Detection of Obstructive Sleep Apnea in Awake Subjects by Exploiting Body Posture Effects on the Speech signal

M. Kriboy, A. Tarasiuk, Y. Zigel

Abstract—**Obstructive sleep apnea (OSA) is a common sleep disorder. OSA is associated with several anatomical and functional abnormalities of the upper airway. It was shown that these abnormalities in the upper airway are also likely to be the reason for increased rate of apneic events in the supine position. Functional and structural changes in the vocal tract can affect the acoustic properties of speech. We hypothesize that acoustic properties of speech that are affected by body position may aid in distinguishing between OSA and non-OSA patients. We aimed to explore the possibility to differentiate OSA and non-OSA patients by analyzing the acoustic properties of their speech signal in upright sitting and supine positions. 35 awake patients were recorded while pronouncing sustained vowels in the upright sitting and supine positions. Using linear discriminant analysis (LDA) classifier, accuracy of 84.6%, sensitivity of 92.7%, and specificity of 80.0% were achieved. This study provides the proof of concept that it is possible to screen for OSA by analyzing and comparing speech properties acquired in upright sitting vs. supine positions. An acoustic-based screening system during wakefulness may address the growing needs for a reliable OSA screening tool; further studies are needed to support these findings.**

Keywords: **OSA, speech signal processing, LDA, KNN.**

I. INTRODUCTION

Obstructive sleep apnea (OSA) is a sleep disorder affecting 3% to 7% of adults [1], characterized by recurrent obstruction of the upper airway and snoring during sleep. OSA severity is defined by the apnea-hypopnea index (AHI), which is the average number of apneas and hypopneas during one hour of sleep. OSA is associated with fragmented sleep, excessive daytime sleepiness, and cardiovascular morbidity [2].

The accepted gold standard diagnostic study for OSA is polysomnography (PSG). During PSG study various biological signals are recorded [3]. PSG is time consuming, expensive, and uncomfortable for the patient; because of these disadvantages many patients remain undiagnosed [1] and alternative cost-effective approaches for OSA diagnosis are needed.

M. Kriboy is with the Department of Biomedical Engineering, Ben-Gurion University of the Negev, Beer–Sheva, Israel (mayakr@post.bgu.ac.il).

A. Tarasiuk is with the Sleep-Wake Disorders Unit, Soroka University Medical Center, and Department of Physiology, Faculty of Health Sciences, Ben-Gurion University of the Negev, Israel (tarasiuk@bgu.ac.il).

Y. Zigel is with the Department of Biomedical Engineering, Ben-Gurion University of the Negev, Beer–Sheva, Israel (corresponding author, phone: +972-8-642-8372; fax: +972-8-642-8371; e-mail: yaniv@bgu.ac.il).

Several studies have confirmed that OSA is associated with anatomical and functional abnormalities of the upper airway [4, 5] that can affect speech [6, 7]. Anatomic and functional differences in the vocal tract components can affect the acoustic properties of speech [8]. Fox et al. [9] hypothesized that some acoustic speech features of patients with OSA may be distinct from those of non-OSA subjects; using a perceptual study it was found that patients with OSA suffer from resonance, phonation, and articulation irregularities. Robb et al. [10] found differences in formant band-width and frequencies of OSA and non-OSA subjects' vowels; Pozo et al. [11] found disparity in the distance between second and third formant frequency of the vowel /i/ in severe OSA patients and non-OSA patients. We recently found that acoustic features from speech signals during wakefulness can detect OSA patients with good specificity and sensitivity; however, in this study speech signals were analyzed only in the upright sitting position [12].

The combination of several elements such as unfavorable airway geometry, increase in collapsibility, gravity, and inadequate dilator muscle compensation, is likely to be the reason for increased frequency and severity of OSA in the supine position [13]. Pae et al. [14] compared cephalograms and tongue EMG recordings of awake healthy subjects and OSA patients, in both supine and upright positions. Statistical differences were found in tongue cross-sectional area, oropharyngeal cross-sectional area, and resting genioglossus EMG activity. Martin et al. [15] studied the effect of posture on upper airway cross-sectional areas using acoustic reflection in awake OSA patients, snorers without OSA, and healthy subjects. OSA patients had smaller decrease in cross-sectional areas when moving from upright sitting position to supine, than both snorers and controls. Montazeri et al. [16] have recorded tracheal breath sounds of awake subjects in the upright sitting and supine positions. Using power spectrum, Kurtosis, and Katz fractal dimensions analysis, classification accuracy of 83% was achieved when classifying subjects into two groups, non-OSA or mild vs. moderate and severe.

Based on these evidences concerning the association between OSA and speech, and OSA and body position, we hypothesize that some of the speech features of OSA patients might be modified when moving from upright sitting position to the supine position in a different manner and extent than those of non-OSA subjects.

Few studies with the objective of OSA/non-OSA classification using speech signals have been conducted [11, 12, 17, 18]. However, the novelty of our study is reflected by

978-1-4244-7929-0/14/$26.00 ©2014 IEEE

exploiting the body position as a tool to highlight the acoustic changes arising from anatomical and functional abnormalities associated with OSA. Moreover, the proposed method uses sustained vowels solely, and therefore has the major benefit of a language-independent method. Using speech signal records of awake patients in the upright sitting and supine positions, we explored the ability to differentiate between OSA and non-OSA subjects using a simple classifier such as linear discriminant analysis (LDA).

II. METHODS

A. Subjects and Data

The database for this study consists of 35 male subjects' speech recordings. Subjects were referred to the Sleep-Wake Unit of Soroka University Medical Center for PSG study in order to evaluate sleep-disordered breathing. Subjects were recorded immediately prior to PSG study, using a digital audio recorder (Handy recorder "H4" by "ZOOM"), pronouncing the sustained vowels /a/ and /e/ for about 2 seconds, in the upright sitting and supine positions. We chose to use sustained vowels only, since several studies [19-21] have indicated that sustained vowels are more affected by body posture; moreover, this scheme reduces the influence of subjects' cooperation and accent. We recorded at a sampling rate of 44.1 kHz. Subjects' age, BMI, and AHI are summarized in Table I. Subjects were divided to two groups using a conventional cutoff value of 10, i.e., OSA patients (AHI>10) and non-OSA subjects (AHI≤10) [22].

TABLE I. SUBJECTS' CHARACTERISTICS

Group	Number of subjects	AHI [events/hr] (range)	BMI [kg/m^2] (range)	Age [yr] (range)
AHI ≤ 10	12	5.33 ± 2.31 (1.40-9.50)	28.00 ± 5.10 (23.00-40.60)	39.64 ± 16.18 (19.1-69.5)
AHI > 10	23	28.30 ± 19.69 (10.20-64.30)	30.01 ± 3.84 (17.00-38.30)	47.64 ± 11.75 (27.10-71.90)

The values are presented as mean ± SD corresponding to the relevant units.

B. Preprocessing and Feature extraction

For each subject, four speech segments were labeled: /a/-supine, /a/-upright, /e/-supine, /e/-upright. Then, a preprocessing procedure of DC removal, normalization, and pre-emphasizing was applied.

Our study focuses on body posture effects; therefore, we decided to use features that can reveal morphological changes. Linear predictive coding coefficients (LPC) represent the speech signals' spectral envelope that models the vocal tract as a linear filtering system; therefore, differences in the vocal tract shape could be conveyed by the LPC. From each segment, 48 LPC coefficients were extracted [8].

In order to explore differences between OSA and non-OSA subjects in the upright position, in the supine position, and finally the differences between the positions, we created 3 feature vectors for each subject: the first, containing 48 LPC coefficients for each vowel in the upright sitting position; the second, containing 48 LPC coefficients for each vowel in the supine position; and the third, containing the subtraction of the second from the first. Finally, we concatenated the 3 feature vectors into one vector.

C. Feature selection

We ran a two-tailed t-test on each of the features in order to find the most discriminative features. Since the number of tested hypotheses is large, we used the weighted Bonferroni correction [23]. Using a large number of features for classification in a small population can cause over-fitting; therefore, the K-best scheme was applied. We chose the most significantly different features between OSA and non-OSA groups (i.e., features with the lowest *p*-value). Since our database is relatively small, we limited the number of selected features to two.

D. Classification

Using the selected features, we experimented and evaluated the LDA classifier. We have also compared the LDA performance to the K nearest neighbors (KNN) classifier, which is also considered as a simple classifier.

LDA classifier: assuming two normal distributions, the Bayes decision rule, is a quadratic function of the observations. The Bayes optimal solution is to classify points as being from the first class if the log-likelihoods ratio (LLR) is above some threshold [24]:

$$LLR(\mathbf{x}) = \frac{1}{2}(\mathbf{x}-\boldsymbol{\mu}_O)^T \boldsymbol{\Sigma}_O^{-1}(\mathbf{x}-\boldsymbol{\mu}_O) - \frac{1}{2}(\mathbf{x}-\boldsymbol{\mu}_{NO})^T \boldsymbol{\Sigma}_{NO}^{-1}(\mathbf{x}-\boldsymbol{\mu}_{NO}) + \frac{1}{2}\ln\frac{|\boldsymbol{\Sigma}_O|}{|\boldsymbol{\Sigma}_{NO}|} > \ln\frac{P_O}{P_{NO}} \tag{1}$$

where $\mathbf{x}$ is the feature vector, $\boldsymbol{\mu}_O, \boldsymbol{\mu}_{NO}$ and $\boldsymbol{\Sigma}_O, \boldsymbol{\Sigma}_{NO}$ are the mean vectors and covariance matrices of the OSA group and the non-OSA group class conditional densities, respectively. P_O, P_{NO} are the a priori probabilities. When assuming that both groups have the same pooled covariance matrix, the Bayes decision rule can be expressed as a linear function of the observations [24]:

$$LLR(\mathbf{x}) = (\boldsymbol{\mu}_{NO}-\boldsymbol{\mu}_O)^T \boldsymbol{\Sigma}^{-1}\mathbf{x} + \frac{1}{2}(\boldsymbol{\mu}_O^T \boldsymbol{\Sigma}^{-1}\boldsymbol{\mu}_O - \boldsymbol{\mu}_{NO}^T \boldsymbol{\Sigma}^{-1}\boldsymbol{\mu}_{NO}) > \ln\frac{P_O}{P_{NO}} \tag{2}$$

KNN classifier: for each test sample, K nearest neighbors are found by calculating the distance of the test sample from each of the samples. The number of neighbors from each class among the K selected samples is calculated. The test sample is then classified by a voting procedure, i.e., by a majority of the KNNs. In this study we chose to use 3 neighbors, with the Euclidean metric; unity standard deviation normalization was applied.

To evaluate classifiers' performances we have used the Leave one out (LOO) and K-fold (K=5) cross-validation schemes; a comparison to the resubstitution optimistic results was taken as well. Classifiers' performance was evaluated using sensitivity, specificity, and accuracy:

$$Sensitivity = \frac{N_{TP}}{N_{TP} + N_{FN}} \times 100 \qquad (3)$$

$$Specificity = \frac{N_{TN}}{N_{TN} + N_{FP}} \times 100 \qquad (4)$$

$$Accuracy = \frac{N_{TP} + N_{TN}}{N_{TN} + N_{FP} + N_{TP} + N_{FN}} \times 100 \qquad (5)$$

III. RESULTS & DISCUSSION

The K-best feature selection procedure resulted in the two features presented in Table II. Fig. 1 presents boxplots of the selected features; good separation ability can be seen. The selected features both belong to the sustained vowel /e/, which might indicate greater differences between OSA and non-OSA patients during the pronunciation of this vowel.

This result is supported by Fiz et al. [25], who found significant differences in the maximum frequency of harmonics of /e/ and /i/, but not in /a/,/o/,/u/.

Due to a relatively small sample size, we have chosen a relatively simple procedure and criterion for the feature selection process. Since we have chosen the K-best procedure, i.e. each feature is tested separately; we had to examine also the difference features, which are actually a linear combination of the upright and supine features. Moreover, from a clinical point of view these features enable us to examine our main hypothesis that some of the speech features of OSA patients might be modified when moving from upright sitting position to the supine position in a different manner and extent than those of non-OSA subjects.

To evaluate the ability to differentiate between OSA and non-OSA subjects, both the KNN classifier and the LDA classifier were tested; both of these classifiers do not prone to over-fit when using a relatively small sample size due to relatively small number of required parameters. Table III summarizes the classifiers' performances as evaluated using the LOO, K-fold, and resubstitution schemes. Overall, one can see that the LDA classifier outperformed the KNN classifier. The best performance was achieved using two features. Using the K-fold scheme sensitivity of 92.7% and specificity of 80% were achieved. Fig. 2 presents a scatter plot of the data and the decision boundary as determined by the LDA classifier using the resubstitution validation scheme. One can see that a linear decision boundary is a suitable choice in this case.

TABLE II. SELECTED FEATURES

Selection order	Vowel	Feature type	LPC coefficient	Symbol
1	/e/	Difference	42	e_diff_42
2	/e/	Upright	47	e_up_47

TABLE III. CLASSIFIERS' PERFORMANCE

Classifier	Number of features	Validation scheme	Performance*
LDA	1	Resubstitution	81.8 (90.5, 66.7)
	2		87.9 (90.5 ,83.3)
	1	K-fold	77.7 (92.7 ,65.0)
	2		84.6 (92.7 ,80.0)
	1	LOO	78.8 (90.5 ,58.3)
	2		81.8 (85.7, 75.0)
KNN	1	K-fold	82.9 (92.7, 75.0)
	2		82.9 (92.7, 75.0)
	1	LOO	71.4 (73.9, 66.7)
	2		80.0 (82.6, 75.0)

*The values are presented as accuracy (sensitivity, specificity) in percentage.

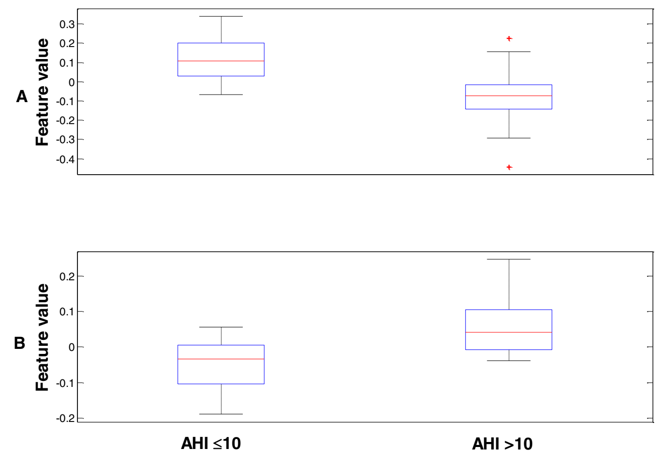

Figure 1. (A) boxplot of e_diff_42 (the difference between the 42nd LPC coefficent of the vowel /e/ in the upright and the supine positions); (B) boxplot of e_up_47 (47th LPC coefficent of the vowel /e/ in the upright position); central marks indicate medians, box edges indicate the 25th and 27th percentiles.

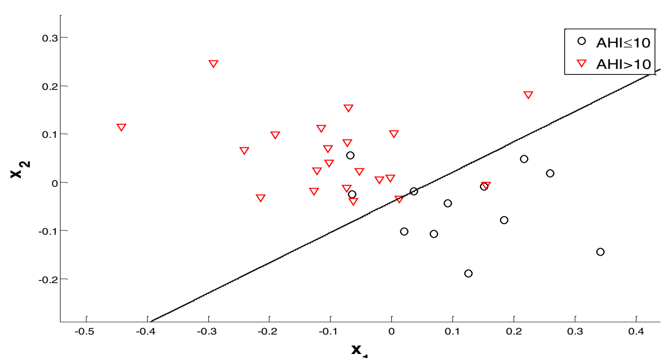

Figure 2. A scatter plot of the data and the decision boundary as determined by the LDA classifier using the resubstitution validation scheme. $\mathbf{x}_1$ and $\mathbf{x}_2$ represent e_diff_42 and e_up_47, respectively. The solid line represents the LDA decision boundary, the circle markers represent subjects with AHI≤10, the triangle markers represent subjects with AHI>10.

IV. CONCLUSIONS & FUTURE WORK

This study provides evidence that using sustained vowel recordings in different body positions can highlight acoustic differences between OSA and non-OSA subjects. Our system is based on sustained vowel recordings; therefore, it may provide an accent- and language-independent, simple to use screening tool for OSA. Further studies are needed to reinforce our findings in a larger database; more complex classifiers should be considered as well.

The proposed method can be used as a screening for OSA. Such a tool may reduce the number of undiagnosed patients and unneeded PSG studies.

ACKNOWLEDGMENT

We would like to thank Mrs. Bruria Freidman from the Sleep Wake Disorder Unit of Soroka University Medical Center, for her support and collaboration.

REFERENCES

[1] N. M. Punjabi, "The epidemiology of adult obstructive sleep apnea," *Proceedings of the American Thoracic Society,* vol. 5, p. 136, 2008.

[2] A. Tarasiuk, S. Greenberg-Dotan, T. Simon, A. Tal, A. Oksenberg, and H. Reuveni, "Low socioeconomic status is a risk factor for cardiovascular disease among adult obstructive sleep apnea syndrome patients requiring treatment," *CHEST Journal,* vol. 130, pp. 766-773, 2006.

[3] H. Reuveni, E. Schweitzer, and A. Tarasiuk, "A cost-effectiveness analysis of alternative at-home or in-laboratory technologies for the diagnosis of obstructive sleep apnea syndrome," *Medical decision making,* vol. 21, pp. 451-458, 2001.

[4] M. R. Mannarino, F. Di Filippo, and M. Pirro, "Obstructive sleep apnea syndrome," *European Journal of Internal Medicine,* vol. 23, pp. 586-593, 2012.

[5] Y. Finkelstein, D. Wexler, E. Horowitz, G. Berger, A. Nachmani, M. Shapiro-Feinberg, *et al.*, "Frontal and lateral cephalometry in patients with sleep-disordered breathing," *The Laryngoscope,* vol. 111, pp. 634-641, 2001.

[6] T. M. Davidson, "The Great Leap Forward: the anatomic basis for the acquisition of speech and obstructive sleep apnea," *Sleep medicine,* vol. 4, pp. 185-194, 2003.

[7] T. M. Davidson, J. Sedgh, D. Tran, and C. J. Stepnowsky Jr, "The anatomic basis for the acquisition of speech and obstructive sleep apnea: evidence from cephalometric analysis supports The Great Leap Forward hypothesis," *Sleep medicine,* vol. 6, pp. 497-505, 2005.

[8] J. R. Deller, J. G. Proakis, and J. H. Hansen, *Discrete-time processing of speech signals*. New York: IEEE Press, 2000.

[9] A. W. Fox, P. K. Monoson, and C. D. Morgan, "Speech dysfunction of obstructive sleep apnea," *Chest,* vol. 96, pp. 589-95, 1989.

[10] M. Robb, J. Yates, and E. Morgan, "Vocal tract resonance characteristics of adults with obstructive sleep apnea," *Acta oto-laryngologica,* vol. 117, pp. 760-763, 1997.

[11] F. P. Rubén, B. M. Jose Luis, H. G. Luis, L. G. Eduardo, A. R. José, and T. Doroteo T, "Assessment of severe apnoea through voice analysis, automatic speech, and speaker recognition techniques," *EURASIP Journal on Advances in Signal Processing,* vol. 2009, 2009.

[12] E. Goldshtein, A. Tarasiuk, and Y. Zigel, "Automatic detection of obstructive sleep apnea using speech signals," *Biomedical Engineering, IEEE Transactions on,* vol. 58, pp. 1373-1382, 2011.

[13] S. A. Joosten, D. M. O'Driscoll, P. J. Berger, and G. S. Hamilton, "Supine position related obstructive sleep apnea in adults: Pathogenesis and treatment," *Sleep medicine reviews,* 2013.

[14] E.-K. Pae, A. A. Lowe, K. Sasaki, C. Price, M. Tsuchiya, and J. A. Fleetham, "A cephalometric and electromyographic study of upper airway structures in the upright and supine positions," *American Journal of Orthodontics and Dentofacial Orthopedics,* vol. 106, pp. 52-59, 1994.

[15] S. E. Martin, I. Marshall, and N. J. Douglas, "The effect of posture on airway caliber with the sleep-apnea/hypopnea syndrome," *American journal of respiratory and critical care medicine,* vol. 152, pp. 721-724, 1995.

[16] A. Montazeri, E. Giannouli, and Z. Moussavi, "Assessment of obstructive sleep apnea and its severity during wakefulness," *Annals of biomedical engineering,* vol. 40, pp. 916-924, 2012.

[17] O. Elisha, A. Tarasiuk, and Y. Zigel, "Detection of obstructive sleep apnea using speech signal analysis," in *Models and analysis of vocal emissions for biomedical applications Conf.*, Florence, 2011, pp. 13-16.

[18] M. Kriboy, A. Tarasiuk, and Y. Zigel, "Obstructive sleep apnea detection using speech signals," in *Afeka-AVIOS Speech processing Conf.*, Tel-Aviv, 2013, pp. 1-4.

[19] M. Stone, G. Stock, K. Bunin, K. Kumar, M. Epstein, C. Kambhamettu, *et al.*, "Comparison of speech production in upright and supine position," *The Journal of the Acoustical Society of America,* vol. 122, p. 532, 2007.

[20] M. K. Tiede, S. Masaki, and E. Vatikiotis-Bateson, "Contrasts in speech articulation observed in sitting and supine conditions," in *Proceedings of the 5th Seminar on Speech Production, Kloster Seeon, Bavaria*, 2000, pp. 25-28.

[21] O. Engwall, "Assessing Magnetic Resonance Imaging Measurements," *Speech Production: Models, Phonetic Processes, and Techniques,* p. 300, 2013.

[22] V. Hoffstein and J. Szalai, "Predictive value of clinical features in diagnosing obstructive sleep apnea," *Sleep,* vol. 16, pp. 118-122, 1993.

[23] Y. Benjamini and Y. Hochberg, "Controlling the false discovery rate: a practical and powerful approach to multiple testing," *Journal of the Royal Statistical Society. Series B (Methodological),* pp. 289-300, 1995.

[24] K. Fukunaga, *Introduction to statistical pattern recognition*: Access Online via Elsevier, 1990.

[25] J. Fiz, J. Morera, J. Abad, A. Belsunces, M. Haro, J. Fiz, *et al.*, "Acoustic analysis of vowel emission in obstructive sleep apnea," *Chest Journal,* vol. 104, pp. 1093-1096, 1993.