Management of Parkinson's Disease Dysarthria: Can Artificial Intelligence Provide the Solution?

Raj Kumar, Manoj Tripathy, Niraj Kumar[1], Radhey Shyam Anand

Department of Electrical Engineering, Indian Institute of Technology, Roorkee, Uttarakhand, [1]Department of Neurology, All India Institute of Medical Sciences, Rishikesh, Uttarakhand, India

Abstract

Speech disorder is a significant problem for people affected with Parkinson's disease (PD) leading to a substantial disability to communicate with others. PD affects the voice, including changes in pitch, intensity, articulation, and syllable rate. We aimed to study the current status of artificial intelligence (AI) using machine learning algorithms (MLAs) in the assessment of speech abnormalities in PD along with the generation of intelligible synthetic speech for voice rehabilitation. We searched the literature for studies focusing on speech/voice disorder in PD and rehabilitation techniques till June 18, 2022. We searched PubMed and Engineering Village (Compendex and Inspec combined) databases. After careful screening of the title and evaluation of abstracts, we used select articles describing the use of AI or its various forms in the management of speech abnormalities in PD to synthesize this review. MLAs classify PD and non-PD patients with an accuracy of more than 90% using only voice features. Non-acoustic sensors can rehabilitate PD patient by converting dysarthric speech to highly intelligible speech using MLAs. MLAs can automatically assess several speech features and quantify the progression of speech abnormalities in PD. PD speech rehabilitation using MLAs may prove superior to other available therapies.

Keywords: Artificial intelligence, machine learning algorithms, Parkinson's disease, random forest method, speech

INTRODUCTION

Parkinson's disease (PD) is a neurodegenerative disorder primarily affecting the population above the age of 60.[1] It involves 0.3% of the general population worldwide.[2] PD manifests with several motor and non-motor symptoms.[1] Rest tremor, rigidity, bradykinesia, and postural instability are its classical motor features. Nearly 90% of PD patients suffer speech impairment,[3] but only 5% of them receive any therapy for speech-related issues.[4] One-third of PD patients who are aware of their speech problems describe it as the most disabling feature of the disease, with many of them losing interest to participate in conversations and suffer depression.[5]

The speech abnormalities in PD patients include a reduction in speech volume, breathiness in voice, fluctuation in pitch, and a rapid rate of word output with incomprehensible speech. Major speech impairments can be categorized as hypophonia, dysarthria, dysphonia, and tachyphemia. Hypophonia is characterized by a soft voice or reduced voice volume, an early motor symptom of PD.[1] Dysarthria is related to articulation difficulties and dysphonia is related to defective use of the voice.[6] In Tachyphemia, an unwanted movement acceleration characterized by high speech rate and rapid stammering occurs that makes speech unintelligible.[7]

The cause of speech impairment in PD can be understood by the speech chain shown in Figure 1. The transmission of a message starts with the formation of words and sentences in the brain, known as the linguistic level, continues on the physiological level with neural and muscular activity, and generates and transmits a sound wave at the acoustic level. Speech problem in PD starts at the linguistic level where the neural signal does not transmit appropriately to the physiological level via motor nerves.

Several therapies have been used to address speech abnormalities in PD, with most of them having significant limitations and none of them providing a long-lasting solution [Table 1]. Levodopa therapy may not improve speech in all PD patients.[8] Moreover, it often results in significant dyskinesias involving orofacial and respiratory muscles involved in speech production.[8] Deep brain stimulation lacks a consistent effect on the improvement of speech in PD patients.[9] Because of the progressive nature of the worsening of speech in PD, vocal cord procedures like vocal fold augmentation may not be helpful.[10] Speech therapy, including Lee Silverman Voice Therapy (LSVT), requires a huge effort

Address for correspondence: Dr. Niraj Kumar, Department of Neurology, All India Institute of Medical Sciences, Rishikesh, Uttarakhand, India.
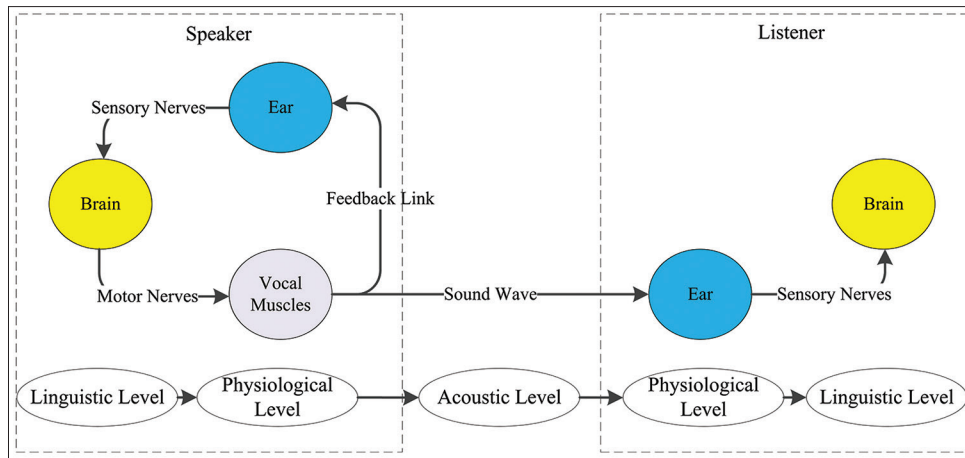E-mail: drnirajkumarsingh@gmail.com

Figure 1: Various levels of human speech chain

Table 1: Treatment modalities used for the improvement of speech intelligibility in PD patients		
Type of therapy	**Characteristics**	**Limitations**
Pharmacological therapy		
Levodopa	May improve speech loudness and intelligibility.[11]	May cause speech complications related to articulation like uncontrolled tongue-rolling and lip-clicking.[11]
Surgical therapy		
Deep Brain Stimulation	May improve several motor activities, including articulatory and phonatory components such as loudness.	Overall intelligibility may worsen[9] Occasionally speech improves after stopping the stimulation.[12]
Vocal Fold Augmentation	May improve speech intensity and quality by reducing space between the vocal folds by injection of implant materials in it.	Temporary improvement.
Speech therapy		
LSVT	An intense intervention emphasizes increasing amplitude and recalibrating vocal loudness.[13]	Requires numerous meetings with speech therapists (16 sessions of 50-60 min duration each in a month along with home practice[14]) Not widely available; only 3-4% PD population receive this treatment.[12]
Expiratory muscle strength training (EMST)	It builds up the muscles[15] that push air out of the lungs for speech. Improve syllables per breath and voice intensity in PD patients. EMST150 and The Breather™ are a few available devices used for EMST. Can be done at home.	It requires regular practice by blowing through the device as hard as possible for effectiveness.
Game-based therapy	This technique helps patients use voice-based features including pitch and phonation time to control parameters in the game, similar to a joystick. Suitable for in-home self-treatment. It provides immediate feedback via speech recognition and patients do not feel dull and monotone. May provide self-motivation and autonomy to PD patients.[16]	It requires regular practice Time consuming.
Portable devices		
Amplifier	It increases vocal loudness.[17]	Cumbersome to carry.
Smartphones	Apps that provide information on volume, rate, and pitch of voice during a conversation and guide users to adjust accordingly.	Only act as an indicator of voice quality but does not improve the voice.
Earpiece	SpeechEasyPD device works on the principle of delayed auditory feedback, which introduces a slightly delayed recording of the speech into the ear, and the feedback loop causes the speaker to slow down and speak more clearly. SpeechVive device plays background sounds in the ear while the patient talks and turns off when he stops talking. This prompts the user to speak louder, slower, and clearer, known as the Lombard effect.[18]	Costly (SpeechVive costs $2500 while SpeechEasyPD costs around $2500-4500) Overall, speech naturalness reduces.[19]

Sl. No.	Feature	Explanation	Trend in PD patients as compared to healthy adults
Table 2: The different voice features along with the trend in PD patients as compared to healthy controls[22,23]			
A	Phonation (sustained phonation of any vowel for ≥5-10 sec)		
	Jitter (%), Jitter (Absolute)	It measures variation in fundamental frequency, that is, the time period between vocal fold opening.	Higher
	Shimmer, Shimmer (decibel), Shimmer	A measure of variation in amplitude, that is, the extent of vocal fold opening.	Lower
	Noise/harmonic ratio and Harmonic/noise ratio	A measure of the ratio of noise to tonal components in the voice.	Lower
B	Prosody (Reading sentences, stories, rhythm, or maybe conversational speech)		
	Average fundamental frequency	An objective correlates of the pitch, measured in hertz.	Lower
	Speech rate	The number of syllables uttered per second.	Lower
	Pauses ratio	The ratio of all pauses duration to the total time duration.	Lower
	Intensity	The sound pressure level in decibel or average of squared magnitude of recorded voice.	Lower
C	Articulation (monosyllabic (vowel) or bisyllabic phonemes (a mixture of vowels and consonants) are spoken at a faster pace repeatedly)		
	DDK rate	The number of/pa/-/ta/-/ka/syllable per second. It measures the ability to articulate quickly and regularly.	Lower
	DDK regularity	The rate of change of DDK rate with time. It indicates the ability to maintain a constant DDK rate.	Higher
	Voice onset time	It measures the average length across/p/,/t/, and/k/consonants extracted from all three/pa/-/ta/-/ka/syllable repetitions.	Higher
	Vowel space area	Measure area of the triangle formed by/a/,/i/, and/u/vowels in the F1-F2 plane (First and second Formant frequency).	Lower
	Vowel articulatory index	vowel formant centralization.	Lower

from PD patients, which may not be easy. Other therapies, such as game-based therapy and portable devices have got their own limitations, including the need for practice, time consumption in the former, and a huge expense for the latter. On the other hand, artificial intelligence (AI) is a non-invasive and non-pharmaceutical technique, hence virtually no side effects. It can train itself with new data and thus adapt to changes over time.

With advancements in sensors and computational technology, extracting and analyzing voice data from PD patients has become quite comprehensive. Researchers have extracted several baseline voice features[20] from PD patients to quantify the disorder. It involves several voice-based tasks, for example, phonation, prosody, or articulation for voice recording. Table 2 shows these tasks and features extracted from each task. The speech characteristic varies a lot if PD voice is compared with the control group. This helps to design a speech-based assessment of PD. In addition, this helps to understand the modification required to make speech intelligible.

The baseline speech features do not capture subtle variations in fundamental frequency and amplitude.[21] To overcome these issues, researchers have used Mel-frequency cepstral coefficients, perceptual linear prediction, and wavelet transform coefficients as voice features.[21] These features characterize the speech signal's time power spectrum envelope, representing the vocal tract. These features have better time

and frequency resolutions, and when combined with baseline features, they characterize PD voice much better.[21] This article reviews machine learning algorithms (MLAs), a sub-set of AI applications for detection, assessment, and voice rehabilitation of PD. We discuss MLAs and their uses in PD diagnosis along with various possibilities for their use in speech rehabilitation.

METHOD

We searched the literature for studies focusing on speech/voice disorder in PD and rehabilitation techniques till June 18, 2022. We searched PubMed and Engineering Village (Compendex and Inspec combined) database. Using search criteria ["Parkinson disease"/exp OR "Parkinson disease" OR (Parkinsons AND disease)] AND ("speech" OR voice) AND (rehabilitation OR improvement OR "diagnosis") AND [machine AND learning OR (artificial AND intelligence) OR automatic OR neural OR (neural AND network) OR probability OR "support vector machine (SVM)" OR (deep AND neural) OR "convolutional neural network" OR "nearest neighbor algorithm" OR (decision AND tree)], we got 335 articles in PubMed database, and 108 articles in Engineering Village database. After careful screening of the title and evaluation of abstracts, we used select articles describing the use of AI or its various forms in the management of speech abnormalities in PD to synthesize this review.

Machine Learning Algorithms

MLAs learn functions to find the relationship between input and output in supervised learning. Input are features derived from the signal, and output can be a discrete (classification) or continuous value (regression). The basic idea of machine learning is to learn functions or boundaries which can predict or classify given test input. MLAs can detect dysarthria resulting from neurological dysfunction and compare the speech parameters to that of healthy individuals[24] through voice assessment with high accuracy. Its ability to detect small changes in voice features is better than speech therapists or audiologists.[25] The possibility of discriminative assessment among patients with PD, progressive supranuclear palsy, and multiple system atrophy has been reported.[23] The selection of adequate voice features and MLAs make voice a possible biomarker for PD diagnosis.

A number of voice features affect the performance of MLAs. Narendra *et al.*[26] have used two feature-set derived consisting of 16 and 39 features, respectively. The results show the accuracy of the classifier method is better when higher numbers of features are utilized, probably due to the fact that more features contain more information. But at the same time, it has to be understood that selected features are independent; otherwise, the performance of the classifier will be affected. Tracy *et al.*[27] have ranked 2330 acoustic features as per their importance and shown that after the first 100 features, importance drops drastically, that is, inclusion or exclusion of lower-ranked features does not affect performance significantly.

Use of MLAs in the Diagnosis of PD

In recent years, MLAs have been widely used in the diagnosis of PD patients. The advantages of MLAs are as follows:

Detection of PD

The parameters described in Table 2 show distinct variations from healthy individuals to PD patients,[28] which can be used for voice-based PD classification,[29] and may facilitate early detection of PD.[30] Generally, vowels are used for sustained phonation, for example, \a\, \e\, \u\, etc., where patients have to speak a vowel for 5–10 sec. Skodda *et al.*[31] and Proença *et al.*[32] have used only two formant frequencies (F1 and F2) of the vowels for PD classification. They used geometrical calculation to see changes from the control group to PD cases, but it does not work in the early stage of PD.

Several MLAs classify patients into PD and non-PD groups by learning a decision boundary in the speech data feature space. Random forest (RF) is the simplest method based on the decision tree concept,[33] with a reported accuracy of 96.8%. A decision tree is a flowchart-like representation of speech features that graphically resembles a tree. The tree's root is a feature connected to other roots through branches. Each branch act as an action based on the root (feature) value that can be taken to move down in the tree, and the tree's leaves (endpoints)

are classes, that is, PD or non-PD class. Naive Bayes classifier is based on Bayes Theorem of conditional probability. SVM optimizes class boundary using a few samples known as support vectors in a region of feature space where samples belong to two different classes. All samples beyond this region are ignored. An SVM cannot handle more than two classes. An accuracy of 85.25% has been reported[34] using SVM for classification. Artificial neural networks (ANN) is the technique adopted from the human brain that can learn any complex decision boundary. It consists of multiple layers with several neurons acting as an activation function. Åström *et al.*[35] have used 9 Parallel neural networks for classification and achieved an accuracy of 91.2% among 8 healthy and 23 PD patients. Sakar *et al.*[21] have combined the tunable Q-factor wavelet transform coefficients as additional features with baseline features for PD classification and showed that accuracy improved for all MLAs, with a maximum accuracy of 86% using combined features compared to baseline features only.

The selection of features plays an important role in the classification accuracy of classifiers. Ashour *et al.*[36] have shown that the accuracy of SVM improves from 88 to 94% by selecting eigenvector corresponding to significant eigenvalues compared to principal component analysis which is based on autocorrelation to discard highly correlated samples. Mostafa *et al.*[37] used five feature evaluators to rank each feature and showed that the accuracy of Decision Tree, Naive Bayes, ANN, RF, and SVM classifiers improved by around 10% when the best 11 features were selected out of 23 features. Optimal feature selection has improved the performance of SVM and RF significantly.[33]

Deep neural networks (DNN) is an extension of ANN having a higher number of neurons in each layer along with several hidden layers, with the capability of in-built feature extraction and feature selection.[38] The major challenge for DNN-based assessment is that it requires a large amount of training data because of its increased complexity.

Remote assessment using smartphones

Nowadays, smartphones are equipped with high-performance processors and sensors. Many researchers have reported the application of smartphones for PD diagnosis and remote assessment. Almeida *et al.*[39] have shown an accuracy of 92.94% in PD classification using a smartphone's microphone speech data, which is closer to 94.55% accuracy achieved using a standard microphone. Rusz *et al.*[22] have shown that hypokinetic dysarthria can be detected in the early stages of PD using smartphone microphone data. Another significant advantage of using smartphones is that it will help screen large populations effectively by avoiding the need for speech recording at the clinic and facilitating telediagnosis of PD.[29] Some challenges to remote assessment using smartphones include degraded voice quality caused by noise, reverberation, and other non-linear distortion.[25] Speech enhancement techniques can improve voice quality and, therefore, PD classification accuracy.[25]

Severity measurement

Neuro-fuzzy system (NFS) and support vector regression are used for the prediction of the total unified Parkinson's disease rating scale (UPDRS) using sustained phonation task of vowel \a\.[40] In design, NFS is similar to ANN, where the activation functions are based on fuzzy logic. In fuzzy logic, the output is a continuous value between 0 and 1, obtained by applying a rule to the input value. This rule varies from neuron to neuron, hence called fuzzy. The estimated UPDRS score can be useful for severity assessment remotely. Bayestehtashk *et al.*[41] have used all three voice tasks mentioned in Table 2, that is, phonation, prosody, and articulation, for UPDRS estimation using the regression method. They also showed that the reading task provides better estimation than the phonation and diadochokinetic (DDK) tasks.

INTELLIGIBLE SYNTHETIC SPEECH GENERATION USING MLAs

This section overlooks the potential of MLAs to map or generate highly intelligible synthetic speech. We surveyed studies using these algorithms to treat dysarthric speech caused by PD. This algorithm involves getting speech data from microphones. The microphone can be both acoustic as well as non-acoustic type. The signals from non-acoustic microphones or sensors do not sound like speech but contain vital vocal excitation and articulation information. Authors have attempted to map this information to phonetic sounds. Supplementary Figure 1 represents the methodology used for the conversion of sensor information into highly intelligible synthetic speech. It can also be referred to as the "silent speech technique," since speech is produced without voice. In the first phase, features are extracted which characterize the articulation during speech like tongues, lips, jaws, and other vocal muscles movements. In the next step, the phonetic sequence is generated using MLAs that map articulation features to phonetic sequences or texts. In the final step, phonetic sequences/texts are converted to speech using natural language processing techniques based on the desired rhythm, intonation, and syntactic information.

Utilizing a similar but more straightforward methodology, words are predicted from dysarthric speech using MLAs.[42] A message is formed from these words by mapping words combination to the most frequently used sentence. At the final stage, the sentence is converted to clear synthetic speech.

A novel approach is proposed for voice rehabilitation, which predicts phonetic sequence based on myoelectric (EMG) signals placed in the neck area using NFS.[43] In a similar work, Janke *et al.*[44] have implemented a facial surface EMG system. They have used several MLAs and showed that DNN is the best choice for mapping sensor data to articulated phonemes.

A variety of non-acoustic sensors can be used as sensor input. These can reveal speech attributes such as low-energy consonant voice bars, nasality, and glottalized excitation, which are not captured by acoustic sensors.[45] The non-acoustic sensors are highly noise-robust as they do not depend on air pressure variation instead vibration from the skin. These sensors can be placed in several places, including around the throat, behind the neck, jawline, and temple [Supplementary Figure 2]. Table 3 shows the merits and demerits of various types of non-acoustic sensors.

Voice features can be extracted from parallel recorded voices using close-talk microphone (placed close to the mouth) as an acoustic sensor and throat microphones (touching the neck area) as a non-acoustic sensor for classification.[47] Although acoustic sensors have been widely used for PD voice rehabilitation, the use of non-acoustic sensors is yet to be explored.

Some researchers have developed devices consisting of magnetic sensors and magnet-implant in the mouth. The received magnetic data is mapped to phonemes using signal processing techniques since each phoneme has specific facial and tongue movements. Gilbert *et al.*[48] have developed such devices and claimed to achieve speech recognition accuracy above 90%. These devices may be helpful for PD patients if they can articulate but cannot speak loud enough.

Table 3: Types of non-acoustic sensor with their advantages and disadvantages

Microphone	Advantages	Disadvantages
Bone conduction microphone	Captures low-frequency information. Fairly robust noise.	Narrowband speech. Position-dependent performance.[46]
Throat microphone	Skin-attached piezo-electric sensors. Measures vocal cord vibration effectively. Significantly more robust to environmental noise.	Lacks intelligibility and sounds unnatural. Lacks higher frequency contents.
Electroglottography	Measures vocal fold contact area. Detect glottal activity. The recognition rate is high when EGG is combined with speech.	Vocal tract characteristics are not captured.
Physiological microphone	Pick up sounds from the body skin. Better quality and intelligibility than a close-talk microphone.	Capture unwanted air combined vibrations.
Non-audible Murmur	Placed on the neck (behind the speaker's ear) can detect very quietly uttered speech. Noise robust. Recognize the whispered and murmur sounds effectively.	Affected by body tissues and lip radiations.

The major advantage of MLAs lies in their ability to generate intelligible speech without any physical side effects or harm to PD patients. Many improvements are expected with the ever-evolving new architectures in MLAs. Nowadays, MLAs are being used in all fields of life, making hardware employing MLAs easily available and accessible by the general public. With the increased demands for MLA-supported hardware, it is expected that costs will keep falling in the future.

CONCLUSION

Speech abnormalities start from an early stage of PD, and these changes become very obvious as the disease progresses. At an early disease stage, minor speech abnormalities are not perceivable by humans over a short period, but MLAs can automatically assess several speech features and quantify the progression in speech abnormalities as well as the stage of PD. PD speech rehabilitation techniques using MLAs may prove superior to medical and surgical therapies as well as to other external aid devices and mobile apps. An amalgamation of MLAs and advanced sensors for speech rehabilitation of PD patients at any disease stage may reduce the burden on audiologists or speech therapists.
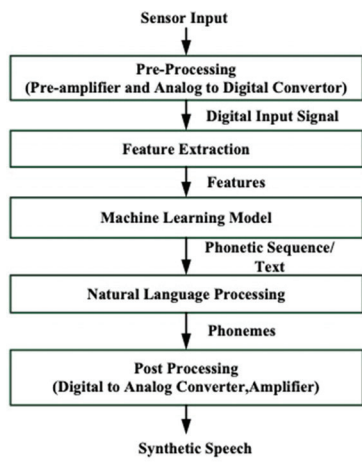
Financial support and sponsorship
Nil.

Conflicts of interest
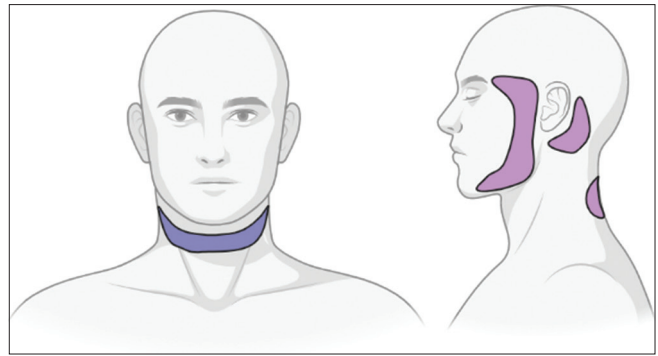There are no conflicts of interest.

REFERENCES

1. Rizek P, Kumar N, Jog MS. An update on the diagnosis and treatment of Parkinson disease. Can Med Assoc J 2016;188:1157-65.
2. de Lau LML, Giesbergen PCLM, de Rijk MC, Hofman A, Koudstaal PJ, Breteler MMB. Incidence of parkinsonism and Parkinson disease in a general population: The Rotterdam study. Neurology 2004;63:1240-4.
3. Ramig L, Fox C, Sapir S. Speech treatment for Parkinson's disease. Expert Rev Neurother 2008;8:297-309.
4. Liotti M, Ramig LO, Vogel D, New P, Cook CI, Ingham RJ, *et al.* Hypophonia in Parkinson's disease: Neural correlates of voice treatment revealed by PET. Neurology 2003;60:432-40.
5. Sunwoo MK, Hong JY, Lee JE, Lee HS, Lee PH, Sohn YH. Depression and voice handicap in Parkinson disease. J Neurol Sci 2014;346:112-5.
6. Sakar BE, Isenkul ME, Sakar CO, Sertbas A, Gurgen F, Delil S, *et al.* Collection and analysis of a Parkinson speech dataset with multiple types of sound recordings. IEEE J Biomed Health Inform 2013;17:828-34.
7. Nolden LF, Tartavoulle T, Porche DJ. Parkinson's disease: Assessment, diagnosis, and management. J Nurse Pract 2014;10:500-6.
8. Rusz J, Tykalová T, Klempíř J, Čmejla R, Růžička E. Effects of dopaminergic replacement therapy on motor speech disorders in Parkinson's disease: Longitudinal follow-up study on previously untreated patients. J Neural Transm (Vienna) 2016;123:379-87.
9. Skodda S. Effect of deep brain stimulation on speech performance in Parkinson's disease. Parkinsons Dis 2012;2012:850596.
10. Allensworth JJ, O'Dell K, Ziegler A, Bryans L, Flint P, Schindler J. Treatment outcomes of bilateral medialization thyroplasty for presbylaryngis. J Voice 2019;33:40-4.
11. Critchley EM. Speech disorders of Parkinsonism: A review. J Neurol Neurosurg Psychiatry 1981;44:751-8.
12. Dashtipour K, Tafreshi A, Lee J, Crawley B. Speech disorders in Parkinson's disease: Pathophysiology, medical management and surgical approaches. Neurodegener Dis Manag 2018;8:337-48.
13. Ramig L, Halpern A, Spielman J, Fox C, Freeman K. Speech treatment in Parkinson's disease: Randomized controlled trial (RCT). Mov Disord 2018;33:1777-91.
14. Sackley CM, Smith CH, Rick CE, Brady MC, Ives N, Patel S, *et al.* Lee Silverman Voice Treatment versus standard speech and language therapy versus control in Parkinson's disease: A pilot randomised controlled trial (PD COMM pilot). Pilot Feasibility Stud 2018;4:1-10.
15. Laciuga H, Rosenbek JC, Davenport PW, Sapienza CM. Functional outcomes associated with expiratory muscle strength training: Narrative review. J Rehabil Res Dev 2014;51:535-46.
16. Mühlhaus J, Frieg H, Bilda K, Ritterfeld U. Game-based speech rehabilitation for people with parkinson's disease. In: Lecture Notes in Computer Science. Springer, Cham; 2017;10279:76-85.
17. Andreetta MD, Adams SG, Dykstra AD, Jog M. Evaluation of speech amplification devices in Parkinson's disease. Am J speech-language Pathol 2016;25:29-45.
18. Adams S, Kumar N, Rizek P, Hong A, Zhang J, Senthinathan A, *et al.* Efficacy and acceptance of a Lombard-response device for hypophonia in Parkinson's disease. Can J Neurol Sci 2020;47:634-41.
19. Brendel B, Lowit A, Howell P. The effects of delayed and frequency shifted feedback on speakers with Parkinson disease. J Med Speech Lang Pathol 2004;12:131-8.
20. Barnish MS, Horton SMC, Butterfint ZR, Clark AB, Atkinson RA, Deane KHO. Speech and communication in Parkinson's disease: A cross-sectional exploratory study in the UK. BMJ Open 2017;7:1-10.
21. Sakar CO, Serbes G, Gunduz A, Tunc HC, Nizam H, Sakar BE, *et al.* A comparative analysis of speech signal processing algorithms for Parkinson's disease classification and the use of the tunable Q-factor wavelet transform. Appl Soft Comput J 2019;74:255-63.
22. Rusz J, Hlavnicka J, Tykalova T, Novotny M, Dusek P, Sonka K, *et al.* Smartphone allows capture of speech abnormalities associated with high risk of developing Parkinson's disease. IEEE Trans Neural Syst Rehabil Eng 2018;26:1495-507.
23. Sachin S, Shukla G, Goyal V, Singh S, Aggarwal V, Gureshkumar, *et al.* Clinical speech impairment in Parkinson's disease, progressive supranuclear palsy, and multiple system atrophy. Neurol India 2008;56:122-6.
24. Rusz J, Novotny M, Hlavnicka J, Tykalova T, Ruzicka E. High-accuracy voice-based classification between patients with Parkinson's disease and other neurological diseases may be an easy task with inappropriate experimental design. IEEE Trans Neural Syst Rehabil Eng 2017;25:1319-21.
25. Poorjam AH, Kavalekalam MS, Shi L, Raykov JP, Jensen JR, Little MA, *et al.* Automatic quality control and enhancement for voice-based remote Parkinson's disease detection. Speech Commun 2021;127:1-16.
26. Narendra NP, Alku P. Automatic assessment of intelligibility in speakers with dysarthria from coded telephone speech using glottal features. Comput Speech Lang 2020;65:101117.
27. Tracy JM, Özkanca Y, Atkins DC, Hosseini Ghomi R. Investigating voice as a biomarker: Deep phenotyping methods for early detection of Parkinson's disease. J Biomed Inform 2020;104:1-10.
28. Arora S, Baghai-Ravary L, Tsanas A. Developing a large scale population screening tool for the assessment of Parkinson's disease using telephone-quality voice. J Acoust Soc Am 2019;145:2871-84.
29. Zhang YN. Can a smartphone diagnose parkinson disease? A deep neural network method and telediagnosis system implementation. Parkinsons Dis 2017;2017:6209703.
30. Upadhya SS, Cheeran AN, Nirmal JH. Thomson multitaper MFCC and PLP voice features for early detection of Parkinson disease. Biomed Signal Process Control 2018;46:293-301.
31. Skodda S, Visser W, Schlegel U. Vowel articulation in parkinson's disease. J Voice 2011;25:467-72.
32. Proença J, Veiga A, Candeias S, Lemos J, Januário C, Perdigão F. Characterizing Parkinson's Disease Speech by Acoustic and Phonetic Features. In: Lecture Notes in Computer Science. Springer, Cham; 2014;8775:24-35.
33. Wu K, Zhang D, Lu G, Guo Z. Learning acoustic features to detect Parkinson's disease. Neurocomputing 2018;318:102-8.
34. Perez C, Roca YC, Naranjo L, Martin J. Diagnosis and Tracking of Parkinson's Disease by using Automatically Extracted Acoustic Features. J Alzheimer's Dis Park 2016;6.

35. Åström F, Koker R. A parallel neural network approach to prediction of Parkinson's disease. Expert Syst Appl 2011;38:12470-4.
36. Ashour AS, Nour MKA, Polat K, Guo Y, Alsaggaf W, El-Attar A. A novel framework of two successive feature selection levels using weight-based procedure for voice-loss detection in Parkinson's disease. IEEE Access 2020;8:76193-203.
37. Mostafa SA, Mustapha A, Mohammed MA, Hamed RI, Arunkumar N, Abd Ghani MK, *et al*. Examining multiple feature evaluation and classification methods for improving the diagnosis of Parkinson's disease. Cogn Syst Res 2019;54:90-9.
38. Lecun Y, Bengio Y, Hinton G. Deep learning. Nature 2015;521:436-44.
39. Almeida JS, Rebouças Filho PP, Carneiro T, *et al*. Detecting Parkinson's disease with sustained phonation and speech signals using machine learning techniques. Pattern Recognit Lett 2019;125:55-62.
40. Nilashi M, Ibrahim O, Ahani A. Accuracy improvement for predicting Parkinson's disease progression. Sci Rep 2016;6:1-18.
41. Bayestehtashk A, Asgari M, Shafran I, McNames J. Fully automated assessment of the severity of Parkinson's disease from speech. Comput Speech Lang 2015;29:172-85.
42. Hawley MS, Cunningham SP, Green PD, Enderby P, Palmer R, Sehgal S, *et al*. A voice-input voice-output communication aid for people with severe speech impairment. IEEE Trans Neural Syst Rehabil Eng 2013;21:23-31.
43. Malcangi M, Felisati G, Saibene A, *et al*. Myo-To-Speech - Evolving Fuzzy-Neural Network Prediction of Speech Utterances from Myoelectric Signals. In: Communications in Computer and Information Science. Springer, Cham; 2018;893:158-68.
44. Janke M, Diener L. EMG-to-Speech: Direct generation of speech from facial electromyographic signals. IEEE/ACM Trans Audio Speech Lang Process 2017;25:2375-85.
45. Quatieri TF, Brady K, Messing D, Campbell JP, Campbell WM, Brandstein MS, *et al*. Exploiting nonacoustic sensors for speech encoding. IEEE Trans Audio, Speech Lang Process 2006;14:533-42.
46. Mcbride M, Tran P, Letowski T, Patrick R. The effect of bone conduction microphone locations on speech intelligibility and sound quality The effect of bone conduction microphone locations on speech intelligibility and sound quality. Appl Ergon 2010;42:495-502.
47. Atzori A, Carullo A, Vallan A, Cennamo V, Astolfi A. Parkinson disease voice features for rehabilitation therapy and screening purposes. In: Medical Measurements and Applications. IEEE; 2019:4-9.
48. Gilbert JM, Rybchenko SI, Hofe R, Ell SR, Fagan MJ, Moore RK, *et al*. Isolated word recognition of silent speech using magnetic implants and sensors. Med Eng Phys 2010;32:1189-97.

Sensor Input

↓

Pre-Processing
(Pre-amplifier and Analog to Digital Convertor)

↓ Digital Input Signal

Feature Extraction

↓ Features

Machine Learning Model

↓ Phonetic Sequence/Text

Natural Language Processing

↓ Phonemes

Post Processing
(Digital to Analog Converter, Amplifier)

↓

Synthetic Speech

Supplementary Figure 1: A methodology used for synthetic speech generation

Supplementary Figure 2: Skin area representing sensors placement for speech production