Formant Frequencies and Bandwidths in Relation to Clinical Variables in an Obstructive Sleep Apnea Population

*Ana Montero Benavides, *José Luis Blanco Murillo, *Rubén Fernández Pozo, *Fernando Espinoza Cuadros, †Doroteo Torre Toledano, ‡José D. Alcázar-Ramírez, and *Luis A. Hernández Gómez, *†Madrid, ‡Málaga, Spain

Summary: Objectives. We investigated whether differences in formants and their bandwidths, previously reported comparing small sample population of healthy individuals and patients with obstructive sleep apnea (OSA), are detected on a larger population representative of a clinical practice scenario. We examine possible indirect or mediated effects of clinical variables, which may shed some light on the connection between speech and OSA.

Study Design. In a retrospective study, 241 male subjects suspected to suffer from OSA were examined. The apnea-hypopnea index (AHI) was obtained for every subject using overnight polysomnography. Furthermore, the clinical variables usually reported as predictors of OSA, body mass index (BMI), cervical perimeter, height, weight, and age, were collected. Voice samples of sustained phonations of the vowels /a/, /e/, /i/, /o/, and /u/ were recorded.

Methods. Formant frequencies F1, F2, and F3 and bandwidths BW1, BW2, and BW3 of the sustained vowels were determined using spectrographic analysis. Correlations among AHI, clinical parameters, and formants and bandwidths were determined.

Results. Correlations between AHI and clinical variables were stronger than those between AHI and voice features. AHI only correlates poorly with BW2 of /a/ and BW3 of /e/. A number of further weak but significant correlations have been detected between voice and clinical variables. Most of them were for height and age, with two higher values for age and F2 of /o/ and F2 of /u/. Only few very weak correlations were detected between voice and BMI, weight and cervical perimeter, wich are the clinical variables more correlated with AHI.

Conclusions. No significant correlations were detected between AHI and formant frequencies and bandwidths. Correlations between voice and other clinical factors characterizing OSA are weak but highlight the importance of considering indirect or mediated effects of such clinical variables in any research on speech and OSA.

Key Words: Formant–Bandwidth–Obstructive sleep apnea–Clinical variables.

INTRODUCTION

Sleep disorders are becoming a focus of study and growing interest as they are known to cause daytime sleepiness and impaired work and are said to increase the risk of traffic accidents.[1] The prevalence of sleep disordered breathing varies across different populations, but on average, it has been estimated to affect 24% and 9% of middle-aged men and women, respectively. Among sleep disorders, sleep apnea is the most frequent one[2] and obstructive sleep apnea (OSA) is its most common type. OSA affects 85% of sleep apnea patients[3] and is characterized by recurring episodes of breathing pauses, greater than 10 seconds at a time, during sleep, caused by a blockage of the upper airway (UA) at the level of the pharynx. In severe cases, more than 30 apnea episodes may occur in a single hour. Most patients suffering from this disorder are not aware of their condition, and those who have visited a physician endure a waiting list that may last for more than 1 year to receive a final diagnosis.[1] At present, the most effective and widespread treatment for OSA is nasal continuous positive airway pressure (CPAP), which prevents apnea episodes by providing a pneumatic splint to the airway and so avoids UA collapse episodes but does not provide a cure.

OSA patients are first evaluated on the basis of a physical examination and questionnaires. The latter include questions on their customs, snoring history, and sleep habits. Clinical parameters in the identification of patients with suspected OSA usually include age, sex, body mass index (BMI, the ratio between weight and squared height), cervical perimeter, waist circumference, nasal examination, pharyngeal examination, snoring habits, and so forth. These cues on their own are not accurate OSA predictors[4–8]; a full overnight sleep study is needed to confirm diagnosis.

The gold standard for sleep apnea diagnosis is the polysomnography (PSG) test.[9] This test requires an overnight stay of the patient at the sleep unit within a hospital to monitor, breathing patterns, heart rhythm, and limb movements. As a result of the test, the apnea-hypopnea index (AHI) is computed as the average number of apnea and hypopnea episodes (partial and total breath cessation episodes, respectively) per hour of sleep. Because of its high reliability, most clinicians rely on AHI to describe the severity of patients' condition: low AHI indicates a healthy subject or mild OSA patient, whereas AHI >30 is associated with severe OSA. Waiting lists for PSG may exceed 1 year.[1] Therefore, faster and less costly screening and stratification techniques may contribute to set patients' priority to access PSG test and to optimize the usage of the available resources.

Accepted for publication January 22, 2015.
From the *Signal, Systems and Radiocommunications Department, Universidad Politécnica de Madrid, Madrid, Spain; †ATVS Biometric Recognition Group, Universidad Autónoma de Madrid, Madrid, Spain; and the ‡Hospital Quirón, Málaga, Spain.
Address correspondence and reprint requests to Ana Montero Benavides, Signal Processing Applications Group, Universidad Politécnica de Madrid, Avenida Complutense, 30, 28040 Madrid, Spain. E-mail: ana.montero@gaps.ssr.upm.es
Journal of Voice, Vol. 30, No. 1, pp. 21-29
0892-1997/$36.00

http://dx.doi.org/10.1016/j.jvoice.2015.01.006

Among current proposals for fast and cost-effective OSA detection, speech-based methods are very attractive as they are nonintrusive and may provide useful information, only requiring a few minutes to record and process a small sample of patient's voice, which could even be recorded just before the interview with a physician.

The rationality of using speech analysis in OSA detection can be found on studies by Davidson et al,[10] in which the acquisition of speech is connected to the appearance of OSA from an anatomic basis. Although the connection between anatomic abnormalities and the presence and severity of sleep apnea are difficult to disentangle,[11–13] OSA involves several physical alterations, such as craniofacial abnormalities, dental occlusion, longer distance between the hyoid bone and the mandibular plane, relaxed soft tissues, large tongue base, and so forth, that generally cause longer and more collapsible UA. Consequently, abnormal or particular speech features in OSA speakers may be expected from an altered structure or function of their UA.

Early approaches to speech-based OSA detection can be found in the studies by Fox et al[14] and Monoson and Fox.[15] In the study by Fox et al,[14] authors used perceptive speech descriptors (related to articulation, phonation, and resonance) to correctly identify 96.3% of normal (healthy) subjects, although only 63.0% of sleep apnea speakers were detected. The use of acoustic analysis of speech for OSA detection was first presented in the studies by Fiz et al[16] and Robb et al.[17] Fiz et al examined the harmonic structure of vowels spectra, finding a narrower frequency range for OSA speakers, which may point at differences in laryngeal behavior between OSA and non-OSA speakers.[16] Later on, Robb et al[17] presented an acoustic analysis of vocal tract formant frequencies and bandwidths, thus focusing on the supra-laryngeal level where, according to the pathogenesis of the disorder, OSA-related alterations should have larger impact. Results in the study by Robb et al,[17] extracted from the analysis of sustained phonations of vowel sounds, showed significant differences in the first and second formants (F1 and F2) and the first and second resonant frequencies (BW1 and BW2).

These early contributions have driven recent proposals for using acoustic analysis of speech in OSA detection.[18–22] Different classification techniques have been proposed for Hebrew[18] and Spanish[19] languages. Results have been reported for different types of speech (ie, sustained and/or continuous speech),[18,20,22] different speech features,[18,19,21] and modeling different linguistic units[20] as well.

Despite the remarkable results obtained from these recent proposals, there are, in our view, two major issues that must be addressed to guide future progress in this field. First, reported results have been based on small databases recorded under nonrealistic clinical practice scenarios. Second, most of the proposed OSA detection techniques rely on applying feature selection techniques over a large number of acoustic features (sometimes on the order of hundreds[18]). These approaches, although computationally efficient, may hinder the study of the connections between basic speech features and OSA that was the main focus in foundational studies.[16,17]

In this work, we seek to contribute in addressing the aforementioned issues by investigating whether the significant lower formant frequencies and wider bandwidths found on a small group of 18 male OSA patients compared with 10 non-OSA subjects in the study by Robb et al[17] are observed over a larger population that is representative of a clinical practice scenario. As pointed out in the study by Robb et al,[17] the review of these initial research results on larger populations is essential to fully realize the applicability of acoustic analysis in the assessment of OSA.

According to the study by Robb et al,[17] two main reasons may support the hypothesis of a relationship between formants and OSA. First, cephalometric research has indicated that the distance from the hyoid bone and the mandibular plane is significantly longer in patients with OSA compared with non-OSA individuals; this suggests that longer vocal tracts in OSA patients can lead to lower formant frequencies compared with non-OSA subjects. A second reason relates to an increase in both velar and pharyngeal compliance commonly found among OSA patients. This increased collapsibility of the pharyngeal airway has been related to an excessive fat deposition[23,24] that would increase the sound damping inside the vocal tract.[17] Consequently, formant bandwidths characterizing OSA voices may be expected to be wider than those for non-OSA subjects.

Another important aim for our research was to investigate to what extent these possible connections between formants and OSA may be affected by the specific clinical characteristics in OSA populations. The objective was to evaluate whether clinical variables such as height, age, cervical perimeter, or BMI can present a stronger impact on vocal tract resonances than specific OSA-related alterations. This is an important issue that has already raised a debate when evaluating recent speech-based detection methods[18,21]: it might be that due to human voice changes with age, height, or BMI, detection methods may be predicting these demographic and anthropometric variables rather than OSA. To contribute to the research on this important topic, for each individual in our database, not only speech records but also clinical features (height, weight, BMI, cervical perimeter, gender, and age) were made available.

From a methodologic point of view, we expanded the analysis in the study by Robb et al[17] to include the first three formant central frequencies (ie, F1, F2, and F3) and their associated bandwidths (ie, B1, B2, and B3), measured over sustained phonations of /a/, /e/, /i/, /o/, and /u/. We decided to extend our analysis to all five Spanish vowels and to follow a class-dependent approach as it has been suggested that acoustic cues for OSA detection can be phonetic dependent.[25] We realize that analyzing connected speech may provide further insights on the connections between speech and OSA; however, in this study, we have followed the same methodologic approach designed in the study by Robb et al,[17] in which sustained phonations were used to avoid spurious deviations due to phonetic variability.[26]

In summary, in this study, we test whether the observed lower formant frequencies values and wider formant bandwidths found in the study by Robb et al[17] among OSA subjects are detected on a large database of subjects suspected to suffer from OSA submitted to PSG test in a sleep unit in Spain. This result will elucidate if plausible explanations of longer vocal tracts in