Objectification of dysarthria in Parkinson's disease using Bayes theorem

JAN RUSZ[1,2], ROMAN CMEJLA[1], HANA RUZICKOVA[2], EVZEN RUZICKA[2]
[1]Department of Circuit Theory, Faculty of Electrical Engineering, Czech Technical University in Prague
Technicka 2, 166 27, Prague 6
CZECH REPUBLIC
[2]Department of Neurology and Centre of Clinical Neuroscience, First Faculty of Medicine, Charles University in Prague
Katerinska 30, 120 00, Prague 2
CZECH REPUBLIC
ruszjan@fel.cvut.cz

Abstract: - This paper presents an assessment of vocal impairment for separating healthy persons from patients with Parkinson's disease (PD). We have recently shown that deterioration of speech performances in PD speakers is notable from an early stage of the disease, even before starting pharmacotherapy. In this study, we present the potential of the simple Bayes rule to reveal changes in degradable speech performance in the course of PD-related dysarthria. The various speech data were recorded from 23 speakers with recently diagnosed PD and 23 healthy speakers. It has been found that 19 various acoustic measurements are able to differentiate PD significantly from healthy speakers. Subsequently, the Bayes theorem was applied to each of these measurements. As a result, the 21 PD patients and 21 healthy people were correctly classified according to their group. The Bayes theorem thus confirms its feasibility for identifying the features of the impaired voice.

Key-Words: Parkinson's disease, Hypokinetic dysarthria, Speech and voice disorders, Acoustic analysis, Speech signal processing, Bayes theorem.

1 Introduction

Parkinson's disease (PD) is a neurological illness characterized by progressive degeneration of dopaminergic neurons, primarily in the midbrain nucleus of the substantia nigra [1]. This progressive dopaminergic loss is associated with a variety of motor deficits and non-motor deficits such as disorders of mood, behaviour, thinking, sensation, and speech characterized as hypokinetic dysarthria. As the second most common neurodegenerative disorder after Alzheimer's disease, PD affects a large part of worldwide population. PD is assumed primarily to affect persons over the age of 50; only approximately 10% of patients report symptoms before the age of 40 [2]. Moreover, PD affects 1.6 % of all persons after the age of 65 [3]. In addition, statistics of the number of persons with PD are expected to increase with the aging of the worldwide population in total [4]. Currently, there is no available causal cure, although medication offers alleviation of some symptoms, especially at early stages of the disease [5]. Thus, early diagnosis of PD has an important role in slowing down or even preventing the degenerative progress of this disease. Moreover, early diagnosis of PD will be crucial when treatment become feasible.

Several previous studies have shown that vocal impairment may be one of the earliest indicators of the disease [6, 7], and deficiencies in speech affect approximately $75-90$ % people with PD [8, 9]. In addition, the measurement of the voice is non-invasive, cheap and simple to administer. Therefore, development of tools capable of performing automatic vocal tests can be very useful for assisting in tracking of the progression of the disease, and thus can partially alleviate the inconvenience and cost of physical visits [10, 11].

The most salient features of PD speech impairment include deficits in the production of vocal sounds (dysphonia), and problems with motor speech disorder (dysarthria) [9]. On the other hand, it has been demonstrated that people with PD may show individual deficits in various speech subsystems such as *phonation*, *articulation*, and *prosody*, starting from early stages of disease [12]. Furthermore, PD individuals may manifest abnormalities in all dimensions of speech including reduced melody, reduced loudness, imprecise articulation, reduced stress, variability of speech rate, speech disfluencies, and others [12, 13].

The deficits in speech related to PD are of a wide scope of interest; however, there is a lack of acoustic characterizations of the extent of vocal impairment in early stages of PD where the progression of symptoms of PD speech is not affected by medication. Thus, we investigated quantitative acoustic parameters to explore the signs of PD-related degradable speech symptoms [12]. In this study, we focus on the effectiveness of the Bayes theorem to assess the extent of vocal impairment in early untreated PD.

Table 1: Summary of the participants' data.

	Subjects	
	PD (n = 23)	HC (n = 23)
Age (year)	61.74 ± 12.60	58.08 ± 12.91
Male	*n* = 19	*n* = 16
Female	*n* = 4	*n* = 7
Duration of PD (month)	30.22±22.21	n/a
H&Y stage	1-2	n/a
UPDRS III score	17.52 ± 7.26	n/a

The values are given in the form mean ± standard deviation. Entries labeled "n/a" are not applicable for HC.

2 Methods

2.1 Speech Data

We used data from the original study Rusz *et al*., in which 23 subjects with recently diagnosed idiopathic PD were recruited [12]. None of these subjects received symptomatic pharmacotherapy or speech treatment. All PD patients were examined in the drug-naive state, before the treatment was started. As a healthy control (HC) group, 23 persons with no history of neurological disorders were matched for the respective age. Table 1 summarizes the details of all participants.

Table 2 details the speech data used; the vocal tasks ranged from producing isolated vowels to reading short sentences and producing a spontaneous monolog on the given subjects. In each vocal task, the best speech performances for every subject were retained.

Table 2: Summary of speech data.

Speech data
Sustained phonation of /i/ on one breath at a comfortable pitch and loudness as constant and long as possible, at least 5-sec.
Rapid steady /pa/-/ta/-/ka/ syllables repetition on one breath as constant and long as possible, repeated at least 5-times.
Approximately 5-sec sustained vowels of /a/, /i/, /u/ on one breath at a comfortable pitch and loudness.
Reading the same standard text of 136 word
Monologue, at least approx. 90-sec.
Reading the same text containing 8 variable sentences of 71 words with varied stress patterns on 10 indicated words.
Reading 10 sentences according specific emotions in a comfortable voice in response to an emotionally neutral sentence.
Rhythmically read text containing 8 rhymes of 34 words following the example set by the examinator.

Table 3: Overview of measurement methods used.

Acoustic features description
F0 SD - variations of fundamental frequency, vibration rate of vocal folds [15].
Jitter - the average absolute difference between a period and the average of it and its four closest neighbours, divided by the average period [15].
Shimmer - average absolute difference between the amplitudes of consecutive periods, divided by the average amplitude [15].
NHR - Noise-to-Harmonics-Ratio, the amplitude of noise relative to tonal components [15].
HNR - Harmonics-to-Noise-Ratio, the amplitude of tonal relative to noise components [15].
Percent pause time - the percent change from the unedited sample length to the edited sample length [12].
Articulation rate - the number of syllables produced per second, after removing silence period exceeding 60 ms [12].
Number of pauses - the number of all pauses compared to total time duration, after removing silence period not lasting more than 60 ms [12].
Intensity SD - variations of average squared amplitude within a predefined time segment ("energy") after removing silence period exceeding 60 ms [12].
DDK rate - the number of /pa/-/ta/-/ka/ syllable vocalizations per second [12].
DDK regularity - the degree of /pa-/ta/-/ka/ syllable vocalizations rate variations in the period [12].
Vowel space area - quantitative measure which involves plotting the three corner vowels in F1/F2 plane [12].
Rhythm - measurement of ability to reproduce perceived rhythm through dynamic time warping [12].
RIRV - Relative Intensity Range Variations, the variations of energy [12].
RRIS - Robust Relative Intensity Slope, the robust linear regression of energy [12].
SDCV - Spectral Distance Change Variations, the variations of spectral distance changes in signal spectrum [12].
RFPC - Robust Formant Periodicity Correlations, the first autocorrelation coefficient of F2 contour [12].

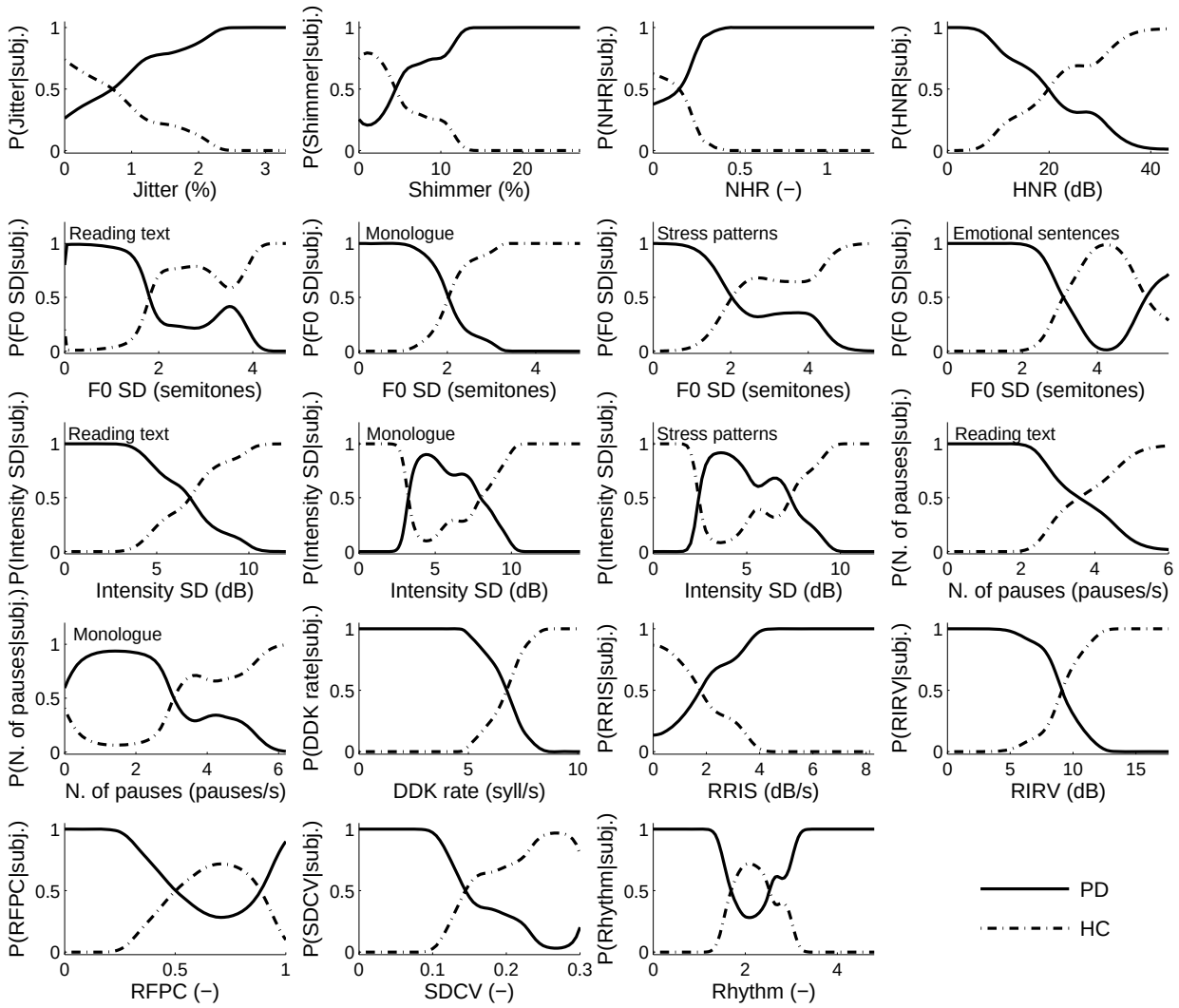
2.2 Feature selection

There are number of acoustic measures that can be applied to selected speech data. We have extracted the fundamental frequency of vocal chords using direct-time domain pitch estimation algorithm [14, 15]. Afterward, we calculated *fundamental frequency variations* (F0 SD) for sustained phonation (for demonstrating defects in phonation) and for text reading, monolog, stress pronouncement, and emotions (for demonstrating reduced melody of speech).

Using the speech analyzer PRAAT [16], we have extracted measures of *shimmer*, *jitter*, *noise-to harmonics ratio* (NHR), and *harmonics-to-noise ratio* (HNR) from the vocal task of sustained phonation. These were used to demonstrate deficits in voice functions.

ISBN: 978-960-474-276-9

Fig. 1: The probability values *P*(*subject*|*measure*) of the acoustic measures used as features for Bayes theorem classification. The dash-dot lines are for HC speakers, the solid lines for Parkinson's speakers.



To assess variability of the speech rate, we calculated *articulation rate*, *percent pause time*, and *number of pauses* using vocal tasks of reading text and monolog.

Intensity of voice was computed from the signal energy contour with relative calibration to the reference 0 dB. Subsequently, *intensity variations* (Intensity SD) were determined using the reading text, monolog, and stress patterns.

Diadochokinetic (DDK) *rate* and *regularity* [17] were determined from repetition of three-syllable items of /pa/-/ta/-/ka/.

The first (F1) and second (F2) formant frequency were obtained from sustained phonations using a robust formant tracker [18]. From the corner vowels of /a/, /i/, /u/, the *vowel space area* was calculating by plotting on an *xy* coordinate plane with F1 on the *x*-axis and F2 on the *y*-axis.

Ability to reproduce perceived *rhythm* was measured using a rhythmically read text, calculated as the similarity between subject performance and a template recording on the basis of dynamic time warping [19].

Several novel acoustic measures of articulation were designed and performed using /pa/-/ta/-/ka/ syllable repetition. These include measurement of the sound pressure level decline calculated as a *robust relative intensity slope* (RRIS), measurement of *relative intensity range variations* (RIRV), *spectral distance change variations* (SDCV) calculated using the Bayesian autoregressive change-point detector [20], and *robust formant periodicity correlation* (RFPC) calculated as the self-similarity of F2 sequence.

Table 3 summarizes the measurements used; a detailed description of all measures can be found in [12].

2.3 Statistics

As not all variables show normal distribution, the non-parametric Wilcoxon signed rank-sum test was used for comparison between the PD and HC group.

For each measurement that obtains a statistically significant difference between the groups, the Bayes theorem is applied. The probability densities for each feature for both PD and HC groups *P*(*measure*|*subject*) are estimated by using the Gaussian kernel density

ISBN: 978-960-474-276-9

method. The number of samples for both groups for each measurement are of the same size, therefore we can consider the same probability of $P(PD) = P(HC) = 0.5$. The probabilities that a person belongs to the PD group in dependence on given feature $P(PD|measure)$ are then obtained using the Bayes theorem

$$P(PD \mid measure) = \frac{P(measure \mid PD) \times P(PD)}{P(measure \mid PD) \times P(PD) + P(measure \mid HC) \times P(HC)}.$$

For each subject, we calculated the average sum of points $P(subject|overall)$ across all retained measurements. The higher value then predicts greater vocal impairment, while a value greater than 0.5 predicts PD speech performance and a value lower than 0.5 predicts HC speech performance.

3 Results

Figure 1 shows probabilities for 19 acoustic measurements that were able significantly to separate PD from HC. These include F0 SD extracted from reading text, monolog, stress patterns, and emotions; jitter, shimmer, NHR, and HNR extracted from sustained phonation; Intensity SD extracted from reading text, monolog, and stress patterns; number of pauses extracted from reading text and monolog; rhythm extracted from rhythmically read text; and DDK rate, RIRS, RIRV, SDCV, RFPC extracted from /pa/-/ta/-/ka/ syllable repetition.

From these results, we can consider that abnormalities in phonation captured by jitter, shimmer, NHR, and HNR may be clinically interpreted as hoarseness, hypophony, and tremolo. Deficits in articulation captured during rapid articulation by DDK rate, RIRS, RIRV, SDCV, and RFPC can manifest occlusive weakening, lowered clarity and accuracy of articulation, and weakness in the production of stable airflow from the lungs. Finally, the defects in prosody captured by F0 SD, Intensity SD, Number of Pauses, and Rhythm can be caused by changed laryngeal tension, decreased breath support, and decreased range of motions.

Table 4 details the results obtained using the Bayes rule. From overall measurements performances, the 42 participants (91.30 %) were classified according to their group. When taking separately, two of the 23 persons with PD (8.7 %) reached a speech performance of HC - e.g. $P(PD|overall) < 0.5$, and two of the 23 healthy persons (8.7 %) reached PD speech performance, - e.g. $P(PD|overall) > 0.5$. The classifier on the basis of Bayes theorem was confirmed to find the sign of disordered or healthy voices according to the subject's speech performance.

Using Bayes rule classification applied to separate measurements, the RIRV in DDK task carry the greatest amount of information for separating both groups of speakers with classification score of 85.71 %. The F0 SD measurement in monolog was the best assessment method and gained 80.42% performance. The lowest score in determining both of group speech performances was found in measurement of jitter. The accuracy of the remaining measurements ranged between 63.04 % and 78.25 %.

Table 4: Summary of the classification results using Bayes theorem.

Measurement	PD correctly classified	HC correctly classified	(No.) Overall classification
Sustained phonation			
01. Jitter	34.78 %	91.30 %	(19.) 63.04 %
02. Shimmer	69.57 %	86.96 %	(3-5.) 78.25 %
03. NHR	30.42 %	95.65 %	(3-5.) 78.25 %
04. HNR	65.22 %	82.61 %	(9-14.) 73.91 %
Reading text			
05. F0 SD	73.91 %	73.91 %	(9-14.) 73.91 %
06. Intensity SD	78.26 %	69.57 %	(9-14.) 73.91 %
07. N. of pauses	56.52 %	73.91 %	(18.) 65.22 %
Monolog			
08. F0 SD	100 %	60.87 %	(2.) 80.42 %
09. Intensity SD	78.26 %	69.57 %	(9-14.) 73.91 %
10. N. of pauses	60.87 %	91.30 %	(6-8.) 76.08 %
Stress patterns			
11. F0 SD	65.22 %	78.26 %	(15.) 71.74 %
12. Intensity SD	82.61 %	65.22 %	(9-14.) 73.91 %
Emotional sentences			
13. F0 SD	78.26 %	78.26 %	(3-5.) 78.25 %
Rhythmic text			
14. Rhythm	52.38 %	85.71 %	(16-17.) 69.05 %
DDK task			
15. DDK rate	76.19 %	76.19 %	(6-8.) 76.19 %
16. RRIS	71.74 %	80.41 %	(6-8.) 76.08 %
17. RRIV	85.71 %	85.71 %	(1.) 85.71 %
18. RFPC	52.38 %	85.71 %	(16-17.) 69.05 %
19. SDCV	57.14 %	90.28 %	(9-14.) 73.91 %
Overall performance	91.30 %	91.30 %	(all) 91.30 %

4 Summary and Conclusion

In this study, we introduce the classifier based on the Bayes theorem for separating healthy people from persons with PD, in assessing of their speech performances. For extraction of the features from speech, we have designed number of new measures that are capable of automatic assessment of the major part of traditionally clinically used methods for quality of speech evaluation [12]. Here, we demonstrate that simple Bayes rule is capable to reliable assess the extent of vocal impairment of each subject and efficiency of each measurement. Considering that only 4 persons were incorrectly classified and overall classification rate was approximately 91 %, we can confirm the Bayes theorem as an available classifier that can be useful in effective

and objective assisting in evaluation of speech and voice disorders.

The results of this study thus support the argument that features of impaired speech can be partially captured from early stages of PD. Acoustic analysis and speech signal processing algorithms have proved to be an excellent tool for voice disorders detection. The use of these techniques combined with classification methods can provide the development of expert aided systems for detection of speech pathology. Acoustic measurements then might serve as useful tool in assessment of vocal impairment, remote tracking of speech progression, and feedback in voice treatment.

Acknowledgment

We are obliged to doctors Jan Roth, Jiri Klempir, Veronika Majerova, and Jana Picmausova for provision of clinical data.

This research was partly supported by the Czech Science Foundation, project GACR 102/08/H008, Grant Agency of the Czech Technical University in Prague, project SGS 10/180/OHK3/2T/13, Czech Ministry of Health, project NT 11331-6/2010 and Czech Ministry of Education, projects MSM 0021620849, and MSM 6840770012.

References:

[1] O. Hornykiewicz, "Biochemical aspects of Parkinson's disease," *Neurology* Suppl. 2, **51**, S2-S9, (1998).

[2] M. Hoehn, "The natural history of Parkinson's disease in the pre-levodopa and post-levodopaeras," *Neurologic Clinics*, **10**, 331-339, (1992).

[3] M. C. de Rijk, C. Tzourio, M. M. Breteler, J. F. Dartiques, L. Amaducci, S. Lopez-Pousa, J. M. Manubens-Bertran, A. Alpérovitch and W. A. Rocca, "Prevalence of parkinsonism and Parkinson's disease in Europe: the EUROPARKINSON Collaborative Study. European Community Concerted Action on the Epidemiology of Parkinson's disease," *J. Neurol. Neurosurg. Psychiatry*, **62**, 10-15, (1997).

[4] S. K. Van Den Eeden, C. M. Tanner, A. L. Bernstein, R. D. Fross, A. Leimpeter, D. A. Bloch, and L. M. Nelson, "Incidence of Parkinson's disease: Variation by Age, Gender, and Race/Ethnicity," *Am. J. Epidem*., **157**, 1015-1022, (2003).

[5] N. Singh, V. Pillay, and Y. E. Choonara, "Advances in the treatment of Parkinson's disease," *Prog. Neurobiol*., **81**, 29-44, (2007).

[6] J. R. Duffy, "Motor Speech Disorders: substrates, differential diagnosis and management. New York: Mosby, 2nd ed.; 2005.

[7] B. Harel, M. Cannizzaro, and P. J. Snyder, "Variability in fundamental frequency during speech in prodromal and incipient Parkinson's disease: A longitudinal case study," *Brain Cognition*, **56**, 24-29, (2004).

[8] A. K. Ho, R. Iansek, C. MariglianC, J. Bradshaw, and S. Gates, Speech impairment in large sample of patients with Parkinson's disease," *Behav. Neurol.*, **11**, 131-137, (1998).

[9] J. A. Logemann, H. B. Fisher, B. Boshes, and E. R. Blonsky, "Frequency and coocurrence of vocal tract dysfunction in the speech of a large sample of Parkinson patients," *J. Speech Hear. Disord*., **43**, 47-57, (1978).

[10] M. A. Little, P. E. McSharry, E. J. Hunter, J. Spielman, and L. O. Ramig, "Suitability of dysphonia measurement for telemonitoring of Parkinson's disease," *IEEE Trans. Biomed. Eng*., **56**, 1015-1022, (2009).

[11] C. Ruggiero, R. Sacile, and M. Giacomini, "Home telecare," *J. Telemed. Telecare*, **5**, 11-17, (1999).

[12] 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," *J. Acoust. Soc. Am*., **129**, 350-367, (2011).

[13] A. M. Goberman, and C. Coelho, "Acoustic analysis of Parkinsonian speech I: Speech characteristics and L-Dopa therapy," *Neurorehab*., **17**, 237-246, (2002).

[14] H. Boril, and P. Pollak, "Direct time domain fundamental estimation of speech in noisy conditions," In: *EUSIPCO Proceedings*, Vienna, Austria; 2004.

[15] H. Boril, "Robust speech recogniton: Analysis and equalization of Lombard effect in Czech corpora," Ph.D. dissertation, Czech Technical University in Prague, Czech Republic, (2008).

[16] P. Boersma, and D. Weenink, "Praat, a system for doing phonetics by computer," *Glot International*, **5,** 341-345, (2001).

[17] S. Fletcher, "Time-by-count measurement of diadochokinetic syllable rate," *J. Speech. Hear. Disord.,* **15**, 757–762, (1972).

[18] K. Mustafa and I. C. Bruce, "Robust formant tracking for continuous speech with speaker variability," *IEEE Trans. Audio, Speech, Lang. Process.*, **14**, 435–444, (2006).

[19] H. Sakoe and S. Chiba, "Dynamic programming algorithm optimization for spoken word recognition," *IEEE Trans. Acoust. Speech Signal Process.*, **26**, 43–49, (1978).

[20] R. Cmejla, and P. Sovka, "Recursive Bayesian autoregressive changepoint detector for sequential signal segmentation," In *EUSIPCO Proceedings*, Wien (2004), pp. 245–248.