Preprints.org

Speech Rate as a Potential Biomarker of Respiratory Health: Evidence from Mandarin Repetitive Articulation Task

Xiaoyu Zhou , Biao Zeng [*] , Fasih Haider , Qianqian Meng , Mengyang Pang , Saturnino Luz

Article

Speech Rate as a Potential Biomarker of Respiratory Health: Evidence from Mandarin Repetitive Articulation Task

Xiaoyu Zhou [1], Biao Zeng [2,*], Fasih Haider [3], Qianqian Meng [4], Mengyang Pang [5] and Saturnino Luz [6]

[1] Department of Respiratory and Critical Care Medicine, The Frist Affiliated Hospital of Bengbu Medical University, Anhui, China

[2] Applied Psychology Research and Innovation Group, Department of Psychology, University of South Wales, UK

[3] Usher Institute, Edinburgh Medical School, the University of Edinburgh, UK

[4] Department of Respiratory and Critical Care Medicine, The Municipal Hospital Affiliated to Xuzhou Medical University, Jiangsu, China

[5] Bengbu Medical University, Anhui, China

[6] Usher Institute, Edinburgh Medical School, the University of Edinburgh, UK

* Correspondence: biao.zeng@southwales.ac.uk

Abstract

This study investigates the relationship between speech rate and respiratory function using a repetitive articulation task in Mandarin Chinese, validating the cross-linguistic applicability of speech-based respiratory assessments. Seventy-four native Mandarin speakers performed a 30-second articulation of the phrase "*pīng pāng qiú*," (table tennis or ping pong in Mandarin) during which speech and breathing metrics were collected and analyzed alongside pulmonary function test results. ANOVA results revealed significant effects of breathing group and articulation run on speech rate, while regression analyses demonstrated that Peak Expiratory Flow (PEF) showed the strongest correlation with speech rate among all respiratory parameters. Unlike previous studies, no significant sex differences were observed in speech rate or breathing patterns, which may be due to the analysis being limited to the first two breathing groups. In addition, machine learning regression models were developed for predicting respiratory indices. Random Forest and AdaBoost, outperformed linear statistical methods in this task, highlighting the potential of machine learning in non-invasive respiratory health monitoring. These findings underscore the utility of speech rate as a proxy for expiratory strength and support the integration of AI and language-specific tasks in speech-based respiratory diagnostics.

Keywords: speech breathing; PEF; speech rate; Mandarin; AI; respiratory health; machine learning; helicopter task

1. Introduction

Research of speech breathing, the specialized pattern of respiration that supports spoken language, has significant implications for evaluating lung function. Traditional pulmonary assessments, such as spirometry, require active patient participation, making them challenging for individuals with respiratory impairments. Speech-based assessments provide a promising alternative by analysing natural speech breathing patterns to infer lung function [1]. Previous studies have shown that speech breathing metrics such as pause ratio, breath group length, and vocal intensity can serve as indirect markers of respiratory efficiency [2,3].

Additionally, speech breathing analysis can be applied to neurological conditions, such as Parkinson's disease and post-stroke dysarthria, to assess respiratory-linguistic coordination and track rehabilitation progress. Speech breathing analysis aids in assessing and tracking rehabilitation in neurological conditions like Parkinson's disease and post-stroke dysarthria. Solomon and Hixon [4] found altered speech breathing in Parkinson's patients due to rib cage rigidity. Arnold et al. [5] demonstrated improved breath support and speech intelligibility in dysarthric patients through respiratory muscle training.

1.1. Breathing and Speech Breathing

Breathing, or ventilation, is the physiological process of moving air into and out of the lungs to facilitate gas exchange with the bloodstream. It consists of two primary phases: inhalation, where oxygen enters the lungs, and exhalation, where carbon dioxide is expelled. The regulation of breathing is primarily controlled by the respiratory centers in the brainstem, responding to fluctuations in carbon dioxide levels, oxygen demand, and metabolic requirements [6]. The diaphragm, intercostal muscles, and accessory muscles work in coordination to expand and contract the thoracic cavity, driving airflow efficiently [7].

The respiratory system comprises the upper (nose, nasal cavity, pharynx, larynx) and lower (trachea, bronchi, lungs, alveoli) respiratory tracts. It plays a fundamental role in gas exchange, allowing oxygen to enter the bloodstream while eliminating carbon dioxide. This exchange occurs at the alveolar-capillary interface and is regulated by central and peripheral chemoreceptors that detect fluctuations in blood oxygen, carbon dioxide, and pH levels [6]. The key processes involved in respiration include pulmonary ventilation, external respiration (gas exchange at the alveoli), transport of gases via the circulatory system, and internal respiration at the cellular level [8].

Speech breathing is a specialized adaptation of the respiratory cycle that supports phonation and spoken communication. Unlike quiet breathing, which primarily serves metabolic demands, speech breathing requires voluntary control to regulate airflow for speech production. This is characterized by a prolonged exhalation phase, during which air is released steadily to maintain airflow for phonation [9]. Additionally, speech breathing requires a higher subglottal pressure to support voiced sound production and precise timing of inspiratory pauses, which align with linguistic and prosodic structures rather than metabolic needs [10,11].

The interplay between autonomic breathing, the respiratory system, and speech breathing is crucial for effective speech production. While autonomic breathing is controlled by metabolic requirements, speech breathing is governed by cortical mechanisms that enable voluntary control over respiratory patterns. Neural mechanisms involved in speech breathing include central pattern generators (CPGs) in the brainstem, which regulate rhythmic respiration, as well as cortical control for voluntary breath adjustments during speech [12]. Furthermore, laryngeal valving mechanisms adjust subglottal pressure, while articulatory structures such as the tongue, lips, and jaw synchronize with respiratory patterns to ensure clear and intelligible speech [13].

Understanding speech breathing is essential for assessing respiratory health, especially in individuals with conditions affecting speech-respiratory coordination. Disorders such as chronic obstructive pulmonary disease (COPD), Parkinson's disease, and dysarthria can significantly impair speech breathing, influencing both phonatory control and speech intelligibility. Research highlights the impact of neurodegenerative diseases on speech breathing and their implications for clinical diagnosis and treatment approaches [14–16].

1.2. Articulation Tasks in Evaluating Respiratory Functions

Speech production and articulation tasks have been widely utilized in assessing respiratory function, as they require precise coordination between the respiratory, phonatory, and articulatory systems.

1.2.1. Types of Speech Breathing Tasks

Maximum phonation time (MPT) measures how long an individual can sustain a vowel sound (e.g., /a/, /i/, /u/) on a single breath. It assesses respiratory support and phonatory efficiency, which can be affected in conditions such as chronic obstructive pulmonary disease (COPD) and neuromuscular disorders [17–20]. Furthermore, the S/Z ratio is a comparative measure of the maximum duration an individual can sustain the voiceless /s/ and voiced /z/ sounds. A higher-than-normal ratio may indicate inefficient glottal closure and respiratory insufficiency [21,22]. Diadochokinesis (DDK) is another widely used method and measures the rate at which a person can repetitively produce syllables like /pa-ta-ka/. It is often used to assess respiratory control and articulation coordination in speech disorders, particularly in patients with Parkinson's disease or amyotrophic lateral sclerosis (ALS) [23–27].

Beyond of syllable level tasks, reading and passage tasks require patients to read a standard passage aloud, such as the "Rainbow Passage" [28]. Respiratory efficiency is analysed by observing breath groups, pause placement, and phonatory effort [29–31]. Counting tasks involve counting as high as possible in a single breath or timed counting tasks, which help assess speech breathing patterns and respiratory endurance [32–34]. Spontaneous speech assessment includes free speech tasks, such as describing a picture or narrating a story, which allow clinicians to analyse respiratory control during natural speech production, detecting abnormalities in breath support [35]. Sustained vowel and consonant production [36] involves instructing patients to hold out specific vowels or consonants at different intensities. The variation in duration and intensity can provide insights into pulmonary function and voice control.

1.2.2. Measures of Speech Rate: Phoneme, Syllable and Word

Speech rate can be measured at different linguistic levels, including phonemes, syllables, and words. Each method has distinct advantages and limitations. Calculating speech rate by phonemes allows for precise articulation analysis, making it useful for phonetic studies and motor speech disorder assessments; however, it requires transcription and lacks standardization across languages [37,38]. Alternatively, a syllable nuclei detection algorithm could be employed which offer a reasonably accurate proxy syllable-based speech rate estimation [39]. Syllable-based speech rate offers a balanced approach, as syllables are more consistent linguistic units across languages and are commonly used in fluency and respiratory studies; nonetheless, difficulties in defining syllable boundaries can introduce variability [29,40,41]. Word-based speech rate, on the other hand, is the most intuitive and clinically applicable method, particularly in evaluating speech efficiency in discourse and spontaneous speech; however, it is less reliable due to variations in word or sentence length and structure [42–44]. To counter such variations, a method based on normalising the time a speaker takes to complete an utterance based on the time taken by a speech synthesizer set at a fixed rate has been employed in applications of speech analysis in Alzheimer's disease patients [45]. The selection of an appropriate measurement method depends on the specific application, whether it is for clinical speech evaluation, research, or respiratory function analysis.

1.3. Speech Rate in Health and Patient Groups

Research indicates that normal spirometry values—such as an FEV_1/FVC ratio greater than 0.70 and both forced expiratory volume in 1s (FEV_1) and forced vital capacity (FVC) above 80% of predicted values—facilitate regular speech patterns. In healthy individuals, efficient respiratory mechanics support normal speech production. Studies have shown that adequate lung volumes and airflow rates are essential for maintaining typical speech rates. Additionally, studies exploring the effects of speaking rate on breathing and voice behaviour have demonstrated that variations in speech rate can influence respiratory patterns, further highlighting the interplay between respiratory function and speech production [29,46,47].

While FVC and FEV1 measure overall lung volume and expiratory force over time, COPD is characterized by airflow limitation due to obstructive lung disease, leading to reduced FEV_1 and FEV_1/FVC ratios. This impairment often results in decreased speech rates, shorter breath groups, and more frequent pauses as patients may experience dyspnea during speech, limiting their ability to produce continuous phrases. The severity of COPD, classified based on FEV_1 percentages, correlates with the extent of speech disruption [48,49]. Furthermore, research has shown that COPD-induced changes subtly affect breathing-related functions, such as speech articulation and the pause time between words, indicating a direct relationship between respiratory impairment and speech characteristics [48,50].

Peak Expiratory Flow (PEF) specifically reflects the maximum flow rate achieved during a forceful exhalation and is particularly sensitive to upper airway resistance. PEF is closely linked to speech production, as effective phonation requires sufficient airflow. In asthma, another obstructive lung disease, reduced PEF during exacerbations correlates with disrupted speech breathing and slower, fragmented speech due to bronchoconstriction and increased airway resistance. Asthma involves reversible airway obstruction, leading to variable reductions in FEV1 and PEF [51,52]. During asthma exacerbations, patients may exhibit slower speech rates or fragmented speech due to bronchoconstriction and increased airway resistance. Monitoring PEF can help assess the severity of asthma and its impact on speech [53]. Additionally, studies have found that respiratory differences in spontaneous and scripted speech among individuals with asthma can influence speech breathing patterns, further affecting speech rate and fluency [54,55].

Quantitative assessments of respiratory functions, such as FVC, FEV_1, and PEF, are crucial for understanding their impact on speech rate. In healthy individuals, normal respiratory parameters support typical speech patterns. Conversely, in conditions like COPD and asthma, impaired respiratory function can lead to reduced speech rates, highlighting the importance of respiratory management in maintaining effective communication.

Overall, the interplay between PEF and speech rate emphasizes the impact of pulmonary function on communicative effectiveness and the clinical value of integrating speech-based metrics in respiratory assessment.

1.4. AI and Speech Breathing

Recent advances in speech-based analysis have enabled the use of deep learning to estimate respiratory parameters from voice recordings. These models can detect subtle changes in breathing patterns and infer limitations that impact speech fluency [56–60]. A wide range of speech features have been investigated in relation to health status. Farrús et al. categorized speech features into acoustic and prosodic information, with prosodic aspects—intonation, stress, and rhythm—being particularly relevant for assessing emotional states and disorders such as bipolar disorder [61]. Zeng et al. proposed a three-tier analysis of speech breathing that includes acoustic, linguistics, and breathing features. This framework examines vowel formants, intensity, and pitch (acoustic), speech rhythm and pause patterns (linguistics), and respiratory rate (breathing) to assess respiratory health through speech [1]. Recent advances in deep learning have further expanded the scope of speech analysis for health applications, enabling automated feature extraction and disease classification [62]. For example, Nallanthighal et al. explored breathing analysis from speech, emphasizing its importance in speech production and potential for respiratory health assessment [56].

1.4.1. Physiological and Psychological Stress Detection

Research has demonstrated that speech-based biomarkers can reflect physiological stress, making speech analysis a valuable tool for stress monitoring. A study examined acoustic and prosodic speech features and their relationship with stress markers, highlighting the correlation between speech characteristics and physiological responses [63]. Similarly, Cummins and his colleagues reviewed the state-of-the-art deep learning approaches for detecting health conditions

based on speech, showcasing the increasing role of artificial intelligence in speech-based diagnostics [62].

1.4.2. Deep Learning for Speech-Based Health Monitoring

Several studies have focused on breathing patterns in speech to estimate respiratory health status. One study proposed a deep-learning model for detecting respiratory pathology through audio analysis [64]. Another research explored deep learning architectures for estimating respiratory signals from speech recordings, demonstrating the feasibility of extracting breathing metrics non-invasively [56]. Recent deep learning techniques have further improved lung sound analysis for intelligent stethoscopes, enhancing the identification of abnormal respiratory patterns [65]. Another study compared different machine learning models for respiratory estimation from conversational speech, reinforcing the possibility of monitoring respiratory health using only speech data [66].

Speech markers have also been investigated for detecting neurodegenerative diseases. A study on machine learning and speech acoustics found that specific speech features could distinguish patients with neurodegenerative conditions from healthy individuals with high accuracy [67]. Similarly, another study developed a non-invasive cognitive impairment detection system based on voice acoustic features, aiming to address dementia detection challenges [68].

1.5. Rationale and Aims of the Research

In summary, the intricate relationship between speech breathing, speech rate, and respiratory underscores the complexity of speech production. Zeng et al. proposed three-tier framework and provided a structured approach to analyse these interactions [1]. In this study, we created one equivalent task called " *pīng pāng qiú*" (Mandarin Chinese for 'table tennis'). Additionally, the potential applications of speech breathing studies in clinical settings highlight the importance of continued research in this area.

This study investigates the relationship between speech breathing and speech production in Mandarin Chinese, with a specific focus on speech breathing tasks, speech rate, and the influence of respiratory function. The research is guided by the following questions:

1. Do the order of breathing runs (e.g., first run vs. second run) and breathing groups affect the speech rate during articulation of "*pīng pāng qiú*"?
2. Is there a correlation between speech rate and lung function measures, particularly PEF ?

2. Materials and Methods

2.1. Participants

A total of 74 healthy, native Mandarin Chinese-speaking participants (34 men, 40 women; mean age: 22.5 years, range: 21–40 years; height:1.69 ± 0.08m; weight: 64.91 ± 13.33kg) were recruited from Bengbu Medical University through random sampling. All participants completed a Clinical Report Form, which included information such as age, sex, height, weight and pulmonary function status. No other general health status, medication use, or physical activity parameters were investigated.

2.2. Apparatus

An Apple iPhone 15 installed with the AVR APP voice application was used for speech recording, and the Praat software [69] was used for analysis. The pulmonary function test was conducted using the Jaeger MS-DIFFUSION Spirometer (Germany).

2.3. Procedure

Each participant was instructed to sit quietly in a room at approximately 10 cm from the microphone of the mobile phone. The task was run with requiring participants to repeat the Chinese word "*pīng pāng qiú*" (Mandarin Chinese for 'table tennis') as quickly and as many times as possible

within a 30-second period. Each participant performed three 30-second runs, with a mandatory 20-second rest period between runs. The experimenter gave a signal at the start of each task and at the start of each rest period. The instructions were as follows (given in Mandarin Chinese):

> *"For this task, please repeat the word 'pīng pāng qiú' as quickly and as many times as possible, each time for 30 seconds. You will perform this task three times, with a short 20-second rest between each session. When I say 'start,' please begin repeating 'pīng pāng qiú' until I say 'stop.' After a 20-second rest, I will say 'start,' and you should continue repeating the word until I say 'stop.' After another rest, I will ask you to do it one more time. I will time and record the entire procedure, so please sit quietly during the rest periods."*

> *"Do you have any questions?"*

> *"Are you ready to begin?"*

The recordings were collected using the AVR APP. The audio recordings were in MP3 format and were later processed using Praat software. The three 30-second recordings were extracted as separate files for acoustic analysis. Only instances of the word produced in full were kept for analysis; instances with disfluencies, prominent background noise, or overlap with the experimenter's instructions were excluded from the acoustic analysis. Audible breaths (inspirations) within each recording were identified and recorded as named pauses, allowing the counting and duration calculation for each breath. The primary data collected included the number and duration of the first and second breaths, as well as the pause time (inspiratory time) between the first and second breaths. Word boundaries and pause boundaries in all audio files were manually segmented by the first annotator, and a second trained phonetician verified and corrected the annotations.

On the same day as the voice data collection, each participant underwent pulmonary function testing using the Jaeger Master Screen PFT System (Germany). Detailed pulmonary function parameters were recorded, including FEV1, FVC, FEV1/FVC, PEF, mean expiratory flow at 75%, 50% and 25% of vital capacity (MEF75, MEF50, MEF25) and maximum voluntary ventilation (MVV).

3. Results

This section presents the findings of the study based on the analytical framework proposed [1]. The data were analysed using SPSS (IBM Corp., 2025), and results are reported in accordance with this framework. First, descriptive statistics, including means and standard deviations, provide an overview of the dataset. Second, ANOVA tests assess the significance of relationships and differences among variables. Finally, various regression analysis explores the interplay of speech breathing, speech production and respiratory system, offering further insights into the underlying patterns. Table 1 and 2 illustrate key descriptive statistics, which will be further interpreted in the Discussion section.

Table 1. Descriptive statistics of age, BMI and pulmonary function parameters (N=74).

Parameters	Mean	SD
Age	22.49	3.52
Height	1.69	0.08
Weight	64.91	13.42
BMI	22.75	3.91
FVC	4.15	0.93
FEV1	3.48	0.72
FF	84.38	6.62
PEF	7.47	1.93

MEF75	6.57	1.58
MEF50	4.26	1.13
MEF25	1.76	0.63
MVV	121.54	30.56

Table 2. Mean and SD of breathing duration, pause duration, number of words and speech rate in one breathing group.

Run	Group	Group Duration	Pause Duration	Speech Rate
1st	1st	7.41 (2.75)	.69 (.46)	2.17 (.28)
	2nd	5.14 (2.68)		2.04 (.25)
2nd	1st	7.50 (2.83)	.67 (.38)	2.06 (.25)
	2nd	5.16 (2.88)		1.92 (.29)
3rd	1st	7.45 (2.96)	.73 (.42)	2.03 (.29)
	2nd	5.26 (2.67)		1.85 (.32)

*each articulation run is 30-second long and consists of various inhalation-exhalation breathing groups. We only selected the 1st and 2nd breathing groups to analyze group duration (seconds), pause duration (seconds), and speech rate (the number of words divided by the breathing group duration, n/second).

3.1. Breathing Effects

3.1.1. Speech Rate

A 2×2×3 mixed-subjects ANOVA was conducted with between-subject independent variable sex (male and female), two within-subject independent variables breathing group (1st and 2nd) and articulation run (1st, 2nd and 3rd) and speech rate as the dependent variable. Results indicated significant main effects of breathing group ($F(1, 72) = 91.58$, $p < .001$, $\eta^2 = 0.56$), articulation run ($F(2, 144) = 41.14$, $p < .001$, $\eta^2 = 0.36$), but no main effect of sex and other interaction effects. Pairwise comparison showed that repeating "*pīng pāng qiú*" was faster (M= 2.09, SE=0.03) in the first breathing group than second group (M=1.94, SE=.03). Furthermore, Pairwise comparison showed that speech rate of repeating "*pīng pāng qiú*" was descending from the first run (M= 2.11, SE=0.03), the second (M=1.99, SE=.03) and the third run (M=1.95, SE=.03).

3.1.2. Breathing group Duration

A 2×2×3 mixed-subjects ANOVA was conducted with between-subject independent variable sex (male and English) breathing group (1st and 2nd) and articulation run (1st, 2nd and 3rd) and breathing group duration as the dependent variable. A significant main effect of breathing group was found, $F(1,72) = 116.47$, $p < .001$, $\eta^2 = 0.62$. It shows that the first breathing group duration (M=7.46, SE=0.31) is longer than the second breathing group duration (M=5.20, SE= 0.29), which indicates the articulation task taxes the respiratory system and shortens the breathing to inhale air. However, sex and articulation run were not significant, and no interaction effects were observed.

3.1.3. Pause Duration

A 2×2 mixed-subjects ANOVA was conducted with between-subject independent variable sex (male and female), within-subject independent variable articulation run (first, second and third runs) as independent variables and pause as the dependent variable. Results revealed no any main or interaction effect.

3.2. Regression Methods and Hyperparameter Optimization

This study evaluates the predictive power of machine learning models across a spectrum of pulmonary function parameters (including FVC, FEV1, FF, PEF, MEF75, MEF50, MEF25, MVV) using four different feature configurations: (i) all features (demographic features: age, sex, height, weight, BMI; breathing features: breathing group duration, pause duration; speech features: number of words, speech rate), (ii) combined speech and breath features, (iii) breath-only features, and (iv) speech-only features. The regression results are discussed below with a focus on RMSE (root mean square error) and Pearson's correlation coefficient as key performance indicators.

We employed a suite of nine regression algorithms ranging from linear to ensemble methods and each rigorously tuned via exhaustive grid search with five-fold cross-validation on the 80% training partition. For stochastic gradient descent (SGD Regressor), the regularization strength α was swept over $[1 \times 10^{-4}, 1 \times 10^{-3}, 1 \times 10^{-2}, 1 \times 10^{-1}]$. Ridge and Lasso were both optimized over $\alpha \in \{2.0, 2.1, 2.2\}$ (with a maximum number of iterations of 1 000 000, to ensure convergence). ElasticNet extended this by additionally varying the lasso penlty (L_1) ratio over $\{0.01, 0.25, 0.5, 0.75, 1.0\}$. Support Vector Regression explored regularisation parameter $C \in \{0.1, 1, 10, 100\}$ and $\varepsilon \in \{0.01, 0.1, 1\}$. Decision trees were configured with maximum depth over $\{3, 5, 7, 10\}$ and minimum samples split in $\{2, 5, 10\}$, while the Random Forest and Gradient Boosting regressors both tuned n estimators of 10, 20, and 30 (the latter also testing a learning rate in $\{0.01, 0.1\}$ and maximum depth of $\{3, 5, 7\}$). Finally, AdaBoost varied the same numbers os estimators and learning rate in $\{0.01, 0.1, 1.0\}$. At each grid point, model performance was assessed by averaging RMSE across folds, and the optimal configuration was then retrained on the full training set. Final evaluation on the 20% hold-out used RMSE and Pearson's r after clipping predictions to the empirical [min, max] bounds of the training targets. The best results are reported in table below.

As a baseline for the regression task, we computed the Root Mean Squared Error (RMSE) between the original target values and a constant prediction equal to the mean of the training set as shown in Table 1. This represents the performance of a naïve model that predicts the average value for all inputs, without using any features. By comparing our regression model's RMSE against this baseline, we can assess whether the model captures meaningful patterns beyond simply fitting to the average.

Table 3. Baseline results.

Score	rmse (cv)↓	rmse (test)↓
FVC	0.986	0.659
FEV1	0.744	0.592
FF	7.783	5.451
PEF	2.057	1.218
MEF75	1.656	1.185
MEF50	1.163	0.936
MEF25	0.651	0.507
MVV	31.243	24.248

3.2.1. All Features

The best performance across most metrics is observed when using all features combined. For key indicators such as FVC and FEV1, Random Forest and Gradient Boosting models yield strong results with low RMSE and high Pearson's r values (e.g., FVC: RMSE = 0.504/0.670, r = 0.886/0.718; FEV1: RMSE = 0.404/0.425, r = 0.853/0.750). This reinforces the importance of multimodal feature integration in modelling complex physiological processes. However, performance for variables such as FF and MEF25 is suboptimal, suggesting that these metrics may be more sensitive to noise or less represented in the feature space.

Table 4. Regression Analysis Results all features.

Score	Method	rmse (cv)↓	rmse (test)↓	Pearson's r(cv)↑	Pearson's r(test)↑
FVC	Random Forest	0.504	0.670	0.886	0.718
FEV1	Gradient Boosting	0.404	0.425	0.853	0.750
FF	ElasticNet	7.451	6.121	0.279	-0.434
PEF	AdaBoost	1.212	1.181	0.811	0.627
MEF75	Random Forest	1.045	1.125	0.764	0.550
MEF50	Random Forest	0.973	1.011	0.522	0.359
MEF25	SVR	0.653	0.504	0.069	-0.618
MVV	AdaBoost	22.460	16.650	0.710	0.723

3.2.2. Speech + Breath Features

A marked drop in performance is observed when isolating the feature set to only speech and breath, despite retaining the same model classes. For instance, FVC prediction drops from $r = 0.718$ (all features) to a negative correlation ($r = -0.345$), and similar degradation is noted for FEV1 and MVV. This decline indicates that while speech and breath features offer a rich, non-invasive proxy, their standalone predictive fidelity is limited without contextual or demographic covariates (i.e., age, sex, weight, BMI).

Table 5. Regression analysis results using speech + breath features.

Score	Method	rmse (cv)↓	rmse (test)↓	Pearson's r (cv)↑	Pearson's r (test)↑
FVC	Gradient Boosting	0.921	0.931	0.458	-0.345
FEV1	AdaBoost	0.670	0.641	0.512	0.04
FF	SVR	7.755	5.462	-0.130	0.204
PEF	Random Forest	1.773	1.303	0.510	0.344
MEF75	Random Forest	1.356	1.251	0.550	0.294
MEF50	Gradient Boosting	1.135	0.950	0.203	0.021
MEF25	Gradient Boosting	0.661	0.51	0.048	-0.342
MVV	ElasticNet	28.180	22.007	0.441	0.463

3.2.3. Breath-Only Features

Isolating to breath features alone yields moderately better results than the speech+breath fusion, particularly for MVV (Ridge Regression: r = 0.484) and MEF50 (Gradient Boosting: $r = 0.465$). This supports prior findings that breath-derived biomarkers such as airflow patterns carry meaningful information reflective of pulmonary effort and obstruction levels. Nevertheless, FVC and FEV1 correlations remain weak or negative, suggesting breath alone may inadequately capture volume-related parameters.

Table 6. Regression analysis breath-only features.

Score	Method	rmse (cv)↓	rmse (test)↓	Pearson's r (cv)↑	Pearson's r (test)↑
FVC	AdaBoost	0.884	0.847	0.538	- 0.222
FEV1	AdaBoost	0.689	0.6690	0.453	- 0.113
FF	SVR	7.747	5.446	0.080	**0.210**
PEF	ElasticNet	1.992	1.145	0.171	0.338

MEF75	AdaBoost	1.590	1.204	0.291	0.261
MEF50	Gradient Boosting	1.163	0.905	-0.154	0.465
MEF25	Gradient Boosting	0.659	0.494	0.0132	-0.007
MVV	Ridge	29.71	21.05	0.264	0.484

3.2.4. Speech-Only Features

Interestingly, speech-only models perform comparably or better than breath-only models in certain parameters. For example, PEF and MEF75 achieve Pearson's *r* of 0.230 and 0.207 respectively using AdaBoost, outperforming their breath-only counterparts. This may be attributed to phonation characteristics (e.g., loudness, pitch variation, articulation speed) being indirectly modulated by respiratory strength and control.

However, performance variability is higher, and some predictions (e.g., MEF25) exhibit negative correlations, underscoring the challenge of isolating pulmonary mechanics from speech patterns alone.

Table 7. Regression analysis speech-only feature.

Score	Method	rmse (cv)↓	rmse (test)↓	Pearson's r (cv)↑	Pearson's r (test)↑
FVC	Adaboost	0.927	0.770	0.436	0.047
FEV1	Adaboost	0.6426	0.627	0.550	0.158
FF	SVR	7.7572	5.462	-0.248	0.180
PEF	Adaboost	1.632	1.593	0.607	0.230
MEF75	Adaboost	1.314	1.397	0.568	0.207
MEF50	Gradient Boosting	1.130	0.972	0.251	-0.121
MEF25	Gradient Boosting	0.6635	0.519	-0.066	-0.178
MVV	Lasso	28.2391	22.630	0.431	0.390

3.2.5. Model Observations

Across tables, ensemble methods such as Random Forest and AdaBoost emerge as robust performers, often outperforming linear models (ElasticNet, Lasso) or SVR. This trend aligns with their ability to handle non-linear feature interactions and resist overfitting when trained on moderate-sized datasets with complex interdependencies.

Compared to the baseline model, which predicts a constant mean and thus yields moderate RMSE but undefined Pearson correlation (due to zero variance in predictions), all feature-based models demonstrate improved performance. Using **all features**, the models consistently outperform the baseline across nearly all pulmonary parameters, achieving lower RMSE and strong positive correlations, for example, FVC (RMSE: 0.670 vs. 0.659 baseline; r: 0.718) and FEV1 (RMSE: 0.425 vs. 0.592 baseline; r: 0.750). In contrast, models trained on **breath-only** or **speech-only** features often show mixed results: they sometimes improve RMSE (e.g., MEF50 and MVV) but frequently exhibit weak or even negative correlations, especially for volume-related parameters like FVC and FEV1. The **speech+breath** configuration performs slightly better than each modality alone but still fails to match the predictive accuracy of the all-feature model. Overall, while non-invasive features offer some predictive value, the inclusion of demographic and physiological context remains critical to consistently outperforming the baseline.

The comparative analysis underscores the superior performance of multimodal feature fusion for pulmonary function regression tasks. While breath and speech signals provide valuable non-invasive input modalities, they require complementary features to approach clinically viable prediction accuracy. Future work could investigate feature selection, deep learning architectures, and

cross-modal representations to further enhance the utility of speech and breath data in respiratory health monitoring.

4. Discussion

The study employed ANOVA and regression analyses to investigate how speech rate relates to respiratory function, particularly in the context of repetitive articulation tasks. ANOVA findings revealed that both breathing group and articulation run significantly influenced speech rate, with faster rates observed during the first breathing group and earlier runs, suggesting increased respiratory load or fatigue as the task progressed. The first breathing group was also notably longer, indicating greater respiratory effort at the onset. Regression analyses demonstrated that models using all available features—including demographic variables (age, weight, height, BMI), speech, and breathing parameters, achieved the highest predictive accuracy for pulmonary measures such as FVC, FEV1, and PEF. Notably, when demographic features were removed, model performance declined substantially, suggesting that respiratory function is highly correlated with demographic factors, especially age and body size. Importantly, among the various respiratory indices, PEF showed a strong positive association with speech rate, reinforcing the potential of speech rate as a non-invasive biomarker for respiratory strength and airflow capacity.

This study expands on the work of Zeng et al. [1] by validating the cross-linguistic applicability of speech-based respiratory assessment through repetitive articulation tasks. Specifically, the "*pīng pāng qiú*" (table tennis in Mandarin Chinese) task employed in this study yielded findings similar to those from the "helicopter" task, originally designed for English. Both tasks demonstrated that speech breathing is sensitive to respiratory load and articulatory fatigue, specifically in the current study with consistent declines in speech rate across repeated groups and runs. This similarity suggests that the core principle, using linguistically simple, repetitive speech to tax the respiratory system and evaluate lung function, is valid across languages, including Mandarin. It highlights the potential for localized versions of speech-based assessments in multilingual clinical settings, particularly for populations where traditional spirometry may be challenging. These findings align with prior research indicating that controlled speech tasks can effectively mirror respiratory mechanics [18,35,70].

Secondly, although FVC and FEV1 are widely accepted as the gold standards in pulmonary function testing, this study emphasizes the value of PEF in the context of speech production. PEF, which measures the maximum speed of expiration, demonstrated a strong correlation with speech rate across the experimental conditions. This suggests that PEF may better reflect the dynamic airflow requirements associated with rapid articulation, unlike FVC and FEV1, which primarily assess volume and sustained airflow. Prior research supports the clinical relevance of PEF in speech analysis, especially in conditions like asthma and COPD, where airflow resistance impacts phonation [4,9]. The positive relationship between speech rate and PEF observed here supports the hypothesis that speech-based tasks can reveal subtle impairments in expiratory strength and airflow coordination, thus offering an alternative lens for respiratory evaluation.

Interestingly, no significant sex differences were found in speech rate or breathing parameters during the task, which contrasts with some previous studies suggesting that males and females exhibit distinct respiratory-linguistic coordination patterns [1,70–75] (but some studies suggest that no significant gender difference or task variables or physical characteristics may play a greater role than gender in determining normative speech breathing behaviors ([76,77]). This discrepancy may be attributable to the data analysis, which focused only on the first two breathing groups in each 30-second run. It is plausible that sex-based differences may emerge more clearly in prolonged speech under fatigue, where anatomical and physiological differences in lung volume and respiratory muscle endurance become more pronounced. Future studies could extend the analysis to full-length articulatory tasks or include additional respiratory parameters such as subglottal pressure and phonatory effort over time to capture such variance. Additionally, incorporating hormonal and

biomechanical factors into the analysis may offer a more nuanced understanding of sex-related variability in speech breathing.

Lastly, the integration of AI and machine learning proved instrumental in advancing the analysis of speech-respiratory data. Regression models using ensemble techniques such as Random Forest and AdaBoost consistently outperformed traditional linear models, particularly when predicting key respiratory indices like FVC, FEV1, and PEF. The ability of these models to capture non-linear interactions and subtle patterns among multimodal features (speech, breathing, and demographic data) demonstrates the value of AI in health diagnostics. These findings resonate with a growing body of research advocating for AI-enhanced speech analysis in respiratory and neurological health assessments [56,62]. As speech datasets grow in scale and complexity, AI-based models will be critical for extracting meaningful health markers that are both non-invasive and scalable, especially in telehealth and low-resource settings. In future work, we intend to investigate the use of machine learning on raw audio data to automate the extraction of speech and breathing features, thus further broadening the scope of application of the methods presented in this paper.

5. Conclusions

In conclusion, the findings support the clinical potential of speech-based respiratory tasks across languages, underscore the diagnostic relevance of PEF in speech, point to new directions for examining sex differences, and highlight the power of AI in modelling complex physiological phenomena from speech data. Further longitudinal and patient studies incorporating diverse populations and conditions are needed to fully harness these insights for real-world healthcare applications.

Abbreviations

The following abbreviations are used in this manuscript:

BMI	Body Mass Index
FVC	Forced Vital Capacity
FEV1	Forced Expiratory Volume in 1 Second
FF	ratio of FEV_1 to FVC
PEF	Peak Expiratory Flow
MEF75	Maximal Expiratory Flow at 75% of FVC
MEF50	Maximal Expiratory Flow at 50% of FVC
MEF25	Maximal Expiratory Flow at 25% of FVC
MVV	Maximum Voluntary Ventilation

References

1. Zeng, B.; Williams E.M.; Owen, C.; Zhang, C.; Davies, S.K.; Evans, K.; Preudhomme S-R. Exploring the acoustic and prosodic features of a lung-function-sensitive repeated-word speech articulation test. *Front. Psychol.* **2023**; 14:1167902. doi: 10.3389/fpsyg.2023.1167902

2. Werner, R.J. The phonetics of speech breathing: pauses, physiology, acoustics, and perception. PhD dissertation, der Universität des Saarlandes, Saarbrücken, 2023.

3. Hoit, J.D.; Hixon, T.J. Age and speech breathing. *Journal of Speech, Language, and Hearing Research*. **1987**,30(3),351-66.

4. Solomon, N.P.; Hixon, T.J. Speech breathing in Parkinson's disease. *Journal of Speech, Language, and Hearing Research*. **1993**,36(2),294-310.

5. Arnold, R.J.; Gaskill, C.S.; Bausek, N. Effect of combined respiratory muscle training (cRMT) on dysphonia following single CVA: a retrospective pilot study. *Journal of Voice*. **2023**,37(4),529-38.

6. Cummins, E.P.; Strowitzki, M.J.; Taylor, C.T. Mechanisms and consequences of oxygen and carbon dioxide sensing in mammals. *Physiological reviews*. **2020**, 100(1), 463-88.

7. Honda, K. Physiological processes of speech production. *Springer handbook of speech processing*. 2008:7-26.

8. Patel, N.; Chong, K.; Baydur, A. Methods and applications in respiratory physiology: Respiratory mechanics, drive and muscle function in neuromuscular and chest wall disorders. *Frontiers in Physiology*. **2022**,14;13:838414.

9. Hixon, T.J.; Weismer, G., Hoit, J.D. *Preclinical speech science: Anatomy, physiology, acoustics, and perception*. Plural Publishing; 2018 Aug 31.

10. Traser, L.; Burk, F.; Özen, A.C.; Burdumy, M.; Bock, M.; Blaser, D.; Richter, B.; Echternach, M. Respiratory kinematics and the regulation of subglottic pressure for phonakenttion of pitch jumps–a dynamic MRI study. *PLoS One*. **2020**, 31;15(12):e0244539.

11. Murray, E.S.; Michener, C.M.; Enflo, L.; Cler, G.J.; Stepp, C.E. The impact of glottal configuration on speech breathing. *Journal of Voice*. **2018**, 1, 32(4):420-7.

12. Krohn, F.; Novello, M.; van der Giessen, R.S.; De Zeeuw, C.I.; Pel, J.J.; Bosman, L.W. The integrated brain network that controls respiration. *Elife*. **2023**, 8;12:e83654.

13. Simonyan, K.; Horwitz, B. Laryngeal motor cortex and control of speech in humans. *The Neuroscientist*. **2011**, 17(2):197-208.

14. Darling-White, M.; Anspach, Z.; Huber, J.E. Longitudinal effects of Parkinson's disease on speech breathing during an extemporaneous connected speech task. *Journal of Speech, Language, and Hearing Research*. **2022**, 4;65(4):1402-15.

15. Subramaniam, N.S.; Bawden, C.S.; Waldvogel, H.; Faull, R.M.; Howarth, G.S.; Snell, R.G. Emergence of breath testing as a new non-invasive diagnostic modality for neurodegenerative diseases. *Brain Research*. **2018**,15;1691:75-86.

16. Ghosh, S. Breathing disorders in neurodegenerative diseases. *Handbook of Clinical Neurology*. 2022, 1;189:223-39.

17. Moreno, E.G.; Calassa, B.T.; Oliveira, D.V.; Silva, M.I.; Albuquerque, L.C.; Freitas-Dias, R.D.; Silva, B.R.; Araújo, R.C.; Costa, E.L.; Costa, E.C.; Correia Junior, M.A. Maximum phonation time in the pulmonary function assessment. *Revista CEFAC*. **2021**,24;23:e9720.

18. Feltrin, T.D.; Gracioli, M.D.; Cielo, C.A.; Souza, J.A.; de Oliveira Moraes, D.A.; Pasqualoto, A.S. Maximum phonation times as biomarkers of lung function. *Journal of Voice*. **2024**, 8.

19. Maslan, J.; Leng, X.; Rees, C.; Blalock, D.; Butler, S.G. Maximum phonation time in healthy older adults. *Journal of Voice*. **2011**,1;25(6):709-13.

20. Vaca, M.; Mora, E.; Cobeta, I. The aging voice: influence of respiratory and laryngeal changes. *Otolaryngology–Head and Neck Surgery*. **2015**,153(3):409-13.

21. Eckel, F.C.; Boone, D.R. The s/z ratio as an indicator of laryngeal pathology. *Journal of speech and hearing disorders*. **1981**,46(2):147-9.

22. Vaca, M.; Cobeta, I.; Mora, E.; Reyes, P. Clinical assessment of glottal insufficiency in age-related dysphonia. *Journal of Voice*. **2017**,1;31(1):128-e1.

23. Yang, C.C.; Chung, Y.M.; Chi, L.Y.; Chen, H.H.; Wang, Y.T. Analysis of verbal diadochokinesis in normal speech using the diadochokinetic rate analysis program. *Journal of Dental Sciences*. **2011**,1;6(4):221-6.

24. Kent, R.D.; Kim, Y.; Chen, L.M. Oral and laryngeal diadochokinesis across the life span: A scoping review of methods, reference data, and clinical applications. *Journal of Speech, Language, and Hearing Research*. **2022**, 9;65(2):574-623.

25. Kent, R.D. Research on speech motor control and its disorders: a review and prospective. *J Commun Disorder*. **2000**,33(5):391-428. doi:10.1016/s0021-9924(00)00023-x

26. Lancheros, M.; Pernon, M.;Laganaro, M. Is there a continuum between speech and other oromotor tasks? evidence from motor speech disorders, *Aphasiology*, **2022**, 37(5), 715–734. doi: 10.1080/02687038.2022.2038367.

27. Bouvier, L.; McKinlay, S.; Truong, J.; Genge, A.; Dupré, N.; Dionne, A.; Kalra, S.; Yunusova, Y. Speech timing and monosyllabic diadochokinesis measures in the assessment of amyotrophic lateral sclerosis in Canadian French. *Int J Speech Lang Pathol*. **2024**, 26(2):267-277. doi: 10.1080/17549507.2023.2214706.

28. Dietsch, A.M.; Mocarski, R.; Hope, D.A.; Woodruff, N.; McKelvey, M. Revisiting the rainbow: Culturally responsive updates to a standard clinical resource. *American Journal of Speech-Language Pathology*. **2023**,11;32(1):377-80.

29. Wang, Y.T.; Green, J.R.; Nip, I.S.; Kent, R.D.; Kent, J.F. Breath group analysis for reading and spontaneous speech in healthy adults. *Folia Phoniatrica et Logopaedica*. **2010**,62(6):297-302.

30. Zhang, Z. Respiratory laryngeal coordination in airflow conservation and reduction of respiratory effort of phonation. *Journal of Voice*. **2016**,30(6):760-e7.

31. Honda, J.; Murakawa, M.; Inoue, S. Effect of averaging time and respiratory pause time on the measurement of acoustic respiration rate monitoring. *JA Clinical Reports*. **2023**, 9(1):61.

32. Dishnica, N.; Vuong, A.; Xiong, L.; Tan, S.; Kovoor, J.; Gupta, A.; Stretton, B.; Goh, R.; Harroud, A.; Schultz, D.; Malycha, J. Single count breath test for the evaluation of respiratory function in myasthenia gravis: a systematic review. *Journal of Clinical Neuroscience*. **2023**,112:58-63.

33. Waage, A.K.; Iwarsson, J. The Effect of Speaking Rate on Voice and Breathing Behavior. *Journal of Voice*. **2024**. https://doi.org/10.1016/j.jvoice.2024.07.004.

34. Ali, S.S.; O'Connell, C.; Kass, L.; Graff, G. Single-breath counting: a pilot study of a novel technique for measuring pulmonary function in children. *The American journal of emergency medicine*. **2011**, 29(1):33-6.

35. Winkworth, A.L.; Davis, P.J.; Adams, R.D.; Ellis, E. Breathing patterns during spontaneous speech. *Journal of Speech, Language, and Hearing Research*. **1995**, 38(1):124-44.

36. Kent, R.D.; Kent, J.F.; Rosenbek, J.C. Maximum performance tests of speech production. *Journal of speech and hearing disorders*. **1987**,52(4):367-87.

37. Shriberg, L.D.; Kwiatkowski, J. Continuous speech sampling for phonologic analyses of speech-delayed children. *Journal of speech and hearing disorders*. **1985**,50(4):323-34.

38. Heselwood, B. Howard, S. Clinical phonetic transcription. *The handbook of clinical linguistics*. 2008,381-99.

39. de Jong, N. H.; Wempe, T. Praat script to detect syllable nuclei and measure speech rate automatically. *Behavior Research Methods*. **2009**, 41(2):385–390.

40. Franke, M.; Hoole, P.; Schreier, R.; Falk, S. Reading Fluency in Children and Adolescents Who Stutter. *Brain Sciences*. **2021**,11(12):1595.

41. Darling-White, M.; Banks, S.W. Speech rate varies with sentence length in typically developing children. *Journal of Speech, Language, and Hearing Research*. **2021**,64(6S):2385-91.

42. Peres, L.M.; Juste, F.; Costa, J.B.; Sassi, F.C.; Ritto, A.P.; Andrade, C.R. Variability of speech rate and articulatory transition ability. *Audiology-Communication Research*. **2024**, 29:e2883.

43. Gordon, J.K.; Clough, S. How do clinicians judge fluency in aphasia?. *Journal of Speech, Language, and Hearing Research*. **2022**, 65(4):1521-42.

44. Tanner, J.; Sonderegger, M.; Stuart-Smith, J.; Kendall, T.; Mielke, J.; Dodsworth, R.; Thomas, E. Exploring the anatomy of articulation rate in spontaneous English speech: relationships between utterance length effects and social factors. *arXiv preprint arXiv*:2408.06732. **2024**,13.

45. Luz, S.; De La Fuente Garcia, S.; Albert, P. A Method for Analysis of Patient Speech in Dialogue for Dementia Detection. In Resources and Processing of linguistic, para-linguistic and extra-linguistic data from people with various forms of cognitive impairment. Paris, France: ELRA. 2018. p. 35-42

46. Tong, J.Y.; Sataloff, R.T. Respiratory function and voice: the role for airflow measures. *Journal of Voice*. **2022**,36(4):542-53.

47. Kuhlmann, L.L.; Iwarsson, J. Effects of speaking rate on breathing and voice behavior. *Journal of Voice*. **2024**,38(2):346-56.

48. 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. *Applied Sciences*. **2021**,11(17):7999.

49. 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(5):379-85.

50. Mayr, W.; Triantafyllopoulos, A.; Batliner, A.; Schuller, B.W.; Berghaus, T.M. Assessing the clinical and functional status of COPD patients using speech analysis during and after exacerbation. *International Journal of Chronic Obstructive Pulmonary Disease*. **2025**,137-47.

51. Llewellin, P.; Sawyer, G.; Lewis, S.; Cheng, S.O; Weatherall, M.; Fitzharris, P.; Beasley, R. The relationship between FEV1 and PEF in the assessment of the severity of airways obstruction. *Respirology*. **2002**, 7(4):333-7.

52. Tayler, N.; Grainge, C.; Gove, K.; Howarth, P.; Holloway, J. Clinical assessment of speech correlates well with lung function during induced bronchoconstriction. *NPJ primary care respiratory medicine*. **2015**,25(1):1-3.

53. Koury, T.G.; Counselman, F.L.; Huff, J.S.; Peebles, J.S.; Kolm, P. Comparison of peak expiratory flow rate with speaking time in ED patients presenting with acute exacerbation of asthma. *The American journal of emergency medicine*. **1998**, 16(6):572-5.

54. Loudon, R.G.; Lee, L.; Holcomb, B.J. Volumes and breathing patterns during speech in healthy and asthmatic subjects. *Journal of Speech, Language, and Hearing Research*. **1988**, 31(2):219-27.

55. Wiechern, B.; Liberty, K.A.; Pattemore, P.; Lin, E. Effects of asthma on breathing during reading aloud. *Speech, Language and Hearing*. **2018**, 21(1):30-40.

56. Nallanthighal, V.S.; Mostaani, Z.; Härmä, A.; Strik, H.; Magimai-Doss, M. Deep learning architectures for estimating breathing signal and respiratory parameters from speech recordings. *Neural Networks*. **2021**,141:211-24.

57. Alam, M.Z.; Simonetti, A.; Brillantino, R.; Tayler, N.; Grainge, C.; Siribaddana, P.; Nouraei, S.R.; Batchelor, J.; Rahman, M.S.; Mancuzo, E.V.; Holloway, J.W. Predicting pulmonary function from the analysis of voice: a machine learning approach. *Frontiers in digital health*. **2022**,4:750226.

58. Ngo, D.; Pham, L.; Phan, H.; Tran, M.; Jarchi, D. A deep learning architecture with spatio-temporal focusing for detecting respiratory anomalies. In2023 IEEE Biomedical Circuits and Systems Conference (BioCAS), 2023 Oct 19 (pp. 1-5). IEEE.

59. Pham, L.; Phan, H.; Palaniappan, R.; Mertins, A.; McLoughlin, I. CNN-MoE based framework for classification of respiratory anomalies and lung disease detection. IEEE journal of biomedical and health informatics. 2021 Mar 8;25(8):2938-47.

60. Jácome, C.; Ravn, J.; Holsbø, E.; Aviles-Solis, J.C.; Melbye, H.; Ailo Bongo, L. Convolutional neural network for breathing phase detection in lung sounds. *Sensors*. **2019**,19(8):1798.

61. Farrús, M.; Codina-Filbà, J.; Escudero, J. Acoustic and prosodic information for home monitoring of bipolar disorder. *Health Informatics Journal*. **2021**,27(1):1460458220972755.

62. Cummins, N.; Baird, A.; Schuller, B.W. Speech analysis for health: Current state-of-the-art and the increasing impact of deep learning. Methods. **2018**,151:41-54.

63. Kappen, M.; Vanhollebeke, G.; Van Der Donckt, J.; Van Hoecke, S.; Vanderhasselt, M.A. Acoustic and prosodic speech features reflect physiological stress but not isolated negative affect: a multi-paradigm study on psychosocial stressors. *Scientific Reports*. **2024**,14(1):5515.

64. Heitmann, J.; Glangetas, A.; Doenz, J.; Dervaux, J.; Shama, D.M.; Garcia, D.H.; Benissa, M.R.; Cantais, A.; Perez, A,; Müller, D.; Chavdarova, T. DeepBreath—automated detection of respiratory pathology from lung auscultation in 572 pediatric outpatients across 5 countries. *NPJ digital medicine*. **2023**,6(1):104.

65. Huang, D.M.; Huang, J.; Qiao, K.; Zhong, N.S.; Lu, H.Z.; Wang, W.J. Deep learning-based lung sound analysis for intelligent stethoscope. *Military Medical Research*. **2023**,10(1):44.

66. Nallanthighal, V.S.; Strik, H. Deep sensing of breathing signal during conversational speech. In: *Proceedings of the 20th Annual Conference of the International Speech Communication Association (Interspeech)*; 2019. p. 4110–4114. doi:10.21437/interspeech.2019-1796.

67. Schultz, B.G.; Joukhadar, Z.; Nattala, U., del Mar Quiroga, M.; Noffs, G.; Rojas, S.; Reece, H.; Van Der Walt, A.; Vogel, A.P. Disease Delineation for Multiple Sclerosis, Friedreich Ataxia, and Healthy Controls Using Supervised Machine Learning on Speech Acoustics. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*. 2023 Oct 4;31:4278-85.

68. Noto, S.; Sekiyama, Y.; Nagata, R.; Yamamoto, G.; Tamura T. Analysis of Speech Features in Alzheimer's Disease with Machine Learning: A Case-Control Study. *Healthcare*. **2024**, Vol. 12, No. 21, p. 2194. MDPI.

69. Boersma, P, Weenink, D. Praat: doing phonetics by computer program. Version 6.4.35. Retrieved from http://www.praat.org/. 2025

70. Winkworth, A.L.; Davis, P.J.; Ellis, E.; Adams, R.D. Variability and consistency in speech breathing during reading: Lung volumes, speech intensity, and linguistic factors. *Journal of Speech, Language, and Hearing Research*. **1994**, 37(3):535-56.

71. LoMauro, A.; Aliverti, A. Sex differences in respiratory function. *Breathe*. **2018**, 14: 131–140.

72. Stathopoulos, E.T.; Sapienza, C. Respiratory and laryngeal function of women and men during vocal intensity variation. *Journal of Speech, Language, and Hearing Research*. **1993**, 36(1):64-75.

73. Pépiot, E. Male and female speech: a study of mean f0, f0 range, phonation type and speech rate in Parisian French and American English speakers. *Speech prosody* 7. 2014; pp 305-309.

74. Yuan, J.; Liberman, M.; Cieri, C. Towards an integrated understanding of speaking rate in conversation. *Interspeech* 2006 Sep 17.

75. Kim, J. Effects of gender, age, and individual speakers on articulation rate in Seoul Korean spontaneous speech. *Phonetics and Speech Sciences*. **2018**,10(4):19-29.

76. Van Borsel, J.; De Maesschalck, D. Speech rate in males, females, and male-to-female transsexuals. *Clinical Linguistics & Phonetics*. **2008**, 22(9):679-85.

77. Hodge, M.M.; Rochet, A.P. Characteristics of speech breathing in young women. *Journal of Speech, Language, and Hearing Research*. **1989**,32(3):466-80.