Available online at www.sciencedirect.com

ScienceDirect

Computer Speech and Language 28 (2014) 434–452

COMPUTER SPEECH AND LANGUAGE

www.elsevier.com/locate/csl

Analysis of voice features related to obstructive sleep apnoea and their application in diagnosis support☆

Ana Montero Benavides[a], Rubén Fernández Pozo[a], Doroteo T. Toledano[b,*], José Luis Blanco Murillo[a], Eduardo López Gonzalo[a], Luis Hernández Gómez[a]

[a] *Signal, Systems and Radiocommunications Department, Universidad Politécnica de Madrid, Spain*
[b] *ATVS Biometric Recognition Group, Universidad Autónoma de Madrid, Spain*

Received 1 March 2012; received in revised form 31 July 2013; accepted 5 August 2013
Available online 22 August 2013

Abstract

Obstructive sleep apnoea (OSA) is a highly prevalent disease affecting an estimated 2–4% of the adult male population that is difficult and very costly to diagnose because symptoms can remain unnoticed for years. The reference diagnostic method, *Polysomnography* (PSG), requires the patient to spend a night at the hospital monitored by specialized equipment. Therefore fast and less costly screening techniques are normally applied for setting priorities to proceed to the polysomnography diagnosis. In this article the use of speech analysis is proposed as an alternative or complement to existing screening methods. A set of voice features that could be related to apnoea are defined, based on previous results from other authors and our own analysis. These features are analyzed first in isolation and then in combination to assess their discriminative power to classify voices as corresponding to apnoea patients and healthy subjects. This analysis is performed in a database containing three repetitions of four carefully designed sentences read by 40 healthy subjects and 42 subjects suffering from severe apnoea. As a result of the analysis, a *linear discriminant model* (LDA) was defined including a subset of eight features (signal-to-disperiodicity ratio, a nasality measure, harmonic-to-noise ratio, jitter, difference between third and second formants on a specific vowel, duration of two of the sentences and the percentage of silence in one of the sentences). This model was tested on a separate database containing 20 healthy and 20 apnoea subjects yielding a sensitivity of 85% and a specificity of 75%, with a F1-measure of 81%. These results indicate that the proposed method, only requiring a few minutes to record and analyze the patient's voice during the visit to the specialist, could help in the development of a non-intrusive, fast and convenient PSG-complementary screening technique for OSA.

Keywords: OSA screening tools; Apnoea discrimination; Voice features; Feature selection; Voice pathology

Abbreviations: OSA, obstructive sleep apnoea; PSG, polysomnography; AHI, apnoea–hypopnoea index; BMI, body mass index; CER, classification error rate; EER, equal error rate; LDA, linear discriminant analysis; MLR, multiple linear regression.

☆ This paper has been recommended for acceptance by 'Dr. Björn Schuller'.

* Corresponding author at: Av. Francisco Tomás y Valiente 11, 28049 Madrid, Spain. Tel.: +34 419 2217; fax: +34 419 2207.
E-mail addresses: ana.montero@gaps.ssr.upm.es (A. Montero Benavides), ruben@gaps.ssr.upm.es (R. Fernández Pozo), doroteo.torre@uam.es (D.T. Toledano), jlblanco@gaps.ssr.upm.es (J.L. Blanco Murillo), eduardo@gaps.ssr.upm.es (E. López Gonzalo), luisalfonso.hernandez@upm.es (L. Hernández Gómez).

0885-2308/$ – see front matter © 2013 Elsevier Ltd. All rights reserved.
http://dx.doi.org/10.1016/j.csl.2013.08.002

1. Introduction

Obstructive sleep apnoea (OSA) is a highly prevalent disease (Puertas et al., 2005), affecting an estimated 2–4% of the male population between the ages of 30 and 60. It is characterized by recurring episodes of sleep-related collapse of the upper airway at the level of the pharynx and it is usually associated with loud snoring and increased daytime sleepiness. Besides OSA, there are two other forms of sleep disease related to apnoeas: *central sleep apnoea* (CSA) and *combined* or *mixed sleep apnoea* that only constitute the 0.4% and 15% of cases respectively (Morgenthaler et al., 2006); therefore their relevance can be considered minor compared to OSA, and for this reason we will focus on the study of OSA. A common criterion for diagnosing OSA is having an *apnoea–hypopnoea index* (AHI) over 15. AHI represents the average number of apnoeas and hypopnoeas per hour of sleep. OSA is a serious threat to an individual's health if not treated. The condition is a risk factor for hypertension and cardiovascular diseases, it is usually related to traffic accidents caused by somnolent drivers, and it can lead to a poor quality of life and impaired work performance. At present, the most effective and widespread treatment for OSA is nasal CPAP (*continuous positive airway pressure*) which prevents apnoea episodes by providing a pneumatic splint to the airway.

OSA can be diagnosed on the basis of a physical examination and a medical history that includes questions about habits or the degree of snoring and the daytime sleepiness, although both screening characteristics by themselves are not always accurate discriminative features for OSA (Teculescu, 1998), and a full overnight sleep study is needed to confirm the disorder. The procedure is known as conventional *polysomnography* (PSG), which involves the recording of neuroelectrophysiological and cardiorespiratory variables. Nevertheless, this diagnostic procedure is expensive and time-consuming, and patients usually have to endure a waiting list of several years before the test is done, since recently the demand for diagnostic studies for OSA has recently increased (Puertas et al., 2005). There is, therefore, a strong need for alternatives methods to PSG of screening apnoea patients in order to detect patients with high risk of OSA and reduce these considerable delays. In the published literature, we found numerous efforts to devise PSG-complementary clinical methods of predicting OSA (Ramachandran and Josephs, 2009). These methods are broadly classified as questionnaires about sleepiness (Ahmadi et al., 2008) and clinical prediction models (Gurubhagavatula et al., 2004; Friedman et al., 2010) using anthropometric characteristics (i.e. BMI or morphology) and epidemiological and medical history parameters (i.e. snoring).

This need of searching for novel OSA screening tools gives a good reason for this contribution, which investigates the acoustical characteristics of the voice in patients with OSA for the purpose of learning whether OSA may be detected using speech analysis. The acoustic properties of voice from speakers suffering OSA are not well understood as not much research has been carried out in this area. However, some studies have suggested that certain abnormalities in voice *articulation*, *phonation* and *resonance* may be connected to the condition (Fox and Monoson, 1989). In order to have a controlled experimental framework to study apnoea voice characterization we have collected a speech database (Fernández et al., 2008) designed following linguistic and phonetic criteria derived from previous research in the field. Our work is focused on continuous speech rather than on sustained vowels, the latter being the standard approach in pathological voice analysis. This is justified because we are interested in the acoustic analysis of the characteristics of the speech signal in specific linguistic and phonetic contexts, and also in the analysis of other features not used previously in the characterization of apnoea voices such as the dynamics of speech, and in particular the patterns of pauses, which we hypothesized as valuable discriminative features. Together with these speech-dynamics and pause features, OSA-related voice characteristics associated to articulation, phonation and nasalization were evaluated over apnoea and non-apnoea voices to have a contrastive study on the acoustic discrimination attainable in our database using either individual features or a combination of some or all of the proposed features. In addition, and due to the crucial dependency of OSA disease with BMI (body mass index) and age factors, we finished with a study about relations between OSA-related voice features and these two important factors.

The rest of this article is organized as follows: Section 2 presents the main physiological characteristics of OSA patients and acoustic characteristics of their voices, as described in the previous literature. Section 3 describes the experimental setup, including the description, design and capture of the corpora used in this study, the definition and selection of features related to apnoea and the details of the experimental setup. Section 4 presents detailed results