Identification of Obstructive Sleep Apnea Patients from Tracheal Breath Sound Analysis during Wakefulness in Polysomnographic Studies

Jordi Solà-Soler; José A. Fiz, *IEEE Member*; Abel Torres, *IEEE Member*; and Raimon Jané, *IEEE Member*

Abstract—**Obstructive Sleep Apnea (OSA) is currently diagnosed by a full nocturnal polysomnography (PSG), a very expensive and time-consuming method. In previous studies we were able to distinguish patients with OSA through formant frequencies of breath sound during sleep. In this study we aimed at identifying OSA patients from breath sound analysis during wakefulness. The respiratory sound was acquired by a tracheal microphone simultaneously to PSG recordings. We selected several cycles of consecutive inspiration and exhalation episodes in 10 mild-moderate (AHI<30) and 13 severe (AHI>=30) OSA patients during their wake state before getting asleep. Each episode's formant frequencies were estimated by linear predictive coding. We studied several formant features, as well as their variability, in consecutive inspiration and exhalation episodes. In most subjects formant frequencies were similar during inspiration and exhalation. Formant features in some specific frequency band were significantly different in mild OSA as compared to severe OSA patients, and showed a decreasing correlation with OSA severity. These formant characteristics, in combination with some anthropometric measures, allowed the classification of OSA subjects between mild-moderate and severe groups with sensitivity (specificity) up to 88.9% (84.6%) and accuracy up to 86.4%. In conclusion, the information provided by formant frequencies of tracheal breath sound recorded during wakefulness may allow identifying subjects with severe OSA.**

I. Introduction

Obstructive Sleep Apnea (OSA) is a widely prevalent sleep disorder in the general population. Its clinical implications range from sleep disruption and excessive daytime sleepiness, to suspected long-term cardiovascular implications [1]. The standard method for OSA diagnosis is a full night polysomnography (PSG), an expensive and labor procedure, which has long wait lists in public health services.

Many alternatives have been reported for the early screening of OSA, ranging from the use of type 3-4 home apnea monitors [2], to the indirect identification of OSA from the analysis of a single channel such as oxygen saturation [3], snoring sound [4] or nocturnal breath sound [5]. Recently, many efforts have been directed to OSA severity detection during wakefulness, which go beyond the early pioneering attempts of our group [6]: advanced speech analysis methods [7], a diurnal negative expiratory pressure test to predict collapsibility during sleep [8], oronasal airway pressure analysis [9], or tracheal breath sound analysis [10].

Several differences are known to exist in the upper airway (UA) of OSA as compared to normal subjects, such as higher airway collapsibility, greater force of the genioglossus and musculus uvulae, smaller cross-sectional area in the retropalatal region [11]. A recent computed tomography study has confirmed statistically significant correlations between the AHI and the UA minimal cross-sectional area, both during sleep and wakefulness [12]. The UA characteristics can be indirectly studied through sound formant frequencies, which reflect the resonances produced by the UA onto the sound. Formant frequencies of speech, snoring and nocturnal breath sound previously allowed the identification of OSA to several degrees [4-7]. Our hypothesis in the present work is that the UA differences between normal and OSA subjects can be reflected in the formant frequencies of breath sound during wakefulness.

II. Material and Methods

A. Signal Acquisition

Respiratory sound was acquired simultaneously to full-night PSG studies at Germans Trias i Pujol Universitary Hospital by a single-channel device (Snoryzer Uno, Sibel S.A., Barcelona, Spain). The external sound was recorded using a unidirectional electret condenser microphone placed over the trachea at the level of the cricoid cartilage and coupled to the skin through a conic air cavity. The external sound signal was amplified and filtered between 70 and 2000 Hz, and then digitized at a sampling rate of 5000 Hz with a 12-bit A/D converter.

B. Sound Synchronization to the PSG channels

In a typical PSG study, the two recording systems (polysomnograph and external sound recorder) are not simultaneously turned on. This fact introduces a variable and

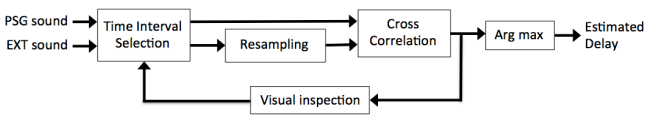


Fig. 1. Overview of the semi-automatic delay estimation algorithm between the PSG channels and the external tracheal sound channel (EXT).

This work was supported in part by Ministerio de Economía y Competitividad from Spanish Government under grant TEC2010-21703-C03-01.

J.Solà-Soler, A.Torres and R.Jané, are with Universitat Politècnica de Catalunya (UPC), Institut de Bioenginyeria de Catalunya (IBEC) and CIBER de Bioengenieria, Biomateriales y Nanomedicina (CIBER-BBN) Baldiri Reixac, 4, Torre I, 9th floor, 08028, Barcelona, Spain (jordi.sola@upc.edu; abel.torres@upc.edu; raimon.jane@upc.edu)

J.A. Fiz is with Hospital Universitari Germans Trias i Pujol, Badalona, Spain, IBEC and CIBER-BBN, (jafiz@msn.com).

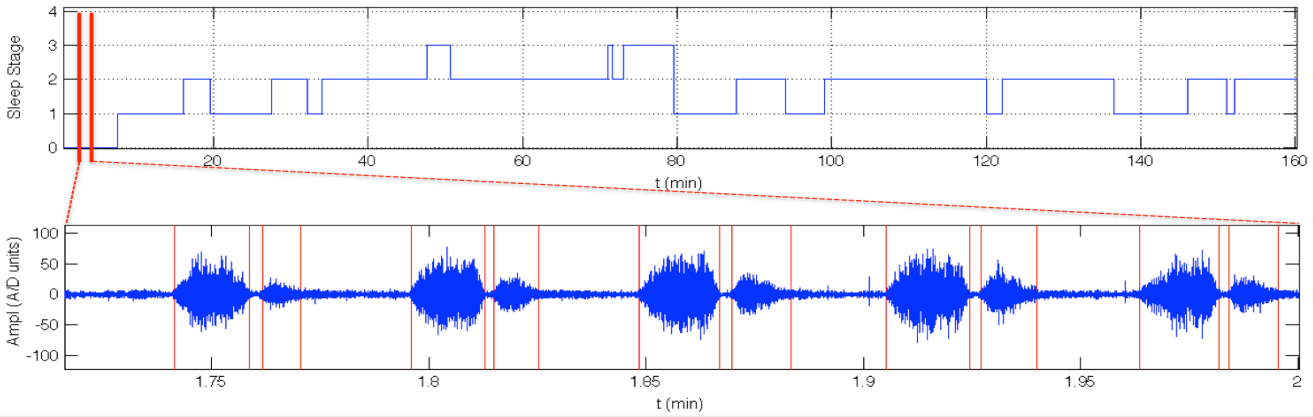


Fig. 2. After being synchronized with the sound signal, the hypnogram signal (upper panel) allowed us to identify the first period of wake state (stage 0) before the subject fell asleep. In this period, we selected several excerpts of consecutive inspiration and exhalation breathing cycles (lower panel).

undesired delay between their respective signal channels. To make sure that we selected breath sound episodes from a subject in the wake stage, a hypnogram signal perfectly synchronized to the sound channel was necessary. The synchronization was made in a semi-automatic way as follows (Fig.1): A sound signal fragment with breathing activity was manually selected. The lower quality PSG sound channel was extracted and cross-correlated with the external sound after a resampling process. The delay was then estimated as the lag of the correlation maximum. In case this maximum was not clearly visible, a new signal fragment was selected. The overall process is schematically depicted in figure 1. Table I shows the range of delay values found in the subjects of our database.

C. Selection of Episodes and Subject Database

Respiratory sound and PSG signals were simultaneously recorded and synchronized in 51 subjects suspected from OSA. For each subject, the hypnogram signal allowed us to identify the period of wake state before getting asleep. In this period, we selected at least one fragment of external sound that contained four to six consecutive inspiration and exhalation episode cycles (Fig. 2). Several subjects had to be discarded because they had a very short wake state (especially those with OSA) and/or their breathing sound had a low signal to noise ratio. The characteristics of the final database with 23 subjects are shown in Table I.

TABLE I
CHARACTERISTICS OF THE DATABASE

		AHI<30	AHI≥30	p_{val}
Subjects (N=23)	M/F	8/2	12/1	----
Age (yr)	m	48.9	56.3	0.025
	s	11.4	7.9	
BMI (kg/m^2)	m	26.5	29.8	0.051
	s	3.2	3.0	
AHI (h^{-1})	m	16.0	52.5	<0.001
	s	9.3	16.0	
Delay (s)	min	16.1	18.4	----
	med	20.7	26.1	
	max	120.0	424.9	

M=Males, F=Females, m=mean, s=standard deviation, min=minimum, max=maximum, med=median, BMI=Body Mass Index, AHI=Apnea Hypopnea Index, p_{val}=Statistical Significance of the Mann-Whitney U test between both groups.

D. Signal Analysis

Each breathing episode was characterized through the formant frequencies of its spectral envelope. Due to its local prediction error property, autoregressive (AR) linear prediction was used for spectral envelope estimation. Linear prediction coefficients were calculated by the autocorrelation method as used previously in sleep [5]. For each formant –the local maxima of the spectral envelope– its frequency (F), relative amplitude (M) and attenuation (L) were calculated, as depicted in Fig.3.

In our previous works with snoring sound, we found that the local variability of the sound features in consecutive episodes enhanced the performance of OSA classification algorithms [4]. Therefore in this work, for each episode's formant parameter $P_i \in \{F_i, L_i, M_i\}$ we have also calculated its breath to breath variability defined as the standard deviation *SdP* of the parameter's first difference $dP_i=P_i-P_{i-1}$ in consecutive breathing cycles.

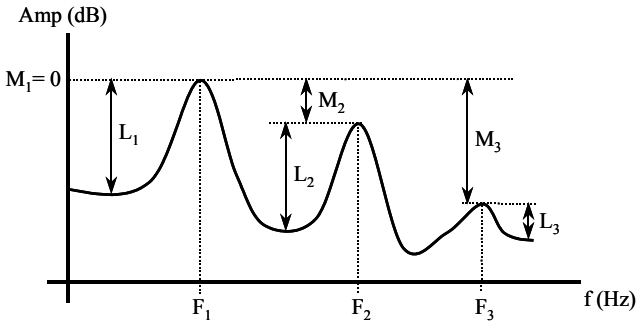
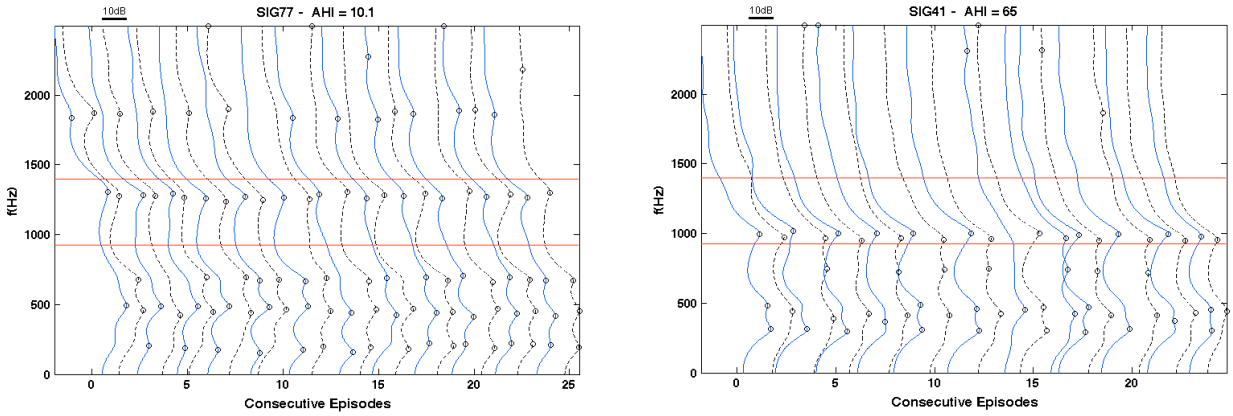


Fig. 3. Each formant in the spectral envelope of the breath sound is characterized by its frequency F, relative amplitude M and attenuation L.



(a) (b)

Fig. 4. Spectral envelope evolution in consecutive inspiration (solid) and exhalation (dashed) episodes of a mild OSA subject (a) and a severe OSA subject (b). We can appreciate a distinct behavior of the formant frequency located between 925Hz and 1400Hz, which is clearly lower in the more severe subject.

III. Results

The spectral envelope of consecutive breathing episodes had a similar morphology. Formant frequencies of inspiration and exhalation appeared in the same frequency bands in most subjects. Each breathing episode contained between four and six formant frequencies. Figure 4 shows the spectral envelope evolution in consecutive inspiration (solid) and exhalation (dashed) episodes of two subjects from the database. A distinct behavior can be observed in the formant frequency located the fourth frequency band B4=(925,1400)Hz between a mild OSA subject and a severe OSA subject (figure 4a,b).

Formant frequencies were studied in the six frequency bands defined in our previous study [5]. For each subject, inspiration and exhalation episodes were analyzed separately. In the inspiration cycle, both the frequency F4 and the attenuation M4 of the formant located in band B4 had a moderate decreasing correlation with OSA severity as measured by the AHI ($r=-0.47$ and $r=-0.53$ respectively, $p<0.05$, see Figure 5a,c). Frequency F4 was significantly lower in the inspiratory episodes of severe OSA patients (AHI≥30) than in mild-moderate ones (F4=1170.66±92.25Hz versus F4=1082.99±51.66Hz, $p=0.036$, see Figure 4 and Figure 5b). In the exhalation cycle the same tendencies were observed, but they did not reach statistical significance.

Subjects in the two severity groups were classified using Linear Discriminant Analysis (LDA). Among all the formant parameters available, the algorithm automatically selected formant frequency F4 and the breath-to-breath variability of its amplitude, SdM4. With only these two parameters, a sensitivity of 76.9%, specificity of 77.8% and accuracy of 77.3% were obtained in the leave one out cross-validation (Table II, Model 1). When the BMI was allowed to enter the model, classification performance raised to 84.6% sensitivity, 88.9% specificity and 86.4% accuracy (Table II, Model 2). In the cross-validation, specificity was maintained, but sensitivity decreased to 76.9% and accuracy to 81.8%.

IV. Discussion

In a previous study, we showed that formant frequencies of normal breath sound were useful to discriminate OSA patients during sleep [4]. The results of the present study indicate that formant frequencies of breath sound may also allow identifying severe OSA patients during wakefulness. This technique seems to capture some of the differences that are known to exist between the airways of OSA as compared to normal subjects [11]. The spectral envelope of sleeping breath sound was previously found to contain a significantly lower formant frequency below 300Hz in subjects with AHI≥10. During wakefulness, this tendency is not observed, but about half of the subjects in the present study did not show any formant frequency below 300Hz. On the other hand, the formant located between 925Hz and 1400Hz shows a significantly lower frequency in subjects with AHI≥30, something that was not previously observed during sleep [5]. Cephalometric studies indicate that the vocal tract is longer in OSA than in non-OSA subjects [12]. This could partially explain why some posterior vocal tract resonances of both speech vowels and breath sound have lower formant frequencies in OSA patients during wakefulness [13].

As mentioned in the introduction, several authors have previously addressed the identification of OSA subjects during wakefulness. Our group conducted an early pioneering study of speech vowels in a group of 18 OSA and 10 controls. Significant differences between both groups were found in the maximum frequency of the spectral harmonics of /i/ and /e/ Spanish vowels, and in the number of harmonics in the /i/ Spanish vowel [6]. Similarly, lower formant frequency values have also been reported in the English speech vowels /a/, /i/ and /u/ of subjects with OSA [13]. Recently, a greater number of acoustic speech features of 67 OSA (AHI≥5) and 26 non-OSA (AHI<5) subjects have been

TABLE II
LDA Classification Results

MODEL 1 (without BMI)

		Predicted Group (%)	
	AHI	<30	≥30
Original Group (%)	<30	77.8	22.2
	≥30	20.0	76.9
ACC = 77.3%			

MODEL 2 (with BMI)

		Predicted Group (%)	
	AHI	<30	≥30
Original Group (%)	<30	88.9	11.1
	≥30	15.4	84.6
ACC = 86.4%			

ACC=Accuracy. LDA = Linear Discriminant Analysis.

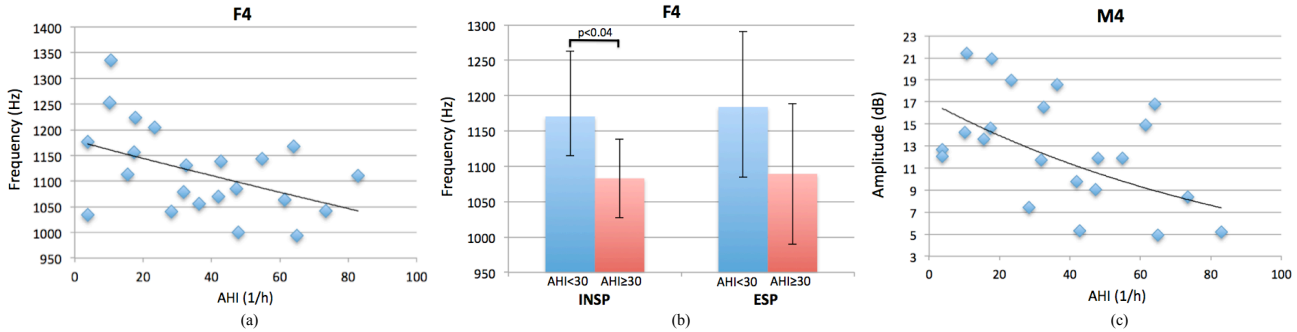


Fig. 5. The amplitude (M4, panel c) and frequency (F4, panels a) of the fourth formant of the breathing sound has a decreasing tendency with the AHI. This formant frequency is significantly lower in the inspiratory episodes of severe OSA subjects as compared to mild-moderate OSA subjects during wakefulness.

studied [7]. Gaussian mixture models allowed the classification of subjects with sensitivity of 79% (84%) and specificity of 83% (86%) for male (female) subjects, respectively [7]. Among the discriminative features, the vocal tract length and the linear predictive coefficients were selected. Therefore, it seems that autoregressive modeling of the sound recorded in the upper airway during wakefulness – be it either speech or breath sound- conveys information relative to the different upper airway dynamics in subjects with and without OSA.

Airway pressure recordings have also been used for the screening of OSA [8,9]. The upper airway collapsibility has been evaluated by a negative expiratory pressure test. Flow drop and expiratory volume in the first 2s were statistically different between 24 normal subjects (AHI<5) and 24 severe OSA patients (AHI≥30). Severe OSA subjects could be identified with sensitivity (specificity) of 95.8% (95.8%) [8]. In this study, both populations were clearly non-overlapping (AHI<5 vs AHI≥30). Short recordings of oronasal airway pressure obtained from PSG during the awake period were also studied in 20 non-OSA (AHI<5) and 21 OSA (AHI≥15) subjects [9]. An index measure based on the Hilbert oscillatory modes allowed the detection of OSA with a sensitivity (specificity) of 81% (95%).

Our present results have been obtained in a more reduced subject database, but the population has a single severity cut-point at AHI=30. Despite this clear overlapping (see Figure 5b), severe OSA subjects have been identified with sensitivity (specificity) up to 88.9% (84.6%). Those figures are expected to reduce as the population increases, as it is the case in a recent study of breath sound intensity in 30 mild OSA (AHI<15) and 23 severe OSA (AHI≥15) [10]. In that study, two sound features of the inspiratory breath sound -the median average power and the kurtosis- allowed the classification of patients with a sensitivity (specificity) of 85% (81.2%), respectively.

In this study the external sound signal was synchronized with the full PSG recordings. This allows a simultaneous analysis of breath sound features during wakefulness and sleep. However, the present study has some limitations: this setup limits the wakefulness analysis to the amount of time that a subject spends initially awake. Therefore, a portion of the originally available recordings (especially among severe OSA) had to be discarded because these subjects fell quickly asleep. The results of the present study need further validation in a database with a greater number of subjects.

V. Conclusion

Some formant frequencies of normal breath sound have a distinct behavior in mild-moderate than in severe OSA subjects during wakefulness. In combination with anthropometric information such as the BMI, and other diurnal parameters such as breath sound intensity, speech or airway pressure measures, they can aid to the screening of OSA, based on the study of the UA acoustic characteristics.

References

[1] J. Duran, S. Esnaola, R. Rubio, A. Iztueta, "Obstructive sleep apnoea-hypopnoea and related clinical features in a population-based sample of subjects aged 30 to 70 yr," *Am. J. Respir. Crit. Care Med*., vol. 63, pp. 685-689, 2001.

[2] W. W. Flemons, M. R. Littner, J. A. Rowley, P. Gay et al., "Home diagnosis of Sleep Apnea: a systematic review of the literature," *Chest*, vol. 124, no. 4, pp. 1543–79, 2003.

[3] D. Alvarez, R. Hornero, J. Victor Marcos, F. Del Campo et al., "Applying time, frequency and nonlinear features from nocturnal oximetry to OSA diagnosis," *30th Conf. Proc. IEEE Eng. Med. Biol. Soc.,* 2008, pp. 3872-5.

[4] J. Solà-Soler, J. A. Fiz, J. Morera and R. Jané, "Multiclass classification of subjects with Sleep Apnoea-Hypopnoea Syndrome through snoring analysis," *Med. Eng. Phys.*, vol. 34, no.9, pp.1213-20, 2012.

[5] J. Solà-Soler, R. Jané, J. A. Fiz and J. Morera, "Formant frequencies of normal breath sounds of snorers may indicate the risk of Obstructive Sleep Apnea Syndrome," *30th Conf. Proc. IEEE Eng. Med. Biol. Soc.,* 2008, pp. 3500-3.

[6] J. A. Fiz, J. Morera, J. Abad, A. Belsunces, J. I. Fiz, R. Jané, P. Caminal and D. Rodenstein "Acoustic analysis of vowel emission in Obstructive Sleep Apnea," *Chest*, vol. 104, no. 4, pp. 1093–6, 1993.

[7] E. Goldshtein, A. Tarasiuk and Y. Zigel, "Automatic detection of Obstructive Sleep Apnea using speech signals," *IEEE Trans. Biomed. Eng.*, vol. 58, no. 5, pp.1373-82, 2011.

[8] S. Romano, A. Salvaggio, R. P. Hirata, A. L. Blue et al., "Upper airway collapsibility evaluated by a negative expiratory pressure test in severe obstructive sleep apnea," *Clinics (Sao Paulo)*, vol. 66, no. 4, pp.567-72, 2011.

[9] P. Caseiro, R. Fonseca-Pinto and A. Andrade, "Screening of obstructive sleep apnea using Hilbert-Huang decomposition of oronasal airway pressure recordings," *Med. Eng. Phys.*, vol. 32, pp.561-8, 2010.

[10] A. Montazeri, E. Giannouli and Z. Moussavi, "Assessment of Obstructive Sleep Apnea and its severity during wakefulness," *Ann. Biomed. Eng.*, vol. 40, no. 4, pp. 916–24, 2012.

[11] I. Ayappa and D. M. Rappaport, "The upper airway in sleep: physiology of the pharynx," *Sleep Med. Rev*., vol. 7, pp. 9-33, 2003.

[12] J. Y. Heo and J. S. Kim, "Correlation between severity of sleep apnea and upper airway morphology: Cephalometry and MD-CT study during awake and sleep states," *Acta Oto-Laringol.*, vol. 131, pp. 84–90, 2011.

[13] M.P. Robb, J. Yates and E. J. Morgan, "Vocal Tract Resonance Characteristics in Adults with Obstructive Sleep Apnea", *Acta Oto-Laringol (Stockh)*, vol. 117, pp. 760–3, 1997.