2021 18th International Conference on Electrical Engineering, Computing Science and Automatic Control (CCE). Mexico City, Mexico. November 10-12, 2021

An Objective Analysis of Human Snoring Based on an Acoustic Technique to Determine the Obstruction Site

Judith Guzman
Electrical Engineering
CINVESTAV, IPN
CDMX, Mexico
jaguzman@cinvestav.mx

Eladio Cardiel
Electrical Engineering
CINVESTAV, IPN
CDMX, Mexico
ecardiel@cinvestav.mx

Laura I. Garay-Jimenez
UPIITA
Instituto Politécnico Nacional
CDMX, Mexico
lgaray@ipn.mx

Pablo Rogelio Hernandez
Electrical Engineering
CINVESTAV, IPN
CDMX, Mexico
pablo.rogeli@cinvestav.mx

Abstract— **Snoring is considered a cause of obstructive sleep apnea syndrome and a significant disturbance of social stability due to the discomfort it can cause in the environment. For these reasons, the study of snoring has generated interest in medicine. It has led to the development of techniques and instrumentation that contribute to improving effective diagnoses and treatments. The acoustic analysis of snoring signals has gained importance in recent years. Consequently, the acoustic characteristics of snoring have been explored as primary medical care alternatives, complemented with polysomnography to study mainly cases of apnea syndrome. However, polysomnography is expensive, both in implementing specialized infrastructure and carrying out the studies, invasive and not very accessible due to waiting times. Using current techniques in speech recognition, the acoustic analysis of snoring is of particular interest and shows promise as a medical alternative. In this work, an application of contemporary speech analysis techniques in the snoring signal, obtained without sedation of the patient, is proposed to determine its relationship with the site of obstruction that originates it that eventually could improve its treatment.**

Keywords— Human Snoring, Acoustical Analysis, Snoring Etiology.

I. Introduction

Snoring is defined as the sound emitted mainly in the inspiratory phase of the respiratory cycle during sleep, with loud and harsh snorts caused by the vibration of the soft palate and or structures of the pharyngeal airway, phenomena considered disturbing of social stability. For the inconvenience, it can cause in the environment [1]. It is regarded as a cause or manifestation of other conditions, such as Obstructive Sleep Apnea Syndrome (OSAS) [2], irritation and inflammation of pharyngeal tissues [3], [4] and some others of occupational nature such as daytime sleepiness [5], [6], or social in marital life.

The sound of snoring varies from person to person, from day to day, and from hour to hour. These variations in its acoustic characteristics, such as volume, duration, and tone, carry information regarding the anatomical or physiological cause that produces it. Objectively evaluating these characteristics can be used to determine their origin, a fundamental factor in specifying surgeries, as well as to assess the evolution of the patient after some treatment to monitor its efficacy [2], since commonly only evaluations are performed subjective in this regard [7], [8].Currently, the diagnosis and evaluation of Obstructive Sleep Apnea Syndrome (OSAS), other conditions, and disorders related to snoring are carried out using nasendoscopy and polysomnography techniques [9]. Although snoring is a subject of sleep medicine, to which no attention has been paid beyond its relation to OSAS, there is evidence to suggest that its impact on health goes beyond the discomfort caused to other people who they share the space with the snorer [2], [9].

Polysomnography is currently the most widely used technique for detecting and investigating snoring besides various other sleep disorders. However, if its employment is for snoring studies only , it results very expensive, near eight thousand dollars. The system requires highly specialized infrastructure and very demanding conditioned spaces. It is invasive and not very accessible, in addition to being used mainly to detect apneas. Additionally, most studies have been carried out with sedated patients and through cannulas, which affects the results obtained. For these reasons, acoustic analysis of snoring is a promising technique as a medical alternative [10], [11], [12]. Agrawal et al. [13] provided the first published evidence that nasendoscopy did not allow natural sleep and reported that induced snores contain a higher frequency component of sound, not evident during natural snoring, consistent with the collapse of the base of the tongue. But further studies under non sedate condition are required to obtain a non-tagged classification [14][15].

This project aims to develop a diagnostic tool based on the digital processing of acoustic signals of snoring in humans, to identify the origin of the obstruction, using snoring produced in natural sleep in a non-clinical environment and applying techniques of speech and paralinguistic analysis. The results obtained are compared with the values of the parameters of the case studies reported by Miyazaki and Agrawal [1], [13], based on clinical studies of induced sleep, whose results provide a relationship between the acoustic parameters of snoring and the obstruction site.

978-1-6654-0029-9/21/$31.00 ©2021 IEEE

II. METHOD AND MATERIALS

A. *Study parameters*

The selection of study parameters detailed in G, was based on the results of the classification of snoring events according to the area of origin of Quian et al. [16], in addition to the work of Miyazaki [1] and Agrawal [13]. This was done after studying in detail the kind of information provided by each of these acoustic analysis techniques, selecting the most relevant for their application in the analysis of snoring signals as shown in Fig.1.

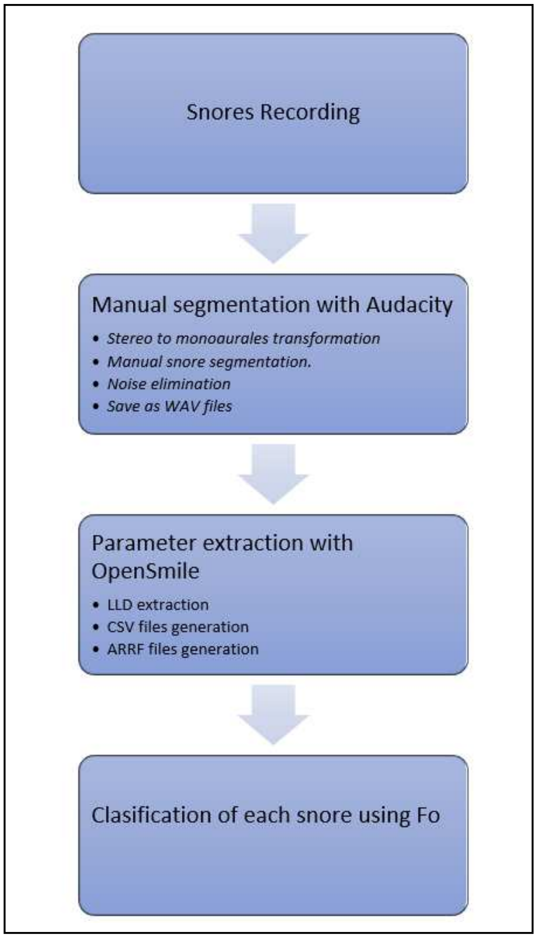

Fig. 1. Methodology used for the analysis of snoring signals

B. *Subjects of study*

Seven snoring volunteers participated in the study, five men and two women, aged between 35 and 65 years old, without exclusion criteria. One of them presented obesity and the other morbid obesity and high blood pressure. The participants declared themselves to be chronic snorers, and only one of them had undergone surgery to remove palatine tonsils and adenoids, in addition to having received treatment in the paranasal fossae.

C. *Recording conditions*

Natural sleep records were obtained for three consecutive days in the environment and the daily sleep schedule of each participant by using a linear PCM recorder (Tascam, DR40 version 2, Tokyo, Japan), configured with dual-mode and .wav format, 44.1ks / sec sampling rate, 120Hz low-pass filter, and -12dB recording level. The device was placed at a distance of 60 cm at the height of the head so as not to interfere with the natural movements of the subject [17]. The temperature of the rooms was between 15 and 20 ° C, 50 and 60% Relative Humidity, in addition to 20 dB ± 10% of ambient noise levels. The position of the subject was predominantly supine dorsal, and all followed their usual circadian cycle.

D. *Signal conditioning*

The digital audio files were manually conditioned using the open-access software Audacity (Carmegie Mellon University, Pittsburgh, PA). The snoring of each record with better intelligibility and lower levels of interference and noise were selected, Fig. 2.

The audio was processed without modifying the sampling frequency (44.1 kHz) and the resolution (16 bits) of the original files. Four actions were carried out to clean and isolate the snoring of each patient: (a) conversion of stereo files into monaural files; (b) trimming or isolating each snoring (an inhalation and its corresponding exhalation); (c) Elimination of noise or unnecessary frequencies (equalization) through a band-pass filter with a lower cut-off frequency of 80 Hz and a higher cut-off frequency of 13 kHz. A quality factor Q of 0.71 and a slope of 48 dB / octave were used for both filter limits. Finally, (d) the processed and separated snoring were exported as .wav files, without compression.

E. *The acoustic signal analysis system*

The open-access platform OpenSmile was used since it allows the extraction of characteristics, pattern recognition specifically designed for audio and directed to academic and scientific projects. This platform extracts and analyzes factors in real-time of the parameters selected for the snoring study, using easy-work templates.

In this sense, a configuration template was developed to extract the characteristics with more significant differentiating potential than those reported by Qian [10], Eiben [12], Miyazaki [1], and Agrawal [18]. The collected audio was processed in batches using OpenSmile software on a computer with Linux Ubuntu 16 operating system. The output files were obtained in CSV format, which included timestamps of each segment of the signal. The ARFF format archives were also obtained, and the data were visualized and analyzed with the open-access software WEKA [17].

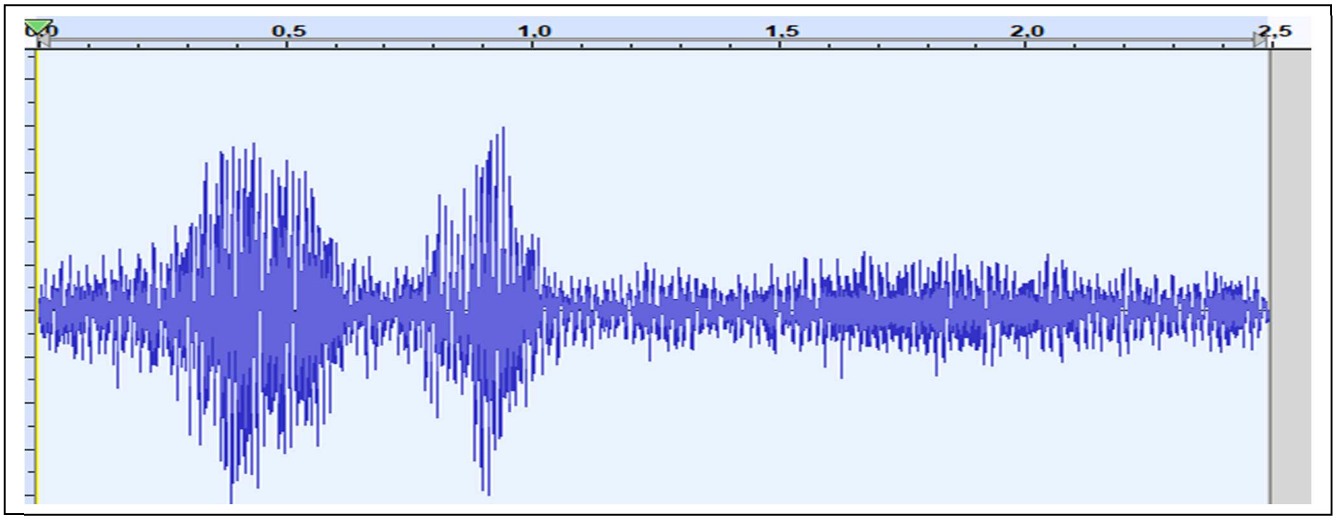

Fig. 2. Example of a manually segmented snoring signals

F. Signal processing

The JAGM_configuracion.conf file was designed, based on the platform of the free open access program openSMILE. This file contains the main configuration template that refers to other templates made modularly for this project, as well as to configuration files included in the OpenSMILE library. The JAGM_configuracion.conf file calls the necessary components to initialize the reading of the input data in .wav format. In turn, reference is made to other configuration files that specifically include the components for obtaining the selected features more easily. The openSMILE components used in the main configuration file are:

1) ../shared/standard_wave_input.conf.inc;
2) calls OpenSMILE configuration file for .wav inputs
3) JAGM_configuracion1_core.lld.conf.inc;
4) calls the low-level descriptor (LLD) configuration file
5) JAGM_configuracion2_core.lld.conf.inc;
6) calls a second LLD configuration file
7) JAGM_configuracion1_core.func.conf.inc;
8) calls the configuration file that obtains the functional values
9) JAGM_configuracion2_core.func.conf.inc;
10) calls a second configuration file that gets more functional
11) ../shared/standard_data_output.conf.inc;
12) calls the OpenSMILE configuration file for data output

G. OpenSmile Algorithm

This configuration file allows OpenSMILE to perform the following actions on the input .wav files:

1) The list of OpenSMILE components to use is initialized. In this case, it is required to define a data memory type component, as established in the OpenSMILE documentation [19].
2) The referenced files are then specified, so that standard_wave_input.conf.inc sets the input file type as a .wav file type.
3) Next, the file that configures low-level feature extraction (LLD) is specified to evaluate individual audio files by performing a window according to what is specified in this file and obtaining the parameters requested for each window.
4) This procedure generates a CSV-type output file with a variable number of lines, depending on the number of windows and the audio length.
5) The following configuration file takes as input the LLD values obtained in the previous block. It performs mathematical operations to get statistical measurements that represent the distribution of the LLD values. The output of this stage is an ARFF type file that will have an equal number of columns for all processed audio segments. This condition allows the results of each audio file to be incrementally added to the same ARFF file. This ARFF file will contain vectors of equal dimensions representing each audio file regardless of their length in time.
6) *JAGM_configuration.conf* is also utilized to concatenate the output vectors being processed,../shared/standard_data_output.conf.inc and make up the CSV and ARFF files.

7) *JAGM_configuracion1_core.lld.conf.inc* has the function of initializing the obtain of the following low level descriptors (LLD)

   a. F_0 (Fundamental frequency)
   b. Espectral flux,
   c. Spectrum on the Mel scale (MFCC)
   d. MFCC coefficients (only for voiced segments)
   e. Logarithmic spectrum
   f. Harmonics H0-H1, H1-A3
   g. Formants F1-F3
   h. Determine if a segment is voiced or not, marking the voiced value with 1.

8) *JAGM_configuracion1_core.lld.conf.inc* also applies contour smoothing to the obtained spectral values.

9) *JAGM_configuracion1_core.func.conf.inc* allows to make temporary statistics of the values obtained for F_0 and obtains summaries of the functional values of F_0 and the voiced segments.

10) *JAGM_configuracion2_core.lld.conf.inc* performs value separation tasks to calculate voiced and non voiced segments separately with their corresponding smoothing.

With these actions, it is possible to execute OpenSMILE by calling the JAGM_configuracion.conf configuration file as an argument through the Linux command line as follows:
opensmile-2.3.0/SMILExtract -C openSMILE-2.3.0/config/JAGM/JAGM_configuracion.conf // -I ENTRADA.wav -appendarff 1 -O SALIDA_func.arff -D SALIDA_LLD.csv -N ETIQUETA.
A Linux Ubuntu 16.06 (headless), 1 vCPU, and 3.75 GB of RAM machine, hosted in the Google Cloud Platform Cloud, was chosen to execute OpenSMILE, specifically in the Compute Engine product. This virtual machine was used to take advantage of the cloud hosting, assign a public IP for remote access, and carry out other applications.

IV. RESULTS

The values in Table 1 of fundamental frequency of snoring obtained in this work were compared with those reported in Miyazaki [1] and Agrawal [13], which were used as a reference since they include a relationship with the structures that produce snoring, even though they were recorded during induced sleep. The values of the other parameters were determined but they could not be compared to any due there is no material reported.
After processing 595 individual snoring audio files with OpenSMILE, the following files were obtained:

• CSV file (comma-separated value) with the values of the low-level descriptors (that is, one value for each point over time). One file is generated for each input audio file to obtain spectrum graphs of the spectral values, corresponding to the values that appeared in the CSV headers.
• ARFF file (Attribute-Relation File Format) with the functional values obtained by applying statistical measures to the total LLDs obtained for each file and thus to represent it with a vector of uniform dimension and compare it with other audio files even if they do not have the same number of windows.
The statistical values of the selected parameters were captured in an ARFF file displayed on WEKA, showing as an example the fundamental frequency for each snore used for the classification done in Table 1. This data format could be used as input for machine learning applications such as classification and clustering for future studies.

TABLE I. FUNDAMENTAL FREQUENCY (F_0) HZ, VALUES OBTAINED IN THIS WORK AND ITS CORRESPONDENCE WITH THE OBSTRUCTION SITES REPORTED BY AGRAWAL ET AL [1] AND MIYAZAKI ET AL [13].

Reported Obstruction site	Reported Central tendency estimator of Fo by [1][13]	Mean and range of Fo for snore classification in this work	Number of snores
Soft palate [1]	102.8(34.9)	102.8(89.27,116.3)	4
Soft palate [13]	137(105,189)	137(105,189)	44
Combination of sites [1]	115.7(58.9)	115(91.7,138.3)	11
Tonsils [13]	170 (85,201)	170 (85,201)	112
Larynx[1]	250 (79.7)	250(166.4,333.6)	164
Tonsils and base of tongue [1]	331.7(144.8)	331.7(248.2,425.3)	117
Epiglottis [13]	490(331,510)	490(331,510)	141
Tongue [13]	1243(2015,1277)	1243(2015,1277)	2

V. DISCUSSION

A method was developed to obtain functional parameters from snorings with the same dimension vectors regardless of their duration. The data obtained with this proposal and the relationship with medical diagnoses can be used as inputs forward to the training of neural networks for classification, allowing access to the same server running audio analysis tasks with OpenSMILE.

There are research reports similar to those of this project but that used snoring produced mainly during induced sleep [11],[12],[20]. In this work, a methodology is presented to extract characteristics of snoring produced during natural sleep.

The fundamental frequency of snoring sounds has been reported between 102.8 Hz and 331.7 Hz, depending on the site of the obstruction [1]. In another work, fundamental frequencies between 137 Hz and 1242 Hz were obtained in clinics under a medical treatment scheme with spaces and conditions different from those that study subjects habitually use to sleep [16].

The results obtained in this work, as shown in Table 1, allows establishing a relationship between the fundamental frequency of the snores studied and the anatomical area that produces them shown in Fig. 3, considering as an approximate reference the study cases reported in the literature [1] [13].

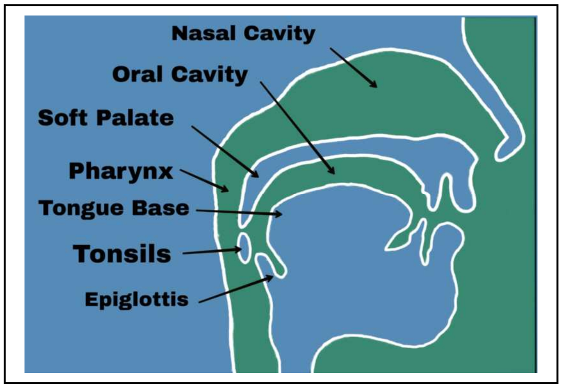

Fig. 3. Origin sites considered in [1]and [13] in the upper airways. Modified from [21]

This relationship, constitute a first approach of medical opinion respect to the possible origin of snoring but the viability of snores being used as a non-invasive preliminary test for several diseases is an open research issue [21],[22],[23],[24].

VI. Conclusions

This proposal will identify by a non-invasive procedure the origin of the airflow obstruction, using snoring sounds produced in natural sleep in a non-clinical environment by applying speech and paralinguistic analysis techniques.

The results obtained with the proposal of this work suggest carrying out a study with more cases under medical monitoring to generate a non-invasive and valuable tool in the clinic to accurately determine the specific areas of obstruction of the airways for adequate medical care.

Acknowledgment

Judith Guzman gratefully acknowledges the scholarship from CONACyT to pursue her postgraduate studies.

References

[1] S. Miyazaki, Y. Itasaka, K. Ishikawa, and K. Togawa, “Acoustic analysis of snoring and the site of airway obstruction in sleep related respiratory disorders,” Acta Otolaryngol., no. 537, pp. 47–51, 1998.

[2] J. Clerk Maxwell, A Treatise on Electricity and Magnetism, 3rd ed., vol. 2. Oxford: Clarendon, 1892, pp.68–73.

[3] J. R. Perez-Padilla, E. Slawinski, L. M. Difrancesco, R. R. Feige, J. E. Remmers, and W. A. Whitelaw, “Characteristics of the Snoring Noise in Patients with and without Occlusive Sleep Apnea,” Am. Rev. Respir. Dis., vol. 147, no. 3, pp. 635–644, Mar. 1993..

[4] H. Nakano, T. Furukawa, and S. Nishima, “Relationship between snoring sound intensity and sleepiness in patients with obstructive sleep apnea,” J. Clin. Sleep Med., vol. 4, no. 6, pp. 551–556, 2008.

[5] J. M. Marin and S. J. Carrizo, Overlap Syndromes of Sleep and Breathing Disorders, Sixth Edit. Elsevier Inc., 2017.

[6] S. E. Al-jawder and A. S. Bahammam, Obstructive Sleep Apnea, Sixth Edit., vol. 2. Elsevier Inc., 2011.

[7] V. Hoffstein, S. Mateika, and D. Anderson, “Snoring: is it in the ear of the beholder?” Sleep, vol. 17, no. 6, pp. 522–6, Sep. 1994.

[8] H. Jin et al., “Acoustic Analysis of Snoring in the Diagnosis of Obstructive Sleep Apnea Syndrome: A Call for More Rigorous Studies,” J. Clin. Sleep Med., vol. 11, no. 7, pp. 765–71, Jul. 2015.

[9] F. Dalmasso and R. Prota, “Snoring: Analysis, measurement, clinical implications and applications,” Eur. Respir. J., vol. 9, no. 1, pp. 146–159, 1996.

[10] K. Qian et al., “Classification of the Excitation Location of Snore Sounds in the Upper Airway by Acoustic Multifeature Analysis,” IEEE Trans. Biomed. Eng., vol. 64, no. 8, pp. 1731–1741, Aug. 2017.

[11] R. Beck, M. Odeh, A. Oliven, and N. Gavriely, “The acoustic properties of snores,” Eur. Respir. J., vol. 8, no. 12, pp. 2120–2128, 1995.

[12] F. Eyben et al., “The Geneva Minimalistic Acoustic Parameter Set (GeMAPS) for Voice Research and Affective Computing,” IEEE Trans. Affect. Comput., vol. 7, no. 2, pp. 190–202, Apr. 2016.

[13] S. Agrawal, P. Stone, K. Mcguinness, J. Morris, and A. E. Camilleri, “Sound frequency analysis and the site of snoring in natural and induced sleep,” Clin. Otolaryngol., vol. 27, no. 3, pp. 162–166, Jun. 2002.

[14] C. Janott et al., “Snoring classified: The Munich-Passau Snore Sound Corpus,” Comput. Biol. Med., vol. 94, no. January, pp. 106–118, 2018, doi: 10.1016/j.compbiomed.2018.01.007.

[15] K. Qian et al., “Can Machine Learning Assist Locating the Excitation of Snore Sound? A Review,” IEEE J. Biomed. Heal. Informatics, vol. 25, no. 4, pp. 1233–1246, 2021, doi: 10.1109/JBHI.2020.3012666.

[16] K. Qian, Z. Xu, H. Xu, and B. P. Ng, “Automatic detection of inspiration related snoring signals from original audio recording,” in 2014 IEEE China Summit and International Conference on Signal and Information Processing, IEEE ChinaSIP 2014 - Proceedings, 2014, pp. 95–99.

[17] M. Herzog, T. Kühnel, T. Bremert, B. Herzog, W. Hosemann, and H. Kaftan, “The impact of the microphone position on the frequency analysis of snoring sounds,” Eur. Arch. Oto-Rhino-Laryngology, vol. 266, no. 8, pp. 1315–1322, 2009.

[18] S. Agrawal, A. E. Camilleri, K. Mcguinness, and P. Stone, “Sound frequency analysis and the site of snoring in natural and sedation induced sleep,” Clin. Otolaryngol. Allied Sci., vol. 23, no. 3, pp. 280–280, Jun. 1998.

[19]] I. H. Witten, E. Frank, M. A. Hall, and C. J. Pal, “The WEKA workbench,” Data Min., pp. 553–571, Jan. 2017.

[20] S. Berg, V. Hoffstein, and T. Gislason, “Acidification of distal esophagus and sleep-related breathing disturbances,” Chest, vol. 125, no. 6, pp. 2101–2106, Jun. 2004.

[21] C. Janott et al., “Snoring classified: The Munich-Passau Snore Sound Corpus,” Comput. Biol. Med., vol. 94, no. January, pp. 106–118, 2018, doi: 10.1016/j.compbiomed.2018.01.007.

[22] K. Qian et al., “Can Machine Learning Assist Locating the Excitation of Snore Sound? A Review,” IEEE J. Biomed. Heal. Informatics, vol. 25, no. 4, pp. 1233–1246, 2021, doi: 10.1109/JBHI.2020.3012666.

[23] C. Wang, J. Peng, and X. Zhang, “A classification method related to respiratory disorder events based on acoustical analysis of snoring,” *Arch. Acoust.*, vol. 45, no. 1, pp. 141–151, 2020, doi: 10.24425/aoa.2020.132490.

[24] H. Alshaer, R. Hummel, M. Mendelson, T. Marshal, and T. D. Bradley, “Objective relationship between sleep apnea and frequency of snoring assessed by machine learning,” J. Clin. Sleep Med., vol. 15, no. 3, pp. 463–470, 2019, doi: 10.5664/jcsm.7676.