








Communication

Speech-Based Support System to Supervise Chronic Obstructive Pulmonary Disease Patient Status

Mireia Farrús ^{1,2,*} , Joan Codina-Filbà ¹ , Elisenda Reixach ³ , Erik Andrés ³ , Mireia Sans ^{4,5} ,
Noemí Garcia ⁴  and Josep Vilaseca ^{4,5,6,7} 

- ¹ Department of Information and Communication Technologies, Universitat Pompeu Fabra, 08018 Barcelona, Spain; joan.codina@upf.edu
 - ² Language and Computation Centre, Universitat de Barcelona, 08007 Barcelona, Spain
 - ³ Fundació TIC Salut Social, Departament de Salut, Generalitat de Catalunya, 08005 Barcelona, Spain; ereixach@ticsalutsocial.cat (E.R.); eandres@ticsalutsocial.cat (E.A.)
 - ⁴ CAP Comte Borrell, Consorci d'Atenció Primària de Salut de Barcelona Esquerra (CAPSBE), 08029 Barcelona, Spain; misans@clinic.cat (M.S.); ngarcia@clinic.cat (N.G.); vilaseca@clinic.cat (J.V.)
 - ⁵ Grup de Recerca Transversal en Atenció Primària, IDIBAPS, 08036 Barcelona, Spain
 - ⁶ Universitat de Vic—Universitat Central de Catalunya, 08500 Vic, Spain
 - ⁷ Department of Medicine, Universitat de Barcelona, 08036 Barcelona, Spain
- * Correspondence: mfarrus@ub.edu or mireia.farrus@upf.edu; Tel.: +34-93-403-56-57

Featured Application: This work represents a proof of concept for COPD patient supervision, which can lead to a potential application related to home monitoring for clinicians working on the respiratory field.



Citation: Farrús, M.; Codina-Filbà, J.; Reixach, E.; Andrés, E.; Sans, M.; Garcia, N.; Vilaseca, J. Speech-Based Support System to Supervise Chronic Obstructive Pulmonary Disease Patient Status. *Appl. Sci.* **2021**, *11*, 7999. <https://doi.org/10.3390/app11177999>

Academic Editor: Michael Döllinger

Received: 6 July 2021

Accepted: 24 August 2021

Published: 29 August 2021

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2021 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

Abstract: Patients with chronic obstructive pulmonary disease (COPD) suffer from voice changes with respect to the healthy population. However, two issues remain to be studied: how long-term speech elements such as prosody are affected; and whether physical effort and medication also affect the speech of patients with COPD, and if so, how an automatic speech-based detection system of COPD measurements can be influenced by these changes. The aim of the current study is to address both issues. To this end, long read speech from COPD and control groups was recorded, and the following experiments were performed: (a) a statistical analysis over the study and control groups to analyse the effects of physical effort and medication on speech; and (b) an automatic classification experiment to analyse how different recording conditions can affect the performance of a COPD detection system. The results obtained show that speech—especially prosodic features—is affected by physical effort and inhaled medication in both groups, though in opposite ways; and that the recording condition has a relevant role when designing an automatic COPD detection system. The current work takes a step forward in the understanding of speech in patients with COPD, and in turn, in the research on its automatic detection to help professionals supervising patient status.

Keywords: chronic obstructive pulmonary disease; COPD; machine learning; prosody; speech analysis

1. Introduction

Acoustic speech features are affected in patients suffering from respiratory diseases such as chronic obstructive pulmonary disease (COPD) or asthma [1–3], especially fundamental frequency (f_0) and voice quality parameters. Some works have reported relevant correlations between f_0 and voice quality parameters with the smoking index and the forced expiratory volume in one second (FEV_1) and the asthma degree [1,2], and increased voice quality parameters or breath sounds in patients with pneumonia and COPD [4–7], so that relevant differences appear between healthy and COPD subjects in acoustic and perceptual voice parameters [8]. Voice changes are also encountered in respiratory sounds in infants [9], bronchodilator response effects in asthma tracheal sounds [10–12], obstructive sleep apnea [13], and voice rise in dyspnea [14].

Most of these studies are focused on the analysis of small segments of controlled speech, or even only on sustained sounds of vowels. However, natural speech, either read or conversational, provides information not contained in short speech segments, such as prosodic elements [15]. Further studies have dealt with the automatic detection of breathing signal during conversational speech from low-level spectral features [16], the prediction of exacerbations in high-risk patients using the COPD assessment test (CAT) [17], the identification of clinical factors that modulate the risk of progression to COPD among asthma patients using electronic medical records [18], or the prediction of non-COPD/COPD patients from natural speech by means of pauses and voice quality parameters [19]. All these works represent a step forward in the automatic assessment of COPD measurements, which can greatly benefit the task of clinicians. Nevertheless, a more fine-grained detection is needed, to see to what extent speech features can predict the exacerbations of COPD associated with impairments reflected in parameters such as FEV₁, mMRC (modified Medical Research Council) dyspnea scale [20,21], and CAT (COPD assessment test) questionnaire [22], and to what extent different recording conditions—such as doing a physical effort or medication intake—can affect the widely reported voice variations and thus the accuracy in automatic detection of COPD measurements.

The main contributions of this paper are:

1. The analysis of prosodic features in long read speech, apart from low-level acoustic features. Based on the hypothesis that prosodic features are also vulnerable to changes in COPD patients, long speech analysis is needed to extract prosodic parameters apart from mere acoustic features.
2. The analysis of how these speech features are affected by physical effort and after medication intake. Based on the hypothesis that physical effort and inhaled medication can affect oral airflow and air pressure during speech production, the speech parameters will then be modified under these conditions.
3. The analysis of to what extent COPD/non-COPD subjects can be predicted from speech, and how variations of speech that feature in different recording conditions affect the automatic prediction of COPD measurements. If COPD is manifested through speech, one should be able to predict—to a certain extent—whether an individual suffers COPD or not. Moreover, if knowledge on COPD measurements can be inferred from speech, and if speech features change due to different recording conditions, then the way in which an automatic detection system is defined from speech recordings is crucial.

2. Materials and Methods

40 COPD patients and 19 non-COPD control subjects were recruited in a primary health care centre (CAP) to collect voice samples in different conditions: in rest, after doing a physical effort, and after an inhaled medication intake (only COPD patients). From the initial number of 40 study and 19 control subjects, only valid data from 35 study (288 recordings) and 13 control individuals (124 recordings) were collected due to several reasons: insufficient number of voice samples, recordings performed incorrectly, lack of availability or interest to repeat or complete the sessions, spirometry test missing, or FEV₁ results not meeting the inclusion criteria (<0.8). Speech features were then extracted from the recordings, and they were used to perform a statistical analysis to see how they varied between different conditions, and to build an automatic detection system to predict FEV₁ values from speech to infer how such variations can affect a classification experiment.

2.1. Sampling and Study Design

This study is an observational study. Participants were selected from a patient population attending at CAP Comte Borrell in Barcelona and were divided into study and control groups. Due to the exploratory character of the pilot study, no formal calculation was performed for sample size. The study group sample size was chosen based on experimental criteria selecting a total of 40 patients. For the control group, a minimum size of 10 patients was chosen. While 50 patients are generally considered a small group, numerous studies in

the field of voice analysis have shown statistically significant correlations with a sample of 50 individuals or inferior [1,10,11,13].

2.1.1. Inclusion and Exclusion Criteria

Inclusion criteria for the study group: people >45 years with diagnosis of COPD according to FEV₁ [23]. Exclusion criteria: patients with diagnosis of COPD and home-administered oxygen therapy; patients with pathologies that affect speech; patients with pharmaceutical prescription that are known to influence speech; patients with reduced mobility.

Inclusion criteria for the control group: people >45 years old not diagnosed with COPD, with a value of FEV₁/FVC > 0.80. Exclusion criteria were applied as in the study group. Potential candidates were selected and contacted by phone by convenience sampling by family doctors in CAP Comte Borrell. Those who agreed to participate in the study gave their written consent. The ethics committee of the Hospital Clínic in Barcelona approved the study design and protocol (HCB/2018/1190).

2.1.2. Study Variables

For the study and control groups, socio-demographic variables (sex, age) and the following clinical variables related to COPD were collected: for the control group, FEV₁ before bronchodilator and smoking habits and history. For the study group, FEV₁ before/after bronchodilator, smoking habits and history, mMRC scale, CAT questionnaire, COPD exacerbations/year, and time since the last episode were also collected. FEV₁ is used to assess COPD and monitor its progression. The mMRC scale for COPD ranges from 0 to 4 grades based on the severity of dyspnea symptoms [24]. CAT questionnaire measures the impact of COPD in daily life, consisting of a list of eight questions rated from 0 to 5, so that the final sum ranges from 0 to 40 [25].

The variables of the study were gathered through the analysis of the health information system (ECAP database) or while interviewing the patient. To rule out undiagnosed COPD, the control group participants underwent a spirometry test performed by trained personnel without using the bronchodilator. During the first visit, the SpO₂ was recorded for each of the subjects before and after a test voice recording.

2.1.3. Voice Recording Conditions

All participants in the study received a voice recorder (Model EVIDA L69), a hard copy user manual, a document with instructions on conditions for recordings, a sample text consisting of two long paragraphs to be used during the recordings, and a recordings table to keep record. Several test recordings were done to ensure that participants understood the complete procedure. The participants made the recordings at home in a completely autonomous way, and they returned the recorder once finished.

Each participant in the study was requested to record ten speech recordings in different conditions and different days. In the study group, the conditions of recording were classified according to intake of medication and the physical exertion performed. Recordings were requested to be made in three different days. The recording conditions were: before using the inhaler (days 1–2–3), one hour after using the inhaler (days 1–2–3), at rest (days 1–2), and at rest and five minutes after exercise (days 1–2). Participants with no inhaler prescription were requested to perform only the recordings related to rest and exercise conditions. In the control group, the voice recordings were requested to be made in five different days, and the conditions of recording were classified for each day: at rest, and five minutes after doing exercise. The physical exercise was also adapted to the fitness of the subject.

2.2. Extraction of Speech Features

The recordings were performed with portable recorders that provided files in DVI ADPCM (digital visual interface/adaptive differential pulse-code modulation) format,

with a sampling rate of 48 kHz and a bit rate of 192 kbps. The original files were converted into PCM WAV (pulse-code modulation/waveform audio file format) for their further processing with the Praat speech analysis software [26].

f_0 , voice quality parameters such as jitter and shimmer, and pauses in COPD analysis were included following existing literature [1,19,27]. Prosodic information was also included, accounting for the variation of f_0 (range and slope), and other rhythm features such as speech rate, articulation rate, and syllable duration. For the extraction of f_0 and its related features, the auto-correlation method in Praat was used with an interval of 10 ms and a Hanning window of length 40 ms. Duration and speech rate features were extracted by adapting the Praat script found in [28], without relying on transcriptions, and thus were language independent.

2.3. Statistical Analysis

To analyse how speech was modified by physical effort and medication, their mean values and standard deviations were obtained for each feature, each subject, and each recording condition. First, a Shapiro–Wilk t test was performed to check normality in distributions, obtaining p values from 0.166 to 0.999 and W values from 0.771 to 0.999. Then, a one-way statistical paired t test with a confidence interval of 95% was performed to obtain p values and check whether differences of speech features between groups are statistically significant.

2.4. Classification Experiments

Machine learning-based classification was used to predict a class (FEV_1) given a set of data points (speech features). FEV_1 was split into intervals specified by the Global Initiative for Chronic Obstructive Lung Disease (GOLD) according to the severity degree [29], as shown in Table 1.

Table 1. Intervals of FEV_1 values according to their severity degree. GOLD classification.

Class	Degree	Range
1	mild	$FEV_1 \geq 80\%$
2	moderate	$50\% \leq FEV_1 < 80\%$
3	severe	$30\% \leq FEV_1 < 50\%$
4	very severe	$FEV_1 < 30\%$

The classification experiments were carried out in two different scenarios:

Scenario 1: To discern between non-COPD/COPD individuals (only in rest condition since non-COPD subjects are not taking medication).

Scenario 2: To discern between different COPD degrees in patients. Due to the severity degree of patients with mMRC = 4 and $FEV_1 < 30\%$, the most severe patients were excluded due to their low representation. Only subjects from GOLD 1 and 2 classes were analysed. Here, two different classification experiments were tested: (a) using only in-rest recording conditions; (b) using all recording conditions, in two different subcases: 1) experiments using only speech information, and 2) experiments providing additional information on the recording condition and subject medication.

The experiments were performed with the Weka workbench [30]. Among several classification algorithms included in Weka, Random Forest was chosen for being the best performative one in terms of accuracy. The whole set of instances were used in a 10-fold cross-validation to get the maximum number of instances in the classification task [19], ensuring that subjects used in the training folds were not present in the testing fold, and using the following hyperparameters: number of iterations (-I) 100, batch size percentage (-P) 100, seed (-S) 1, and a maximum path length of 0.

3. Results

3.1. Characteristics of The Participants

Table 2 shows the main characteristics of participants from the study and control groups.

Table 2. Characteristics of the participants.

Descriptor	Study Group (N _s = 35)	Control Group (N _c = 13)	p Value *
Age in years, average (SD)	74.6 (7.8)	67.4 (8.9)	0.01
Female, n (%)	8 (22.9)	4.0 (30.8)	0.2
FEV ₁ without BD (%), average (SD)	67.6 (20.2) ¹	90.5 (7.6)	<0.01
FEV ₁ with BD (%), average (SD) ¹	71.6 (19.8)	n/a	n/a
SPO2 (%), average (SD)	96.6 (2.0) ¹	97.8 (1.28)	0.02
Smoking habits, n (%)			
Smoker	8 (22.9)	2 (15.4)	0.05
Non smoker	27 (77.1)	11 (84.6)	n/a
Ex-smoker	25 (92.6)	7 (63.6)	0.3
Packages per year, average (SD)	71.6 (103.2) ²	9.8 (9.7)	0.01
Years of tobacco use, average (SD)	37.9 (16.5)	24.3 (12.6)	0.09
Years since quitting tobacco, average (SD) ²	17.7 (15.3)	31.9 (9.3)	0.01
mMRC, average (SD) ³	1.1 (0.9)	n/a	n/a
CAT score, average (SD) ⁴	11.3 (7.5)	n/a	n/a
Number of exacerbations/year, average (SD) ⁵	1.3 (1.1)	n/a	n/a

¹ no information in $n = 1$; ² only ex-smokers; ³ no information in $n = 2$; ⁴ no information in $n = 3$; ⁵ no information in $n = 9$; * significant differences at $p < 0.05$.

3.2. Voice Samples and Statistical Analysis

Table 3 shows the statistically significant changes in each of the speech features due to physical effort in both study (S) and control (C) groups, and Table 4 shows the statistically significant changes in each of the speech features due to medication in the study group. Corresponding FEV₁, mMRC, and CAT values are provided along with the statistical measures. In both tables, the smoking habits information has been included next to the subject ID: smoker (s), non-smoker (ns, never smoked), and ex-smoker (es). Only p values < 0.05 are shown for simplicity. Positive (+) and negative (−) changes in the mean values between rest and physical effort/medication conditions are shown after the p values. In the control and study groups, the significant differences after doing physical effort showed absolute t values ranging from $t(2) = 2.13$ to $t(2) = 3.96$, and $t(2) = 2.38$ to $t(2) = 24.10$, respectively. In the study group, the significant differences after taking medication showed t values ranging from $t(2) = 3.10$ to $t(2) = 24.51$. The percentage of subjects in which each feature exhibited a statistically significant modification is also shown. Values with strongest statistical significance ($p < 0.010$) are marked in bold.

To avoid data imbalance in scenario 1, the non-COPD instances were oversampled to equal both groups and use a baseline of 50% in accuracy. A total of 314 instances (157 for each class) were used to classify COPD/non-COPD classes from the set of eleven speech features. Then, an automatic classification using a random forest (RF) with 10-fold cross-validation in rest condition was performed, achieving an accuracy of 72.0% of COPD detection in the dataset and using only those instances recorded in rest conditions, over a baseline of 50%. Sensitivity and specificity are 70.0% and 74.5%, respectively.

Table 3. Changes in speech features due to physical effort (study and control groups).

Subject, Smoking	f0	Range f0	Slope f0	Jitter_(loc)	Jitter_(abs)	Shimmer_(loc)	Shimmer_(abs)	Pauses_Ratio	Speaking_Rate	Articulation_Rate	Syllable_Duration	preFEV ₁	postFEV ₁	mMRC	CAT
Study group															
S-1 es	0.040+		0.036–		0.042–		0.037–		0.042–			95	100	1	11
S-2 es				0.013–								63	63	2	9
S-4 s	0.045+		0.037+									64	65	1	8
S-5 es			0.036–							0.014+	0.014–	78	83	2	8
S-6 es												67	73	2	6
S-7 es												56	60	1	13
S-8 es												56	57	0	10
S-9 s	0.002+		0.026+		0.019–	0.015+	0.014+					83	88	2	16
S-10 s												30	40	2	15
S-11 es	0.011–		0.025–			0.026–	0.027–		0.026–			36	40	1	16
S-12 es							0.041–					49	64	2	5
S-13 es							0.017–					55	60	1	10
S-14 s		0.049–				0.018+	0.001+					42	57	1	19
S-16 s												83	82	1	5
S-17 es									0.032–			47	49	1	n/a
S-20 es												33	32	3	11
S-21 es												56	68	1	14
S-22 s				0.034+								77	77	2	31
S-23 es	0.002+	0.012+						0.028+	0.034–	0.013–	0.030+	40	42	1	8
S-24 es												97	103	0	8
S-25 es	0.043+											95	103	1	8
S-26 ns				0.012–								78	92	0	0
S-27 s												69	71	1	13
S-28 es												55	62	1	10
S-29 es	0.014+											n/a	n/a	3	29
S-30 es									0.045+			88	89	0	0
S-31 es												56	56	n/a	n/a
S-33 es				0.001–				0.028–				99	102	0	4
S-34 es												52	54	1	31
S-36 es	0.001+		0.048–						0.011+			73	73	2	4
S-37 es				0.004–	0.011–	0.046–	0.035–					70	83	0	3
S-39 es												86	90	0	10
S-40 s %	24.24	0.027– 9.09	18.18	15.15	9.09	15.15	21.21	6.06	18.18	6.06	6.06	83	80	0/1	9
Control group															
C-1 es												92		0	0
C-2 ns					0.035–	0.048–		0.049–		0.036–	0.035+	81		0	0
C-9 es	0.037+	0.043–						0.008+				102		0	0
C-10 s										0.038–	0.035+	87		0	0
C-11 es												84		0	0
C-12 ns												82		0	0
C-13 es												95		0	0
C-14 s	0.022+	0.046–			0.035+	0.030–	0.047–					91		0	0
C-15 ns				0.013–	0.040–							103		0	0
C-16 ns			0.044–	0.012–	0.028–							87		0	0
C-17 es	0.019+											81		0	0
C-18 es				0.001–		0.029+	0.043+					91		0	0
C-19 es		0.015+	0.005–	0.027–	0.013–			0.049+				99		0	0
%	23.08	23.08	15.38	30.77	38.46	23.08	15.38	23.08	0.00	15.38	15.38				

Table 4. Changes in speech features due to medication (study group).

Subject, Smoking	f0	Range f0	Slope f0	Jitter_(loc)	Jitter_(abs)	Shimmer_(loc)	Shimmer_(abs)	Pauses_Ratio	Speaking_Rate	Articulation_Rate	Syllable_Duration	preFEV ₁	postFEV ₁	mMRC	CAT
S-4 s		0.009–			0.036–							64	65	1	8
S-5 es									0.002+	0.003+	0.003–	78	83	2	8
S-6 es												67	73	2	6
S-7 es												56	60	1	13
S-8 es												56	57	0	10
S-9 s												83	88	2	16
S-10 s									0.021+			30	40	2	15
S-11 es												36	40	1	16
S-12 es			0.045–									49	64	2	5
S-13 es	0.021+			0.032–	<0.001–							55	60	1	10
S-14 s			0.034+									42	57	1	19
S-17 es												47	49	1	n/a
S-20 es												33	32	3	11
S-22 s		0.031–							0.040+			77	77	2	31
S-23 es	<0.001+	0.011+								0.011+	0.012–	40	42	1	8
S-27 s			0.022+	0.022+				0.019–	0.008+			69	71	1	13
S-28 es		0.003+										55	62	1	10
S-29 es					0.014–	0.027–						n/a	n/a	3	29
S-30 es												88	89	0	0
S-31 es				0.025+								56	56	n/a	n/a
S-32 ns				0.010+								102	103	0	4
S-34 es						0.018+	0.014+		0.006+	0.036+	0.042–	52	54	1	31
S-36 es												73	73	2	4
S-37 es												70	83	0	3
%	8.33	16.67	12.50	16.67	16.67	8.33	4.17	4.17	20.83	12.50	12.50				

In scenario 2, after excluding the most severe cases, the average values of FEV₁ before the bronchodilator (preFEV₁) and after the bronchodilator (postFEV₁) were computed and used to set up a threshold to define the classes in the classification experiments, ensuring the maximum balance of instances per class. In preFEV₁, class1 was assigned to instances with preFEV₁ < 75%, and class2 to instances with preFEV₁ ≥ 75%. Similarly, the threshold for postFEV₁ was fixed to 77%. In total, 222 instances were used in case (a)—rest conditions—, and 288 instances in case (b)—all recording conditions—to classify two classes for preFEV₁ and two classes for postFEV₁ from the set of speech features. In a third experiment, the classification was fed with additional information regarding which condition is related to each speech feature, and whether the patient is taking medication or not. As in scenario 1, the groups are balanced, and baseline is of 50% (Table 5). Considering class1 as negative condition and class2 the positive condition, the following measurements are shown: accuracy (percentage of correct instances), sensitivity (true positive rate), and specificity (true negative rate).

Table 5. Accuracy (Acc), sensitivity (Sens) and specificity (Spec) values for the automatic classification of FEV₁ given a reference threshold.

	preFEV ₁			postFEV ₁		
	Acc (%)	Sens (%)	Spec (%)	Acc (%)	Sens (%)	Spec (%)
baseline		50.0			50.0	
in-rest condition	62.1	61.4	62.8	59.6	59.7	59.5
all conditions (rest, effort, medication)	56.5	58.1	55.6	56.1	55.9	56.2
all conditions + additional information	75.0	81.1	71.0	73.9	77.7	71.1

4. Discussion

From the results in Tables 3 and 4, based on the experimental dataset, it can be inferred that, with regards to speech features, they all exhibit statistically significant differences between rest and physical/medication conditions, ranging from 4% to 38% (except for speaking rate) depending on the speech feature and the group of analysis, although no significant correlation is observed between the number of relevant modifications in speech features and COPD variables (FEV₁, mMRC and CAT). However, when validating the significant modification in f0 values after taking the medication by computing the correlation of such changes (i.e., f0 after medication—f0 before medication) with four different COPD indicators: the FEV₁ value without bronchodilator, the FEV₁ value with bronchodilator, the modified Medical Research Council (mMRC) scale, and the COPD assessment test (CAT), the obtained results showed that high FEV₁ values and low mMRC and CAT indicators (which are related to healthier subjects), clearly showed a tendency to smaller f0 modifications. This means that f0 modifications after medication are larger in those subjects with higher severity COPD degrees.

Regarding the effect of physical effort, in the study group, the less significant differences are observed in the prosodic features related to rhythm, and the most significant differences are in f0 and shimmer. Effort tends to be more relevant in the control group, with voice quality, f0, f0 range, and pause ratio being the most affected speech features. Regarding the effect of medication, the greatest effect is observed on the speaking rate.

The difference of the physical effort effect on both the study and control groups is significant. While a larger effect would have been expected in the COPD group due to their pulmonary and respiratory impairments, the results show the opposite. This could be explained by the fact that COPD patients might not have the same capability of making a large effort as the healthy group, and the resulting physical effort is lighter, so is their effect; or perhaps their voice parameters are always impaired, so no differences are perceived before and after exercise. A detailed look at a feature basis reveals that physical effort in general causes higher f0 values, more pauses, longer syllables, a more flattened intonation (lower values of f0 range and slope), lower voice quality features (probably due to a lower

air pressure), and lower speech rates. The results also suggest that, unlike the physical effort effect, medication produces a less flattened intonation (f_0 range and slope), and faster rhythm characteristics: higher speech rates, fewer pauses, and lower segment durations. By contrasting the results with the smoking habits of the subjects (currently smoker (s), never smoked (ns), and ex-smoker (es)), it transpired that there is no direct correspondence between the number and the degree of changes in speech features neither in the physical effort nor in the medication conditions. Even for the two smokers of the control group, their behaviour is similar, and they do not differ with respect to the non-smoking subjects.

The differences encountered can be relevant when designing a COPD prediction experiment, and the following conclusions can be inferred from Table 5. First, preFEV₁ can be detected more accurately than postFEV₁ when only the speech information extracted from the rest condition is used. This makes sense, since the *in-rest* samples have been recorded without any kind of medication, while postFEV₁ is obtained after use of the bronchodilator, which can cause changes not reflected in the corresponding speech samples. Second, by using all three conditions, the detection of both FEV₁ values decreases, since the aim is to predict the same value (preFEV₁ or postFEV₁) from speech samples that suffer a degree of variability due to the different recording conditions. However, in this case, the difference in the prediction of both preFEV₁ and postFEV₁ is not significant. Third, by using speech samples extracted from different conditions but also feeding the classifier with information about what recording condition corresponds to each speech sample and whether that patient is usually taking medication, the accuracy is highly increased. This can be explained by the fact that the system is provided with information about the source of speech variation, which is thus learned by the system. Furthermore, patients taking medication are usually associated with lower FEV₁ values, which also helps the system to discern between different COPD degrees. Last, it can be observed that sensitivity and specificity remain similar for each set of experiments, except for the last case (all conditions plus additional information) in which the rate of true positives detected (sensitivity) clearly outperforms the rate of true negatives (specificity).

5. Conclusions

The presented work explores the effect of physical exertion and medication on speech features in subjects with COPD. The effects of physical effort have been compared to those produced by a control group. The results have shown that several speech features, ranging from f_0 , voice quality, to prosodic features based on intonation and rhythm, might suffer significant changes under both physical effort and inhaled medication, and that the differences of the physical effort and medication effects are consistent in intonation and rhythm prosodic features. While effort causes a more flattened intonation and lower rhythm characteristics, medication seems to help patients to achieve more expressive speech and a higher rhythm, reflected as higher speech rates, a smaller number of pauses, and lower segment durations.

The current study also explored the performance of a classification system to discern between COPD/non-COPD subjects, and to predict FEV₁. Very few studies deal with classification of COPD variables. The results achieved up to 72% of classification accuracy over a baseline of 50%, which are comparable to—and even outperform—the ones obtained in previous studies such as [19], where COPD and non-COPD individuals were found differentiable with an accuracy of 68%. Moreover, it was shown that beyond the specific accuracy values obtained, modifications in speech feature values due to different conditions can affect the performance of a COPD automatic detection system, so that its design regarding the recording conditions is important for a successful performance.

This work represents a proof of concept for COPD patient supervision using machine learning algorithms and takes a step forward to the understanding of COPD, and to the benefits of automatic detection of COPD variables from speech, which could be useful for clinicians working in the respiratory field.

Author Contributions: M.F., J.C.-F., E.R., E.A., M.S. and N.G. conceived and designed the study. E.R. and E.A. monitored the planning and execution of the study. E.R., M.S. and N.G. obtained the ethical approval for the study. M.F., J.C.-F., E.R., E.A., M.S., J.V. and N.G. participated in the data collection. M.F., J.C.-F., E.R., E.A. and J.V. participated in the data analysis and explanation. M.F., E.R. and E.A. wrote the first draft, and J.C.-F. and J.V. helped also with the critical revision. All authors contributed to the final version, and all of them have read and agreed to the published version of the manuscript.

Funding: This research has been funded by GlaxoSmithKline, S.A. The first author has been funded by the Agencia Estatal de Investigación (AEI), Ministerio de Ciencia, Innovación y Universidades and the Fondo Social Europeo (FSE) under grant RYC-2015-17239 (AEI/FSE, UE).

Institutional Review Board Statement: The ethics committee of the Hospital Clínic in Barcelona approved the study design and protocol (HCB/2018/1190).

Informed Consent Statement: Informed consent was obtained from all subjects involved in the study.

Data Availability Statement: The data presented in this study are available on request from the corresponding author. The data are not publicly available due to privacy and ethical issues.

Acknowledgments: The authors would like to thank all the participants that took part in this study and the following professionals from CAPSBE: Amparo Hervas Docon, Cristina Colungo Francia, Joan Clos Soldevila, Núria Sánchez Ruano, Emma Magraner Oliver, Anna Peña Sanromà, and Laia Gené Huguet.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Mohamed, E.E.; El Maghraby, R.A. Voice changes in patients with chronic obstructive pulmonary disease. *Egypt. J. Chest Dis. Tuberc.* **2014**, *63*, 561–567. [[CrossRef](#)]
2. Kutor, J.; Balapangu, S.; Adofo, J.K.; Dellor, A.A.; Nyakpo, C.; Brown, G.A. Speech signal analysis as an alternative to spirometry in asthma diagnosis: Investigating the linear and polynomial correlation coefficient. *Int. J. Speech Technol.* **2019**, *22*, 611–620. [[CrossRef](#)]
3. Saeed, A.M.; Riad, N.M.; Osman, N.M.; Khattab, A.N.; Mohammed, S.E. Study of voice disorders in patients with bronchial asthma and chronic obstructive pulmonary disease. *Egypt. J. Bronchol.* **2018**, *12*, 20–26. [[CrossRef](#)]
4. Neili, Z.; Fezari, M.; Abdeghani, R. Analysis of Acoustic Parameters from Respiratory Signal in COPD and Pneumonia patients. In Proceedings of the 2018 International Conference on Signal, Image, Vision and their Applications (SIVA), Guelma, Algeria, 26–27 November 2018; Institute of Electrical and Electronics Engineers (IEEE): New York, NY, USA, 2018; pp. 1–4.
5. Hashemi, A.; Arabalibiek, H.; Agin, K. Classification of Wheeze Sounds Using Wavelets and Neural Networks. In *International Conference on Biomedical Engineering and Technology*; IACSIT Press: Singapore, 2011.
6. Taplidou, S.A.; Hadjileontiadis, L.J. Analysis of Wheezes Using Wavelet Higher Order Spectral Features. *IEEE Trans. Biomed. Eng.* **2010**, *57*, 1596–1610. [[CrossRef](#)]
7. Song, I. Diagnosis of pneumonia from sounds collected using low cost cell phones. In Proceedings of the 2015 International Joint Conference on Neural Networks (IJCNN), Killarney, Ireland, 12–17 July 2015; Institute of Electrical and Electronics Engineers (IEEE): New York, NY, USA, 2015; pp. 1–8.
8. Shastry, A.; Balasubramaniam, R.K.; Acharya, P.R. Voice Analysis in Individuals with Chronic Obstructive Pulmonary Disease. *Int. J. Phonosurg. Laryngol.* **2014**, *4*, 45–49. [[CrossRef](#)]
9. Elphick, H.E.; Lancaster, G.; Solis, A.; Majumdar, A.; Gupta, R.; Smyth, R.L. Validity and reliability of acoustic analysis of respiratory sounds in infants. *Arch. Dis. Child.* **2004**, *89*, 1059–1063. [[CrossRef](#)] [[PubMed](#)]
10. Fiz, J.A.; Jané, R.; Salvatella, D.; Izquierdo, J.; Lores, L.; Caminal, P.; Morera, J. Analysis of Tracheal Sounds During Forced Exhalation in Asthma Patients and Normal Subjects. *Chest* **1999**, *116*, 633–638. [[CrossRef](#)] [[PubMed](#)]
11. Bhalla, R.K.; Watson, G.; Taylor, W.; Jones, A.S.; Roland, N.J. Acoustic Analysis in Asthmatics and the Influence of Inhaled Corticosteroid Therapy. *J. Voice* **2009**, *23*, 505–511. [[CrossRef](#)] [[PubMed](#)]
12. Dogan, M.; Eryuksel, E.; Kocak, I.; Celikel, T.; Sehitoglu, M.A. Subjective and Objective Evaluation of Voice Quality in Patients With Asthma. *J. Voice* **2007**, *21*, 224–230. [[CrossRef](#)]
13. Fiz, J.A.; Morera, J.; Abad, J.; Belsunces, A.; Haro, M.; Caminal, P.; Rodenstein, D. Acoustic Analysis of Vowel Emission in Obstructive Sleep Apnea. *Chest* **1993**, *104*, 1093–1096. [[CrossRef](#)]
14. Binazzi, B.; Lanini, B.; Romagnoli, I.; Garuglieri, S.; Stendardi, L.; Bianchi, R.; Gigliotti, F.; Scano, G. Dyspnea during Speech in Chronic Obstructive Pulmonary Disease Patients: Effects of Pulmonary Rehabilitation. *Respiration* **2011**, *81*, 379–385. [[CrossRef](#)] [[PubMed](#)]
15. Farrús, M. Fusing prosodic and acoustic information for speaker recognition. *Int. J. Speech Lang. Law* **2009**, *16*, 169–171. [[CrossRef](#)]
16. Nallanthighal, V.S.; Härmä, A.; Strik, H. Deep Sensing of Breathing Signal During Conversational Speech. *Interspeech 2019* **2019**. [[CrossRef](#)]

17. Lee, S.-D.; Huang, M.-S.; Kang, J.; Lin, C.-H.; Park, M.J.; Oh, Y.-M.; Kwon, N.; Jones, P.W.; Sajkov, D. The COPD assessment test (CAT) assists prediction of COPD exacerbations in high-risk patients. *Respir. Med.* **2014**, *108*, 600–608. [[CrossRef](#)]
18. Himes, B.E.; Dai, Y.; Kohane, I.S.; Weiss, S.; Ramoni, M.F. Prediction of Chronic Obstructive Pulmonary Disease (COPD) in Asthma Patients Using Electronic Medical Records. *J. Am. Med. Inform. Assoc.* **2009**, *16*, 371–379. [[CrossRef](#)] [[PubMed](#)]
19. Nathan, V.; Vatanparvar, K.; Rahman, M.; Nemati, E.; Kuang, J. Assessment of Chronic Pulmonary Disease Patients Using Biomarkers from Natural Speech Recorded by Mobile Devices. In Proceedings of the 2019 IEEE 16th International Conference on Wearable and Implantable Body Sensor Networks (BSN), Chicago, IL, USA, 19–22 May 2019; Institute of Electrical and Electronics Engineers (IEEE): New York, NY, USA, 2019; pp. 1–4.
20. Mahler, D.A.; Wells, C.K. Evaluation of Clinical Methods for Rating Dyspnea. *Chest* **1988**, *93*, 580–586. [[CrossRef](#)]
21. Hajiro, T.; Nishimura, K.; Tsukino, M.; Ikeda, A.; Koyama, H.; Izumi, T. Analysis of clinical methods used to evaluate dyspnea in patients with chronic obstructive pulmonary disease. *Am. J. Respir. Crit. Care Med.* **1998**, *158*, 1185–1189. [[CrossRef](#)] [[PubMed](#)]
22. Jones, P.W.; Harding, G.; Berry, P.; Wiklund, I.; Chen, W.H.; Leidy, N.K. Development and first validation of the COPD Assessment Test. *Eur. Respir. J.* **2009**, *34*, 648–654. [[CrossRef](#)]
23. Spector, N.; Connolly, M.A.; Carlson, K.K. Dyspnea: Applying Research to Bedside Practice. *AACN Adv. Crit. Care* **2007**, *18*, 45–60. [[CrossRef](#)] [[PubMed](#)]
24. MMRC Dyspnea Severity Scale: MediCalc®. Available online: <http://www.scymed.com/en/smnxpr/prwck220.htm> (accessed on 26 May 2020).
25. COPD Assessment Test (CAT)—MDCalc'. Available online: <https://www.mdcalc.com/copd-assessment-test-cat> (accessed on 26 May 2020).
26. Boersma, P. Praat, a system for doing phonetics by computer. *Glott Int.* **2002**, *5*, 341–345.
27. Farrús, M. Voice disguise in automatic speaker recognition. *ACM Comput. Surv.* **2018**, *51*, 1–22. [[CrossRef](#)]
28. De Jong, N.H.; Wempe, T. Praat script to detect syllable nuclei and measure speech rate automatically. *Behav. Res. Methods* **2009**, *41*, 385–390. [[CrossRef](#)] [[PubMed](#)]
29. Guía de Bolsillo para el Diagnóstico, Manejo y Prevención de la EPOC. *Una Guía Para Profesionales de la Asistencia Sanitaria*; Global Initiative for Chronic Obstructive Lung Disease (GOLD) Inc., 2017. Available online: <https://goldcopd.org/wp-content/uploads/2016/04/wms-spanish-Pocket-Guide-GOLD-2017.pdf> (accessed on 26 August 2021).
30. Frank, E.; Hall, M.A.; Witten, I.H.; Kaufmann, M. *WEKA Workbench Online Appendix for “Data Mining: Practical Machine Learning Tools and Techniques”*; Morgan Kaufmann: Burlington, MA, USA, 2016.