Analyzing the Probability Density Distribution of Sustained Phoneme Voice Features in the PC-GITA Dataset for Parkinson's Disease Identification

Nemuel Daniel Pah[1], Veronica Indrawati[1], Dinesh K. Kumar[2], and Mohammod A. Motin[3]

[1] Universitas Surabaya, Surabaya 60293, Indonesia,
nemuelpah@staff.ubaya.ac.id
[2] RMIT University, Melbourne, VIC 3000, Australia
[3] Rajshahi University of Engineering and Technology, Rajshahi 6204, Bangladesh

Abstract. One of the possibilities for developing computerized diagnostic tools for Parkinson's disease (PD) is to utilize the voice change known as Parkinsonian dysarthria. Voice features extracted from sustained phonemes have been statistically investigated as parameters for this purpose. However, the commonly used statistical presentation methods often obscure interpretations. This paper introduces an alternative approach using probability density distribution analysis to analyze voice features. The analysis was applied to recordings of sustained phonemes from the PC-GITA dataset. The findings reveal a significant overlap between the distributions of PD and healthy subjects (HC), with PD features exhibiting a wider distribution compared to HC. This result suggests the potential use of these features to identify PD, but it should be noted that a considerable number of PD cases may have voice features similar to HC.

Keywords: Parkinsonian dysarthria, voice features, probability density distribution.

1 Introduction

Parkinson's disease (PD) is the second most common neurodegenerative disorder after Alzheimer's disease [7]. Its prevalence is predicted to rise due to an aging population. PD is characterized by a range of motor and non-motor impairments [20]. The current gold standard for diagnosing PD involves clinical evaluation using the Unified Parkinson Disease Rating Scale (UPDRS) [3] or the Hoehn and Yahr (H&Y) scale [16]. These assessments examine the presence of symptoms such as tremors, rigidity, bradykinesia, or postural impairment, as well as non-motor symptoms including dysarthria, functional impairment, and cognitive impairment [19].

The development of portable, automatic, and smart PD diagnostic tools or applications requires the efficacy of PD biomarkers that are non-invasively accessible. One of the early symptoms of PD is the change in voice, known as Parkinsonian hypokinetic dysarthria [17]. This symptom is observed in approximately 90% of individuals with PD and is associated with reduced voice intensity, increased voice nasality, heightened acoustic noise, speech prosody, imprecise articulation, narrower pitch range, mono loudness, longer pauses, vocal tremor, harsh and breathy voice quality, and disfluency [22].

Voice features extracted from sustained phonemes have been investigated as parameters for diagnosing and monitoring the progression of PD [15, 9, 4, 18, 2]. These features encompass aspects such as glottal vibration impairments, harmonics-to-noise ratio, the respiratory mechanism for glottal pressure control, and vocal tract control [14]. The effectiveness of these features has been examined using various statistical analyses, including mean value, standard deviation, effect size, p-values from statistical tests such as t-tests and ANOVA, and the performance of classification tools like linear regression, Support Vector Machine, k-Nearest Neighbours, and random forest [12–14]. However, the presentation of results using tabulated data or box plots has often been visually unintuitive and has obscured interpretations due to the limitations of these methods in presenting multidimensional statistical distributions with sparse data. Furthermore, the relatively small sample size in available PD voice datasets has posed challenges for these analysis methods.

This paper presents an analysis of PD voice features using probability density distribution analysis. The analysis was conducted in a two-dimensional space and addressed the issue of sparse data by employing a Gaussian kernel to interpolate the probability density of the scattered data. The study focused on analyzing sustained phonemes of the five vowels recorded in the PC-GITA dataset [10]. The findings of this analysis serve as a basis for the future development of an automatic computerized tool for diagnosing individuals with PD.

2   Methods

2.1   Dataset of Recordings

The voice features analyzed in this study were extracted from the publicly available PC-GITA dataset, provided by Rafael Orozco et al. [10]. This dataset includes recordings of the sustained vowels /a/, /e/, /i/, /o/, and /u/ obtained from 100 native speakers of Colombian-Spanish. Among the participants, 50 individuals had been diagnosed with Parkinson's disease (PD), while the remaining 50 served as healthy control (HC) subjects, matched in terms of age and gender. A comprehensive overview of the demographic and clinical information for both the PD and HC groups can be found in Table 1. The calculated $p-values$, determined through ANOVA or t-test [8], indicate no significant demographic differences between the two groups. The dataset complied with the principles outlined in the Helsinki Declaration and was approved by the Ethics Committee

of the Clinica Noel in Medellin, Colombia. Three recordings of each sustained vowel were collected from each subject.

Table 1. The Demographics of PC-GITA database

	PD		HC		p-value
	Male	Female	Male	Female	
# Subjects	25	25	25	25	
Age (years)	61.56 ± 11.63	60.72 ± 7.27	60.36 ± 11.56	61.44 ± 6.98	0.966
UPDRS	35.92 ± 22.77	37.56 ± 14.03			0.760
H&Y	2.30 ± 0.94	2.28 ± 0.54			0.927
Years diagnosed	8.86 ± 5.88	12.58 ± 11.52			0.157

2.2   The Voice Features

The voice features were extracted from the recordings using a code developed on Praat [1], a publicly available speech analysis software. A total of 10 voice features were extracted from the recordings. The features were related to three aspects of speech production controls that are hypothetically related to Parkinsonian dysarthria: (i) the glottal vibration control stability, (ii) the lung control stability, and (iii) the vocal tract control stability.

The stability of glottal vibration was assessed through jitter absolute, shimmer absolute, harmonic-to-noise ratio (HNR), and noise-to-harmonic ratio (NHR) [13]. Mean and standard deviation of voice intensity features were extracted to evaluate lung performance, while the standard deviation of the formant frequencies $F1$ and $F2$, along with the apparent vocal tract length (VTL) [11] were utilized to capture the subjects' ability to control the vocal tract. Table 2 provides the detail of the features.

2.3   Probability Density Distribution

The probability density distribution function [5] is an estimation of the probability density for two-dimensional scattered data. The estimation is based on two dimensions ($d = 2$) normal Gaussian kernel function, K, scaled at h and is evaluated at equally-spaced points that cover the range of the data in the x and y-axis. In this study, the probability density distribution function, $\hat{f}(x)$, was calculated using the *'kdensity'* function in Matlab 2022 that performs the kernel density estimation as in equation (1).

$$\hat{f}_h(x) = \frac{1}{nh^d(n)} \sum_{i=1}^{n} K\left(\frac{x - x_i}{h(n)}\right), x \in \mathbb{R}^d \tag{1}$$

Table 2. List of The Extracted Features

Voice Production Aspect	Features	unit	Descriptions
Glottal Vibration Stability	Jitter-abs	s	Absolute time perturbation of glottal pulses
	Shimmer-abs	dB	Absolute amplitude perturbation of glottal pulses
	HNR	dB	Harmonics-to-Noise Ratio
	NHR	-	Noise-to-Harmonics Ratio
Lung contol stability	mean(Intensity)	dB	Mean of voice intensity
	SD(Intensity)	dB	Standard deviation of voice intensity
Vocal Tract Stability	SD(F1)	Hz	Standard deviation of formant F1
	SD(F2)	Hz	Standard deviation of formant F2
	VTL(F1)	cm	Apparent Vocal Tract Length based on F1
	VTL(F2)	cm	Apparent Vocal Tract Length based on F2
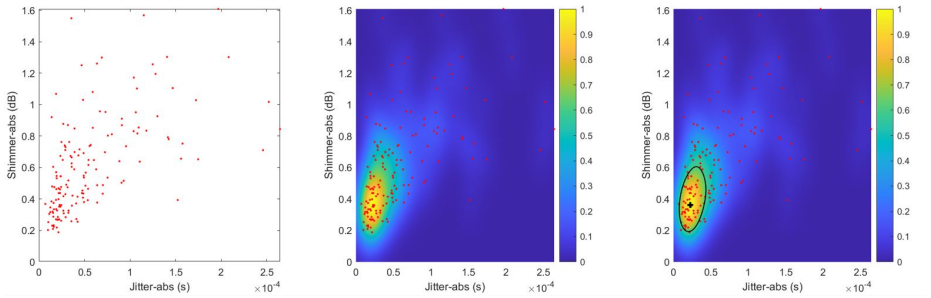
Fig. 1. The probability density distribution function of Jitter-abs versus Shimmer-abs. (a) the scatter plot; (b) the probability density distribution; (c) the contour line of 60.65% probability density and the centroid of the density distribution '+'

Fig. 1 illustrates the process of calculating and displaying the probability density distribution of the Jitter-abs and Shimmer-abs data extracted from recordings of 50 individuals with PD. Fig. 1a presents a scatter plot of the data points. The probability density of the features was estimated using the 'kdensity' function, which normalized the estimation and represented it in a colormap display (Fig. 1b). Fig. 1c illustrates the 60.65% contour line, along with the centroid, indicating the region where the feature density exceeds 60.65%. To simplify the presentation, the subsequent figures in this paper omit the colormap display and only display the 60.65% contour line and the '+' mark to represent the distribution centroid. The 60.65% contour line correlates to the standard deviation of the distribution.

3   Results and Discusions

3.1   Distribution of Glottal Vibration Stability Features

The probability density distributions of the features related to glottal vibration stability are depicted in Fig. 2 and Fig. 3. These figures demonstrate a significant overlap between the PD and HC distributions, with centroids that are closely located to each other. The distribution of HC shows a higher concentration within a smaller region compared to the distribution of PD.
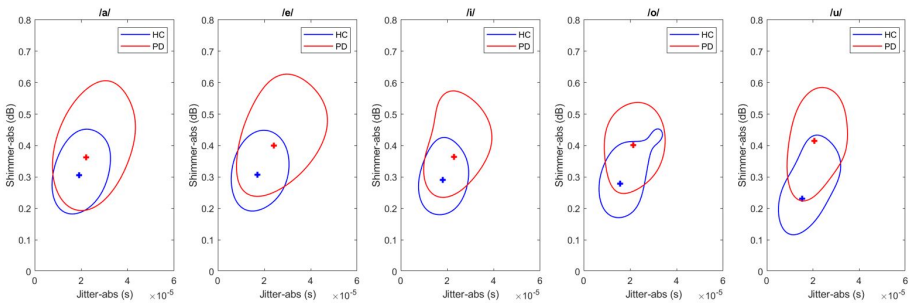

Fig. 2. The 60.65% probability density distribution contour line of Jitter-abs vs Shimmer-abs for PD and HC.
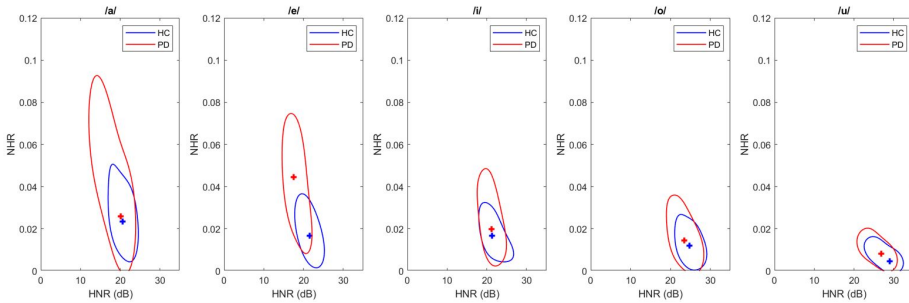
Fig. 3. The 60.65% probability density distribution contour line of HNR vs NHR for PD and HC.

The distribution of time and amplitude perturbation of glottal vibration (Jitter-abs and Shimmer-abs) in the PD group was wider compared to the HC group and showed an overlap with the HC distribution. Individuals with PD may exhibit higher values of jitter and shimmer, indicating increased time and amplitude perturbations in glottal vibration. However, it is worth noting that

a significant number of individuals with PD have jitter or shimmer features that are similar to those of healthy individuals. This observation is particularly observed in the distribution of /a/, /e/, and /i/, as depicted in Fig. 2.

Harmonics-to-noise ratio (HNR) and noise-to-harmonics ratio (NHR) are the features that represent the relative strength of periodic (voiced) and non-periodic (noise) components in glottal vibration [13]. These ratios indicate the extent of harmonic reduction associated with decreased glottal vibration. A low HNR (or high NHR) is indicative of dysarthria, which can be caused by pathological conditions such as PD. Fig. 3 confirms that individuals with PD may exhibit lower HNR and higher NHR values. However, this change is not consistently observed in all individuals with PD. A significant proportion of individuals with PD produce sustained phonemes with HNR and/or NHR values that are similar to those of healthy individuals, as indicated by the overlap between the PD and HC distributions in the figure. Among the vowel phonemes, the features extracted from /a/ were more effective in capturing the PD-related changes, while the features extracted from /u/ were found to be the least effective.

3.2   Distribution of Lung Control Stability Features

Fig. 4 presents the probability density distribution of voice intensity features, which reflect the lung's ability to control voice production. The majority of the figures indicate that there were no significant differences observed between individuals with PD and healthy individuals.
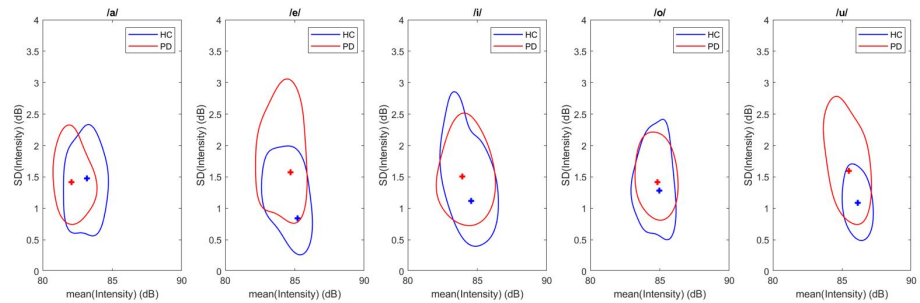

Fig. 4. The 60.65% probability density distribution contour line of voice intensity features for PD and HC.

The distribution of voice intensity features extracted from /e/ and /u/ demonstrates the potential to differentiate individuals with PD from the HC group. People with PD exhibited a higher standard deviation of voice intensity, indicating a greater degree of instability in voice intensity.

3.3   Distribution of Vocal Tract Stability Features

Fig. 5 and 6 illustrate the alterations in the frequencies of formants ($F1$ and $F2$) and the length of the vocal tract associated with PD. The standard deviation of $F1$ and $F2$ during the pronunciation of /a/, /i/, and /o/ exhibited certain variations between PD and HC. While the majority of individuals with PD demonstrated similar instability in $F1$ and $F2$, the figure suggests a possibility that individuals with PD may exhibit higher standard deviations of $F1$ (SD($F1$)) and $F2$ (SD($F2$)).
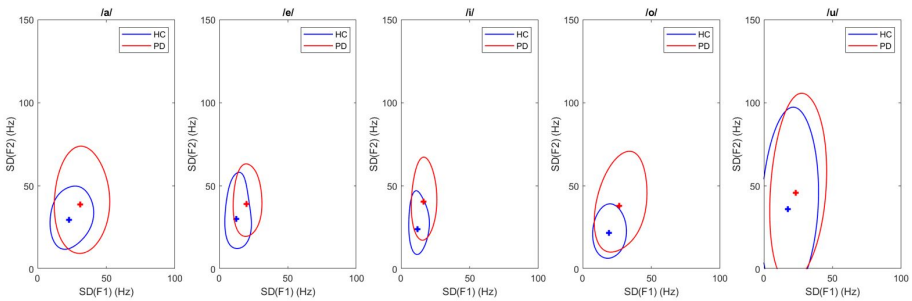
Fig. 5. The 60.65% probability density distribution contour line of SD($F1$) vs SD($F2$) for PD and HC.
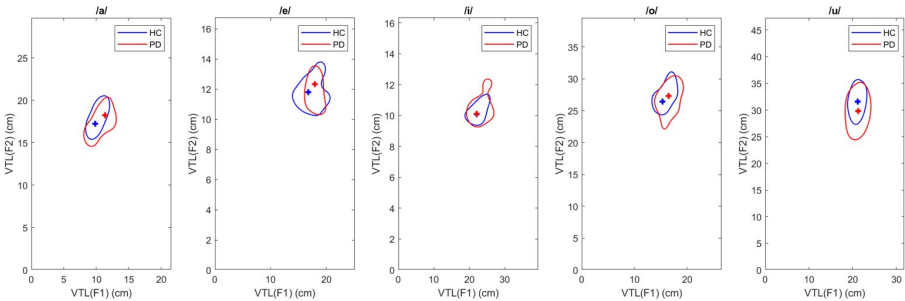

Fig. 6. The 60.65% probability density distribution contour line of apparent vocal tract length features for PD and HC.

The vocal tract length refers to the measurement of the midline length along the vocal tract, from the glottis to the lips. It is a crucial factor influencing variability in speech production [6]. Studies have shown a strong correlation between vocal tract length and formant frequencies [21]. The apparent vocal tract length is an estimation of the physical length of an individual's vocal tract

while producing a specific sound, based on the formant frequencies. Pah et al. [12] reported the relation of VTL to sustained phonemes features of Parkinson's disease. Fig. 6 displays the probability distribution of the apparent vocal tract length, extracted from recordings in the PC-GITA dataset. The figure indicates that there were no significant changes observed in the vocal tract length as a result of Parkinson's disease.

4   Conclusions

Parkinsonian hypokinetic dysarthria, an early symptom of PD, holds promise as a potential biomarker for PD. Voice features obtained from sustained phonemes have been investigated as a potential diagnostic marker for PD. This study employs the probability density distribution function to assess the effectiveness of sustained phoneme features associated with glottal vibration, lung functionality, and vocal tract control in distinguishing individuals with PD from healthy subjects (HC) using the PC-GITA dataset.

The analysis reveals a wider spread in the probability density distribution of PD compared to HC across most features. The presence of significant overlap between the PD and HC distributions suggests that a considerable number of individuals with PD exhibit voice features that are similar to those of HC subjects. Notably, the glottal vibration features extracted from /a/, /e/, and /i/ demonstrate effectiveness in capturing this distinction. Moreover, lung control features derived from /e/ and /u/ have the potential to differentiate individuals with PD. The distribution of vocal tract stability features highlights an increased instability of $F1$ and $F2$ frequencies in individuals with PD. Conversely, no significant changes are observed in vocal tract length features.

In summary, this study underscores the usefulness of probability density distribution to analyze voice features extracted from sustained phonemes as a biomarker for PD. Further research is required to fully leverage these findings and develop a reliable diagnostic tool for Parkinsonian hypokinetic dysarthria in PD.

Acknowledgment

References

1. Boersma, B.P., Heuven, V.V.: Speak and unSpeak with PRAAT. Glot International **5**(9-10), 341–347 (2001)

2. Braga, D., Madureira, A.M., Coelho, L., Ajith, R.: Automatic detection of Parkinson's disease based on acoustic analysis of speech. Engineering Applications of Artificial Intelligence **77**, 148–158 (2019). https://doi.org/10.1016/j.engappai.2018.09.018,   https://doi.org/10.1016/j.engappai.2018.09.018

3. Goetz, C.G., Tilley, B.C., Shaftman, S.R., Stebbins, G.T., Fahn, S., Martinez-martin, P., Poewe, W., Sampaio, C., Stern, M.B., Dodel, R., Dubois, B., Holloway, R., Jankovic, J., Kulisevsky, J., Lang, A.E., Lees, A., Leurgans, S., Lewitt, P.A., Nyenhuis, D., Olanow, C.W., Rascol, O., Schrag, A., Teresi, J.A., Hilten, J.J.V., Lapelle, N.: Movement Disorder Society-Sponsored Revision of the Unified Parkinson's Disease Rating Scale (MDS-UPDRS): Scale Presentation and Clinimetric Testing Results. Movement Disorders **23**(15), 2129–2170 (2008). https://doi.org/10.1002/mds.22340

4. Goyal, J., Khandnor, P., Aseri, T.C.: Engineering Applications of Artificial Intelligence Classification, Prediction, and Monitoring of Parkinson's disease using Computer Assisted Technologies: A Comparative Analysis . Engineering Applications of Artificial Intelligence **96**(August), 103955 (2020). https://doi.org/10.1016/j.engappai.2020.103955, https://doi.org/10.1016/j.engappai.2020.103955

5. Jarnicka, J.: Multivariate kernel density estimation with a parametric support. Opuscula Mathematica **29**(1), 41 (2009). https://doi.org/10.7494/opmath.2009.29.1.41

6. Lammert, A.C., Narayanan, S.S.: On short-time estimation of vocal tract length from formant frequencies. PLoS ONE **10**(7), 1–23 (2015). https://doi.org/10.1371/journal.pone.0132193

7. de Lau, L.M., Breteler, M.M.: Epidemiology of Parkinson's disease. Lancet Neurol **5**(6), 525–535 (2006). https://doi.org/10.1016/S1474-4422(06)70471-9, https://www.ncbi.nlm.nih.gov/pubmed/16713924

8. McDonald, J.H.: Handbook of Biological Statistics. Sparky House Publishing, Baltimore, Maryland, USA, 3rd edn. (2014)

9. Moro-Velázquez, L., Gómez-García, J.A., Godino-Llorente, J.I., Villalba, J., Orozco-Arroyave, J.R., Dehak, N.: Analysis of speaker recognition methodologies and the influence of kinetic changes to automatically detect Parkinson's Disease. Applied Soft Computing **62**, 649–666 (2018). https://doi.org/10.1016/j.asoc.2017.11.001,   http://dx.doi.org/10.1016/j.asoc.2017.11.001

10. Orozco-Arroyave, J.R., Arias-Ledono, J.D., Vargas-Bonilla, J.F., Gonzalez-Rativa, M.C.: New Spanish speech corpus database for the analysis of people suffering from Parkinson's disease. In: International Conference on Language Resources and EvaluationAt: Reykjavik, Iceland. No. May 2014 (2014)

11. Pah, N.D., Indrawati, V., Kumar, D.K.: Voice features of sustained phoneme as covid-19 biomarker. IEEE Journal of Translational Engineering in Health and Medicine **10**, 1–9 (2022). https://doi.org/10.1109/JTEHM.2022.3208057

12. Pah, N.D., Motin, M.A., Kumar, D.K.: Voice Analysis for Diagnosis and Monitoring Parkinson's Disease. In: Arjunan, S.P., Kumar, D.K. (eds.) Techniques for Assessment of Parkinsonism for Diagnosis and Rehabilitation, pp. 119–134. Springer Nature Singapore (2022)

13. Pah, N.D., Motin, M.A., Kumar, D.K.: Detecting Effect of Levodopa in Parkinson's Disease Patients Using Sustained Phonemes. IEEE Journal of Translational Engineering in Health and Medicine **9**, 1–9 (2021). https://doi.org/10.1109/JTEHM.2021.3066800

14. Pah, N.D., Motin, M.A., Kumar, D.K.: Phonemes based detection of parkinson's disease for telehealth applications. Scientific Reports **12**(1), 1–9 (2022). https://doi.org/10.1038/s41598-022-13865-z, https://doi.org/10.1038/s41598-022-13865-z

15. Parisi, L., RaviChandran, N., Manaog, M.L.: Feature-driven machine learning to improve early diagnosis of Parkinson's disease. Expert Systems with Applications **110**, 182–190 (2018)

16. Rabey, J.M., Korczyn, A.D.: The Hoehn and Yahr Rating Scale for Parkinson's Disease. Instrumental Methods and Scoring in Extrapyramidal Disorders (2), 7–17 (1995). https://doi.org/10.1007/978-3-642-78914-4_2

17. Rusz, J., Cmejla, R., Ruzickova, H., Ruzicka, E.: Quantitative acoustic measurements for characterization of speech and voice disorders in early untreated Parkinson's disease. The Journal of the Acoustical Society of America **129**(1), 350–367 (2011). https://doi.org/10.1121/1.3514381, https://www.ncbi.nlm.nih.gov/pubmed/21303016

18. Sakar, C.O., Serbes, G., Gunduz, A., Tunc, H.C., Nizam, H., Sakar, B.E., Tutuncu, M., Aydin, T., Isenkul, M.E., Apaydin, H.: 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 Journal **74**, 255–263 (2019). https://doi.org/10.1016/j.asoc.2018.10.022, https://doi.org/10.1016/j.asoc.2018.10.022

19. Simonet, C., Schrag, A., Lees, A.J., Noyce, A.J.: The motor prodromes of parkinson's disease: from bedside observation to large-scale application. J Neurol pp. 1–10 (2019). https://doi.org/10.1007/s00415-019-09642-0, https://www.ncbi.nlm.nih.gov/pubmed/31802219

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

21. Turner, R.E., Walters, T.C., Monaghan, J.J.M., Patterson, R.D.: A statistical, formant-pattern model for segregating vowel type and vocal-tract length in developmental formant data. The Journal of the Acoustical Society of America **125**(4), 2374–2386 (2009). https://doi.org/10.1121/1.3079772

22. Yang, S., Wang, F., Yang, L., Xu, F., Luo, M., Chen, X., Feng, X., Zou, X.: The physical significance of acoustic parameters and its clinical significance of dysarthria in Parkinson's disease. Scientific Reports **10**(11776), 1–9 (2020). https://doi.org/10.1038/s41598-020-68754-0, https://doi.org/10.1038/s41598-020-68754-0