Machine learning- and statistical-based voice analysis of Parkinson's disease patients: A survey

(Article begins on next page)

Machine Learning- and Statistical-based Voice Analysis of Parkinson's Disease Patients: A Survey

Federica Amato[1], Giovanni Saggio[2], Valerio Cesarini[*3], Gabriella Olmo[4], Giovanni Costantini[5]

[1]*E-mail:* federica.amato@polito.it - *dept. Control and Computer Engineering, Polytechnic University of Turin, Turin, Italy*
[2]*E-mail:* saggio@uniroma2.it - *dept. Electronic Engineering, University of Rome Tor Vergata, Rome, Italy*
[3]*E-mail:* valerio.cesarini@uniroma2.it - *dept. Electronic Engineering, University of Rome Tor Vergata, Rome, Italy*
[4]*E-mail:* gabriella.olmo@polito.it - *dept. Control and Computer Engineering, Polytechnic University of Turin, Turin, Italy*
[5]*E-mail:* costantini@uniroma2.it - *dept. Electronic Engineering, University of Rome Tor Vergata, Rome, Italy*

* **Corresponding author**. Tel.: +39 3336116177
E-mail address: valerio.cesarini@uniroma2.it
Address: Via del Politecnico 1, 00133 Rome, Italy

Abstract

The preliminary diagnosis and evaluation of the presence and/or severity of Parkinson's disease is crucial in controlling the progress of the disease. Real-time, non-invasive methodologies based on machine learning-enhanced voice analysis are gathering more interest as the potential of this field unveils. Specifically, acoustic features are employed in many machine learning techniques, and could also function as indicators of the overall state of the subjects' voice: this review aims at identifying the most widely employed and promising feature-based machine learning methodologies, evidencing baselines and state-of-the-art solutions. A total of 102 works plus 5 review articles were selected from the IEEE Xplore, PubMed, Elsevier, and Web of Science electronic databases. A statistical assessment is performed identifying the most frequently used features as well as those deemed as most effective; an overview of algorithms, public datasets, toolboxes, and general metadata is also performed. According to our results, Jitter, Shimmer, Harmonic-to-Noise Ratio, Fundamental Frequency, and Mel Frequency Cepstral Coefficients are the mostly adopted features. In addition, it is worth noting a fair prevalence of glottal-like models and additional filtering options, such as Detrended Fluctuation Analysis.

Keywords*: Parkinson's disease, acoustic features, Machine Learning, Voice analysis*

Abbreviations

Only acronyms present in the written text are reported

AI	Artificial Intelligence
AT	Amplitude Tremor
BBE	Bark Band Energies
CNN	Convolutional Neural Network
D2	Correlation Dimension
DBS	Deep Brain Stimulation
DDK	Diadochokinetic
DFA	Detrended Fluctuation Analysis
F0	Fundamental Frequency
FT	Frequency Tremor
HNR	Harmonic to Noise Ratio
LLE	Largest Lyapunov Exponent
LPC	Linear Prediction Coefficients
MFCC	Mel-Frequency Cepstral Coefficients
ML	Machine Learning
NHR	Noise to Harmonic Ratio

PD	Parkinson's Disease
PLP	Perceptual linear prediction
RASTA	Relative Spectral Transform
RBD	REM sleep behavior disorders
RPDE	Recurrence Period Density Entropy
SVM	Support Vector Machine
UCI	University of California Irvine
UPDRS	Unified Parkinson's Disease Rating Scale
VOT	Voice Onset Time
VSA	Vowel Space Area

1. Introduction

Approaches based on Machine Learning (ML) have been proving to be a reliable tool for diagnosing, monitoring, and supporting the therapy assessments of a number of pathologies (Asci et al., 2020) (Suppa et al., 2021) (Asci et al., 2021) (Mulfari et al., 2021). Along with the historically employed chemical, physiological, or electrical inputs, vocal signals, gathered by recording tasks spoken by subjects, are slowly gaining importance as ML-based voice analysis advances and the effects of several pathologies on the voice are explored (Saggio, 2020). In addition to the potential of voice sounds revealing physio-pathological information, such a methodology leads to non-invasive, real-time and cost-effective assessments which could be an asset in various steps of the study of a disease. Voice production is a complex motor activity performed by humans with more than 100 muscles involved (Kouba et al., 2022), each vocal sound resulting from the interaction of many systems and sub-systems (e.g., lungs, glottis, oral cavity, nasal cavity, trachea), all ruled by the brain activity (Saggio & Costantini, 2020).

Parkinson's disease (PD) is the second most common brain disease worldwide, with its dynamics being a gradual neuronal death that leads to an overall impairment of motor functions (Ricci et al., 2020) with clinical and sub-clinical axial signs (Di Lazzaro et al., 2020), which often include speech articulation impairment. Generally, the alert for non-genetic PD stems from an evidence of symptoms, and subsequent diagnosis comes after neuropathological examinations, with the Unified Parkinson's Disease Rating Scale (UPDRS) (Gelb et al., 1999) being the mostly employed indicator of severity. Noticeable symptoms which lead to a PD diagnosis include tremor, muscle stiffness, impaired balance and coordination, as well as alleged depression, dysphonia, and dysarthria (Plouvier et al., 2015). Advantageously, recent studies suggest the feasibility of a prodromal diagnosis through early indicators such as REM sleep behavior disorders (RBD) and voice disorders (Suppa et al., 2022) (Rusz et al., 2018), the evolution of which can also be an indicator of the disease progression. With RBD being linked to an exceptionally high probability of developing the disease years or decades later, vocal disorders have also been observed as powerful indicators of the presence of PD even at the very first stage of the disease. This applies to de-novo patients (i.e., previously undetected and untreated) (Fayad et al., 2021), but patients in middle or advanced stages of PD, who are usually being treated with L-dopa (Gandhi & Saadabadi, 2022) (Ricci et al., 2022) and other possible treatments such as deep brain stimulation (Ehlen et al., 2020) (Ricci et al., 2019), also show peculiar vocal characteristics and could benefit from vocal analysis, especially with regards to the control of medicine dosages.

The voice impairment that patients with PD experience is usually generalized as "hypokinetic dysarthria", mainly involving articulation and breathing difficulties as well as a voice quality empirically described as "trembly" and "unstable". Although other ML-based methodologies have been proposed for the detection of PD, such as the study of EEG (Loh et al., 2021) or MRI (Kaplan et al., 2022), vocal analysis has proven its worth as a reliable detector

(Galaz et al., 2018; Moro-Velazquez et al., 2021; Tuncer et al., 2020; Tuncer & Dogan, 2019) that also involves a completely non-invasive approach. Around 90% of PD patients suffer from dysarthric symptoms, which also usually include other voice-impairing conditions such as body tremor or dysphonia (Baumann et al., 2018). Vocal impairments and changes in voice quality are usually assessed by ear by speech therapists or neurologists, and recently, efforts have been devoted to objectify the PD condition on the basis of measurable parameters gathered from vocal tasks assessed by means of Artificial Intelligence (AI)-enhanced algorithms. “Traditional” ML pipelines involve the study of acoustic features, which could function as objective biomarkers of the state of a subjects’ voice. ML-based voice analysis has also been successfully adopted for other speech-impairing pathologies such as dysphonia (Suppa et al., 2020) (Tulics & Vicsi, 2019) (Tulics & Vicsi, 2018), COVID-19 (Robotti et al., 2021) (Costantini, Dr., et al., 2022) (Han et al., 2021), and, with regards to depression being a symptom of PD, even for emotion recognition (Costantini, Parada-Cabaleiro, et al., 2022) (Parada-Cabaleiro et al., 2020) (Costantini et al., 2021) (Yoon et al., 2018) (Issa et al., 2020).

Deep learning models, especially convolutional neural networks (CNN) applied to spectrographic images, are gaining more and more consideration as classification tools for voice analysis. However, they do not include a proper feature extraction step and have a black box-like behavior that makes them very hard to interpret. Conversely, ML methodologies are commonly characterized by a transparent pipeline which usually results in highly interpretable results. Said pipeline could be generalized as: audio recording, data pre-processing, feature extraction, and classification (or regression). Feature selection is sometimes performed, especially when the extracted features are numerous and widespread.

Since PD can be related to other conditions such as RBD, hypokinetic dysarthria, and PD-induced tremor, features somehow related to these conditions should also be taken into account, in order to identify possible trends and effects on the voice.

As for the commonly employed acoustic features, the most frequently reported ones in literature are based on the Fundamental Frequency (F0), or related to prosody, such as jitter, shimmer, and Mel-Frequency Cepstral Coefficients (MFCC) (Bogert, 1963a). Time and frequency are the most common domains, and lead to various related macroscopic features being employed, often accompanied by phonatory-specific attributes that stem from modeling physical characteristics related to the vocal folds. Since PD also entails an impaired enunciation capability, diadochokinetic (DDK) features are often included as they can measure the ability to produce a series of fast and alternating sounds (syllables). Specific information on the acoustic features used for the automatic identification of PD will be detailed in the following sections.

Even if automatic speech analysis techniques have been widely employed to investigate neurological conditions that lead to speech impairment, many researchers remark the absence of a validated acoustic model (Tăuţan et al., 2021), (Gómez-García et al., 2021), (Moro-Velazquez et al., 2021).

This work aims to fill this gap and to identify the most common and effective feature extraction techniques, acoustic features, and ML models generally adopted for studying early and advanced stages of PD. With the aim of building a baseline, the state-of-the-art is explored by evidencing and discussing different datasets, toolboxes, and libraries for feature extraction. Moreover, statistics on the most commonly employed features and algorithms are evaluated.

2. Related Works

A number of recent studies assessed the existing acoustic modeling strategies in speech analysis with particular emphasis devoted on features and speech-impairing pathologies, including PD.

Tăuţan et al. (Tăuţan et al., 2021) studied the computational approaches used in the entire neurodegenerative spectrum, involving motor and non-motor symptoms. As for PD assessments through vocal analysis, the authors pointed out the presence of a plethora of elements that concur when choosing the optimal set of features. This inevitably leads to a complexity increasing of the model, with added criticalities being the different possible inference options (e.g., disease assessment, severity monitoring, progression monitoring) and the dimension of the speech studied (e.g., phonation, articulation, or prosody analysis). Despite these limitations, the authors still proposed a list of the most frequent features, including time and frequency measures.

Gómez-García et al. (Gómez-García et al., 2021) reviewed the most common measures employed in automatic vocal analysis tools and gathered routines for their extraction in a freely available so termed AVCA toolbox, evidencing amplitude perturbation, frequency perturbation, and noise measures as the most common features (presumably because they included some of the most widespread tools for feature extraction).

Similar evidences were also reported in (Brabenec et al., 2017) and in (A. Ma et al., 2020). In the former, the authors devoted efforts for analyzing phonatory and prosodic changes in PD with particular attention to the pathophysiology behind the vocal alterations and the features used for their description. According to their findings, voice disturbances appear to stem from larynx asymmetric rigidity and incomplete glottic closure. This was also confirmed by laryngoscopy and stroboscopic investigations that revealed the asymmetric position of the vocal fold, phase asymmetry, increased glottal opening time, vocal fold bowing, and incomplete glottic closure. As for the most used features, they included Harmonic to Noise Ratio (HNR), jitter, shimmer, intensity, and F0.

In (Brabenec et al., 2017), the authors reviewed 86 articles dealing with acoustic analysis of PD, with a specific focus on early diagnosis and monitoring, functional imaging studies exploring neural correlates, and the effect of dopaminergic medication and brain stimulation.

Interestingly, researchers usually have been focusing on “conventional” features such as jitter, shimmer, and F0. However, despite having the advantage of being interpretable, they may be insufficient for revealing complex mechanisms or performing advanced analysis (Brabenec et al., 2017), which can involve Detrended Fluctuation Analysis (DFA), Correlation Dimension (D2), and Recurrence Period Density Entropy (RPDE), proved to be effective (Gómez-García et al., 2021).

The absence of common guideline for vocal task selection, recording, analysis was emphasized by 12/27/2022 6:45:00 AM, by stressing the importance of employing validated toolboxes, such as the Praat Software (by Paul Boersma and David Weenink, Phonetic Sciences, University of Amsterdam) for feature extraction. According to their results, the optimal set of features should include intensity, F0, jitter, shimmer, Vowel Space Area (VSA), and Voice Onset Time (VOT), as well as noise, timing, and articulatory measures.

An extended analysis was developed in (Moro-Velazquez et al., 2021) focused on articulatory and phonatory aspect comparison. Evidence from 192 investigated papers highlighted the importance of articulatory analysis with the most frequently adopted features, which include amplitude and frequency perturbation values, noise, complexity, and timing features, in addition to some sets of coefficients such as MFCC, Bark Band Energies (BBE), and Linear Prediction Coefficients (LPC).

The main gaps that led to this review work and that we would like to address include a certain lack of focus on the specific disease of PD and its stages, with other less-than-related motor-impairing pathologies often being included (such as essential tremor); an incomplete analysis of all the viable acoustic features, with some authors only focusing on articulatory and phonatory aspects, or even smaller subsets comprised only of the most commonly employed ones (such as F0, jitter or shimmer). Conversely, possibly also due to the fast pace of AI-based research for healthcare, the conclusions of many review works are ultimately contrasting, and the picture remains incomplete.

With these premises, our work aims at reviewing the vocal measures used for the automatic assessment of PD impairment and underlying the most effective parameterization techniques. To this purpose, we devote efforts to identifying and evaluating the most common features, taking into account feature selection, statistical and post-hoc analysis, as deployed in the selected works. Moreover, particular emphasis is reserved to the distinctive data of investigated patients (e.g., years from diagnosis, the severity of the disease), to the recording equipment (e.g., professional, or low-cost), and to the toolboxes used for feature extraction.

3. Methods

Literature research eligible for our work was carried out in March 2022. Four electronic databases, namely IEEE Xplore, Pubmed, Elsevier, and Web of Science were screened using the following search string:

((Parkinson OR Parkinson's disease) AND (speech OR voice OR vocal) AND (feature OR biomarker OR marker))

We selected peer-reviewed original articles and reviews published from January 2017. Only relevant journals (e.g., medical, biomedical, and engineering) were included (excluding conference papers, and letters). Duplicate results were removed manually.

The database search yielded a total of 1190 articles, whose relevance was evaluated based on their title, abstracts, and keywords. The studies were selected when dealing with the automatic evaluation of speech alterations in PD patients through the analysis of interpretable acoustic features, excluding works considering deep learning-based techniques due to the absence of a proper feature extraction procedure and the impossibility of gathering information on acoustic features (basilary for our work).

With the aim of identifying trends within concrete and interpretable acoustic features, we remark that "black box"-like deep learning methods are out of the scope of the present work. On the other hand, we did include papers that applied handcrafted feature extraction, with deep learning eventually being used only for the classification model. Additionally, we excluded studies reporting datasets with a cardinality of less than 25 subjects per class.

As a result, 102 articles are chosen and organized according to: (i) study ID (authorship and year), (ii) aim of the work, (iii) recording modality, (iv) dataset cardinality, (v) participants characteristics, (vi) set of features employed, (vii) toolboxes used for feature extraction, and (viii) model employed for classification step. The outcomes of each study

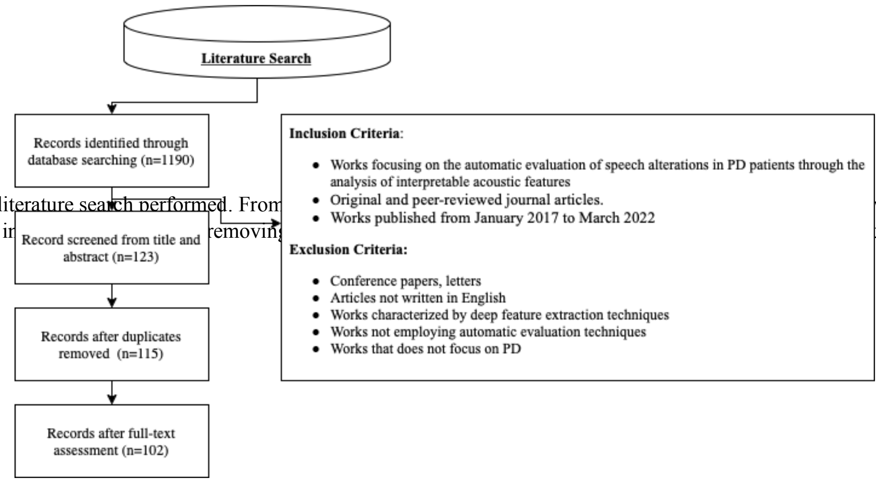

Figure 1 Workflow of the literature search performed. From ... were retained due to the fulfilment of the i... removing ... final of 102 papers were left.

and the subset of features that best performed are reported. The flow of the number of records (papers) identified after each selection step is reported in Figure 1, starting from the initial results of the search on online databases, ending up with 102 suitable studies after the application of inclusion/exclusion criteria, the removal of duplicates and a full-text assessment.

4. Results

4.1 Literature review

Table 1, Table 2, and Table 3 report the most relevant information, the toolboxes and libraries for speech feature extraction, and the freely available corpora related to PD speech, respectively, gathered from the reviewed works. We found that distinctive data of participants were not specified in most of the investigated works, so that no mention is devoted in Table 1 for the sake of conciseness (but are accessible in the Table 1 in Appendix).

To overcome the aforementioned lack of knowledge about validated toolboxes for voice analysis, here we report the most interesting and available ones. Moreover, freely available corpora to test the algorithms in different languages, recording techniques, or noise conditions are rarely reported in the literature, so that in Table 3 we fill this lack including their relevant information (with the caveat that the datasets provided by the Center for Machine Learning and Intelligent Systems at the University of California Irvine (UCI), namely: Little, Sakar18, Sakar13, Naranjo, Tsanas, LSVT, are provided as pre-extracted feature vectors.)

4.2 Common features

Table 4 lists the frequently adopted features (i.e., used by at least 5% of the papers) present in the reviewed literature, together with the definition, a brief description, their physical significance, and the evaluation techniques adopted (Table 2 in Appendix A completes the frame).

Table 1 Relevant information from the reviewed papers. **Ear**: Early; **Adv**: Advanced**; FOG**: Freezing Of Gait; **Mic**: Professional microphones, **n.r.**: Not Reported, **n.a.**: Not Applicable; **c.r**.: custom routines., **T1**: Vowel, **T2**: DDK, **T3**: Reading, **T4**: Monologue, **VAT**: Voice Analysis Toolbox;

REF	Aim	Task	Dataset	Recording modality	Features	Toolboxes	Classifier
(Parisi et al., 2021)	PD vs HC	T1	(1) Little (2) Sakar18	(1) Mic (2) n.r.	DFA, EMD, F0, Formants, GNE, GQ, HNR, Intensity, Jitter, MFCC, NHR, PPE, RPDE, Shimmer, TQWT, VFER	Praat, VAT	SVM
(Nahar et al., 2021)	PD vs HC	T1	Naranjo	n.r.	DFA, GNE, HNR, Jitter, MFCC, NHR, PPE, RPDE, Shimmer	n.r.	Bagging trees
(Jeancolas et al., 2022)	PD vs HC, RBD vs PD	T1, T2, T3, T4	117 PD, 41 RBD, 98 HC	Mic, Laptop, telephone calls	DPI, F0, Intensity, Jitter, NHR, NP, NSP, Rhythm, Shimmer	Praat, c.r.	SVM
(Hoq et al., 2021)	PD vs HC	T1	Sakar18	n.r.	DFA, EMD, F0, Formants, GNE, GQ, HNR, Intensity, Jitter, MFCC, NHR, PPE, RPDE, Shimmer, TQWT, VFER, WT	Praat, VAT	SVM
(Rusz, Hlavnička, et al., 2021)	PD vs HC, RBD vs PD, staging	T1, T2, T3, T4	149 PD, 150 RBD, 149 HC	Mic	HNR, DDK, DPI, F0, Intensity, NSR, VOT	c.r.	n.a.
(Sahandi Far et al., 2021)	PD vs HC	T1	mPower subset: 893 PD, 1291 HC	Smartphone (crowd-sourced)	A, DFA, Entropy, F0, GQC, GQO, HNR, Jitter, MFCC, OQ, RPDE, Shimmer	c.r.	n.a.
(Rusz, Tykalová, et al., 2021)	PD vs HC, Ear vs Lat	T1, T2, T3, T4	24 PD (Early onset), 24 PD (Late onset), 24 HC (Young), 24 HC (Old)	Mic	CPP, DDK, DPI, F0, HNR, Intensity, MPT, RFA, RLR, VOT	Dysarthria analyser	n.a.
(Šimek & Rusz, 2021)	PD vs HC, Ear vs Adv, RBD vs PD	T1, T3, T4	60 PD (Ear), 30 PD (Adv), 60 RBD, 60 HC	Mic	CPP, CPPS	c.r.	n.a.
(Yu et al., 2022)	PD vs HC, FOG vs no FOG	T1	40 PD (with FOG), 40 PD (no FOG), 40 HC	Mic	Autocorrelation, F0, HNR, Jitter, MPT, NHR, Pulse, Shimmer, Voicing	Praat	SVM
(Gaballah et al., 2021)	Medication	T1	51 PD (ON and OFF state), 11 HC	Mic	CPP, GFCC, HNR, Jitter, LPC, MS Area, RPDE, Shimmer, SRMR	Praat, DARTH-VAT, c.r.	n.a.

(Carrón et al., 2021)	PD vs HC	T1	(1) 30 PD, 30 HC (2) mPower subset: 30 PD, 30 HC	(1) Smartphone, (supervised) (2) Smartphone (crowd-sourced)	CPP, D2, Mutual inf. – first min, Corr. – first zero, Entropy, GNE, GQ, GQC, GQO, HNR, Hurst, Jitter, LZ-2, MFCC, MFSW, PPE, Shimmer, ZCR	c.r.	SVM, Passive Aggressive, ANN
(W. Rahman et al., 2021)	PD vs HC	T4	262 PD, 464 HC	Laptops	DFA, F0, HNR, Jitter, MFCC, PPE, RPDE, Shimmer	Praat, c.r.	XGBoost
(Tougui et al., 2020)	PD vs HC	T1	mPower subset: 453 PD, 1027 HC	Smartphone	Chroma Features, Energy, Entropy of energy, MFCC, Spectral variations, ZCR	pyAudioAnalysis library	XGBoost
(Rusz et al., 2022)	PD vs HC, staging	T1, T2, T3	100 PD, 100 HC	Mic	CPP, DDK, DPI, F0, HNR, Intensity, NSR, PSI, RFA, VOT	**c.r.**	n.a.
(Arora & Tsanas, 2021)	PD vs HC	T1	1078 PD, 5453 HC	Telephone calls	DFA, EMD, F0, GNE, GQ, HNR, Jitter, MFCC, NHR, PPE, RPDE, Shimmer, VFER, WT	DARTH-VAT	SVM
(Sajal et al., 2020)	PD vs HC	T1	54 PD, HC n.r.	Smartphone	D2, DFA, F0, HNR, Jitter, NHR, PPE, RPDE, Shimmer, Spread	c.r.	KNN
(Suphinnapong et al., 2021)	PD vs HC, Ear vs Adv	T1	48 PD (Ear), 52 PD (Adv), 101 HC	Mic	DUV, F0, Jitter, MP, NHR, PFR, Shimmer, SPI, Voice irregularity, VTI	c.r.	n.a.
(Kodrasi & Bourlard, 2020)	PD vs HC	T3	45 PD, 45 HC	Mic	Gini index, SHP – CHI, SHP – WEIBULL, Spectral variations	n.r.	SVM
(Amato, Borzi, et al., 2021)	PD vs HC	T3	(1) Dimauro (2) 26 PD 18 HC	(1) Mic (2) Smartphone, Laptop (crowd-sourced)	DFA, Duration ratio, ET, Intensity difference, MFCC, RASTA-PLP, Spectral parameters	n.r.	SVM
(Arora et al., 2021)	PD vs HC, prodromal PD	T1	335 PD, 112 RBD, 92 HC	Smartphone	DFA, EMD, F0, GNE, GQ, HNR, Jitter, MFCC, NHR, PPE, RPDE, Shimmer, VFER, WT	DARTH-VAT	RF
(Rusz et al., 2018)	PD vs HC, RBD vs PD	T1, T2, T3	30 PD, 30 RBD, 30 HC	Mic, Smartphone	DDK, DPI, F0, HNR, Intensity, Jitter, RFA, RST, Shimmer, VOT	Praat, c.r.	n.a.
(Demir et al., 2021)	PD vs HC	T1	Sakar18	n.r.	DFA, EMD, F0, Formants, GNE, GQ, HNR, Intensity, Jitter, MFCC, NHR, PPE, RPDE, Shimmer, TQWT, VFER, WT	Praat, VAT	SVM
(Palacios-Alonso et al., 2020)	PD vs HC	T1, T2	54 PD, HC n.r.	Smartphone	DDK, Energy, F0, FCR, FTA, Jitter, NTA, PTA, Shimmer, TrB, VF, VSA	BioMetRTools	n.a.

(Liu, Penttilä, et al., 2021)	PD vs HC, RBD vs PD, staging	T3	(1) 35 PD, 32 HC (2) 50 PD, 50 HC (3) 8 PD, 7 HC	(1), (2) Mic (3) n.r.	F2i/F2u, FCR, VAI, VSA	c.r.	n.a.
(Xiong & Lu, 2020)	PD vs HC	T1	Sakar18	n.r.	DFA, EMD, F0, Formants, GNE, GQ, HNR, Intensity, Jitter, MFCC, NHR, PPE, RPDE, Shimmer, TQWT, VFER, WT	Praat, DARTH-VAT	NN
(Gunduz, 2019)	PD vs HC	T1	Sakar18	n.r.	DFA, EMD, F0, Formants, GNE, GQ, HNR, Intensity, Jitter, MFCC, NHR, PPE, RPDE, Shimmer, TQWT, VFER, WT	Praat, DARTH-VAT	CNN
(Chen et al., 2020)	PD vs HC	T1	37 PD, 35 HC	Smartphone	MFCC	c.r.	Elastic net
(Ali et al., 2019)	PD vs HC	T1	Sakar13	Mic	Autocorrelation, F0, HNR, Jitter, NHR, Pulse, Shimmer, Voicing	Praat	NN
(Reddy & Alku, 2021)	PD vs HC	T1	(1) 110 PD, 93 HC (2) 50 PD, 50 HC	Mic	LFCC, MFCC	c.r.	SVM
(Quan et al., 2021)	PD vs HC	T1, T3	30 PD, 15 HC	Smartphone	BBE, MFCC	Neurospeech	Bi-LSTM
(Ashour et al., 2020)	PD vs HC	T1	Sakar18	n.r.	DFA, EMD, F0, Formants, GNE, GQ, HNR, Intensity, Jitter, MFCC, NHR, PPE, RPDE, Shimmer, TQWT, VFER, WT	Praat, DARTH-VAT	SVM
(Saxon et al., 2020)	Hypernasality	T3	38 PD, 37 dysarthric	n.r.	MFCC, PLP	c.r.	n.a.
(Tunc et al., 2020)	staging	T1	(1) 86 PD (2) Tsanas	(1) n.r. (2) Low-cost equipment	CPP, DFA, EMD, Energy, F0, Formants, GNE, GQ, HNR, Intensity, Jitter, MFCC, NHR, PPE, RPDE, Shimmer, SoE, TQWT, VFER, WT	Praat, DARTH-VAT, Voice Sauce	XGBoost
(X. Zhang et al., 2021)	PD vs HC	T1	(1) Sakar13 (2) 31 PD	Mic	Autocorrelation, F0, HNR, Jitter, NHR, Pulse, Shimmer, Voicing	Praat	SVM
(J. Ma et al., 2021)	PD vs HC, staging	T1	(1) Sakar13 (2) LSVT	Mic	Autocorrelation, DFA, EMD, F0, GNE, GQ, HNR, Jitter, MFCC, NHR, NHR, PPE, Pulse, RPDE, Shimmer, VFER, Voicing, WT	Praat, c.r.	Ensemble classifiers
(Naranjo et al., 2021)	Staging	T1	36 PD	Laptop, Mic	CPP, D2, MFCC, RPDE	n.r.	n.a.
(Klobusiakova et al., 2021)	PD vs HC	T2, T3	34 PD, 25 HC	n.r.	DDK, F0, Formants, Intensity, Loudness, SPIR	Praat, c.r.	n.a.

(Gunduz, 2021)	PD vs HC	T1	Sakar18	n.r.	DFA, EMD, F0, Formants, GNE, GQ, HNR, Intensity, Jitter, MFCC, NHR, PPE, RPDE, Shimmer, TQWT, VFER, WT	Praat, DARTH-VAT	SVM
(Solana-Lavalle & Rosas-Romero, 2021)	PD vs HC	T1	Sakar18	n.r.	DFA, EMD, F0, Formants, GNE, GQ, HNR, Intensity, Jitter, MFCC, NHR, PPE, RPDE, Shimmer, TQWT, VFER, WT	Praat, DARTH-VAT	KNN, RF, MLP, SVM
(Liu, Li, et al., 2021)	PD vs HC, Medication	T1	(1) Sakar13 (2) Little (3) 36 PD Drug I, 54 PD after medication	Mic	Autocorrelation, D2, DFA, F0, HNR, Jitter, NHR, PPE, Pulse, RPDE, Shimmer, Spread, Voicing	Praat, c.r.	SVM, RD, Extreme Learning Machine
(Karan et al., 2020)	PD vs HC	T1, T3	(1) 50 PD, 50 HC (2) 20 PD, 20 HC	Mic	DFA, Formants, GNE, GQC, HNR, IEDCC, IMFCC, Jitter, MFCC, NHR, PPE, RPDE, Shimmer, VFER	n.r.	SVM
(Tracy et al., 2020)	PD vs HC	T1	mPower subset: 246 PD, 2023 HC	Smartphone (crowd-sourced)	BE, Energy, FER, F0, Formants, Formants relative energy, Hammarberg Index, Harmonicity, HNR, Jitter, Loudness, MFCC, Shimmer, Spectral parameters, Voicing, ZCR	OpenSmile	GB
(Vásquez-Correa et al., 2019)	PD vs HC	T2	50 PD 50 HC	Mic	EMD	n.r.	SVM
(Despotovic et al., 2020)	PD vs HC, staging	T1, T4	(1) Little (2) Tsanas	Mic	F0, D2, DFA, HNR, Jitter, NHR, PPE, RPDE, Shimmer, Spread	Praat, c.r.	Gaussian Process Classification
(Pramanik et al., 2022)	PD vs HC	T1	(1) Sakar18 (2) Sakar13 (3) Naranjo	(1) n.r. (2) Mic (3) n.r.	DFA, EMD, F0, Formants, GNE, GQ, HNR, Intensity, Jitter, MFCC, NHR, PPE, RPDE, Shimmer, TQWT, VFER, WT	Praat, DARTH-VAT, n.r.	DT
(Lamba et al., 2022)	PD vs HC	T1	Sakar18	n.r.	DFA, EMD, F0, Formants, GNE, GQ, HNR, Intensity, Jitter, MFCC, NHR, PPE, RPDE, Shimmer, TQWT, VFER, WT	Praat, DARTH-VAT	XGBoost
(Karan et al., 2021)	PD vs HC	T1, T3	(1) 50 PD, 50 HC (2) 20 PD, 20 HC	Mic	Entropy, GNE, GQ, HNR, Jitter, MFCC, NHR, NMF, Shimmer, NMF, VFER	c.r.	SVM
(García et al., 2021)	PD vs HC	T3, T4	40 PD (Ear), 40 HC	Mic	BBE, DPI, DUV, Duration ratio, DVI, ET, F0, GFCC, Jitter, MFCC, NVS, Posterior probabilities	c.r.	SVM

(Masud et al., 2021)	PD vs HC	T1	Sakar18	n.r.	DFA, EMD, F0, Formants, GNE, GQ, HNR, Intensity, Jitter, MFCC, NHR, PPE, RPDE, Shimmer, TQWT, VFER, WT	Praat, DARTH-VAT	LDA
(Karan & Sekhar Sahu, 2021)	PD vs HC	T1, T3	1) 50 PD, 50 HC 2) 20 PD, 20 HC	Mic	HCC, VMD	c.r.	MLP, RF, SVM
(A. Rahman et al., 2021)	PD vs HC	T1	(1)160 PD, 100 HC (2) Sakar13	(1) Smartphone (2) Mic	MFCC	n.r.	SVM
(Cavallieri et al., 2021)	Medication	T1, T4	50 PD	Mic	F0, Intensity, Jitter, MPT, NHR, Shimmer, Voicing	Praat	n.a.
(Jenei et al., 2021)	PD vs HC, Diff. Analysis	T3	80 PD, 140 HC, various databases with different vocal alterations	Mic	Formants, MBE, MFCC	Praat	CNN
(Narendra et al., 2021)	PD vs HC	T1, T2, T3, T4	50 PD, 50 HC	Mic	AQ, BBE, ClQ, Energy, F0, Formants, Harmonicity, HRF, Jitter, MFCC, NAQ, OQ, PSP, QOQ, Shimmer, SQ	Neurospeech, APARAT toolbox	SVM
(T. Zhang et al., 2021)	PD vs HC	T1	(1) 28 PD, 25 HC (Extended from Sakar13) (2) 40 HC, 40 PD.	Mic	EMD	c.r.	SVM, RF
(Goyal et al., 2021)	PD vs HC	T1	(1) Sakar18 (2) Sakar13	(1) n.r. (2) Mic	Autocorrelation, DFA, EMD, F0, Formants, GNE, GQ, HNR,Intensity, Jitter, MFCC, NHR, PPE,Pulse, RPDE, Shimmer, TQWT, VFER, Voicing, WT	Praat, DARTH-VAT, n.r.	LDA
(Karabayir et al., 2020)	PD vs HC	T1	Naranjo	n.r.	DFA, GNE, HNR, Jitter, MFCC, NHR, PPE, RPDE, Shimmer	n.r.	GB
(Polat & Nour, 2020)	PD vs HC	T1	Naranjo	n.r.	DFA, GNE, HNR, Jitter, MFCC, NHR, PPE, RPDE, Shimmer	n.r.	KNN
(Arias-Vergara et al., 2017)	PD vs HC, Age	T1	50 PD, 50 HC, 50 HC (young)	Mic	CHNR, FCR, Formants, GNE, HNR, Jitter, MFCC, NNE, Shimmer, VSA	n.r.	SVM
(Travieso et al., 2017)	PD vs HC, Diff. Analysis	T1	50 PD, 50 HC, databases with different vocal alterations	Mic	Entropy, D2, DFA, Entropy, Hurst, LLE, LZ-2, RPDE	c.r.	SVM
(Benba et al., 2017)	PD vs HC, Diff. Analysis	T1	30 PD, 20 different neurological disorders	Smartphone	Autocorrelation, F0, Formants, HNR, Intensity, Jitter, MFCC, NHR, Pulse, RASTA-PLP, Shimmer, Voicing	Praat	KNN

(Erdogdu Sakar et al., 2017)	PD vs HC, Mild vs Adv	T1	(1) Tsanas (2) Little	(1) Smartphone (2) Mic	DFA, HNR, Jitter, NHR, PPE, RPDE, Shimmer, Spread	Praat, c.r.	SVM
(Naranjo, Pérez, Martín, et al., 2017)	PD vs HC	T1	Naranjo	n.r.	DFA, GNE, HNR, Jitter, MFCC, NHR, PPE, RPDE, Shimmer	n.r.	LR, SVM, KNN
(Benmalek et al., 2017)	PD vs HC, staging	T1	55 HC, 178 PD (Ear), 118 PD (Mid), 24 PD (Adv)	Smartphone, Laptop	DFA, F0, GQ, HNR, Jitter, NHR, PPE, RPDE, Shimmer	DARTH-VAT	Subspace discriminant
(Naranjo, Pérez, & Martín, 2017)	UPDRS	T1	Tsanas	Smartphone	DFA, HNR, Jitter, NHR, PPE, RPDE, Shimmer, Spread	Praat, c.r.	LR
(Hlavnička et al., 2017)	PD vs HC, RBD vs PD	T3, T4	30 PD (Ear), 50 HC, 50 RBD.	Mic	AST, DPI, DUF, DUS, DVI, EST, GBIV, LRE, PIR, RLR, RSR, RST	c.r.	n.a.
(Gómez et al., 2017)	PD vs HC	T1	94 PD, 8 HC	Mic	Formants	n.r.	n.a.
(Y. N. Zhang, 2017)	PD vs HC	T1	(1) Little (2) Sakar13	Mic	Autocorrelation, F0, HNR, Jitter, NHR, Pulse, Shimmer, Voicing	n.r.	KNN
(Oung et al., 2018)	PD vs HC, staging	T1	40 PD, 15 HC	Mic	Entropy, WT	c.r.	ELM
(Godino-Llorente et al., 2017)	PD vs HC	T2	50 PD, 50 HC	Mic	D2, Entropy, Hurst, LLE, PE, RPDE	n.r.	SVM
(Montaña et al., 2018)	PD vs HC	T2	27 PD, 27 HC	Mic	VOT	c.r.	SVM
(Yücelbaş, 2021)	PD vs HC	T1	Sakar18	n.r.	TQWT	Praat, DARTH-VAT	KNN
(Demir et al., 2022)	PD vs HC	T1	Sakar18	n.r.	DFA, EMD, F0, Formants, GNE, GQ, HNR, Intensity, Jitter, MFCC, NHR, PPE, RPDE, Shimmer, TQWT, VFER, WT	Praat, DARTH-VAT	KNN, SVM, DT
(Amato, Borzì, et al., 2021)	PD vs HC	T3	(1) 50 PD, 50 HC (2) 20 PD, 20 HC	Mic	BBE, ET, F0, MFCC, PTS, Spectral Variations, ZCR	n.r.	SVM

(Kacha et al., 2017)	PD vs HC	T1	205 PD, 74 HC	Mic	HNR, NHR	c.r.	n.a.
(Sheibani et al., 2019)	PD vs HC	n.r.	44 PD, HC n.r.	n.r.	F0, D2, DFA, HNR, Jitter, NHR, Spread PPE, PPQ, RPDE, Shimmer	n.r.	MLP, DT, SVM, NB, KNN
(L. Zhang et al., 2020)	PD vs HC, staging	T1	(1) Little (2) Tsanas (3) 48 PD, 20 HC (4) 4 PD	n.r.	F0, D2, DFA, HNR, Jitter, NHR, PPE, RPDE, Shimmer		SVM, LR, MLP
(Arora et al., 2019)	PD vs HC	T1	1483 PD, 8300 HC	Telephone calls	DFA, EMD, F0, GQ, GNE, HNR, Jitter, MFCC, NHR, PPE, RPDE, Shimmer, VFER, WT	c.r.	n.a.
(Vásquez-Correa et al., 2018)	PD vs HC	T1, T2, T3	86 PD, 50 HC	Mic	BBE, Duration ratio, DTW Similarity, DUV, DVI, FLUF, F0, Formants, Jitter, MFCC, NU, NV, RatioDurV/Sig, Shimmer, WA	c.r.	Various regressors
(Singh & Xu, 2020)	PD vs HC	T1	mPower subset: 140 PD, 860 HC	Smartphone	MFCC	c.r.	SVM
(Prince et al., 2019)	PD vs HC	T1	mPower subset: 766 PD, 439 HC	Smartphone	DFA, F0, HNR, Jitter, MFCC, PPE, RPDE, Shimmer	DARTH-VAT	LR, RF, Ensemble, CNN
(Hlavnička et al., 2020)	PD vs HC, Diff. Analysis	T1	40 PD, 40 HC, 200 other pathologies	Mic	AT, FT	c.r.	n.a.
(Cernak et al., 2017)	PD vs HC	T3	50 PD 50 HC	Mic	Energy, MFCC	n.r.	GMM
(Brückl et al., 2018)	PD vs HC	T1	234 PD, 50 HC	n.r.	Amplitude, AT, F0, FT	Praat	n.a.
(Upadhya & Cheeran, 2018a)	PD vs HC	T1	35 PD, 45 HC	n.r.	Autocorrelation, HNR, Jitter, MFCC, Shimmer	n.r.	SVM
(Upadhya et al., 2018)	PD vs HC	T1	45 PD, 45 HC	Mic	MFCC, PLP	n.r.	ANN
(Wu et al., 2018)	PD vs HC	T1	27 PD, 446 HC	n.r.	DFA, GQ, HNR, Jitter, MFCC, NHR, PPE, RPDE, Shimmer	DARTH-VAT, c.r.	SVM, RF
(Burk & Watts, 2019)	PD vs HC	T1, T3	32 PD, 10 HC	Mic	CPP	Analysis of dysphonia in speech and voice (ADSV)	n.a.

(Gillivan-Murphy et al., 2019)	PD vs HC	T1	30 PD, 20 HC	Mic	Amplitude, AT, F0, FT	n.r.	n.a.
(Rodríguez-Pérez et al., 2019)	PD vs HC	T1, T3	30 PD, 32 HC	Mic	F0, FER	n.r.	n.a.
(Benmalek et al., 2018)	PD vs HC, staging	T1	178 PD (Ear), 118 PD (Mid), 24 PD (Adv), 55 HC	n.r.	MFCC, PLP	n.r.	n.a.
(Galaz et al., 2018)	PD vs HC, staging	T1	51 PD	Mic	BE, FER, F0, FLUF, Formants, GNE, HNR, Jitter, NNE, Shimmer	n.r.	XGBoost
(Lahmiri et al., 2018)	PD vs HC	T1	147 PD, 48 HC	n.r.	F0, Jitter, PPQ, Shimmer, HNR, NHR, RPDE, DFA, D2, PPE, Spread	n.r.	SVM, LDA, KNN, NB, RT, RBF-NN, MDC
(Moro-Velázquez et al., 2018)	PD vs HC	T1, T2, T3	(1) 50 PD, 50 HC (2) 164 HC	Mic	LPC, MFCC, RASTA-PLP	n.r.	GMM, GPLDA
(Parisi et al., 2018)	PD vs HC	T1	(1) Sakar13 (2) 28 PD	(1) Mic (2) n.r.	F0, Autocorrelation, HNR, Jitter, NHR, Pulse, Shimmer	n.r.	SVM
(Upadhya & Cheeran, 2018c)	PD vs HC	T1	45 PD, 45 HC	n.r.	Autocorrelation, Energy, HNR, Jitter, NHR, Shimmer	n.r.	Regression (various), NN
(Upadhya & Cheeran, 2018b)	PD vs HC	T1	40 PD, 40 HC	Low-cost equipment	GNE, HNR, NHR	Praat, c.r.	n.a.
(Benba et al., 2019)	PD vs HC, RBD vs PD	T3	30 PD (Ear), 50 RBD, 50 HC	Mic	AST, DPI, DUF, DUS, DVI, EST, GBIV, LRE, PIPR, RLR, RSR, RST	n.r.	SVM
(Gómez-Vilda et al., 2019)	PD vs HC	T1	50 PD, 50 HC	n.r.	Formants	n.r.	n.a.
(Karlsson & Hartelius, 2019)	PD vs HC	T2	38 PD, 38 HC	n.r.	DDK, Jitter, Shimmer	n.r.	LR

(Lahmiri & Shmuel, 2019)	PD vs HC	T1	147 PD, 48 HC	Mic	D2, DFA, F0, HNR, Jitter, NHR, PPE, RPDE, Shimmer, Spread	n.r.	SVM
(Upadhya & Cheeran, 2018a)	PD vs HC	T1	54 PD, 45 HC	Mic	MFCC, Modified MFCC	n.r.	RBFN
(Suppa et al., 2022)	PD vs HC	T1, T3	58 PD (Ear), 57 PD (Adv), 108 HC	Mic	BE, Energy, F0, Harmonicity, HNR, Jitter, Loudness, MFCC, RASTA-PLP, Shimmer, Spectral Variations, Voicing, ZCR	OpenSmile	SVM

Table 2 Toolboxes and libraries for speech feature extraction from the reviewed papers. The '*' denotes additional toolboxes.

Toolbox name	**REF**	**Notes**	**Link for download**
Praat	(Boersma & van Heuven, 2001)	C-based software package for speech analysis, which allows the extraction of various metrics. It can be embedded in Python through the Parselmouth library (Jadoul et al., 2018)	https://www.fon.hum.uva.nl/praat/
Dysarthria Analyser	(Rusz et al., 2018), (Švec & Granqvist, 2010)	Provides an automated acoustic analysis of various dysarthric speech patterns based on selected algorithms for the detection of features extracted from tasks including sustained phonation, syllables, and connected speech	http://dysan.cz
pyAudioAnalysis library	(Giannakopoulos, 2015)	Comprehensive Python library for the extraction of features (MFCC, spectrogram, chromagram) and also construction of classifiers and automatic segmentation.	https://github.com/tyiannak/pyAudioAnalysis
DARTH Voice Analysis Toolbox (DARTH-VAT)	(Tsanas et al., 2010), (Tsanas et al., 2011), (Tsanas, 2012.)	The toolbox runs in MATLAB, and it has only been validated in settings with the sustained vowel /a/. Contains mainly features related to F0, Jitter, Shimmer, MFCC, RPDE, DFA, glottal modeling.	https://www.darth-group.com/software
BioMetR©Tools	(Gómez et al., 2013)	GUI-based toolbox for the extraction of high-level glottal features from the modeling of speech signals.	Not freely available
Voice Sauce	(Shue, 2010),(Shue et al., 2009)	MATLAB toolbox for the extraction of frequency and, especially, harmonic-related content from audio signals.	http://www.phonetics.ucla.edu/voicesauce/

OpenSmile	(Eyben et al., 2009),(Schuller et al., 2007), (Schuller et al., 2009)	Comprehensive software by Audeering, which allows for the extraction of 6000+ parameters based on custom configuration files.	https://www.audeering.com/research/opensmile/
Neurospeech	(Orozco-Arroyave et al., 2018)	Software platform designed to perform speech analysis of people with neurodegenerative disorders, particularly PD. It computes several measures to evaluate phonation, articulation, prosody, and intelligibility.	https://github.com/ElsevierSoftwareX/SOFTX-D-17-00058
APARAT Toolbox	(Airas et al., 2005)	Software package for use in MATLAB environment that implements glottal inverse filtering and several time-based parameters of the voice source in a graphical user interface	https://sourceforge.net/projects/aparat/
Analysis of dysphonia in speech and voice (ADSV)	/	Extracts spectral and cepstral dysphonia-related features from various, selectable speech tasks, including sustained phonation, sentences, or syllables.	Not freely available : https://www.pentaxmedical.com/pentax/en/99/1/Analysis-of-Dysphonia-in-Speech-and-Voice-ADSV
Prosodic impairment analysis tool*	(Galaz, Mekyska, et al., 2016)	MATLAB GUI-based toolbox for the extraction of prosodic features: Monopitch, Monoloudness (squared energy operator- SEO).	http://splab.cz/en/download/software/software-pro-analyzu-prozodickych-vad
Dysarthric speech quantification tool*	(Galaz, Mekyska, et al., 2016), (Galáž, et al.,2016a), (Galáž, et al., 2016b)	Open-source Python library comprising a variety of speech features conventionally used in the field of acoustic analysis of dysarthric speech.	http://splab.cz/en/download/software/software-pro-kvantifikaci-dyzartricke-reci

Table 3 Freely available corpora related to PD speech from the reviewed papers.

Name	**REF**	**Numerosity**	**Task**	**Subjects Characteristics**	**Recording Equipment**
Little	(Little et al., 2009)	23 PD, 8 HC	Sustained vowel /a/. 6 recordings per subject.	0-28 YFD. No info about therapy	Head-mounted microphone (AKG C420) positioned at 8 cm from the lips, in a treated sound booth. The voice signals were recorded directly on computer using Computerized Speech Laboratory (CSL) 4300B hardware (Kay Elemetrics). Sampling rate = 44.1kHz.
Sakar18	(C. O. Sakar et al., 2019)	188 PD, 64 HC	Sustained vowel /a/. 3 recordings per subject.	No info about therapy or disease duration	No info about microphones.

Sakar13	(B. E. Sakar et al., 2013)	Train: 20PD, 20HC; Test: 14PD, 14HC	Sustained vowels /a/, /o/, /u/; numbers from 1 to 10; short sentence; single words. 3 recordings per subject.	Train: 0-6 YFD. years UPDRS3 mean reported only for training data (from 5 to 55). Test: 0-13 YFD.	Trust MC-1500 low-end computer microphone placed at 10 cm distant from the subject. Frequency range = 50-13k Hz; Sampling frequency = 96kHz.
Naranjo	(Naranjo et al., 2016)	40 PD, 40 HC	Sustained vowel /a/. 3 recordings per subject, one a week for 6 months.	>5 YFD. No info about therapy.	No info about microphones.
Tsanas	(Tsanas et al., 2010)	42 PD	Sustained vowel /a/; running speech. Multiple recordings per patient during a 6-months trial.	Early PD. Diagnosis within 5 years from trial. UPDRS3 at 'rial's start = 19.4 ± 8.12. All subjects remained unmedicated for the six-month duration of the study.	Data collected using the Intel Corporation's at-home testing device (AHTD), a telemonitoring system which includes, among the other sensors, a high-quality microphone headset and a universal serial bus (USB) data stick to store test data. Sampling rate = 24 KHz; 16-bit resolution.
LSVT	(Tsanas et al., 2010)	14 PD with multiple recordings during rehabilitation.	Sustained vowel /a/.	No info about therapy or disease duration.	The voice signals were sampled at 44.1 kHz with 16 bits of resolution, and were recorded using the Audacity software package. Soundswere recorded in a double-walled, sound-attenuated room at the National Center for Voice and Speech-Denver (NCVS)
m-Power	(Bot et al., 2016),(Prince et al., 2018)	1060 PD, 5357 HC	Sustained vowel /a/.	Multiple information provided for each subject, including UPDRS, PDQ8, gait and various questionnaires.	Crowdsourced database: iPhone-recorded audio data from English volunteers.
Vaiciukynas	(Vaiciukynas et al., 2017)	64 PD, 35 HC	Sustained vowel /a/; sentence.	No info about therapy or disease duration.	Cardioid microphone (AKG Perception 220) and a smartphone (Samsung Galaxy Note 3) used simultaneously. Both microphones were located at ~10 cm distance from the mouth. PCM wav; Sampling rate = 44.1kHz; 16-bit resolution.
Dimauro	(Dimauro et al., 2017)	28 PD, 22 HC	Read text; sustained vowels /a/, /e/, /i/, /o/, /u/; syllables (/pa/, /ta/); words; sentences.	UPDRS and speech-related characteristics.	Professional microphone (unspecified).

Table 4 List of the most frequent features with related relevant information. N: Number of occurrences

N	Name	Feature definition	Feature description	Reference
62	Jitter	/	Variation of F0 between cycles.	(Eyben et al., 2016), (Gómez-García et al., 2021), (Karan et al., 2021)
60	Shimmer	/	Variation of the peak amplitude between cycles.	(Eyben et al., 2016), (Gómez-García et al., 2021), (Karan et al., 2021)
59	HNR	Harmonic-to-Noise Ratio	Ratio between the harmonic content (estimated from F0) and the non-harmonic, noise-like components.	(Eyben et al., 2016), (Gómez-García et al., 2021),
56	F0	Fundamental Frequency	Fundamental Frequency estimated with various algorithms evaluated on frames.	(Tsanas, 2010), (Little et al., 2009)
52	MFCC	Mel Frequency Cepstral Coefficients	Obtained by inverse-transforming the Mel scaled spectrum of the signal, with each coefficient pertaining to a Mel window.	(Orozco-Arroyave et al., 2013) (Gómez-García et al., 2021), (Bogert et al., 1963)
47	NHR	Noise-to-Harmonics Ratio	Ratio between noise-like estimated components and harmonic content.	
41	RPDE	Recurrence Period Density Entropy	Entropy of the period density function, which quantifies the extent of changes in the periodicity of the vocal fold oscillations.	(Tsanas, 2010), (Gómez-García et al., 2021),
39	DFA	Detrended Fluctuation Analysis	Measure of the stochastic self-similarity of a signal through de-trending over increasingly numerous windows. It can quantify noise in speech and turbulent breath.	(Little et al., 2009), (Gómez-García et al., 2021),
38	PPE	Pitch Period Entropy	Entropy of the period computed from pitch estimation. Quantifies the control of stationary voicing during a sustained phonation.	(Arora et al., 2019), (Little et al., 2009)
28	GNE	Glottal to Noise Excitation Ratio	Extent of noise in speech modelled using linear and non-linear energy concepts.	(Tsanas, 2010), (Gómez-García et al., 2021),
25	Formants	/	Various indicators related to the algorithmically estimated Formant Frequencies, which define the resonances of the vocal tract. Includes centre frequencies and bandwidths.	/
22	GQ	Glottis Quotient	Changes in the duration of the vocal fold cycle estimated from a modelling of the vocal trait.	(Tsanas, 2010), (Karan et al., 2021)
22	Intensity	/	Power of speech signal in dB	
20	EMD	Empirical Mode Decomposition	Decomposition of a signal into a sum of intrinsic mode functions (IMFs) obtained through cubic spline-based envelopes, with increasingly stringent residuals. Lower order IMFs allow higher frequencies. Includes most of the general time-domain features (SNR, duration, etc.) re-applied to IMFs as well as the number of IMFs and their central frequency and widths.	(Despotovic et al., 2020)
20	VFER	Vocal Folds Excitation Ratio	Measure of noise in speech within a glottal-like model of the vocal cords as a system needing an excitation. It is computed using energy, nonlinear energy, and entropy.	(Tsanas, 2010), (Karan et al., 2021)
18	WT	Wavelet Transform	Amplitude, scale, energy and envelope fluctuations quantified using the Wavelet transform.	(Tsanas, 2010), (Arora & Tsanas, 2021)

15	TQWT	Tunable-Q Wavelet Transform	Features related to a Wavelet transform whose Q (quality factor of the wavelets) is controlled externally by the user.	
11	Autocorrelation		Autocorrelation coefficients on each frame	(Mathieu et al., 2010)
11	D2	Correlation Dimension	Measure of the complexity of the system, obtained via estimation of the correlation dimension as an exponent to a base that is a ball. For small scales, it approximates the correlation sum, computed ignoring samples that are knowingly dependent (such as successive ones).	(Little et al., 2009), (Gómez-García et al., 2021),, (Henriquez et al., 2009), (Hegger et al., 1999)
11	Voicing		Probability of voicing for a given signal, computed algorithmically with a pipeline such as the one defined by Yeldener (Patented [REF]).	(Yeldener et al., 1998)
9	Pulse		Parameters related to the number, duration and std. dev of pulses and periods in the signal.	/
8	CPP	Cepstral peak prominence	Difference in amplitude between the cepstral peak and the trend line directly below it.	(Gómez-García et al., 2021),
8	DPI	Duration of Pause Intervals	Median duration of pauses in speech, longer than a certain threshold. It is related to abrupt changes in the speech.	(Jeancolas et al., 2022), (García et al., 2021), (Hlavnička et al., 2017)
7	DDK		Diadochokinetic features, related to the vocalization of syllables. Includes: rate, regularity.	(Vásquez-Correa et al., 2018)
7	Energy		Time-domain energy of the signal.	(Mathieu et al., 2010)
7	Spread		Linear and non-linear variations of F0 within the speech segment.	
6	Entropy	Approximate entropy	Entropy of the signal computed with various algorithms (Fuzzy, Gaussian kernel approximate, Gaussian kernel sample, Permutation, Renyi, Sample, Shannon).	(Gómez-García et al., 2021),
6	Spectral Variations		Various features related to the variation in the frequency spectrum. Includes: Crest Factor, Flux, Flatness, Decrease, Flatness, Rolloff, Centroid, Spread, Skewness, Kurtosis, Slope, etc.	(Mathieu et al., 2010)
5	BBE	Bark Band Energies	Energy within a specific frequency band over a Bark-weigthed spectrum.	/
5	VOT	Voice Onset Time	Length of the entire consonant from initial burst to vowel onset. Duration of the part of a syllable (/pa/, /ta/ or /ka/) between initial burst and vowel onset. Requires preliminary segmentation of the consonant syllables.	(Rusz et al., 2018)
5	ZCR	Zero Crossing Rate	Number of passes of the signal through zero (x-axis) within a frame.	(Tougui et al., 2020)

4.1 Most effective features

Besides the most common features, we also investigated the most effective ones as reported in Table 5 that includes features selected in at least 5% of the 102 selected articles (Table 3 in Appendix A completes the picture).

Table 5 The most effective features for PD speech impairment assessment according to the reviewed papers.

Number of occurrences	Name
28	MFCC
14	HNR
13	F0
12	Jitter
9	Shimmer
9	RPDE
6	DFA
5	DPI

5. Discussion

5.1 Most common and most effective features

We aimed at identifying a comprehensive set of acoustic features of dysarthric speech in PD, especially based on a ML-approach. According to our findings, the investigated papers mostly focus on phonatory and articulatory features, while only a small subset includes prosodic measures of speech. In this regard, authors in (Arias-Vergara et al., 2017) and (Vásquez-Correa et al., 2018) compared the performance of models based on phonatory and articulatory features agreeing that the introduction of the latter boosts the capability of the system to represent vocal alterations of PD patients.

As for the specific type of features, Table 4 and Table 5 show the most frequent and the most effective features. As expected, features describing F0 variations (i.e., Jitter, Shimmer, and F0 statistics), noise measures (e.g., HNR, GNE…), and MFCC result as the most common and, in a few cases, DFA and RPDE too.

MFCCs proved to be the most effective features in terms of classifier performance for PD speech parametrization. According to (Suphinnapong et al., 2021), this set of coefficients is particularly important not only for differentiating healthy control subjects (HC) versus PD patients, but also for determining a correlation with the UPDRS scale, due to their characteristic to capture meaningful reductions in the temporal variability between consecutive speech frames. Moreover, MFCCs also retain information about the pitch and its intelligibility, which makes them somehow related to F0. Perceptually, slow-moving MFCCs result in slow and monotone speech patterns, which often characterize dysarthric patients. Within this context, according to the findings reported in (Upadhya et al., 2018) and (Karan et al., 2021), MFCCs outperform Perceptual Linear Prediction (PLP) coefficients in evaluating dysarthric impairment, but authors in (Upadhya & Cheeran, 2018a) partially disagree. From a macroscopic point of view, the prevalence of MFCC among the most effective features points out the importance of the Cepstral domain, which results from an additional transformation after time to frequency switching. Alternative algorithms to MFCC's have been proposed, such as those based on linear (non-Mel) or Bark-scaled spectra. It is noteworthy that the Cepstral domain has often been employed for the characterization of pitch-related discrepancies in complex and noisy audio data. Interestingly, all the feature selection procedures applied to the mPower dataset (Bot et al., 2016), (Prince et al., 2018), which contains data of over 6000 subjects recorded in sub-optimal and unsupervised conditions, yielded MFCC as a relevant group of features.

Features associated with noise analysis in speech, particularly HNR, come soon after in effectiveness. Indeed, although they are strictly related to F0 evaluation algorithms, these features are often linked to hoarseness in voice, a well-known characteristic of PD patients stemming from the turbulent airflow brought by the asymmetry in vocal folds and movement impairment (Suphinnapong et al., 2021).

The fundamental frequency, F0, is often reported together with its standard deviation, also referred to as *monopitch*, generally gathered from speech tasks (sentences or monologues) (Jeancolas et al., 2022), (Rusz, Hlavnička, et al., 2021), (Yu et al., 2022), (Rusz et al., 2022). Along with the Cepstral domain and noise measures, features related to the F0 resulted as the most prevalent in modeling complex artifacts of the phonatory system due to hoarseness and tremor in voice. A possible physical explanation of this capability was proposed in (Rusz et al., 2018), where the authors revealed that PD speech characteristics, linked to F0 and monopitch, are related to nigral dopamine loss, especially in male subjects. In the context of F0 analysis, Jitter and Shimmer are also frequent, although different authors revealed different opinions on their effectiveness, some of them reporting efficacy (Arias-Vergara et al., 2017), (Benmalek et al., 2017), (Upadhya & Cheeran, 2018a), but some others inconsistency (Carrón et al., 2021), (Despotovic et al., 2020).

Recently, non-linear features are gaining more and more attention too (Carrón et al., 2021), (Amato, Borzi, et al., 2021), (Travieso et al., 2017), (Erdogdu Sakar et al., 2017), (Godino-Llorente et al., 2017), (L. Zhang et al., 2020), leading to parametrization sets that include RPDE, DFA, D2, HURST exponent, and Largest Lyapunov Exponent (LLE). These features are related to physiological mechanisms of the phonatory system, namely abnormal vocal fold vibration, non-linear pressure flow in the glottis, and strain in the vocal fold tissues (Benba et al., 2017). However, although they often result very effective (RPDE in particular), two of the investigated papers (Travieso et al., 2017), (L. Zhang et al., 2020) agreed on the need to include them in a broader set of features from different domains to accurately discriminate between healthy and pathological speakers.

Apart from the quality of the voice, also speech rhythms were found relevant in discriminating PD patients, as the duration of pauses between words (Jeancolas et al., 2022), (Rusz, Hlavnička, et al., 2021), (Rusz, Tykalová, et al., 2021), (Rusz et al., 2018), (Hlavnička et al., 2017), or the time rate (Rusz et al., 2018), (Hlavnička et al., 2017), (Benba et al., 2019). However, in (Benba et al., 2019), the authors did not find significant correlation between rhythmic characteristics and UPDRS staging.

Literature gaps in the assessed papers could thus be summarised with a general tendency to only use partial subsets of features, missing out on other possibly effective domains.

Comparisons between Tables 4 and 5 show no effectiveness in some features, such as NHR, GNE, FORMANTS, INTENSITY, and GQ, even if frequently employed.

5.2 Influence of covariates

As reported in Table 1, the majority of the investigated studies rely on acoustic features fed into ML- and statistical-based methods. The most commonly adopted classifier, resulting with the best performances on average, is the Support Vector Machine (SVM), and its kernel variants. This is consistent to SVM as a powerful classifier for small datasets, such as those normally employed in biomedical AI assessments, and a state-of-the-art method for audio classification.

5.2.1 Distinctive data of subjects

Subjects' distinctive data, namely gender, age, language and years from diagnosis, can determine the importance of some features with respect to others to some extent, as evidenced by some of the selected papers (Jeancolas et al., 2022), (W. Rahman et al., 2021), (Sajal et al., 2020), (Cavallieri et al., 2021), (Amato, Borzì, et al., 2021), (Rodríguez-Pérez et al., 2019). Moreover, some studies pointed out features less influenced by this aspect, such as the monopitch (Rusz et al., 2022), the Frequency Tremor (FT) and the Amplitude Tremor (AT) (Brückl et al., 2018). Interestingly, a statistical analysis performed in (Rusz et al., 2022) suggested that de-novo patients bear fewer differences between genders in PD-related voice characteristics.
Although the majority of papers do include basic information regarding gender and age of the corpora, there is a general lack in reporting crucial metadata that refer to the status of the disease, such as the years from diagnosis, the level of impairment (e.g. UPDRSIII) and the status of the treatment (details are reported in Appendix A). Several

authors proposed to include non-speech covariates to the feature sets as a kind of correction factor: with this regard, (Carrón et al., 2021), (Tougui et al., 2020), (Tunc et al., 2020), (Naranjo et al., 2021), (Galaz, Mzourek, et al., 2016), (Polat & Nour, 2020), (Naranjo, Pérez, Martín, et al., 2017) proved effective for gender, while (Tougui et al., 2020), (Tunc et al., 2020) and (L. Zhang et al., 2020) proved effective for age.

5.2.2 Recording setup and conditions

Also the recording setup and conditions play a rule on the effectiveness of ML-based voice analysis. Ideally, the setup has to be low-cost and easily available (as a smartphone can be) and the conditions have to be unsupervised (as in home environment without special operators) for making the techniques cost-effective and ubiquitous in time and space, as investigated in a work devoted to speech recognition (Hermansky & Morgan, 1994). Among all selected papers, 57 employed at least one database recorded with high-quality equipment (professional condenser microphones often with a cardioid polar pattern), and 25 used at least one database collected with smartphones (omnidirectional electret microphones) laptops or telephone recordings, in both supervised and unsupervised environments (Table 1). In (Jeancolas et al., 2022) and (Amato et al., 2021) no significant decreases in performance were observed when moving from professional microphones to low-cost equipment. The quality of the recordings in terms of medium, format and supervision, is of course a huge element in determining the veridicity of the obtained results: for this reason, studies that do not report the recording conditions for their datasets, as well as recordings that have not been supervised, or that involve self-reportings of any sort, could be safely seen as limitations. Supporting this thesis, (Carrón et al., 2021) observed a drastic reduction in performances when employing recordings under unsupervised conditions.

5.2.3 Feature extraction toolboxes

As for the feature extraction toolboxes, Table 1 reveals that Praat and DARTH-VAT are the most frequently employed features extraction toolboxes. Moreover, our review identified three studies ((Ali et al., 2021), (Vaiciukynas et al., 2017) and (Almeida et al., 2019)) that explored different feature extraction techniques, in turn addressing the feasibility of various evaluation algorithms for the same of those feature, F0 being a notable example.

It is worth noting that some papers, such as (Hlavnička et al., 2020) or (Sajal et al., 2020) reportedly employed custom routines that, albeit allowing for the extraction of peculiar indicators, are often unspecified within the papers and not publicly available, which hinders the reproducibility of the study. On the other hand, a relevant number of studies do not specify their extraction methods at all.

These works compared features extracted by means of several available toolboxes and different types of recorders (e.g., smartphone and professional microphones) included in the Vaiciukynas dataset (Vaiciukynas et al., 2017). Despite using the same corpus, these studies do not agree on the best set of parameters to employ; however, they all acknowledge that the YAAFEE toolbox (Mathieu et al., 2010) always provides good classification performance, and agree about the need of considering different sets of features when gathering vocal tasks with professional microphones with respect to smartphones.

6. Conclusion

This work overviews the current state-of-the-art methodologies for speech analysis correlated to PD conditions, with focusing on signal processing, ML aspect, and available algorithmic tools. With the aim of identifying a baseline of effective methodologies and shedding some light on the existing domains of audio analysis, we examined 102 studies published between 2017 and 2022 using ML methods or acoustic features statistics.

Our findings revealed that the mostly adopted acoustic features for PD are the same employed for common speech analysis, such as F0, Jitter, Shimmer, MFCC, HNR, and NHR, the most effective being MFCC-related. However, complexity measures such as DFA and entropy-related features (e.g., RPDE) are also frequent, followed by parametrization techniques of phonatory and glottal aspects.

According to our review, the information most linkable to the disease can be related to pitch and temporal variations in speech, mainly due to tremor and steadiness, the latter depending to the alteration of movements typical of PD subjects. Moreover, additional well-known pathological characteristics (such as hoarseness) were empirically found and validated by the presence of features like HNR or glottal model related.

Only a limited percentage of the works investigated the influence of covariates such as gender, age, or specific symptomatology, but significantly underline as the inclusion of those covariates in the analysis may increase the model accuracy. Moreover, differences in datasets and recording techniques are found relevant, the recording apparatuses influencing the audio quality and so the effectiveness of the single audio features.

With regarding to the feature extraction procedure, we found Praat and DARTH-VAT toolboxes as the most adopted. In Table 3 a comprehensive list of all freely available software and publicly available data sets.

As a point of weakness, a really ultimate baseline cannot be furnished since exhaustive information on recording techniques, most effective features, additional signal processing techniques and employed toolboxes, are reported in a minority of works only, and so the characteristics of patients and, their symptomatology. In addition, the lack of data and the usage of pre-defined datasets with pre-assembled feature vectors, such as those by UCI, may have biased to some extent the results.

7. Funding

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

8. References

Airas, M., Pulakka, H., Bäckström, T., & Alku, P. (2005). A toolkit for voice inverse filtering and parametrisation. *Interspeech 2005*, 2145–2148. https://doi.org/10.21437/Interspeech.2005-434

Ali, L., He, Z., Cao, W., Rauf, H. T., Imrana, Y., & Bin Heyat, M. B. (2021). MMDD-Ensemble: A Multimodal Data–Driven Ensemble Approach for Parkinson's Disease Detection. *Frontiers in Neuroscience*, *15*, 754058. https://doi.org/10.3389/fnins.2021.754058

Ali, L., Zhu, C., Zhang, Z., & Liu, Y. (2019). Automated Detection of Parkinson's Disease Based on Multiple Types of Sustained Phonations Using Linear Discriminant Analysis and Genetically Optimized Neural Network. *IEEE Journal of Translational Engineering in Health and Medicine*, *7*, 1–10. https://doi.org/10.1109/JTEHM.2019.2940900

Almeida, J. S., Rebouças Filho, P. P., Carneiro, T., Wei, W., Damaševičius, R., Maskeliūnas, R., & de Albuquerque, V. H. C. (2019). Detecting Parkinson's disease with sustained phonation and speech signals using machine learning techniques. *Pattern Recognition Letters*, *125*, 55–62. https://doi.org/10.1016/j.patrec.2019.04.005

Amato, F., Borzi, L., Olmo, G., Artusi, C. A., Imbalzano, G., & Lopiano, L. (2021). Speech Impairment in Parkinson's Disease: Acoustic Analysis of Unvoiced Consonants in Italian Native Speakers. *IEEE Access*, *9*, 166370–166381. https://doi.org/10.1109/ACCESS.2021.3135626

Amato, F., Borzì, L., Olmo, G., & Orozco-Arroyave, J. R. (2021). An algorithm for Parkinson's disease speech classification based on isolated words analysis. *Health Information Science and Systems*, *9*(1), 32. https://doi.org/10.1007/s13755-021-00162-8

Arias-Vergara, T., Vásquez-Correa, J. C., & Orozco-Arroyave, J. R. (2017). Parkinson's Disease and Aging: Analysis of Their Effect in Phonation and Articulation of Speech. *Cognitive Computation*, *9*(6), 731–748. https://doi.org/10.1007/s12559-017-9497-x

Arora, S., Baghai-Ravary, L., & Tsanas, A. (2019). Developing a large scale population screening tool for the assessment of Parkinson's disease using telephone-quality voice. *The Journal of the Acoustical Society of America*, *145*(5), 2871–2884. https://doi.org/10.1121/1.5100272

Arora, S., Lo, C., Hu, M., & Tsanas, A. (2021). Smartphone Speech Testing for Symptom Assessment in Rapid Eye Movement Sleep Behavior Disorder and Parkinson's Disease. *IEEE Access*, *9*, 44813–44824.

https://doi.org/10.1109/ACCESS.2021.3057715

Arora, S., & Tsanas, A. (2021). Assessing Parkinson's Disease at Scale Using Telephone-Recorded Speech: Insights from the Parkinson's Voice Initiative. *Diagnostics*, *11*(10), 1892. https://doi.org/10.3390/diagnostics11101892

Asci, F., Costantini, G., Di Leo, P., Zampogna, A., Ruoppolo, G., Berardelli, A., Saggio, G., & Suppa, A. (2020). Machine-Learning Analysis of Voice Samples Recorded through Smartphones: The Combined Effect of Ageing and Gender. *Sensors*, *20*(18), 5022. https://doi.org/10.3390/s20185022

Asci, F., Costantini, G., Saggio, G., & Suppa, A. (2021). Fostering Voice Objective Analysis in Patients with Movement Disorders. *Movement Disorders*, *36*(4), 1041–1041. https://doi.org/10.1002/mds.28537

Ashour, A. S., Nour, M. K. A., Polat, K., Guo, Y., Alsaggaf, W., & El-Attar, A. (2020). A Novel Framework of Two Successive Feature Selection Levels Using Weight-Based Procedure for Voice-Loss Detection in Parkinson's Disease. *IEEE Access*, *8*, 76193–76203. https://doi.org/10.1109/ACCESS.2020.2989032

Baumann, A., Nebel, A., Granert, O., Giehl, K., Wolff, S., Schmidt, W., Baasch, C., Schmidt, G., Witt, K., Deuschl, G., Hartwigsen, G., Zeuner, K. E., & van Eimeren, T. (2018). Neural Correlates of Hypokinetic Dysarthria and Mechanisms of Effective Voice Treatment in Parkinson Disease. *Neurorehabilitation and Neural Repair*, *32*(12), 1055–1066. https://doi.org/10.1177/1545968318812726

Benba, A., Jilbab, A., & Hammouch, A. (2017). Voice assessments for detecting patients with neurological diseases using PCA and NPCA. *International Journal of Speech Technology*, *20*(3), 673–683. https://doi.org/10.1007/s10772-017-9438-9

Benba, A., Jilbab, A., Sandabad, S., & Hammouch, A. (2019). Voice signal processing for detecting possible early signs of Parkinson's disease in patients with rapid eye movement sleep behavior disorder. *International Journal of Speech Technology*, *22*(1), 121–129. https://doi.org/10.1007/s10772-018-09588-0

Benmalek, E., Elmhamdi, J., & Jilbab, A. (2017). Multiclass classification of Parkinson's disease using different classifiers and LLBFS feature selection algorithm. *International Journal of Speech Technology*, *20*(1), 179–184. https://doi.org/10.1007/s10772-017-9401-9

Benmalek, E., Elmhamdi, J., & Jilbab, A. (2018). Multiclass classification of Parkinson's disease using cepstral analysis. *International Journal of Speech Technology*, *21*(1), 39–49. https://doi.org/10.1007/s10772-017-9485-2

Boersma, P., & van Heuven, V. (2001). *Speak and unSpeak with PRAAT*. *5*(9), 8.

Bogert, B. P. (1963a). The quefrency alanysis of time series for echoes ; Cepstrum, pseudo-autocovariance, cross-cepstrum and saphe cracking. *Time Series Analysis*, 209–243.

Bogert, B. P. (1963b). *The quefrency analysis of time series for echoes: Cepstrum, pseudo-autocovariance, cross-cepstrum, and saphe cracking*.

Bot, B. M., Suver, C., Neto, E. C., Kellen, M., Klein, A., Bare, C., Doerr, M., Pratap, A., Wilbanks, J., Dorsey, E. R., Friend, S. H., & Trister, A. D. (2016). The mPower study, Parkinson disease mobile data collected using ResearchKit. *Scientific Data*, *3*(1), 160011. https://doi.org/10.1038/sdata.2016.11

Brabenec, L., Mekyska, J., Galaz, Z., & Rektorova, I. (2017). Speech disorders in Parkinson's disease: Early diagnostics and effects of medication and brain stimulation. *Journal of Neural Transmission*, *124*(3), 303–334. https://doi.org/10.1007/s00702-017-1676-0

Brückl, M., Ghio, A., & Viallet, F. (2018). Measurement of Tremor in the Voices of Speakers with Parkinson's Disease. *Procedia Computer Science*, *128*, 47–54. https://doi.org/10.1016/j.procs.2018.03.007

Burk, B. R., & Watts, C. R. (2019). The Effect of Parkinson Disease Tremor Phenotype on Cepstral Peak Prominence and Transglottal Airflow in Vowels and Speech. *Journal of Voice*, *33*(4), 580.e11-580.e19.

https://doi.org/10.1016/j.jvoice.2018.01.016

Carrón, J., Campos-Roca, Y., Madruga, M., & Pérez, C. J. (2021). A mobile-assisted voice condition analysis system for Parkinson's disease: Assessment of usability conditions. *BioMedical Engineering OnLine*, *20*(1), 114. https://doi.org/10.1186/s12938-021-00951-y

Cavallieri, F., Budriesi, C., Gessani, A., Contardi, S., Fioravanti, V., Menozzi, E., Pinto, S., Moro, E., Valzania, F., & Antonelli, F. (2021). Dopaminergic Treatment Effects on Dysarthric Speech: Acoustic Analysis in a Cohort of Patients With Advanced Parkinson's Disease. *Frontiers in Neurology*, *11*, 616062. https://doi.org/10.3389/fneur.2020.616062

Cernak, M., Orozco-Arroyave, J. R., Rudzicz, F., Christensen, H., Vásquez-Correa, J. C., & Nöth, E. (2017). Characterisation of voice quality of Parkinson's disease using differential phonological posterior features. *Computer Speech & Language*, *46*, 196–208. https://doi.org/10.1016/j.csl.2017.06.004

Chen, O. Y., Lipsmeier, F., Phan, H., Prince, J., Taylor, K. I., Gossens, C., Lindemann, M., & Vos, M. de. (2020). Building a Machine-Learning Framework to Remotely Assess Parkinson's Disease Using Smartphones. *IEEE Transactions on Biomedical Engineering*, *67*(12), 3491–3500. https://doi.org/10.1109/TBME.2020.2988942

Costantini, G., Dr., V. C., Robotti, C., Benazzo, M., Pietrantonio, F., Di Girolamo, S., Pisani, A., Canzi, P., Mauramati, S., Bertino, G., Cassaniti, I., Baldanti, F., & Saggio, G. (2022). Deep learning and machine learning-based voice analysis for the detection of COVID-19: A proposal and comparison of architectures. *Knowledge-Based Systems*, *253*, 109539. https://doi.org/10.1016/j.knosys.2022.109539

Costantini, G., Parada-Cabaleiro, E., & Casali, D. (2021). Automatic Emotion Recognition from DEMoS Corpus by Machine Learning Analysis of Selected Vocal Features: *Proceedings of the 14th International Joint Conference on Biomedical Engineering Systems and Technologies*, 357–364. https://doi.org/10.5220/0010392503570364

Costantini, G., Parada-Cabaleiro, E., Casali, D., & Cesarini, V. (2022). The Emotion Probe: On the Universality of Cross-Linguistic and Cross-Gender Speech Emotion Recognition via Machine Learning. *Sensors*, *22*(7), Art. 7. https://doi.org/10.3390/s22072461

Demir, F., Sengur, A., Ari, A., Siddique, K., & Alswaitti, M. (2021). Feature Mapping and Deep Long Short Term Memory Network-Based Efficient Approach for Parkinson's Disease Diagnosis. *IEEE Access*, *9*, 149456–149464. https://doi.org/10.1109/ACCESS.2021.3124765

Demir, F., Siddique, K., Alswaitti, M., Demir, K., & Sengur, A. (2022). A Simple and Effective Approach Based on a Multi-Level Feature Selection for Automated Parkinson's Disease Detection. *Journal of Personalized Medicine*, *12*(1), 55. https://doi.org/10.3390/jpm12010055

Despotovic, V., Skovranek, T., & Schommer, C. (2020). Speech Based Estimation of Parkinson's Disease Using Gaussian Processes and Automatic Relevance Determination. *Neurocomputing*, *401*, 173–181. https://doi.org/10.1016/j.neucom.2020.03.058

Di Lazzaro, G., Ricci, M., Al-Wardat, M., Schirinzi, T., Scalise, S., Giannini, F., Mercuri, N. B., Saggio, G., & Pisani, A. (2020). Technology-Based Objective Measures Detect Subclinical Axial Signs in Untreated, de novo Parkinson's Disease. *Journal of Parkinson's Disease*, *10*(1), 113–122. https://doi.org/10.3233/JPD-191758

Dimauro, G., Di Nicola, V., Bevilacqua, V., Caivano, D., & Girardi, F. (2017). Assessment of Speech Intelligibility in Parkinson's Disease Using a Speech-To-Text System. *IEEE Access*, *5*, 22199–22208. https://doi.org/10.1109/ACCESS.2017.2762475

Ehlen, F., Al-Fatly, B., Kühn, A. A., & Klostermann, F. (2020). Impact of deep brain stimulation of the subthalamic nucleus on natural language in patients with Parkinson's disease. *PloS One*, *15*(12), e0244148. https://doi.org/10.1371/journal.pone.0244148

Erdogdu Sakar, B., Serbes, G., & Sakar, C. O. (2017). Analyzing the effectiveness of vocal features in early telediagnosis of Parkinson's disease. *PLOS ONE*, *12*(8), e0182428. https://doi.org/10.1371/journal.pone.0182428

Eyben, F., Scherer, K. R., Schuller, B. W., Sundberg, J., Andre, E., Busso, C., Devillers, L. Y., Epps, J., Laukka, P., Narayanan, S. S., & Truong, K. P. (2016). The Geneva Minimalistic Acoustic Parameter Set (GeMAPS) for Voice Research and Affective Computing. *IEEE Transactions on Affective Computing*, *7*(2), 190–202. https://doi.org/10.1109/TAFFC.2015.2457417

Eyben, F., Wollmer, M., & Schuller, B. (2009). OpenEAR — Introducing the munich open-source emotion and affect recognition toolkit. *2009 3rd International Conference on Affective Computing and Intelligent Interaction and Workshops*, 1–6. https://doi.org/10.1109/ACII.2009.5349350

Fayad, R., Hajj-Hassan, M., Costantini, G., Zarazadeh, Z., Errico, V., Pisani, A., Di Lazzaro, G., Ricci, M., & Saggio, G. (2021). Vocal test Analysis for Assessing Parkinson's Disease at Early Stage. *2021 Sixth International Conference on Advances in Biomedical Engineering (ICABME)*, 171–174. https://doi.org/10.1109/ICABME53305.2021.9604891

Gaballah, A., Parsa, V., Cushnie-Sparrow, D., & Adams, S. (2021). Improved Estimation of Parkinsonian Vowel Quality through Acoustic Feature Assimilation. *The Scientific World Journal*, *2021*, 1–11. https://doi.org/10.1155/2021/6076828

Galáž, Z. (2016a). Potential Of Prosodic Features To Estimate Degree Of Parkinson's Disease Severity. *Proceedings of the 22nd Conference STUDENT EEICT 2016*. s. 533-537. ISBN 978-80-214-5350-05.

Galáž, Z. (2016b). Quantification Of Prosodic Impairment In Patients With Idiopathic Parkinson's Disease. *Proceedings of the 22nd Conference STUDENT EEICT 2016*. ISBN 978-80-214-5350-05.

Galaz, Z., Mekyska, J., Mzourek, Z., Smekal, Z., Rektorova, I., Eliasova, I., Kostalova, M., Mrackova, M., & Berankova, D. (2016). Prosodic analysis of neutral, stress-modified and rhymed speech in patients with Parkinson's disease. *Computer Methods and Programs in Biomedicine*, *127*, 301–317. https://doi.org/10.1016/j.cmpb.2015.12.011

Galaz, Z., Mekyska, J., Zvoncak, V., Mucha, J., Kiska, T., Smekal, Z., Eliasova, I., Mrackova, M., Kostalova, M., Rektorova, I., Faundez-Zanuy, M., Alonso-Hernandez, J., & Gomez-Vilda, P. (2018). Changes in Phonation and Their Relations with Progress of Parkinson's Disease. *Applied Sciences*, *8*(12), 2339. https://doi.org/10.3390/app8122339

Galaz, Z., Mzourek, Z., Mekyska, J., Smekal, Z., Kiska, T., Rektorova, I., Orozco-Arroyave, J. R., & Daoudi, K. (2016). Degree of Parkinson's disease severity estimation based on speech signal processing. *2016 39th International Conference on Telecommunications and Signal Processing (TSP)*, 503–506. https://doi.org/10.1109/TSP.2016.7760930

Gandhi, K. R., & Saadabadi, A. (2022). Levodopa (L-Dopa). In *StatPearls*. StatPearls Publishing. http://www.ncbi.nlm.nih.gov/books/NBK482140/

García, A. M., Arias- Vergara, T., Vasquez- Correa, J., Nöth, E., Schuster, M., Welch, A. E., Bocanegra, Y., Baena, A., & Orozco- Arroyave, J. R. (2021). Cognitive Determinants of Dysarthria in Parkinson's Disease: An Automated Machine Learning Approach. *Movement Disorders*, *36*(12), 2862–2873. https://doi.org/10.1002/mds.28751

Gelb, D. J., Oliver, E., & Gilman, S. (1999). Diagnostic Criteria for Parkinson Disease. *Archives of Neurology*, *56*(1), 33–39. https://doi.org/10.1001/archneur.56.1.33

Giannakopoulos, T. (2015). pyAudioAnalysis: An Open-Source Python Library for Audio Signal Analysis. *PLOS ONE*, *10*(12), e0144610. https://doi.org/10.1371/journal.pone.0144610

Gillivan-Murphy, P., Miller, N., & Carding, P. (2019). Voice Tremor in Parkinson's Disease: An Acoustic Study.

Journal of Voice, *33*(4), 526–535. https://doi.org/10.1016/j.jvoice.2017.12.010

Godino-Llorente, J. I., Shattuck-Hufnagel, S., Choi, J. Y., Moro-Velázquez, L., & Gómez-García, J. A. (2017). Towards the identification of Idiopathic Parkinson's Disease from the speech. New articulatory kinetic biomarkers. *PLOS ONE*, *12*(12), e0189583. https://doi.org/10.1371/journal.pone.0189583

Gómez, P., Mekyska, J., Gómez, A., Palacios, D., Rodellar, V., & Álvarez, A. (2017). Monitoring Parkinson Disease from speech articulation kinematics. *Loquens*, *4*(1), 036. https://doi.org/10.3989/loquens.2017.036

Gómez, P., Rodellar, V., Nieto, V., Álvarez, A., Scola, B., Ramírez, C., Poletti, D., & Fernández, M. (2013). *BioMet®Phon: A System to Monitor Phonation Quality in the Clinics*. 6.

Gómez-García, J. A., Moro-Velázquez, L., Arias-Londoño, J. D., & Godino-Llorente, J. I. (2021). On the design of automatic voice condition analysis systems. Part III: Review of acoustic modelling strategies. *Biomedical Signal Processing and Control*, *66*, 102049. https://doi.org/10.1016/j.bspc.2020.102049

Gómez-Vilda, P., Galaz, Z., Mekyska, J., Vicente, J. M. F., Gómez-Rodellar, A., Palacios-Alonso, D., Smekal, Z., Eliasova, I., Kostalova, M., & Rektorova, I. (2019). Vowel Articulation Dynamic Stability Related to Parkinson's Disease Rating Features: Male Dataset. *International Journal of Neural Systems*, *29*(02), 1850037. https://doi.org/10.1142/S0129065718500375

Goyal, J., Khandnor, P., & Aseri, T. C. (2021). A Comparative Analysis of Machine Learning classifiers for Dysphonia-based classification of Parkinson's Disease. *International Journal of Data Science and Analytics*, *11*(1), 69–83. https://doi.org/10.1007/s41060-020-00234-0

Gunduz, H. (2019). Deep Learning-Based Parkinson's Disease Classification Using Vocal Feature Sets. *IEEE Access*, *7*, 115540–115551. https://doi.org/10.1109/ACCESS.2019.2936564

Gunduz, H. (2021). An efficient dimensionality reduction method using filter-based feature selection and variational autoencoders on Parkinson's disease classification. *Biomedical Signal Processing and Control*, *66*, 102452. https://doi.org/10.1016/j.bspc.2021.102452

Han, J., Brown, C., Chauhan, J., Grammenos, A., Hasthanasombat, A., Spathis, D., Xia, T., Cicuta, P., & Mascolo, C. (2021). Exploring Automatic COVID-19 Diagnosis via Voice and Symptoms from Crowdsourced Data. *ICASSP 2021 - 2021 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, 8328–8332. https://doi.org/10.1109/ICASSP39728.2021.9414576

Hegger, R., Kantz, H., & Schreiber, T. (1999). Practical implementation of nonlinear time series methods: The TISEAN package. *Chaos: An Interdisciplinary Journal of Nonlinear Science*, *9*(2), 413–435. https://doi.org/10.1063/1.166424

Henriquez, P., Alonso, J. B., Ferrer, M. A., Travieso, C. M., Godino-Llorente, J. I., & Diaz-de-Maria, F. (2009). Characterization of Healthy and Pathological Voice Through Measures Based on Nonlinear Dynamics. *IEEE Transactions on Audio, Speech, and Language Processing*, *17*(6), 1186–1195. https://doi.org/10.1109/TASL.2009.2016734

Hermansky, H., & Morgan, N. (1994). RASTA processing of speech. *IEEE Transactions on Speech and Audio Processing*, *2*(4), 578–589. https://doi.org/10.1109/89.326616

Hlavnička, J., Čmejla, R., Tykalová, T., Šonka, K., Růžička, E., & Rusz, J. (2017). Automated analysis of connected speech reveals early biomarkers of Parkinson's disease in patients with rapid eye movement sleep behaviour disorder. *Scientific Reports*, *7*(1), 12. https://doi.org/10.1038/s41598-017-00047-5

Hlavnička, J., Tykalová, T., Ulmanová, O., Dušek, P., Horáková, D., Růžička, E., Klempíř, J., & Rusz, J. (2020). Characterizing vocal tremor in progressive neurological diseases via automated acoustic analyses. *Clinical Neurophysiology*, *131*(5), 1155–1165. https://doi.org/10.1016/j.clinph.2020.02.005

Hoq, M., Uddin, M. N., & Park, S.-B. (2021). Vocal Feature Extraction-Based Artificial Intelligent Model for Parkinson's Disease Detection. *Diagnostics*, *11*(6), 1076. https://doi.org/10.3390/diagnostics11061076

Issa, D., Fatih Demirci, M., & Yazici, A. (2020). Speech emotion recognition with deep convolutional neural networks. *Biomedical Signal Processing and Control*, *59*, 101894. https://doi.org/10.1016/j.bspc.2020.101894

Jadoul, Y., Thompson, B., & de Boer, B. (2018). Introducing Parselmouth: A Python interface to Praat. *Journal of Phonetics*, *71*, 1–15. https://doi.org/10.1016/j.wocn.2018.07.001

Jeancolas, L., Mangone, G., Petrovska-Delacrétaz, D., Benali, H., Benkelfat, B.-E., Arnulf, I., Corvol, J.-C., Vidailhet, M., & Lehéricy, S. (2022). Voice characteristics from isolated rapid eye movement sleep behavior disorder to early Parkinson's disease. *Parkinsonism & Related Disorders*, *95*, 86–91. https://doi.org/10.1016/j.parkreldis.2022.01.003

Jenei, A. Z., Kiss, G., Tulics, M. G., & Sztahó, D. (2021). Separation of Several Illnesses Using Correlation Structures with Convolutional Neural Networks. *Acta Polytechnica Hungarica*, *18*(7), 47–66. https://doi.org/10.12700/APH.18.7.2021.7.3

Kacha, A., Mertens, C., Grenez, F., Skodda, S., & Schoentgen, J. (2017). On the harmonic-to-noise ratio as an acoustic cue of vocal timbre of Parkinson speakers. *Biomedical Signal Processing and Control*, *37*, 32–38. https://doi.org/10.1016/j.bspc.2016.09.004

Kaplan, E., Altunisik, E., Ekmekyapar Firat, Y., Datta Barua, P., Dogan, S., Baygin, M., Burak Demir, F., Tuncer, T., Palmer, E., Tan, R.-S., Yu, P., Soar, J., Fujita, H., & Rajendra Acharya, U. (2022). Novel nested patch-based feature extraction model for automated Parkinson's Disease symptom classification using MRI images. *Computer Methods and Programs in Biomedicine*, *224*, 107030. https://doi.org/10.1016/j.cmpb.2022.107030

Karabayir, I., Goldman, S. M., Pappu, S., & Akbilgic, O. (2020). Gradient boosting for Parkinson's disease diagnosis from voice recordings. *BMC Medical Informatics and Decision Making*, *20*(1), 228. https://doi.org/10.1186/s12911-020-01250-7

Karan, B., Sahu, S. S., Orozco-Arroyave, J. R., & Mahto, K. (2020). Hilbert spectrum analysis for automatic detection and evaluation of Parkinson's speech. *Biomedical Signal Processing and Control*, *61*, 102050. https://doi.org/10.1016/j.bspc.2020.102050

Karan, B., Sahu, S. S., Orozco-Arroyave, J. R., & Mahto, K. (2021). Non-negative matrix factorization-based time-frequency feature extraction of voice signal for Parkinson's disease prediction. *Computer Speech & Language*, *69*, 101216. https://doi.org/10.1016/j.csl.2021.101216

Karan, B., & Sekhar Sahu, S. (2021). An improved framework for Parkinson's disease prediction using Variational Mode Decomposition-Hilbert spectrum of speech signal. *Biocybernetics and Biomedical Engineering*, *41*(2), 717–732. https://doi.org/10.1016/j.bbe.2021.04.014

Karlsson, F., & Hartelius, L. (2019). How Well Does Diadochokinetic Task Performance Predict Articulatory Imprecision? Differentiating Individuals with Parkinson's Disease from Control Subjects. *Folia Phoniatrica et Logopaedica*, *71*(5–6), 251–260. https://doi.org/10.1159/000498851

Klobusiakova, P., Mekyska, J., Brabenec, L., Galaz, Z., Zvoncak, V., Mucha, J., Rapcsak, S. Z., & Rektorova, I. (2021). Articulatory network reorganization in Parkinson's disease as assessed by multimodal MRI and acoustic measures. *Parkinsonism & Related Disorders*, *84*, 122–128. https://doi.org/10.1016/j.parkreldis.2021.02.012

Kodrasi, I., & Bourlard, H. (2020). Spectro-Temporal Sparsity Characterization for Dysarthric Speech Detection. *IEEE/ACM Transactions on Audio, Speech, and Language Processing*, *28*, 1210–1222. https://doi.org/10.1109/TASLP.2020.2985066

Kouba, T., Illner, V., & Rusz, J. (2022). Study protocol for using a smartphone application to investigate speech biomarkers of Parkinson's disease and other synucleinopathies: SMARTSPEECH. *BMJ Open*, *12*(6), e059871. https://doi.org/10.1136/bmjopen-2021-059871

Lahmiri, S., Dawson, D. A., & Shmuel, A. (2018). Performance of machine learning methods in diagnosing Parkinson's disease based on dysphonia measures. *Biomedical Engineering Letters*, *8*(1), 29–39. https://doi.org/10.1007/s13534-017-0051-2

Lahmiri, S., & Shmuel, A. (2019). Detection of Parkinson's disease based on voice patterns ranking and optimized support vector machine. *Biomedical Signal Processing and Control*, *49*, 427–433. https://doi.org/10.1016/j.bspc.2018.08.029

Lamba, R., Gulati, T., & Jain, A. (2022). A Hybrid Feature Selection Approach for Parkinson's Detection Based on Mutual Information Gain and Recursive Feature Elimination. *Arabian Journal for Science and Engineering*. https://doi.org/10.1007/s13369-021-06544-0

Little, M. A., McSharry, P. E., Hunter, E. J., Spielman, J., & Ramig, L. O. (2009). Suitability of Dysphonia Measurements for Telemonitoring of Parkinson's Disease. *IEEE Transactions on Biomedical Engineering*, *56*(4), 1015–1022. https://doi.org/10.1109/TBME.2008.2005954

Liu, Y., Li, Y., Tan, X., Wang, P., & Zhang, Y. (2021). Local discriminant preservation projection embedded ensemble learning based dimensionality reduction of speech data of Parkinson's disease. *Biomedical Signal Processing and Control*, *63*, 102165. https://doi.org/10.1016/j.bspc.2020.102165

Liu, Y., Penttilä, N., Ihalainen, T., Lintula, J., Convey, R., & Räsänen, O. (2021). Language-Independent Approach for Automatic Computation of Vowel Articulation Features in Dysarthric Speech Assessment. *IEEE/ACM Transactions on Audio, Speech, and Language Processing*, *29*, 2228–2243. https://doi.org/10.1109/TASLP.2021.3090973

Loh, H. W., Ooi, C. P., Palmer, E., Barua, P. D., Dogan, S., Tuncer, T., Baygin, M., & Acharya, U. R. (2021). GaborPDNet: Gabor Transformation and Deep Neural Network for Parkinson's Disease Detection Using EEG Signals. *Electronics*, *10*(14), Art. 14. https://doi.org/10.3390/electronics10141740

Ma, A., Lau, K. K., & Thyagarajan, D. (2020). Voice changes in Parkinson's disease: What are they telling us? *Journal of Clinical Neuroscience*, *72*, 1–7. https://doi.org/10.1016/j.jocn.2019.12.029

Ma, J., Zhang, Y., Li, Y., Zhou, L., Qin, L., Zeng, Y., Wang, P., & Lei, Y. (2021). Deep dual-side learning ensemble model for Parkinson speech recognition. *Biomedical Signal Processing and Control*, *69*, 102849. https://doi.org/10.1016/j.bspc.2021.102849

Masud, M., Singh, P., Gaba, G. S., Kaur, A., Alghamdi, R. A., Alrashoud, M., & Alqahtani, S. A. (2021). CROWD: Crow Search and Deep Learning based Feature Extractor for Classification of Parkinson's Disease. *ACM Transactions on Internet Technology*, *21*(3), 1–18. https://doi.org/10.1145/3418500

Mathieu, B., Essid, S., Fillon, T., Prado, J., & Richard, G. (2010). YAAFE, An Easy To Use And Efficient Audio Feature Extraction Software. *International Society for Music Information Retrieval Conference, 6*.

Montaña, D., Campos-Roca, Y., & Pérez, C. J. (2018). A Diadochokinesis-based expert system considering articulatory features of plosive consonants for early detection of Parkinson's disease. *Computer Methods and Programs in Biomedicine*, *154*, 89–97. https://doi.org/10.1016/j.cmpb.2017.11.010

Moro-Velazquez, L., Gomez-Garcia, J. A., Arias-Londoño, J. D., Dehak, N., & Godino-Llorente, J. I. (2021). Advances in Parkinson's Disease detection and assessment using voice and speech: A review of the articulatory and phonatory aspects. *Biomedical Signal Processing and Control*, *66*, 102418. https://doi.org/10.1016/j.bspc.2021.102418

Moro-Velázquez, L., Gómez-García, J. A., Godino-Llorente, J. I., Villalba, J., Orozco-Arroyave, J. R., & Dehak, N.

(2018). Analysis of speaker recognition methodologies and the influence of kinetic changes to automatically detect Parkinson's Disease. *Applied Soft Computing*, *62*, 649–666. https://doi.org/10.1016/j.asoc.2017.11.001

Mulfari, D., Meoni, G., Marini, M., & Fanucci, L. (2021). Machine learning assistive application for users with speech disorders. *Applied Soft Computing*, *103*, 107147. https://doi.org/10.1016/j.asoc.2021.107147

Nahar, N., Ara, F., Neloy, Md. A. I., Biswas, A., Hossain, M. S., & Andersson, K. (2021). Feature Selection Based Machine Learning to Improve Prediction of Parkinson Disease. In M. Mahmud, M. S. Kaiser, S. Vassanelli, Q. Dai, & N. Zhong (A c. Di), *Brain Informatics* (Vol. 12960, pp. 496–508). Springer International Publishing. https://doi.org/10.1007/978-3-030-86993-9_44

Naranjo, L., Pérez, C. J., & Campos-Roca, Y. (2021). Monitoring Parkinson's disease progression based on recorded speech with missing ordinal responses and replicated covariates. *Computers in Biology and Medicine*, *134*, 104503. https://doi.org/10.1016/j.compbiomed.2021.104503

Naranjo, L., Pérez, C. J., Campos-Roca, Y., & Martín, J. (2016). Addressing voice recording replications for Parkinson's disease detection. *Expert Systems with Applications*, *46*, 286–292. https://doi.org/10.1016/j.eswa.2015.10.034

Naranjo, L., Pérez, C. J., & Martín, J. (2017). Addressing voice recording replications for tracking Parkinson's disease progression. *Medical & Biological Engineering & Computing*, *55*(3), 365–373. https://doi.org/10.1007/s11517-016-1512-y

Naranjo, L., Pérez, C. J., Martín, J., & Campos-Roca, Y. (2017). A two-stage variable selection and classification approach for Parkinson's disease detection by using voice recording replications. *Computer Methods and Programs in Biomedicine*, *142*, 147–156. https://doi.org/10.1016/j.cmpb.2017.02.019

Narendra, N. P., Schuller, B., & Alku, P. (2021). The Detection of Parkinson's Disease From Speech Using Voice Source Information. *IEEE/ACM Transactions on Audio, Speech, and Language Processing*, *29*, 1925–1936. https://doi.org/10.1109/TASLP.2021.3078364

Orozco-Arroyave, J. R., Arias-Londoño, J. D., Vargas-Bonilla, J. F., & Nöth, E. (2013). Analysis of Speech from People with Parkinson's Disease through Nonlinear Dynamics. In T. Drugman & T. Dutoit (A c. Di), *Advances in Nonlinear Speech Processing* (Vol. 7911, pp. 112–119). Springer Berlin Heidelberg. https://doi.org/10.1007/978-3-642-38847-7_15

Orozco-Arroyave, J. R., Vásquez-Correa, J. C., Vargas-Bonilla, J. F., Arora, R., Dehak, N., Nidadavolu, P. S., Christensen, H., Rudzicz, F., Yancheva, M., Chinaei, H., Vann, A., Vogler, N., Bocklet, T., Cernak, M., Hannink, J., & Nöth, E. (2018). NeuroSpeech: An open-source software for Parkinson's speech analysis. *Digital Signal Processing*, *77*, 207–221. https://doi.org/10.1016/j.dsp.2017.07.004

Oung, Q. W., Muthusamy, H., Basah, S. N., Lee, H., & Vijean, V. (2018). Empirical Wavelet Transform Based Features for Classification of Parkinson's Disease Severity. *Journal of Medical Systems*, *42*(2), 29. https://doi.org/10.1007/s10916-017-0877-2

Palacios-Alonso, D., Melendez-Morales, G., Lopez-Arribas, A., Lazaro-Carrascosa, C., Gomez-Rodellar, A., & Gomez-Vilda, P. (2020). MonParLoc: A Speech-Based System for Parkinson's Disease Analysis and Monitoring. *IEEE Access*, *8*, 188243–188255. https://doi.org/10.1109/ACCESS.2020.3031646

Parada-Cabaleiro, E., Costantini, G., Batliner, A., Schmitt, M., & Schuller, B. W. (2020). DEMoS: An Italian emotional speech corpus: Elicitation methods, machine learning, and perception. *Language Resources and Evaluation*, *54*(2), 341–383. https://doi.org/10.1007/s10579-019-09450-y

Parisi, L., RaviChandran, N., & Manaog, M. L. (2018). Feature-driven machine learning to improve early diagnosis of Parkinson's disease. *Expert Systems with Applications*, *110*, 182–190. https://doi.org/10.1016/j.eswa.2018.06.003

Parisi, L., Zaernia, A., Ma, R., & Youseffi, M. (2021). M-ark-Support Vector Machine for Early Detection of Parkinson's Disease from Speech Signals. *International Journal of Mathematics and Computers in Simulation*, *15*, 34–41. https://doi.org/10.46300/9102.2021.15.7

Plouvier, A. O. A., Olde Hartman, T. C., Boots, L. P. W., Bloem, B. R., van Weel, C., & Lagro-Janssen, A. L. M. (2015). Time intervals in diagnosing Parkinson's disease: The patients' views. *Patient Education and Counseling*, *98*(6), 777–782. https://doi.org/10.1016/j.pec.2015.02.010

Polat, K., & Nour, M. (2020). Parkinson disease classification using one against all based data sampling with the acoustic features from the speech signals. *Medical Hypotheses*, *140*, 109678. https://doi.org/10.1016/j.mehy.2020.109678

Pramanik, M., Pradhan, R., Nandy, P., Bhoi, A. K., & Barsocchi, P. (2022). The ForEx++ based decision tree ensemble approach for robust detection of Parkinson's disease. *Journal of Ambient Intelligence and Humanized Computing*. https://doi.org/10.1007/s12652-022-03719-x

Prince, J., Andreotti, F., & De Vos, M. (2019). Multi-Source Ensemble Learning for the Remote Prediction of Parkinson's Disease in the Presence of Source-Wise Missing Data. *IEEE Transactions on Biomedical Engineering*, *66*(5), 1402–1411. https://doi.org/10.1109/TBME.2018.2873252

Prince, J., Arora, S., & de Vos, M. (2018). Big data in Parkinson's disease: Using smartphones to remotely detect longitudinal disease phenotypes. *Physiological Measurement*, *39*(4), 044005. https://doi.org/10.1088/1361-6579/aab512

Quan, C., Ren, K., & Luo, Z. (2021). A Deep Learning Based Method for Parkinson's Disease Detection Using Dynamic Features of Speech. *IEEE Access*, *9*, 10239–10252. https://doi.org/10.1109/ACCESS.2021.3051432

Rahman, A., Rizvi, S. S., Khan, A., Afzaal Abbasi, A., Khan, S. U., & Chung, T.-S. (2021). Parkinson's Disease Diagnosis in Cepstral Domain Using MFCC and Dimensionality Reduction with SVM Classifier. *Mobile Information Systems*, *2021*, 1–10. https://doi.org/10.1155/2021/8822069

Rahman, W., Lee, S., Islam, M. S., Antony, V. N., Ratnu, H., Ali, M. R., Mamun, A. A., Wagner, E., Jensen-Roberts, S., Waddell, E., Myers, T., Pawlik, M., Soto, J., Coffey, M., Sarkar, A., Schneider, R., Tarolli, C., Lizarraga, K., Adams, J., … Hoque, E. (2021). Detecting Parkinson Disease Using a Web-Based Speech Task: Observational Study. *Journal of Medical Internet Research*, *23*(10), e26305. https://doi.org/10.2196/26305

Reddy, M. K., & Alku, P. (2021). A Comparison of Cepstral Features in the Detection of Pathological Voices by Varying the Input and Filterbank of the Cepstrum Computation. *IEEE Access*, *9*, 135953–135963. https://doi.org/10.1109/ACCESS.2021.3117665

Ricci, M., Di Lazzaro, G., Pisani, A., Mercuri, N. B., Giannini, F., & Saggio, G. (2020). Assessment of Motor Impairments in Early Untreated Parkinson's Disease Patients: The Wearable Electronics Impact. *IEEE Journal of Biomedical and Health Informatics*, *24*(1), 120–130. https://doi.org/10.1109/JBHI.2019.2903627

Ricci, M., Di Lazzaro, G., Pisani, A., Scalise, S., Alwardat, M., Salimei, C., Giannini, F., & Saggio, G. (2019). Wearable Electronics Assess the Effectiveness of Transcranial Direct Current Stimulation on Balance and Gait in Parkinson's Disease Patients. *Sensors*, *19*(24), 5465. https://doi.org/10.3390/s19245465

Ricci, M., Lazzaro, G. D., Errico, V., Pisani, A., Giannini, F., & Saggio, G. (2022). The Impact of Wearable Electronics in Assessing the Effectiveness of Levodopa Treatment in Parkinson's Disease. *IEEE Journal of Biomedical and Health Informatics*, *26*(7), 2920–2928. https://doi.org/10.1109/JBHI.2022.3160103

Robotti, C., Costantini, G., Saggio, G., Cesarini, V., Calastri, A., Maiorano, E., Piloni, D., Perrone, T., Sabatini, U., Ferretti, V. V., Cassaniti, I., Baldanti, F., Gravina, A., Sakib, A., Alessi, E., Pietrantonio, F., Pascucci, M., Casali, D., Zarezadeh, Z., … Benazzo, M. (2021). Machine Learning-based Voice Assessment for the Detection of Positive and Recovered COVID-19 Patients. *Journal of Voice*. https://doi.org/10.1016/j.jvoice.2021.11.004

Rodríguez-Pérez, P., Fraile, R., García-Escrig, M., Sáenz-Lechón, N., Gutiérrez-Arriola, J. M., & Osma-Ruiz, V. (2019). A transversal study of fundamental frequency contours in parkinsonian voices. *Biomedical Signal Processing and Control*, *51*, 374–381. https://doi.org/10.1016/j.bspc.2019.02.021

Rusz, J., Hlavnička, J., Novotný, M., Tykalová, T., Pelletier, A., Montplaisir, J., Gagnon, J., Dušek, P., Galbiati, A., Marelli, S., Timm, P. C., Teigen, L. N., Janzen, A., Habibi, M., Stefani, A., Holzknecht, E., Seppi, K., Evangelista, E., Rassu, A. L., … Šonka, K. (2021). Speech Biomarkers in Rapid Eye Movement Sleep Behavior Disorder and Parkinson Disease. *Annals of Neurology*, *90*(1), 62–75. https://doi.org/10.1002/ana.26085

Rusz, J., Hlavnicka, J., Tykalova, T., Novotny, M., Dusek, P., Sonka, K., & Ruzicka, E. (2018). Smartphone Allows Capture of Speech Abnormalities Associated With High Risk of Developing Parkinson's Disease. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, *26*(8), 1495–1507. https://doi.org/10.1109/TNSRE.2018.2851787

Rusz, J., Tykalová, T., Novotný, M., Růžička, E., & Dušek, P. (2021). Distinct patterns of speech disorder in early-onset and late-onset de-novo Parkinson's disease. *Npj Parkinson's Disease*, *7*(1), 98. https://doi.org/10.1038/s41531-021-00243-1

Rusz, J., Tykalová, T., Novotný, M., Zogala, D., Růžička, E., & Dušek, P. (2022). Automated speech analysis in early untreated Parkinson's disease: Relation to gender and dopaminergic transporter imaging. *European Journal of Neurology*, *29*(1), 81–90. https://doi.org/10.1111/ene.15099

Rusz, J., Tykalova, T., Ramig, L. O., & Tripoliti, E. (2021). Guidelines for Speech Recording and Acoustic Analyses in Dysarthrias of Movement Disorders. *Movement Disorders*, *36*(4), 803–814. https://doi.org/10.1002/mds.28465

Saggio, G. (2020). Are Sensors and Data Processing Paving the Way to Completely Non-invasive and Not-painful Medical Tests for Widespread Screening and Diagnosis Purposes?: *Proceedings of the 13th International Joint Conference on Biomedical Engineering Systems and Technologies*, 207–214. https://doi.org/10.5220/0009098002070214

Saggio, G., & Costantini, G. (2020). Worldwide Healthy Adult Voice Baseline Parameters: A Comprehensive Review. *Journal of Voice*, S0892199720303283. https://doi.org/10.1016/j.jvoice.2020.08.028

Sahandi Far, M., Eickhoff, S. B., Goni, M., & Dukart, J. (2021). Exploring Test-Retest Reliability and Longitudinal Stability of Digital Biomarkers for Parkinson Disease in the m-Power Data Set: Cohort Study. *Journal of Medical Internet Research*, *23*(9), e26608. https://doi.org/10.2196/26608

Sajal, Md. S. R., Ehsan, Md. T., Vaidyanathan, R., Wang, S., Aziz, T., & Mamun, K. A. A. (2020). Telemonitoring Parkinson's disease using machine learning by combining tremor and voice analysis. *Brain Informatics*, *7*(1), 12. https://doi.org/10.1186/s40708-020-00113-1

Sakar, B. E., Isenkul, M. E., Sakar, C. O., Sertbas, A., Gurgen, F., Delil, S., Apaydin, H., & Kursun, O. (2013). Collection and Analysis of a Parkinson Speech Dataset With Multiple Types of Sound Recordings. *IEEE Journal of Biomedical and Health Informatics*, *17*(4), 828–834. https://doi.org/10.1109/JBHI.2013.2245674

Sakar, C. O., Serbes, G., Gunduz, A., Tunc, H. C., Nizam, H., Sakar, B. E., Tutuncu, M., Aydin, T., Isenkul, M. E., & Apaydin, H. (2019). A comparative analysis of speech signal processing algorithms for Parkinson's disease classification and the use of the tunable Q-factor wavelet transform. *Applied Soft Computing*, *74*, 255–263. https://doi.org/10.1016/j.asoc.2018.10.022

Saxon, M., Tripathi, A., Jiao, Y., Liss, J., & Berisha, V. (2020). Robust Estimation of Hypernasality in Dysarthria with Acoustic Model Likelihood Features. *IEEE/ACM Transactions on Audio, Speech, and Language Processing*, *28*, 2511–2522. https://doi.org/10.1109/TASLP.2020.3015035

Schuller, B., Eyben, F., & Rigoll, G. (2007). Fast and Robust Meter and Tempo Recognition for the Automatic

Discrimination of Ballroom Dance Styles. *2007 IEEE International Conference on Acoustics, Speech and Signal Processing - ICASSP '07*, I-217-I–220. https://doi.org/10.1109/ICASSP.2007.366655

Schuller, B., Steidl, S., & Batliner, A. (2009). The INTERSPEECH 2009 Emotion Challenge. *Proc. Interspeech. 312-315.*

Sheibani, R., Nikookar, E., & Alavi, S. (2019). An ensemble method for diagnosis of Parkinson's disease based on voice measurements. *Journal of Medical Signals & Sensors*, *9*(4), 221. https://doi.org/10.4103/jmss.JMSS_57_18

Shue, Y.-L. (2010). The voice source in speech production: Data, analysis and models. *Doctoral Dissertation from the University of California, Los Angeles.*

Shue, Y.-L., Keating, P., & Vicenik, C. (2009). VOICESAUCE: A program for voice analysis. *The Journal of the Acoustical Society of America*, *126*(4), 2221. https://doi.org/10.1121/1.3248865

Šimek, M., & Rusz, J. (2021). Validation of cepstral peak prominence in assessing early voice changes of Parkinson's disease: Effect of speaking task and ambient noise. *The Journal of the Acoustical Society of America*, *150*(6), 4522–4533. https://doi.org/10.1121/10.0009063

Singh, S., & Xu, W. (2020). Robust Detection of Parkinson's Disease Using Harvested Smartphone Voice Data: A Telemedicine Approach. *Telemedicine and E-Health*, *26*(3), 327–334. https://doi.org/10.1089/tmj.2018.0271

Solana-Lavalle, G., & Rosas-Romero, R. (2021). Analysis of voice as an assisting tool for detection of Parkinson's disease and its subsequent clinical interpretation. *Biomedical Signal Processing and Control*, *66*, 102415. https://doi.org/10.1016/j.bspc.2021.102415

Suphinnapong, P., Phokaewvarangkul, O., Thubthong, N., Teeramongkonrasmee, A., Mahattanasakul, P., Lorwattanapongsa, P., & Bhidayasiri, R. (2021). Objective vowel sound characteristics and their relationship with motor dysfunction in Asian Parkinson's disease patients. *Journal of the Neurological Sciences*, *426*, 117487. https://doi.org/10.1016/j.jns.2021.117487

Suppa, A., Asci, F., Saggio, G., Di Leo, P., Zarezadeh, Z., Ferrazzano, G., Ruoppolo, G., Berardelli, A., & Costantini, G. (2021). Voice Analysis with Machine Learning: One Step Closer to an Objective Diagnosis of Essential Tremor. *Movement Disorders*, *36*(6), 1401–1410. https://doi.org/10.1002/mds.28508

Suppa, A., Asci, F., Saggio, G., Marsili, L., Casali, D., Zarezadeh, Z., & Costantini, G. (2020). Voice analysis in adductor spasmodic dysphonia: Objective diagnosis and response to botulinum toxin. *Parkinsonism & Related Disorders, Elsevier*.

Suppa, A., Costantini, G., Asci, F., Di Leo, P., Al-Wardat, M. S., Di Lazzaro, G., Scalise, S., Pisani, A., & Saggio, G. (2022). Voice in Parkinson's Disease: A Machine Learning Study. *Frontiers in Neurology*, *13*, 831428. https://doi.org/10.3389/fneur.2022.831428

Švec, J. G., & Granqvist, S. (2010). Guidelines for Selecting Microphones for Human Voice Production Research. *American Journal of Speech-Language Pathology*, *19*(4), 356–368. https://doi.org/10.1044/1058-0360(2010/09-0091)

Tăuţan, A.-M., Ionescu, B., & Santarnecchi, E. (2021). Artificial intelligence in neurodegenerative diseases: A review of available tools with a focus on machine learning techniques. *Artificial Intelligence in Medicine*, *117*, 102081. https://doi.org/10.1016/j.artmed.2021.102081

Tougui, I., Jilbab, A., & Mhamdi, J. E. (2020). Analysis of Smartphone Recordings in Time, Frequency, and Cepstral Domains to Classify Parkinson's Disease. *Healthcare Informatics Research*, *26*(4), 274–283. https://doi.org/10.4258/hir.2020.26.4.274

Tracy, J. M., Özkanca, Y., Atkins, D. C., & Hosseini Ghomi, R. (2020). Investigating voice as a biomarker: Deep phenotyping methods for early detection of Parkinson's disease. *Journal of Biomedical Informatics*, *104*, 103362.

https://doi.org/10.1016/j.jbi.2019.103362

Travieso, C. M., Alonso, J. B., Orozco-Arroyave, J. R., Vargas-Bonilla, J. F., Nöth, E., & Ravelo-García, A. G. (2017). Detection of different voice diseases based on the nonlinear characterization of speech signals. *Expert Systems with Applications*, *82*, 184–195. https://doi.org/10.1016/j.eswa.2017.04.012

Tsanas, A., Little, M. A., McSharry, P. E., & Ramig, L. O. (2010). Accurate Telemonitoring of Parkinson's Disease Progression by Noninvasive Speech Tests. *IEEE Transactions on Biomedical Engineering*, *57*(4), 884–893. https://doi.org/10.1109/TBME.2009.2036000

Tsanas, A., Little, M. A., McSharry, P. E., & Ramig, L. O. (2011). Nonlinear speech analysis algorithms mapped to a standard metric achieve clinically useful quantification of average Parkinson's disease symptom severity. *Journal of The Royal Society Interface*, *8*(59), 842–855. https://doi.org/10.1098/rsif.2010.0456

Tsanas, A. (2012). Accurate telemonitoring of Parkinson's disease symptom severity using nonlinear speech signal processing and statistical machine learning. Doctoral Thesis from the Oxford University.

Tulics, M. G., & Vicsi, K. (2018). Automatic classification possibilities of the voices of children with dysphonia. *Infocommunications journal*, *3*, 30–36. https://doi.org/10.36244/ICJ.2018.3.5

Tulics, M. G., & Vicsi, K. (2019). The automatic assessment of the severity of dysphonia. *International Journal of Speech Technology*, *22*(2), 341–350. https://doi.org/10.1007/s10772-019-09592-y

Tunc, H. C., Sakar, C. O., Apaydin, H., Serbes, G., Gunduz, A., Tutuncu, M., & Gurgen, F. (2020). Estimation of Parkinson's disease severity using speech features and extreme gradient boosting. *Medical & Biological Engineering & Computing*, *58*(11), 2757–2773. https://doi.org/10.1007/s11517-020-02250-5

Tuncer, T., & Dogan, S. (2019). A novel octopus based Parkinson's disease and gender recognition method using vowels. *Applied Acoustics*, *155*, 75–83. https://doi.org/10.1016/j.apacoust.2019.05.019

Tuncer, T., Dogan, S., & Acharya, U. R. (2020). Automated detection of Parkinson's disease using minimum average maximum tree and singular value decomposition method with vowels. *Biocybernetics and Biomedical Engineering*, *40*(1), 211–220. https://doi.org/10.1016/j.bbe.2019.05.006

Upadhya, S. S., & Cheeran, A. N. (2018a). Discriminating Parkinson and Healthy People Using Phonation and Cepstral Features of Speech. *Procedia Computer Science*, *143*, 197–202. https://doi.org/10.1016/j.procs.2018.10.376

Upadhya, S. S., & Cheeran, A. N. (2018b). Investigation Of Pitch And Noise Features Extracted From Voice Samples Of Healthy And Parkinson Affected People Using Statistical Tests. *Journal of Engineering Science and Technology* 13(9):2792-2804.

Upadhya, S. S., & Cheeran, A. N. (2018c). Performance Comparison Of Regression Techniques In Predicting Parkinson Disease Severity Score Using Speech Features. *Biomedical Engineering: Applications, Basis and Communications*, *30*(04), 1850025. https://doi.org/10.4015/S1016237218500254

Upadhya, S. S., Cheeran, A. N., & Nirmal, J. H. (2018). Thomson Multitaper MFCC and PLP voice features for early detection of Parkinson disease. *Biomedical Signal Processing and Control*, *46*, 293–301. https://doi.org/10.1016/j.bspc.2018.07.019

Vaiciukynas, E., Verikas, A., Gelzinis, A., & Bacauskiene, M. (2017). Detecting Parkinson's disease from sustained phonation and speech signals. *PLOS ONE*, *12*(10), e0185613. https://doi.org/10.1371/journal.pone.0185613

Vásquez-Correa, J. C., Orozco-Arroyave, J. R., Bocklet, T., & Nöth, E. (2018). Towards an automatic evaluation of the dysarthria level of patients with Parkinson's disease. *Journal of Communication Disorders*, *76*, 21–36. https://doi.org/10.1016/j.jcomdis.2018.08.002

Vásquez-Correa, J. C., Rios-Urrego, C. D., Rueda, A., Orozco-Arroyave, J. R., Krishnan, S., & Nöth, E. (2019). Articulation and Empirical Mode Decomposition Features in Diadochokinetic Exercises for the Speech Assessment of Parkinson's Disease Patients. In I. Nyström, Y. Hernández Heredia, & V. Milián Núñez (A c. Di), *Progress in Pattern Recognition, Image Analysis, Computer Vision, and Applications* (Vol. 11896, pp. 688–696). Springer International Publishing. https://doi.org/10.1007/978-3-030-33904-3_65

Wu, K., Zhang, D., Lu, G., & Guo, Z. (2018). Learning acoustic features to detect Parkinson's disease. *Neurocomputing*, *318*, 102–108. https://doi.org/10.1016/j.neucom.2018.08.036

Xiong, Y., & Lu, Y. (2020). Deep Feature Extraction From the Vocal Vectors Using Sparse Autoencoders for Parkinson's Classification. *IEEE Access*, *8*, 27821–27830. https://doi.org/10.1109/ACCESS.2020.2968177

Yeldener, S., de Martin, J. C., & Viswanathan, V. (1998). A mixed sinusoidally excited linear prediction coder at 4 kb/s and below. *Proceedings of the 1998 IEEE International Conference on Acoustics, Speech and Signal Processing, ICASSP '98 (Cat. No.98CH36181)*, *2*, 589–592. https://doi.org/10.1109/ICASSP.1998.675333

Yoon, S., Byun, S., & Jung, K. (2018). Multimodal Speech Emotion Recognition Using Audio and Text. *2018 IEEE Spoken Language Technology Workshop (SLT)*, 112–118. https://doi.org/10.1109/SLT.2018.8639583

Yu, Q., Zou, X., Quan, F., Dong, Z., Yin, H., Liu, J., Zuo, H., Xu, J., Han, Y., Zou, D., Li, Y., & Cheng, O. (2022). Parkinson's disease patients with freezing of gait have more severe voice impairment than non-freezers during "ON state". *Journal of Neural Transmission*, *129*(3), 277–286. https://doi.org/10.1007/s00702-021-02458-1

Yücelbaş, C. (2021). A new approach: Information gain algorithm-based k-nearest neighbors hybrid diagnostic system for Parkinson's disease. *Physical and Engineering Sciences in Medicine*, *44*(2), 511–524. https://doi.org/10.1007/s13246-021-01001-6

Zhang, L., Qu, Y., Jin, B., Jing, L., Gao, Z., & Liang, Z. (2020). An Intelligent Mobile-Enabled System for Diagnosing Parkinson Disease: Development and Validation of a Speech Impairment Detection System. *JMIR Medical Informatics*, *8*(9), e18689. https://doi.org/10.2196/18689

Zhang, T., Zhang, Y., Sun, H., & Shan, H. (2021). Parkinson disease detection using energy direction features based on EMD from voice signal. *Biocybernetics and Biomedical Engineering*, *41*(1), 127–141. https://doi.org/10.1016/j.bbe.2020.12.009

Zhang, X., Ma, J., Li, Y., Wang, P., & Liu, Y. (2021). Few-shot learning of Parkinson's disease speech data with optimal convolution sparse kernel transfer learning. *Biomedical Signal Processing and Control*, *69*, 102850. https://doi.org/10.1016/j.bspc.2021.102850

Zhang, Y. N. (2017). Can a Smartphone Diagnose Parkinson Disease? A Deep Neural Network Method and Telediagnosis System Implementation. *Parkinson's Disease*, *2017*, 1–11. https://doi.org/10.1155/2017/6209703