Computers and Electrical Engineering 68 (2018) 310–321

Contents lists available at ScienceDirect

Computers and Electrical Engineering

journal homepage: www.elsevier.com/locate/compeleceng

A new database of healthy and pathological voices☆

Ugo Cesari[a], Giuseppe De Pietro[b], Elio Marciano[c], Ciro Niri[d], Giovanna Sannino[*,b], Laura Verde[e]

[a] Department of Otorhinolaryngology, University Hospital (Policlinico) Federico II of Naples, Via S.Pansini, 5 Naples, Italy
[b] Institute of High Performance Computing and Networking (ICAR-CNR), Via Pietro Castellino, 111, Naples, Italy
[c] Area of Audiology, Department of Neurosciences, Reproductive and Odontostomatological Sciences, University of Naples Federico II, Via S.Pansini, 5, Naples, Italy
[d] Independent Doctor Surgeon Specialized in Audiology and Phoniatrics, Naples, Italy
[e] Department of Engineering, University of Naples Parthenope, Centro Direzionale di Napoli, Isola C4, Naples, Italy

ARTICLE INFO

Keywords:
Database
Voice disorders
Clinical study
Smart healthcare systems

ABSTRACT

In the era of Edge-of-Things computing for the accomplishment of smart healthcare systems, the availability of accurate and reliable databases is important to provide the right tools for researchers and business companies to design, develop and test new techniques, methodologies and/or algorithms to monitor or detect the patient's healthcare status. In this paper, the study and building of the VOice ICar fEDerico II (VOICED) database are presented, useful for anybody who needs voice signals in her/his research activities. It consists of 208 healthy and pathological voices collected during a clinical study performed following the guidelines of the medical SIFEL (Società Italiana di Foniatria e Logopedia) protocol and the SPIRIT (Standard Protocol Items: Recommendations for Interventional Trials) 2013 Statement. For each subject, the database contains a recording of the vowel /a/ of five seconds in length, lifestyle information, the medical diagnosis, and the results of two specific medical questionnaires.

1. Introduction

Research activities, especially related to the accomplishment of smart healthcare systems, may require databases. In order to produce high-quality research results, three critical features must be examined. Firstly, the *quality of the database*, which helps guarantee that the results are accurate and generalizable. Researchers need data that is correctly labeled and similar to the real world or originates from the real world. Secondly, the *quantity of the database*, which ensures that there is sufficient data to train and validate approaches/tools, a fact which is especially important when utilizing artificial intelligence techniques. Finally, the *availability of data*, which is critical as it allows the research to commence and ensures reproducible results, helping to improve the state of the art. In fact, the reproduction and replication of determined scientific results provides the potential to improve the efficiency of smart healthcare systems and guide the development of new techniques [1].

In the scientific community, a comparison and improvement of results is only possible if identical input data sources are used. In the last few years, the importance of available databases has also been addressed by granting organizations, governments and other agencies. For example, *"The Obama Administration is committed to the proposition that citizens deserve easy access to the results of research their tax dollars have paid for"* [2]. Additionally, as also remarked by Penrose et al. [3] *"in the scientific method it is important that results*

☆ Reviews processed and recommended for publication to the Editor-in-Chief by Guest Editor Dr. M. M. Hassan.

* Corresponding author.
E-mail addresses: ugo.cesari@unina.it (U. Cesari), giuseppe.depietro@icar.cnr.it (G. De Pietro), marciano@unina.it (E. Marciano), ciro.niri@gmail.com (C. Niri), giovanna.sannino@icar.cnr.it (G. Sannino), laura.verde@uniparthenope.it (L. Verde).

https://doi.org/10.1016/j.compeleceng.2018.04.008
Received 21 December 2017; Received in revised form 13 April 2018; Accepted 13 April 2018

be reproducible. An independent researcher should be able to repeat the experiment and achieve the same results. Most research has been done with private or irreproducible corpora generated by random searches on the WWW".

In this work, we present a new reliable database of several voice recordings useful for the design, development, testing and evaluation of the performance of new algorithms and systems for different purposes, such as for example the detection of voice disorders through an analysis of the voice signal. Additionally, the proposed database could be embedded at the Edge level in a BodyCloud system [4,5] to support smart healthcare applications. The database has been realized thanks to a clinical study that we have performed during 2016 and 2017, in which all participants have been examined by medical experts. We have collected recordings of patients, some with vocal fold disorders and some without, conditions verified after the appropriate medical examination.

This paper describes in detail the design of the study and the construction of the **VO**ice **IC**ar f**ED**erico II (VOICED) database, realised by the "Institute of High Performance Computing and Networking of the National Research Council of Italy (ICAR-CNR)" and the Hospital University of Naples "Federico II".

The rest of the paper is organized as follows. In Section 2 we report some background information useful for the reader in the introduction of the problem of dysphonia and the state of the art related to the currently available databases specific to this particular disease. Section 3 presents the methods and procedures adopted to perform the study, collect the data, acquire the voice signals and save all information of interest. A discussion of the collected data is presented in Section 4, while some conclusions are provided in Section 5.

2. Background and related work

Nowadays, the use of technology in the healthcare sphere is increasing rapidly. Technology can play a crucial role in improving and promoting health and healthcare. It can, in fact, encourage healthy behaviours to prevent and reduce pathologies, facilitate continuous health monitoring and support the early detection of specific diseases, such as dysphonia.

Dysphonia is an alteration of voice production due to a morphological or functional alteration of the pneumo-articulatory apparatus. It is a disease that affects a great number of people, with about one third of adults suffering from this disorder at least once in their lifetime [6]. Unfortunately, people often underestimate the symptoms of dysphonia not taking the appropriate countermeasures to avoid a worsening of their health. A computer-based system can be an appropriate instrument to attract the interest of people through the screening and early detection of risk factors and specific symptoms of the pathology, as well as to support the diagnosis of the disease.

The assessment of voice disorders is a multidimensional and multiparametric investigation, composed of several clinical-instrumental analyses. A team instituted by the Italian Society of Phoniatrics and Logopaedics compiled a protocol, called the SIFEL (Società Italiana di Foniatria e Logopedia) protocol [7], of basic clinical-instrumental investigations for the subjective and objective assessment of dysphonia according to the guidelines recommended by the Committee of Phoniatrics of the European Laryngological Society. The protocol recommends several examinations to diagnose dysphonia, such as laryngoscopy, an invasive analysis useful to observe the morphological and functional alterations of the vocal tract, and the logopaedic evaluation of the pneumophonic-articulatory accordance, posture and muscular tensions. Another important examination is the acoustic analysis, thanks to which it is possible to estimate the state of health of the voice evaluating characteristic parameters extracted from the voice signal, which can be indexes of possible laryngeal alterations.

Among the indicated examinations, the acoustic analysis offers several advantages due to its non-invasive nature and its potential to provide quantitative data about the clinical state of the vocal tract, with an appropriate analysis time. Smart computer-based systems, able to perform a correct acoustic analysis, can offer a valid instrument for the automatic evaluation of dysphonia. The developed algorithms and/or systems must be tested with realistic data, and for their refinement the tests must be repeated and reproducible. For this reason, the available databases are essential resources for developers and evaluators of algorithms for the analysis of physiological data.

Currently, there are few databases of voice signals existing in literature and often they are not easily available. One of the most used databases is the Massachusetts Eye and Ear Infirmary (MEEI) Voice and Speech Laboratory [8]. It contains voice recordings of subjects, healthy or pathological, with a wide variety of organic, neuralgic, traumatic and psychogenic voice disorders. Even if it is considered as the basis of many studies of voice pathology assessment [9], it has some limitations, such as for example the fact that it is not freely available for the scientific community, and that different environments and sample frequencies are used to record the healthy and pathological voices.

Another frequently used database is the Saarbruecken Voice Database (SVD) [10–12], collected by the Institute of Phonetics of the University of Saarland in collaboration with the Department of Phoniatrics and Ear, Nose and Throat (ENT) at the Caritas clinic St. Theresia in Saarbruecken. The database contains a collection of voice recordings of subjects suffering from several pathologies, including both functional and organic ones.

The main features of the different voice databases currently on the market are summarised in Table 1.

Based on these considerations, VOICED could be considered as the second freely available voice database, soon to be accessible on physionet.org, and a new element in research on automatic voice disorder detection and classification. Although the proposed database consists of a smaller sample size than the other indicated databases, it is important to highlight that this smaller size might be useful in the detection of relevant findings in the planning and interpretation of testing studies. Additionally, there is no statistical reason why a result achieved in a trial including a wider sample size should be given more credibility than a trial including a smaller sample size. However, it should be remarked here, that VOICED is the first database that contains, in addition to the voice signals and their diagnosis, also information about life habits (voice use, smoking and alcohol abuse), the patient's character and previous or

Table 1
Main features of different voice databases.

Database	# of voices (pathological + healthy)	Signal characteristics	Information included	Availability
MEEI [8]	631 (573 + 58)	25 kHz and 50 kHz 16 bit resolution	clinical diagnosis gender and age	not freely available
SVD [10]	2041 (1354 + 687)	50 kHz 16 bit resolution	clinical diagnosis, gender and age	freely available
VOICED	208 (150 + 58)	8 kHz 32 bit resolution	clinical diagnosis, gender, age lifestyle habits, VHI and RSI	freely available

concomitant diseases, that can have a relationship with voice disorders (such as gastro-esophageal reflux), and the results of two popular questionnaires used for the self-assessment of the voice and the psycho-social consequences of voice disorders, namely the Voice Handicap Index (VHI) [13] and the Reflux Symptom Index (RSI) [14].

3. Methods

The study conducted for the accomplishment of the VOICED database has been performed following a specific trial protocol developed in accordance with the SPIRIT (Standard Protocol Items: Recommendations for Interventional Trials) 2013 Statement [15], which provides guidelines on the performance of a clinical study, and in accordance with the indications of medical specialists and the instructions of the SIFEL protocol [7].

The designed protocol consists of several sections, and has been approved by the Federico II University Ethics Committee. As indicated in Table 2, it is composed of: an *Administrative information* section in which all administrative information is specified; a *Methods* section where all procedures to build the database are described; an *Ethics and Dissemination* section where plans for the ethical approval are indicated; and finally the *Appendix* section, where all documents, such as the consent and anamnestic forms, are included.

The study started on May 16, 2016 and ended on May 15, 2017. The medical phonatric examinations, the voice signal acquisitions, and the completion of the medical questionnaires were performed at the ambulatories of Phoniatrics and Videolaryngoscopy of the Hospital University of Naples “Federico II”, or at the medical room of the “Institute of High Performance Computing and Networking (ICAR-CNR)”.

The steps of the study are summarized in Fig. 1. We conducted five appropriate phases, namely:

- **Recruitment of subjects**: several strategies of recruitment were adopted to involve an adequate number of participants in the study. Recruitment difficulties are commonly encountered in clinical trials [15] and different recruitment methods can substantially influence the number and type of participants. Therefore, in this case, the recruitment was carried out by promoting our study through information campaigns, using tools such as posters, websites, brochures, well-known social networks and meetings with people to inform them about the aims of the study.
  Interested subjects had to conform with appropriate *inclusion criteria*, such as they had to be adults aged between 18 and 70 years and had to be able to follow several phases expected in the process. Subjects aged under 18 and over 70, or with diseases, such as upper respiratory tract infections or with neurological disorders, were excluded. People who met the inclusion criteria were then invited to the registration phase.
- **Registration of participants**: in the registration phase we delivered an information document to all participants. In this document the subjects were informed about the purpose of the study, the examination methods, the benefits and possible risks arising from participation, and the possibility of suspending and/or interrupting the study participation whenever they wanted.
  Successively, all subjects had to sign two forms. The first was the informed consent, by means of which they agreed to participate in the study. The second form was the authorization for the processing of personal data, essential to ensure the protection of the participant’s privacy, personal information and collected data.

Table 2
Protocol sections.

Section	Description
Administrative information	In this section all administrative information is included, such as the title of the process, and the roles and responsibilities of the contributors to the study
Methods	This section specifies all procedures to be followed, identifying the location of the process, the eligibility criteria and recruitment of the participants, and the plans for the assessment and collection of the data
Ethics and Dissemination	Plans for the ethical approval, who will obtain the informed consent and how, are indicated
Appendix	In this section all documents, such as the consent and anamnestic forms, are included.

Fig. 1. Procedures of Methods section.

- **Medical phoniatric examination**: the doctor first investigated the anamnesis of the participant, and then collected information about the participant's lifestyle, such as for example data about the use of the voice, smoking and alcohol abuse, or diet habits, and about previous or concomitant diseases that could have a relationship with voice disorders, such as for example gastro-esophageal reflux.
  Successively, the medical expert performed the laryngoscopy, one of the main diagnostic techniques used in phoniatrics [16]. This allows a visualization of the anatomical structure of the vocal folds and their possible alterations. This examination must be performed following appropriate precautions to limit inconveniences for the patients which could compromise the correct examination performance.
  Therefore, all laryngoscopies were performed according to appropriate guidelines: the patient sat down in front of the doctor with his/her chest slightly inclined forward and head moderately backwards. The doctor, after having calibrated the device and focused the camera of the laryngoscope, introduced the instrument into the patient's mouth stopping the tongue, where the laryngoscope was supported. For more sensitive patients, the laryngoscope was introduced via the nose. Successively, the patient was required to emit an extended vowel /i/, for a visualisation of the glottal plane and of its alterations.
  The laryngoscopy was performed by using the Henke Sass Wolf Laryngoscope 6.0 mm autoclavable 70° stiff model, shown in Fig. 2. The flexible model 2.8 mm was used to perform laryngoscopies via the nose, as shown in Fig. 3.
  To visualise and analyse the anatomical structure of the vocal folds and their possible alterations, all images acquired with the laryngoscopy were displayed on the doctor's monitor by using the software MedicalEDA [17].
  At the end of the medical phoniatric examination, the doctor diagnosed the presence or not of a voice disorder.

Fig. 2. The used laryngoscope 6.0 mm autoclavable 70° stiff model.

Fig. 3. The used laryngoscope 2.8 mm flexible model.

- **Voice signal acquisition**: the acquired signal consists of a recording of a vocalization of the vowel /a/ five seconds in length without any interruption of sound. All samples were recorded in a quiet (< 30 dB of background noise) and not too dry (humidity greater than 30–40%) room.
  The voice recordings were made using an appropriate m-health system [11], able to acquire in real time the voice signal by using the microphone of a mobile device. This system was installed on a Samsung Galaxy S4, Android version 5.0.1. It was held at a distance of about 20 cm from the patient at an angle of about 45°. All recordings were saved in wave format, sampled at 8000 Hz and their resolution was 32-bit. Additionally, each recording was filtered with an appropriate filter [18] to remove any noise accidentally added during the acquisition.
  The participants were instructed to articulate the vocal sample, with a constant voice intensity, as they would during a normal conversation. For each subject certain training tests were performed about two/three times before the recording. The collected data were stored in an anonymous way and identified by a univocal code.
- **Completion of two self-perception questionnaires**: the participants were requested to fill in two questionnaires, the Voice Handicap Index (VHI) [13] and the Reflux Symptom Index (RSI) [14]. The first is used to evaluate the patient's self-assessment of his/her voice and the psycho-social consequences of voice disorders. The RSI questionnaire, instead, helps to estimate the self-perception of the presence and severity of extra-esophageal reflux.
  Each participant had to fill in these questionnaires. Later they were informed about the obtained results and their interpretation.

4. Data collection and discussion

A total of 208 voice recordings were collected in the VOICED Database. In detail, there were 73 male and 135 female participants. There is a prevalence of pathological voices compared to healthy ones, the former numbering 150 (52 male and 98 female), the latter 58 (21 male and 37 female). Women are more likely than men to participate in this type of study and constitute also the category with the higher incidence of voice disorders, as demonstrated by several research works [19,20]. In Table 3 details about the number of participants, distinguishing between healthy and pathological subjects, are provided. In particular, for each category we have indicated the number and percentage of female and male subjects involved in the study, divided into age groups.

The average age of the subjects involved is about 40 years both for the women and men, as shown in Fig. 4. People with an age between 40 and 60 years represent the category of subjects that suffer the most from voice disorders, as shown in Fig. 5. In this figure the state of health is indicated as healthy or pathological. In particular, the pathological condition consists of three types of voice diseases, hyperkinetic or hypokinetic dysphonia and reflux laryngitis, pathologies described in more detail below.

All the collected data, namely the state of health as well as the anamnestic information, were provided in the proposed database. In detail, each item of database is composed of:

- the wave file, that contains the audio signal;
- the .txt and csv files, that include samples of each signal;
- the .hea and a .dat files, necessary for the PhysioBank-compatible recording. The first one contains the ID of the sample, information about the storage format and sampling frequency, and the age, gender and diagnosis of the subject. The .dat file includes the digitized samples of the signal; and
- the info.txt file, that contains all the information about the subject, which includes age, gender, diagnosis, lifestyle habits (i.e. smoking status, alcohol abuse and eating habits) and the VHI and RSI results.

Table 3
Study population.

	Female	Male	Total
Number of participants, *n*(%)	135 (64.9%)	73 (35.1%)	208 (100%)
Age, *n*(%)			
18–34 years	42 (31.1%)	18 (24.7%)	60 (28.8%)
35–49 years	44 (32.6%)	25 (34.2%)	69 (33.2%)
≥ 50 years	49 (36.3%)	30 (41.1%)	79 (38.0%)
Number of healthy voices, *n*(%)	37 (63.8%)	21 (36.2%)	58 (100%)
Age, *n*(%)			
18–34 years	22 (59.5%)	7 (33.3%)	29 (50%)
35–49 years	9 (24.3%)	8 (38.1%)	17 (29.3%)
≥ 50 years	6 (16.2%)	6 (28.6%)	12 (20.7%)
Number of pathological voices, *n*(%)	98 (65.3%)	52 (34.7%)	150 (100%)
Age, *n*(%)			
18–34 years	20 (20.4%)	11 (21.2%)	31 (20.7%)
35–49 years	35 (35.7%)	17 (32.7%)	52 (34.7%)
≥ 50 years	43 (43.9%)	24 (46.1%)	67 (44.6%)

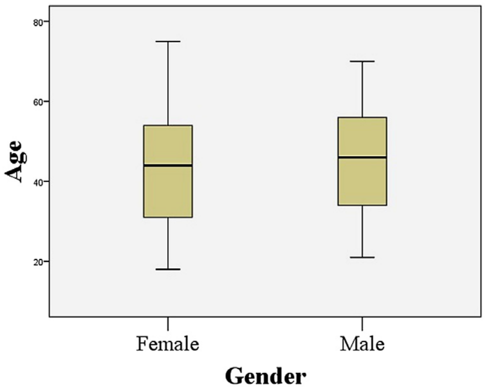

Fig. 4. Distribution of subjects by age and gender.

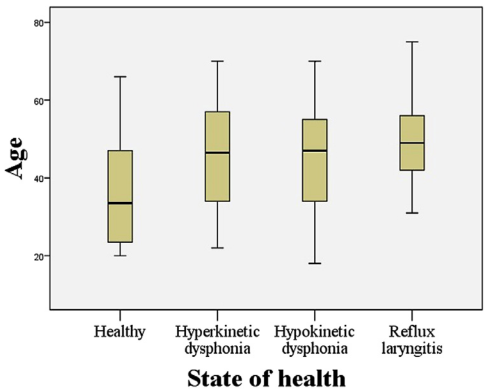

Fig. 5. Distribution of subjects by age and state of health.

In the following subsections, in addition to the description of voice pathologies, we introduce the anamnestic data collected in the study and the results obtained from the completion of the two self-assessment questionnaires.

4.1. Pathologies

The database contains recordings of voices affected by different types of pathology. These pathologies are classified into:

- **Hyperkinetic dysphonia**: this disorder is a common pathology in clinical practice, particularly among people with voice intensive occupations, characterized by muscular hypercontraction of the pneumo-phonic apparatus. The voice is striving and shrieking, the frequency modulation is reduced and the speech is characterized by a hard vocal attack. The great glottic resilience to the expiratory air stream makes the phonation more tiring and causes the alteration of the respiratory dynamics. Several diseases belong to this category, such as vocal fold nodules, Reinke's edema, chorditis, rigid vocal folds, polyps and prolapse.
- **Hypokinetic dysphonia**: hypokinetic dysphonia is characterized by a reduced adduction of the vocal folds during the respiratory cycle (especially during the inspiratory phase) that produces an air flow obstruction at the level of the larynx. The incomplete closure of the vocal folds leads to a weak and breathless voice. In the case of hypokinetic dysphonia, the voice improves with an increase in vocal intensity and this phenomenon can induce incorrect vocal abuse. Voice disorders that belong to hypokinetic dysphonia include, for example, dysphonia of the chordal groove, adduction deficit, presbiphonia, glottic insufficiency, vocal fold paralysis, conversion dysphonia, laryngitis and extraglottic air leak.
- **Reflux laryngitis**: this disorder is an inflammation of the larynx caused by stomach acid backing up in the esophagus. The most common symptom is a chronic hoarseness, but other symptoms can be more or less pronounced, such as pharyngitis, dizzy cough and/or night-time coughing, asthma, night laryngeal spasms and halitosis.

As illustrated in Table 4, the most prevalent disorder in the collected samples is hyperkinetic dysphonia followed by hypokinetic dysphonia and reflux laryngitis. The diagnosis of these disorders increases in frequency during young adulthood, people over the age of 50 being principally afflicted, while, principally, men over the age of 35 suffer from reflux laryngitis. In the hyperkinetic and hypokinetic categories several disorders were included, such as polyps or nodules for hyperkinetic dysphonia, or vocal fold paralysis or glottic insuffiency for hypokinetic dysphonia. In Figs. 6 and 7 the incidence of the different disorders is shown.

Table 4
Pathologies of the Study Population.

	Female	Male	Total
Number of hyperkinetic voices, *n*(%)	47 (67.1%)	23 (32.9%)	70 (100%)
Age, *n*(%)			
18–34 years	10 (21.3%)	7 (30.4%)	17 (24.3%)
35–49 years	16 (34.0%)	7 (30.4%)	23 (32.9%)
≥ 50 years	21 (44.7%)	9 (39.2%)	30 (42.8%)
Number of hypokinetic voices, *n*(%)	32 (78.1%)	9 (21.9%)	41 (100%)
Age, *n*(%)			
18–34 years	9 (28.1%)	2 (22.2%)	11 (26.8%)
35–49 years	10 (31.3%)	2 (22.2%)	12 (29.3%)
≥ 50 years	13 (40.6%)	5 (55.6%)	18 (43.9%)
Number of voices suffering from reflux laryngitis, *n*(%)	19 (48.7%)	20 (51.3%)	39 (100%)
Age, *n*(%)			
18–34 years	1 (5.2%)	2 (10.0%)	3 (7.7%)
35–49 years	9 (47.4%)	8 (40.0%)	17 (43.6%)
≥ 50 years	9 (47.4%)	10 (50.0%)	19 (48.7%)

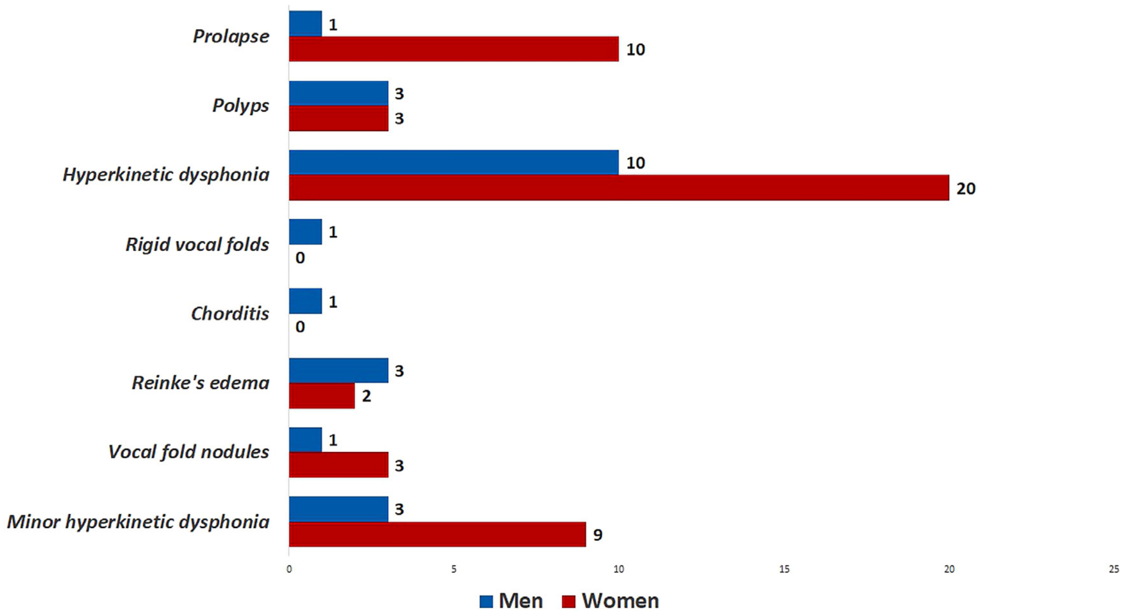

Fig. 6. Incidence of hyperkinetic disorders in the study population.

4.2. Anamnestic information

As specified in Section 3, during the medical phoniatric examination certain anamnestic information was collected to relate the patient's disorders, perceptions and habits with the onset of the pathology. Data about lifestyle habits (e.g. smoking, alcohol and coffee consummation), character (e.g. the patient is anxious, irritable or tranquil), and other pathologies that can influence voice production (e.g. gastrointestinal diseases) were recorded.

As indicated in Table 5, we have collected information about:

- **Smoking status**: smoking is one of the most important factors associated with laryngeal carcinoma and other pathologies. Tobacco smoke is considered a chronic stimulus that causes laryngeal alterations [21], producing, for example, changes in the vibrational behaviours of the vocal folds and consequent pathologies such as Reinke's edema in women and laryngitis in men.
  Participants were divided into non-smokers, hardened smokers (people who smoke more than five cigarettes every day) and casual ones (people who smoke less than five cigarettes every day). They were mainly non-smokers with a percentage of about 73% of all participants. The women non-smokers amounted to about 70% while the hardened ones to about 25%, these

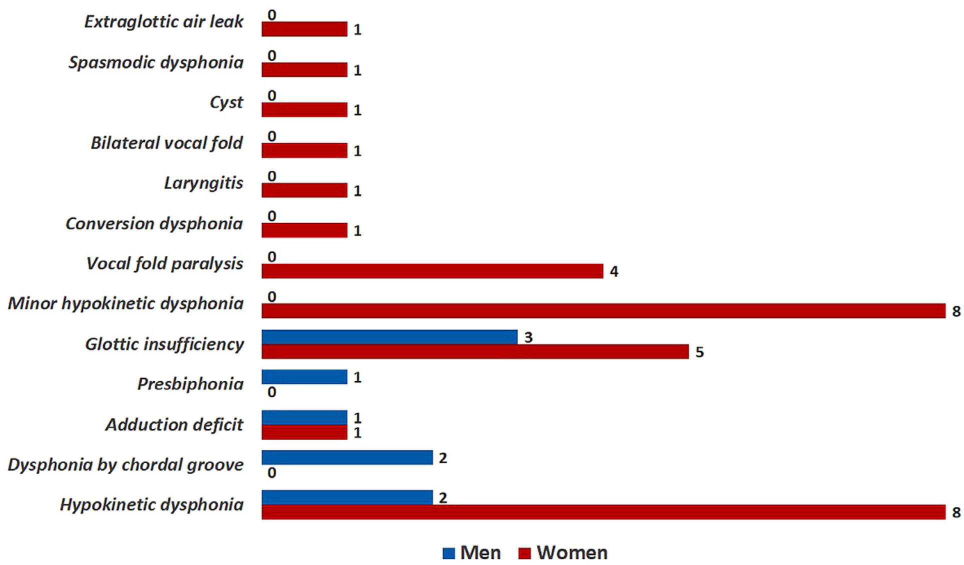

Fig. 7. Incidence of hypokinetic disorders in the study population.

participants smoking on average 11.5 cigarettes every day. The number of cigarettes smoked by the male participants was similar, the average being about 10.5, while the percentage of hardened smokers was lower than for the women (15.1%).

- **Alcohol consumption**: acute alcohol consumption can cause a decrease in the control and coordination of speech articulation, phonation and respiration [22].
  The subjects were divided into three categories. People who drink one glass of an alcoholic beverage (e.g. wine or beer) a day were defined as habitual drinkers, people who occasionally drink alcoholic beverages belong to the casual drinker category and, finally, there are non-drinkers. Most participants belonged to the casual drinker category (about 48%). The women were mainly non-drinkers, while the men were casual drinkers.
- **Hydration**: existing voice literature suggests a relationship between hydration and voice production [23]. An altered water balance can modify the structure and function of the vocal folds and voice production and influence the development of disorders. We asked the participants how many litres of water they drank every day, the result being on average of one litre. There were no great differences between the behavior of the men and women.
- **Eating habits**: Several gastrointestinal diseases, in particular gastroesophageal reflux, can cause an irritation of the vocal folds. Although in the past, reflux was primarily a disease of over-weight middle-aged people, now, it is common in thin, athletic and young people. Chemicals such as caffeine, ethanol, theobromine (in chocolate) and other substances can lead to reflux [24]. Therefore, we asked the participants about their eating habits. In detail, we collected information about the consumption of carbonated beverages, tomatoes, coffee, chocolate, soft cheeses and citrus fruits.
  We also asked the participants if they consumed a specific product during a day, either never, almost never, sometimes, almost always or always. Among the listed food and beverages coffee was the one most consumed. An average of 2.87 cups of coffee were consumed by the participants every day.

Another important factor was the occupational status of the subjects. There is a particular group of people that present a high vocal demand, these being voice professionals, subjects that use, habitually, the voice for their occupational activity. This category includes singers, teachers, telemarketing operators, lawyers and consultants. Prolonged voice use and excessive voice fatigue are the main factors that can cause voice disorders [25], increasing the risk of falling ill from dysphonia. Those afflicted can manifest several signs and symptoms, such as fatigue when speaking, hoarseness, voice instability or tremor, and loss of vocal efficiency, all of which can reduce or compromise the sufferer's social and working life.

In Fig. 8, the distribution of the study population according to occupational status is shown. In this figure, the most frequent occupational categories and those of most interest, the voice professionals, have been indicated. In the category *"other"* we have indicated people with an occupational status not related to the use of voice during their professional activity and also other less common jobs. This category includes drivers, chefs and other occupations.

4.3. Self-perception questionnaires

Finally, all participants in the study were required to complete the self-perception questionnaire. In detail, they had to compile the Voice Handicap Index and Reflux Symptom Index, as indicated in Section 3.

The first questionnaire, the VHI [13], was developed to quantify the subject's perception of his/her impairment due to voice disorders. Each subject answered the 30 questions submitted, to which a specific score depending on the severity of the symptom

Table 5
Anamnestic information.

	Female	Male	Total
Number of participants, *n*(%)	135 (64.9%)	73 (35.1%)	208 (100%)
Smoking status, *n*(%)			
Non-smokers	95 (70.4%)	57 (78.1%)	152 (73.1%)
Hardened smokers	34 (25.2%)	11 (15.1%)	45 (21.6%)
Casual smokers	6 (4.4%)	5 (6.8%)	11 (5.3%)
Alcohol consumption, *n*(%)			
Non-drinkers	64 (47.4%)	19 (26.0%)	83 (39.9%)
Habitual drinkers	10 (7.4%)	14 (19.2%)	24 (11.5%)
Casual drinkers	61 (45.2%)	40 (54.8%)	101 (48.6%)
Water consumption, *n*(%)			
0,25–1 lt	5 (3.7%)	4 (5.5%)	9 (4.3%)
1–2 lt	88 (65.2%)	48 (65.7%)	136 (65.4%)
$\geq$ 2 lt	42 (31.1%)	21 (28.8%)	63 (30.3%)
Carbonated beverages, *n*(%)			
Never	26 (19.3%)	19 (26%)	45 (21.6%)
Almost never	55 (40.7%)	21 (28.8%)	76 (36.5%)
Sometimes	43 (31.8%)	26 (35.6%)	69 (33.2%)
Almost always	9 (6.7%)	5 (6.9%)	14 (6.8%)
Always	2 (1.5%)	2 (2.7%)	4 (1.9%)
Tomatoes, *n*(%)			
Never	1 (0.7%)	2 (2.7%)	3 (1.5%)
Almost never	7 (5.2%)	7 (9.6%)	14 (6.7%)
Sometimes	97 (71.9%)	52 (71.2%)	149 (71.6%)
Almost always	29 (21.5%)	11 (15.1%)	40 (19.2%)
Always	1 (0.7%)	1 (1.4%)	2 (1.0%)
Coffee, *n*(%)			
Never	10 (7.4%)	3 (4.1%)	13 (6.3%)
Almost never	4 (3.0%)	7 (9.6%)	11 (5.3%)
Sometimes	14 (10.4%)	6 (8.2%)	20 (9.6%)
Almost always	11 (8.1%)	9 (12.3%)	20 (9.6%)
Always	96 (71.1%)	48 (65.8%)	144 (69.2%)
Chocolate, *n*(%)			
Never	2 (1.5%)	1 (1.4%)	3 (1.4%)
Almost never	23 (17.0%)	23 (31.5%)	46 (22.1%)
Sometimes	83 (61.5%)	42 (57.5%)	125 (60.1%)
Almost always	20 (14.8%)	1 (1.4%)	21 (10.1%)
Always	7 (5.2%)	6 (8.2%)	13 (6.3%)
Soft cheese, *n*(%)			
Never	3 (2.2%)	1 (1.4%)	4 (1.9%)
Almost never	17 (12.6%)	12 (16.4%)	29 (14.0%)
Sometimes	98 (72.6%)	42 (57.5%)	140 (67.3%)
Almost always	12 (8.9%)	14 (19.2%)	26 (12.5%)
Always	5 (3.7%)	4 (5.5%)	9 (4.3%)
Citrus fruits, *n*(%)			
Never	6 (4.4%)	7 (9.6%)	13 (6.3%)
Almost never	21 (15.6%)	14 (19.2%)	35 (16.8%)
Sometimes	74 (54.8%)	36 (49.3%)	110 (52.9%)
Almost always	28 (20.8%)	13 (17.8%)	41 (19.7%)
Always	6 (4.4%)	3 (4.1%)	9 (4.3%)

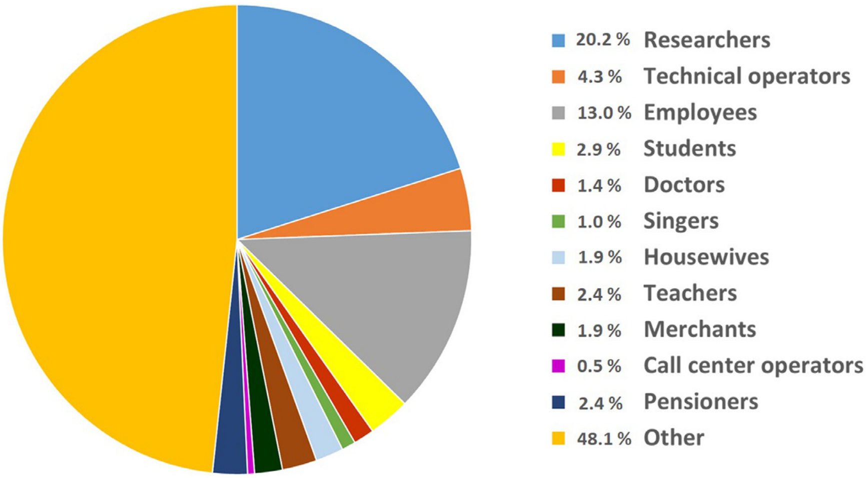

Fig. 8. Distribution of the study population according to occupational status.

corresponds. The self-perception of the disorder was classified as mild, moderate or severe. We obtained the following results:

- 155 subjects did not perceive any voice disorder;
- 19 subjects perceived a mild effect of their voice disorder on their life;
- 16 subjects perceived a moderate effect of their voice disorder on their life;
- 18 subjects perceived a severe effect of their voice disorder on their life.

The Reflux Symptom Index [14], instead, is a nine-item questionnaire for the self-evaluation of laryngopharyngeal reflux using a scale for each item from 0 (no problem) to 5 (severe problem). There were 96 subjects who perceived a laryngopharyngeal disorder, while 112 people did not discern any disturbance.

In Figs. 9 and 10 the percentages of the results obtained, respectively, in the VHI and RSI completion are shown.

In most cases the participants in the study did not perceive any disorder, neither a voice disease evaluated with the VHI questionnaire (about 75%) nor laryngopharyngeal reflux evaluated with the RSI questionnaire (about 54%). However, the number of people who discerned laryngopharyngeal reflux was greater than those who perceive an unwanted effect of their voice disorder on their life, respectively 46% and 26%.

Comparing the results obtained from the completion of these two self-evaluation questionnaires and the diagnosis clinically verified we can observe that for about 51% of the people involved in the study, the results obtained with the VHI questionnaire

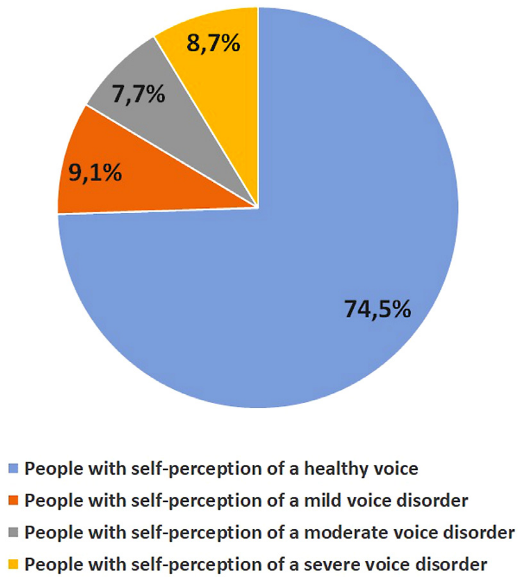

Fig. 9. Distribution of the study population according to the VHI results.

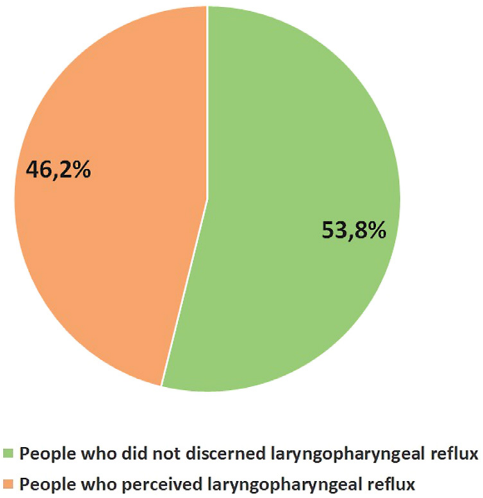

Fig. 10. Distribution of the study population according to the RSI results.

coincided with the diagnosis. For example, in 25% of cases in which a subject perceived a disorder, this disorder was clinically diagnosed. Moreover, for about 26% of people who did not perceive a voice disorder, the diagnosis of a healthy voice was confirmed by the medical examination. This percentage is lower if the results of the RSI questionnaire are considered. In this case, in fact, about 17% of people who did not perceive a laryngopharyngeal reflux disorder did not suffer from this pathology, while in 7 percent of cases in which a subject discerned a reflux disease, laryngitis reflux was indeed diagnosed. The results obtained confirm that in order to perform an accurate diagnosis of voice disorders an investigation, composed of several clinical and instrumental analyses, is necessary. Self-evaluation questionnaires, such as the VHI or RSI, cannot diagnose a pathology alone. Instead, voice assessment through other examinations, such as the acoustic analysis, is indispensable.

5. Conclusions

In recent years, the utilization of Edge-of-Things computing in smart healthcare systems has been increasing. Algorithms, methodologies, intelligent applications and smart healthcare monitoring systems can contribute to an improvement in the assessment and evaluation of appropriate pathologies, and to the remote monitoring. To develop, test and refine these techniques high-quality databases are necessary.

Due to the limited availability in the scientific community of appropriate databases composed of voice signals, a clinical study for the building of a new voice collection have been performed to make available a new database for both academia and industry.

The presented database consists of 208 voice samples. The recordings contain the vowel /a/ of five seconds in length, as required by the clinical protocol for the diagnosis of the main voice pathologies. Each voice sample was classified as healthy or pathological thanks to the diagnosis performed by the medical specialists of the Hospital University of Naples “Federico II”. This database can be considered as the second freely available voice database in the scientific literature, and a new element in research into automatic voice disorder detection and classification. It is, in fact, the first database that contains, in addition to the voice signals and their diagnosis, also information about life habits (i.e. smoking and alcohol abuse), the patient’s character, previous or concomitant diseases, that can have a relationship with voice disorders, and the results of two popular questionnaires used for the self-assessment of the voice and the psychosocial consequences of voice disorders.

References

[1] Ledet J, Teran-Somohano A, Butcher Z, Yilmaz L, Smith AE, Oğuztüzün H, Dayıbaş O, Görür BK. Toward model-driven engineering principles and practices for model replicability and experiment reproducibility. Proceedings of the symposium on theory of modeling & simulation-DEVS integrative. 27. Society for Computer Simulation International; 2014. p. 1–27.
[2] Stodden V, Leisch F, Peng RD. Implementing reproducible research. CRC Press; 2014.
[3] Penrose P, Macfarlane R, Buchanan WJ. Approaches to the classification of high entropy file fragments. Digit Invest 2013;10(4):372–84.
[4] Fortino G, Parisi D, Pirrone V, Fatta GD. Bodycloud: a saas approach for community body sensor networks. Futur Gener Comput Syst 2014;35:62–79.
[5] Gravina R, Ma C, Pace P, Aloi G, Russo W, Li W, Fortino G. Cloud-based Activity-aaService cyber–physical framework for human activity monitoring in mobility. Futur Gener Comput Syst 2017;75:158–71.
[6] Roy N, Merrill RM, Gray SD, Smith EM. Voice disorders in the general population: prevalence, risk factors, and occupational impact. Laryngoscope 2005;115(11):1988–95.
[7] Maccarini AR, Lucchini E. La valutazione soggettiva ed oggettiva della disfonia. il protocollo sifel. Acta Phoniatrica Latina 2002;24(1/2):13–42.
[8] Eye M, Voice EI, Laboratory S. Disorder database model 4337. Boston, MA.
[9] Markaki M, Stylianou Y. Voice pathology detection and discrimination based on modulation spectral features. IEEE Trans Audio Speech Lang Process 2011;19(7):1938–48.
[10] Pützer M, Koreman J. A german database of patterns of pathological vocal fold vibration. Phonus 1997;3:143–53.
[11] Verde L, De Pietro G, Veltri P, Sannino G. An m-health system for the estimation of voice disorders. Multimedia & expo workshops (ICMEW), 2015 IEEE

international conference on. IEEE; 2015. p. 1–6.
[12] Verde L, De Pietro G, Sannino G. A methodology for voice classification based on the personalized fundamental frequency estimation. Biomed Signal Process Control 2018;42:134–44.
[13] Jacobson BH, Johnson A, Grywalski C, Silbergleit A, Jacobson G, Benninger MS, Newman CW. The voice handicap index (VHI): development and validation. Am J Speech-Lang Pathol 1997;6(3):66–70.
[14] Belafsky PC, Postma GN, Koufman JA. Validity and reliability of the reflux symptom index (RSI). JVoice 2002;16(2):274–7.
[15] Chan AW, Tetzlaff JM, Gøtzsche PC, Altman DG, Mann H, Berlin JA, Dickersin K, Hróbjartsson A, Schulz KF, Parulekar WR, Krleža-Jerić K, Laupacis A, Moher D. Spirit 2013 explanation and elaboration: guidance for protocols of clinical trials. BMJ 2013;346:e7586.
[16] Paul BC, Chen S, Sridharan S, Fang Y, Amin MR, Branski RC. Diagnostic accuracy of history, laryngoscopy, and stroboscopy. Laryngoscope 2013;123(1):215–9.
[17] Medicaleda - bildsoftware. Available at: http://www.uptodent-digital.de/produkte/bildsoftware/medicaleda/.
[18] Verde L, De Pietro G, Veltri P, Sannino G. A noise-aware methodology for a mobile voice screening application. Advances in sensors and interfaces (IWASI), 2015 6th IEEE international workshop on. IEEE; 2015. p. 193–8.
[19] Lopes LW, Simões LB, da Silva JD, da Silva Evangelista D, da Norbega E Ugulino AC, Silva POC, Vieira VJD. Accuracy of acoustic analysis measurements in the evaluation of patients with different laryngeal diagnoses. J Voice 2017;31(3):382.e15–26.
[20] Coyle SM, Weinrich BD, Stemple JC. Shifts in relative prevalence of laryngeal pathology in a treatment-seeking population. J Voice 2001;15(3):424–40.
[21] Berg M, Fuchs M, Wirkner K, Loeffler M, Engel C, Berger T. The speaking voice in the general population: normative data and associations to sociodemographic and lifestyle factors. J Voice 2017;31(2):257.e13–24.
[22] Tisljár-Szabó E, Rossu R, Varga V, Pléh C. The effect of alcohol on speech production. J Psycholinguist Res 2014;43(6):737–48.
[23] Hartley NA, Thibeault SL. Systemic hydration: relating science to clinical practice in vocal health. J Voice 2014;28(5):652.e1–652.e20.
[24] Koufman JA, Wright S, Rubin J, Sataloff R, Korovin G. Laryngopharyngeal reflux and voice disorders, diagnosis and treatment of voice disorders. 2014. 419–430.
[25] Remacle A, Petitfils C, Finck C, Morsomme D. Description of patients consulting the voice clinic regarding gender, age, occupational status, and diagnosis. Eur Arch Oto-Rhino-Laryngol 2017;274(3):1567–76.

Ugo Cesari is Ear Nose & Throat Specialist and Phoniatrician as well, Adjunct Professor in otorhinolaryngology at University of Naples Federico II. Member of COMET (Collegium Medicorum Theatri), chief of Post-degree master in Physiopathology of the singing voice, expert in Laryngostroboscopy and Phonosurgery.

Giuseppe De Pietro is Director of the Institute of High Performance Computing and Networking (ICAR-CNR), and Adjunct Professor at Temple University (Philadelphia-USA). His research interests are focused on clinical Decision Support Systems, pervasive and mobile computing, and e-health. He is author of over 200 scientific papers, and he is involved in many program committees and journal editorial boards.

Elio Marciano is Director of the Area of Audiology of the Department of Neurosciences, Reproductive and Odontostomatological Sciences, and Adjunct Professor at University of Naples Federico II. He is responsible of many research projects, and author of a lot of scientific papers published in international journals.

Ciro Niri is a Phoniatrician Specialist and Audiologist as well, adjunct professor in Hearing Aid at University of Naples Federico II. Health Director of the Diocletian Rehabilitation Center in Naples, and expert in Laryngostroboscopy.

Giovanna Sannino is a researcher at the Institute of high-Performance Computing and Networking (ICAR-CNR), and Adjunct Professor of Informatics at University of Naples Federico II. Her research activities are focused on mobile Health, Signal Processing, and Artificial Intelligence for Healthcare.

Laura Verde is currently a Ph.D. student in Information Engineering at the University of Naples Parthenope, and collaborates with the Institute of High-Performance Computing and Networking (ICAR-CNR). Her research interests include Biomedical Signal Processing, m-Health Systems, and Artificial Intelligence for Healthcare.