2023 45th Annual International Conference of the IEEE Engineering in Medicine & Biology Society (EMBC) | 979-8-3503-2447-1/23/$31.00 ©2023 IEEE | DOI: 10.1109/EMBC40787.2023.10340507

Monitoring the Effect of Levodopa Using Sustained Phonemes in Parkinson's Disease Patients

Mohammod A. Motin, *Senior Member IEEE,* Nemuel D. Pah, *Member IEEE*, and Dinesh K. Kumar, *Senior Member, IEEE*

***Abstract*—Parkinson's disease (PD) is a neurological disease identified by multiple symptoms, and levodopa is one of the most effective medications for treating the disease. To determine the dosage of levodopa, it is necessary to meet on a regular basis and observe motor function. The early detection and progression of the disease have been proposed using hypokinetic dysarthria. However, previous studies have not examined the effects of levodopa on speech rigorously and have provided inconsistent results. In this study, three sustained phonemes of PD patients were investigated for the effect of medication. A set of features characterizing vocal fold dynamics as well as the vocal tract coordinators were extracted from the sustained phonemes /of 28 PD patients during levodopa medication off and on states. All the features were statistically investigated and classified using a linear discriminant analysis (LDA) classifier. LDA classifier identified medication on from medication off based on the combined features from phoneme /a/, /o/ and /m/ with the accuracy=82.75% and F1-score=82.18%. Voice recording of PD patients during sustained phonemes /a/, /o/ and /m/ has the potential for identifying whether the patients are in On state or Off state of medication.**

***Clinical Relevance*— The outcomes of this study have the potential to monitor the effect and progress of levodopa on PD patients.**

I. Introduction

Parkinson's disease (PD) is a chronic neurodegenerative disorder that manifests with both motor and non-motor symptom [1]. It is the second most common neurodegenerative disorder after Alzheimer's, and it is expected that its prevalence increases with the aging population, leading to a burden on society and healthcare. The diagnosis of PD is based on clinical assessment of motor symptoms such as tremor, stiffness, bradykinesia, and postural difficulties [2], or non-motor symptoms such as dysarthria, functional impairment, and cognitive impairment [3]. Abnormal speech production, known as Parkinsonian hypokinetic dysarthria, is a common early symptom of PD [4]. An array of symptoms such as reduced voice intensity, increased nasality, increased acoustic noise, and imprecise articulation, are associated with it [5]. Voice testing has been proposed as a tool for early monitoring of PD [6, 7].

The management of Parkinson's disease symptoms involves medications, with levodopa being the most widely used. Patients can be in an "off state" or "on state" depending on the effectiveness of the medication. However, the dosing plan for levodopa can be imprecise and may require frequent visits to a neurologist [8]. To facilitate remote monitoring, speech analysis could provide insight into the medication's effect on patients. Despite this potential, previous studies evaluating the impact of levodopa on speech in Parkinson's disease have produced inconsistent or conflicting results [9]. Some studies have found that levodopa improves consonant articulation in the early stages of PD, while others have found only partial improvement in speech prosody [10] [11]. Compared with a drug-naive state, there has been an increase in speech disfluency after three to six years of dopaminergic therapy [12]. In contrast, several studies have investigated the effect of medication therapy on speech parameters and have found a weak or minor correlation [13-15].

The study aims to examine the use of the phonatory parameters to distinguish PD patients with medication (PD-on) from without medication (PD-*off*). We investigated the change in phonatory parameters by using an analysis of variance test and the LDA algorithm to separate the PD-*on* and PD-*off*. Three different sustained phonemes were considered: /a/, /o/, and /m/. We selected these phonemes based on their specific production process [16-18]. The vowel /a/ is produced by opening the jaw widely and placing the tongue low in the mouth. So, the production of vowel /a/ requires the control of vocal cord muscles in conjunction with lung muscles. Voiced nasal phonemes like /m/ are produced by the vibration of the vocal cord in conjunction with moving air in the nasal cavity. The vowel /o/ is classified as a closed-mid-back vowel due to its position mid-high toward the palate and its rounded lips. As well as the vocal cords, the mouth muscles are required to produce this vowel. Various muscles are involved in this process, including those in the vocal cord, the lung, and the velum. Phonatory parameters of each phoneme will shed light on how PD and its medications affect lungs, vocal cords, voice boxes, and velums.

Mohammod Abdul Motin is with the Department of Electrical & Electronic Engineering, Rajshahi University Engineering & Technology, Bangladesh. He is also with the School of Engineering, RMIT University, Melbourne, Australia (E-mail: m.a.motin@ieee.org).

Nemuel Daniel Pah is with the Electrical Engineering Department, Universitas Surabaya, Surabaya, Indonesia. He is also with the School of Engineering, RMIT University, Melbourne, Australia (Email: nemuelpah@staff.ubaya.ac.id)

Dinesh Kant Kumar is with School of Engineering, RMIT University, Melbourne, VIC 3000, Australia (E-mail: dinesh.kumar@rmit.edu.au).

II. Dataset And Methods

A. Parkinson's Disease Participants

The dataset used in this study has been reported in previous studies [19], and a summary is provided below. Twenty eight PD patients (UPDRS > 24) were recruited from the Movement Disorder Clinic at Monash Medical Centre. These PD patients have all been diagnosed with PD within the last ten years according to Queen Square Brain Bank criteria for idiopathic PD. Any neurological symptoms associated with advance PD was an exclusion criterion. The PD off state (PD-off) is defined as PD participants without anti-parkinsonian medication for at least 12 hours. The PD on state (PD-on) is described as the PD participants within 1.5 hours after subjects' parkinsonian medication and there is clinically observable reduction in the symptoms. An experienced neurologist scored the MDS-UPDRS-III for both off- and on-state motor function. The demographics, MDS-UPDRS-III scores, and Montreal cognitive assessment (MoCA) are presented in Table I. The study protocol was approved by Monash Health, Australia (LNR/16/MonH/319) and RMIT University Human Research Ethics Committee (BSEHAPP22-15KUMAR). A written consent was taken from all PD participants before the conducting the experiment.

TABLE I
Participants' Demographics

Demographic Information	PD Subjects
Total Number of PD subjects	28
Age (years)	71.83 ± 7.67
PD-*off* MDS-UPDRS-III score	25.54 ± 8.78
PD-*on* MDS-UPDRS-III score	19.33 ± 9.30
MoCA	27.58 ± 2.48
Duration of disease (years)	5.63 ± 3.00

B. Sustained Phonemes Recording

Three sustained phonemes /a/, /o/, and /m/ were recorded from twenty-eight PD patients. Between each recording, there were 30 to 60 seconds of relaxation period. The recording duration of each sustained phoneme varies from 8 to 15 seconds. The detail is shown in Table II. The sustained phonemes were recorded both on-*state* and *off-state* of medication for the PD participants. An omnidirectional head-worn microphone, the Samson-SE50, was used to capture the phonemes. To preserve the quality of the recordings, they were saved in uncompressed WAV format, sampling at 48 kHz, and 16-bit resolution. The data were conducted in a noise-free room and the de-identified data can be accessed from RMIT website and was reported earlier [19].

TABLE II
Duration of the Sustained Phonemes

Phoneme	Duration (sec) min – max (mean ±std)
/a/	5.3 – 14.4 (9.8±2.5)
/o/	5.2 – 16.8 (11.3±3.1)
/m/	5.6 – 15.3 (10.2±2.4)

C. Feature Extraction

All the statistical analysis and the features investigated in this work was computed using MATLAB2021b (MathWorks). Before extracting any features from the recording of the sustained phonemes, all the unwanted portion such as instructor's voice before and after recording was removed. Then, a set of jitter and shimmer features were extracted from phonemes. To compute these features, firstly the peaks of the glottal period were identified by searching the maximum point within a moving window of a size of 1.25 of pitch period, T_0. The peak magnitudes (A_i) and its corresponding time instances (t_i) were computed. The instantaneous period of the glottal wave (T_i) was calculated as the difference between subsequent instances of the peak, $T_i = t_{i+1} - t_i$.

Five jitter parameters were computed from each phoneme: absolute jitter (*abs*), relative jitter (*rel*), relative average perturbation (*rap*), and period perturbation quotient *(ppq)*. These jitter features are the perturbation of the difference between T_i and its moving average with a different window size. These five features are computed using the following equations:

$$Jitter(abs) = \frac{1}{N-1}\sum_{i=1}^{N-1}|T_{i+1} - T_i| \quad (1)$$

$$Jitter(rel) = \frac{\frac{1}{N-1}\sum_{i=1}^{N-1}|T_{i+1}-T_i|}{\frac{1}{N}\sum_{i=1}^{N}T_i} \quad (2)$$

$$rap = \frac{\frac{1}{N-2}\sum_{i=2}^{N-1}\left|T_i-\left(\frac{1}{3}\sum_{n=i-1}^{i+1}T_n\right)\right|}{\frac{1}{N}\sum_{i=1}^{N}T_i} \quad (3)$$

$$ppq5 = \frac{\frac{1}{N-4}\sum_{i=3}^{N-2}\left|T_i-\left(\frac{1}{5}\sum_{n=i-2}^{i+2}T_n\right)\right|}{\frac{1}{N}\sum_{i=1}^{N}T_i} \quad (4)$$

$$ppq11 = \frac{\frac{1}{N-10}\sum_{i=6}^{N-5}\left|T_i-\left(\frac{1}{11}\sum_{n=i-5}^{i+5}T_n\right)\right|}{\frac{1}{N}\sum_{i=1}^{N}T_i} \quad (5)$$

Five shimmer parameters: absolute shimmer (abs), the relative shimmer (rel), *apq3*, *apq5*, and *apq11* were computed from each phoneme. These shimmer features are the perturbation of the difference between A_i and its moving average with a different window size. These five shimmer features are computed using the following equations:

$$Shimmer(abs, dB) = \frac{1}{N-1}\sum_{i=1}^{N-1}\left|20 * \log\left(\frac{A_{i+1}}{A_i}\right)\right| \quad (6)$$

$$Shimmer(rel) = \frac{\frac{1}{N-1}\sum_{i=1}^{N-1}|A_{i+1}-A_i|}{\frac{1}{N}\sum_{i=1}^{N}A_i} \quad (7)$$

$$apq3 = \frac{\frac{1}{N-2}\sum_{i=2}^{N-1}\left|A_i-\left(\frac{1}{3}\sum_{n=i-1}^{i+1}A_n\right)\right|}{\frac{1}{N}\sum_{i=1}^{N}A_i} \quad (8)$$

$$apq5 = \frac{\frac{1}{N-4}\sum_{i=3}^{N-2}\left|A_i-\left(\frac{1}{5}\sum_{n=i-2}^{i+2}A_n\right)\right|}{\frac{1}{N}\sum_{i=1}^{N}A_i} \quad (9)$$

$$apq11 = \frac{\frac{1}{N-10}\sum_{i=6}^{N-5}\left|A_i-\left(\frac{1}{11}\sum_{n=i-5}^{i+5}A_n\right)\right|}{\frac{1}{N}\sum_{i=1}^{N}A_i} \quad (10)$$

To measure the energy of the phonemes, Teager-Kaiser energy operator (TKEO) is used. The average, standard deviation, and percentile values of TKEO for the contour T_0 and A_0 were measured.

Vocal fold excitation ratio (VFER) is a measure to detect dysphonia. When the vocal fold cycle is not working properly there is turbulence and uncorrelated asynchronous excitation on a different frequency band that is reflected by VFER.

The above-mentioned jitter, shimmer, TEKO and the VFER features were mainly aimed for understanding of vocal fold dynamics as it is affected in PD patients. Since, PD also affect vocal tract coordinators, a set of mel-frequency cepstral coefficients (MFCCs) features were incorporated for better characterizing the vocal tract coordinators [19].

Since the tongue, jaw, and lips which are the coordinators of articulators of the vocal tract are also impacted by PD [20], it is hypothesized that MFCC will be different for PD-on and PD-off. MFCCs were computed using the equation 11.

$$MFCC_n = \sum_{k=1}^{K} E_k \cos\left[n(k-0.5)\frac{\pi}{K}\right] \quad (11)$$

where n = 0,….., L. L is the number of MFCC. We have chosen L=12. E_k is the average energy of kth frequency band. As well as MFCCs, the first-order (delt) and second-order time derivative (delta-delta) of MFCC were computed [21, 22]. We computed a set of 28 features each from MFCC, delta, and delta delta-delta.

D. Machine Learning Based Classification

A discriminant analysis classifier with linear kernel was used in this study. The performance of the classifier is verified using the leave one subject out cross validation to avoid any subject bias and data leakage during training and testing. To reduce the model complexity and computational cost we only used significant features to train the model. The model was trained using the features of each phoneme individually and later significant features from any two phonemes were combined to train the model. Finally, all significant features from three phonemes were combined to train and test the model. The detail of selected features and the performance of each sustained phonemes are shown in Table III and IV respectively. Four performance metrics: accuracy, sensitivity, specificity, and F1-score were used to assess the model performance.

III. Experimental Results

A. Statistical Analysis

Since the voice features of the two group were not normally distributed and the distribution free non-parametric Kruskal-Wallis's test was used to test for group differences in each of the features. Table III presents the result of the p-value of each feature for PD-*on* vs PD-*off*, and thus identifies the features that are significantly changed by medication. In the analysis, 95% confidence interval or p-value<0.05 were considered, meaning that there was a significant difference in the means between the groups. It shows that p<0.05 was for three, four, and three features from sustained phonemes /a/, /m/, and /o/ respectively. The detail of selected significant features each sustained phonemes are shown in Table III.

TABLE III
THE RESULT OF P-VALUE FOR EACH FEATURES EXTRACTED FROM PHONEMES /A/, /M/, AND /O/.

Phoneme	Feature Description	*p*-value
a	avg_VFER	0.022
a	avg_10th delta	0.025
a	shimmer_abs	0.003
m	shimmer_apq11	0.018
m	shimmer_apq11_GS	0.010
m	shimmer_log	0.024
m	std_shimmer_TKEO	0.004
o	avg_12th delta	0.022
o	std_MFCC_4th_coef	0.004
o	Shimmer_log	0.007

avg: average value, std: standard deviation

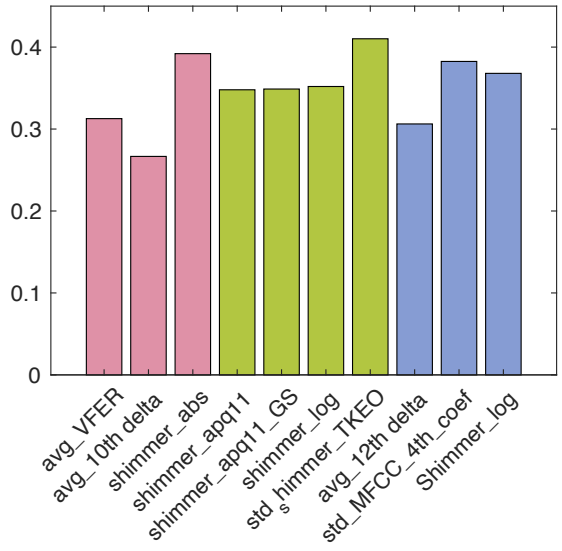

Fig 1. Pearson correlation coefficients values of the significant features for PD-on vs PD-off classification. The pink, olive and indigo color represent features from phoneme /a/, /m/, and /o/ respectively.

B. Classification Results

The results of classification by LDA of the using three, phonemes are shown in Table IV. It is seen that the accuracy and F1-score of the model using features obtained from individual phoneme to classify the PD-*on* and PD-*off* varies from 65.52% to 67.24%, and 62.61% to 67.20% respectively. The inclusion of feature from any two phonemes increased the model performance up to 8. 62% and 10.53% in term of accuracy and F1-score respectively. Finally, the combination of features from three phonemes improved the performance 12.13% and 17.24% compared to features from two phonemes and single phoneme respectively. The classification between PD-on and PD-*off* shows the best result was with the combination of the three phonemes, with accuracy of 82.76%, sensitivity of 89.66%, specificity of 75.86% and F1-score of 82.18%. The detail of the

performance of each sustained phonemes along with their combination are shown in IV respectively.

TABLE III
PERFORMANCE OF EACH PHONEMES TO CLASSIFY PD-ON VS PD-OFF STATE OF MEDICATION.

Sustained Phonemes	Acc (%)	Sen (%)	Spe (%)	F1-Score
/a/	67.24	68.97	65.52	67.20
/m/	65.52	79.31	51.72	62.61
/o/	67.24	72.41	62.07	66.84
/a/ + /m/	74.14	82.76	65.52	73.14
/m/ + /o/	70.69	79.31	62.07	69.64
/a/ + /o/	72.41	75.86	68.97	72.25
/a/ + /o/ + /m/	82.76	89.66	75.86	82.18

Acc: accuracy, Sen: Sensitivity, Spe: Specificity

IV. CONCLUSION

In this study, the difference between the levodopa based On and Off state Parkinson's disease patients using the sustained phonemes /a/, /o/, and /m/ was investigated. The results indicated that On state due to medication had a significant effect on shimmer, vocal fold excitation ratio, and mel-frequency cepstral coefficients of sustained phonemes. Additionally, significant features of the three phonemes were effective for detecting PD On-states vs PD-Off over features from individual phonemes. The classification shows that PD On-states and PD-Off using the combined features from phoneme /a/, /m/, and /o/ could be differentiated with the accuracy of 82.75% and F1-score of 82.18%. The potential of this study is that people with Parkinson's disease can be monitored for their medication state using voice, which thus has the potential for being used for regular monitoring of these patients.

REFERENCES

[1] L. M. de Lau and M. M. Breteler, "Epidemiology of Parkinson's disease," *Lancet Neurol,* vol. 5, no. 6, pp. 525-35, Jun 2006, doi: 10.1016/S1474-4422(06)70471-9.

[2] W. Poewe *et al.*, "Parkinson disease," *Nature reviews Disease primers,* vol. 3, no. 1, pp. 1-21, 2017.

[3] C. Simonet, A. Schrag, A. J. Lees, and A. J. Noyce, "The motor prodromes of parkinson's disease: from bedside observation to large-scale application," *J Neurol,* pp. 1-10, Dec 4 2019, doi: 10.1007/s00415-019-09642-0.

[4] M. Trail, C. Fox, L. O. Ramig, S. Sapir, J. Howard, and E. C. Lai, "Speech treatment for Parkinson's disease," *NeuroRehabilitation,* vol. 20, no. 3, pp. 205-221, 2005.

[5] E. Vaiciukynas, A. Verikas, A. Gelzinis, and M. Bacauskiene, "Detecting Parkinson's disease from sustained phonation and speech signals," *PLOS ONE,* vol. 12, no. 10, pp. 1-16, 2017, doi: 10.1371/journal.pone.0185613.

[6] C. O. Sakar *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," *Applied Soft Computing,* vol. 74, pp. 255-263, 2019.

[7] B. E. Sakar *et al.*, "Collection and Analysis of a Parkinson Speech Dataset With Multiple Types of Sound Recordings," *IEEE Journal of Biomedical and Health Informatics,* vol. 17, no. 4, pp. 828-834, 2013, doi: 10.1109/JBHI.2013.2245674.

[8] M. Pieterman, S. Adams, and M. Jog, "Method of Levodopa Response Calculation Determines Strength of Association With Clinical Factors in Parkinson Disease," *Frontiers in Neurology,* vol. 9, no. May, 2018, doi: 10.3389/fneur.2018.00260.

[9] L. Brabenec, J. Mekyska, Z. Galaz, and I. Rektorova, "Speech disorders in Parkinson's disease: early diagnostics and effects of medication and brain stimulation," *Journal of neural transmission,* vol. 124, pp. 303-334, 2017.

[10] J. Rusz, R. Cmejla, H. Ruzickova, and E. Ruzicka, "Quantitative acoustic measurements for characterization of speech and voice disorders in early untreated Parkinson's disease," *The Journal of the Acoustical Society of America,* vol. 129, no. 1, pp. 350-367, 2011, doi: 10.1121/1.3514381.

[11] N. Elfmarkova, M. Gajdos, M. Mrackova, J. Mekyska, M. Mikl, and I. Rektorova, "Impact of Parkinson's disease and levodopa on resting state functional connectivity related to speech prosody control," *Parkinsonism Relat Disord,* vol. 22 Suppl 1, no. 1, pp. S52-5, Jan 2016, doi: 10.1016/j.parkreldis.2015.09.006.

[12] T. Tykalova *et al.*, "Effect of dopaminergic medication on speech dysfluency in Parkinson's disease : a longitudinal study," *Journal of Neural Transmission,* vol. 122, no. 8, 2015, doi: 10.1007/s00702-015-1363-y.

[13] J. Sanabria *et al.*, "The Effect of Levodopa on Vocal Function in Parkinson's Disease," *Clinical Neuropharmacology,* vol. 24, no. 2, pp. 99-102, 2001, doi: 10.1097/00002826-200103000-00006.

[14] A. Goberman, C. Coelho, and M. Robb, "Phonatory characteristics of Parkinsonian speech before and after morning medication : the ON and OFF states," *Journal of Communication Disorders,* vol. 35, pp. 217-239, 2002, doi: 10.1016/s0021-9924(01)00072-7.

[15] M. Fabbri *et al.*, "Speech and Voice Response to a Levodopa Challenge in Late-Stage Parkinson's Disease," *Front Neurol,* vol. 8, no. 8, p. 432, 2017, doi: 10.3389/fneur.2017.00432.

[16] N. D. Pah, M. A. Motin, P. Kempster, and D. K. Kumar, "Detecting effect of levodopa in Parkinson's disease patients using different sustained phonemes," *IEEE Journal of Translational Engineering in Health and Medicine (Accepted). ,* 2021.

[17] M. A. Motin, N. D. Pah, S. Raghav, and D. K. Kumar, "Parkinson's Disease Detection Using Smartphone Recorded Phonemes in Real World Conditions," *IEEE Access,* vol. 10, pp. 97600-97609, 2022.

[18] N. D. Pah, M. A. Motin, and D. K. Kumar, "Phonemes based detection of parkinson's disease for telehealth applications," *Scientific Reports,* vol. 12, no. 1, p. 9687, 2022.

[19] R. Viswanathan, S. P. Arjunan, A. Bingham, and B. Jelfs, "Complexity Measures of Voice Recordings as a Discriminative Tool for Parkinson's Disease," *Biosensors,* vol. 10, no. 1, 2019, doi: 10.3390/bios10010001.

[20] A. K. Ho, R. Iansek, C. Marigliani, J. L. Bradshaw, and S. Gates, "Speech impairment in a large sample of patients with Parkinson's disease," *Behavioural neurology,* vol. 11, no. 3, pp. 131-137, 1998.

[21] J. I. Godino-Llorente, P. Gomez-Vilda, and M. Blanco-Velasco, "Dimensionality reduction of a pathological voice quality assessment system based on Gaussian mixture models and short-term cepstral parameters," *IEEE transactions on biomedical engineering,* vol. 53, no. 10, pp. 1943-1953, 2006.

[22] A. Tsanas, M. A. Little, P. E. McSharry, and L. O. Ramig, "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,* vol. 8, no. 59, pp. 842-855, 2011.