data was a two-factor mixed effects model. Subjects were treated as a random factor, while method of assessment was treated as a mixed factor allowing random effects for the 3 PSG technicians and a fixed effect for the acoustic device.

The statistical model used in the analysis of the data was:

\[
Y_{ij} = \begin{cases} 
\mu + \tau_i + \beta_j + \varepsilon_{ij} & i = 1, 2, 3 \text{ (technicians)} \\
\tau_4 + \beta_j + \varepsilon_{ij} & i = 4 \text{ (acoustic device)} 
\end{cases}
\]

The following assumptions were made:

1. $\tau_1, \tau_2, \tau_3$ are random effects associated with the 3 technicians and are i.i.d. $N(0, \sigma^2_\tau)$

2. $\mu$ is the common mean AHI value for the 3 technicians and is an unknown constant

3. $\tau_4$ is the mean AHI value for the acoustic device, and is an unknown constant
4. \( \beta_j \) is the random subject effect associated with subject \( j, j = 1, 2, \ldots, r \), where \( r \) is the number of subjects. The \( \beta_j \) are i.i.d. \( \mathcal{N}(0, \sigma^2_{\beta}) \)

5. \( \varepsilon_{ij} \) is the random/experimental error associated with the \( i \)-th assessment for the \( j \)-th subject and they are i.i.d. \( \mathcal{N}(0, \sigma^2) \)

1. \( \tau_i, \beta_j, \varepsilon_{kl} \) are mutually statistically independent for \( i=1,2,3; j; k; l \)

The statistical model was used to derive estimates of the interrater reliability between the 3 PSG technicians as well as model-based Bland-Altman limits of agreement.

### A.2 ANOVA Table

Because the method of assessment factor includes both random and fixed components, unlike the usual model which would include one or the other, the ANOVA table is shown.

<table>
<thead>
<tr>
<th>Source of Variation</th>
<th>Degrees of Freedom</th>
<th>Mean Square</th>
<th>Expected Mean Square</th>
</tr>
</thead>
<tbody>
<tr>
<td>Method of Assessment</td>
<td>3</td>
<td>( MS_{assess} )</td>
<td>( \sigma^2 + \frac{3r}{4}\sigma^2_{\tau} + \frac{r}{4}(\tau_4 - \mu)^2 )</td>
</tr>
<tr>
<td>Technicians</td>
<td>(2)</td>
<td>( MS_{technician} )</td>
<td>( \sigma^2 + r\sigma^2_{\tau} )</td>
</tr>
<tr>
<td>Difference Technicians and Device</td>
<td>(1)</td>
<td>( MS_{diff} )</td>
<td>( \sigma^2 + \frac{r}{4}\sigma^2_{\tau} + \frac{3r}{4}(\tau_4 - \mu)^2 )</td>
</tr>
<tr>
<td>Subjects</td>
<td>( r-1 )</td>
<td>( MS_{subject} )</td>
<td>( \sigma^2 + 4\sigma^2_{\beta} )</td>
</tr>
<tr>
<td>Experimental Error</td>
<td>( 3(r-1) )</td>
<td>( MS_E )</td>
<td>( \sigma^2 )</td>
</tr>
<tr>
<td>Total</td>
<td>( 4r-1 )</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
A.3 Interrater Reliability

The interrater reliability is:

$$IRR = \frac{\sigma^2_\beta}{\sigma^2 + \sigma^2_\tau + \sigma^2_\beta}$$

An estimate of the IRR was obtained from the model as:

$$\hat{I RR} = \frac{\frac{1}{r} MS_{subject} - \frac{1}{r} MS_E}{\frac{1}{r} MS_{technician} + \frac{1}{r} MS_{block} + (\frac{3}{4} - \frac{1}{r}) MS_E}$$

An approximate confidence interval for the IRR was obtained using the approach of Fleiss and Shrout (1978). Let $F_{tech} = MS_{technician}/MS_E$, and let

$$q = \frac{3(r - 1)\{4(IRR) F_{tech} + r (1 + 3 IRR) - 4 IRR\}^2}{24 (r - 1) IRR^2 F_{tech}^2 + \{r (1 + 3 IRR) - 4 IRR\}^2}$$

Use $F^*$ and $F_*$ to denote the upper and lower critical values, respectively, of an $F$ distribution with $r - 1$ and $q$ degrees of freedom. Then the lower (L) and upper (U) bounds of the confidence interval for the interrater reliability are:

$$L = \frac{r(MS_{subject} - F^* MS_E)}{r MS_{subject} + 4 F^* MS_{technician} + (3r - 4) F^* MS_E}$$

$$U = \frac{r(MS_{subject} - F_* MS_E)}{r MS_{subject} + 4 F_* MS_{technician} + (3r - 4) F_* MS_E}$$
Bibliography


sleep apnea: Implications for cardiac and vascular disease. *Chest*, 133:793–804,

104, Jan 1996.


[10] H. Schneider, A. Boudewyns, P. L. Smith, C. P. O’Donnell, S. Canisius, A. Stamm-
nitz, L. Allan, and A. R. Schwartz. Modulation of upper airway collapsibility


[14] T. Young, P. E. Peppard, and D. J. Gottlieb. Epidemiology of obstructive sleep
1239, May 2002.


[22] T. Young, J. Blustein, L. Finn, and M. Palta. Sleep-disordered breathing and


[37] H. M. Engleman, S. E. Martin, I. J. Deary, and N. J. Douglas. Effect of continuous


