QUANTIFICATION OF PROSODIC IMPAIRMENT IN PATIENTS WITH IDIOPATHIC PARKINSON'S DISEASE

Zoltán Galáž
Doctoral Degree Programme (2), FEEC BUT
E-mail: z.galaz@phd.feec.vutbr.cz

Supervised by: Jiří Mekyska
E-mail: mekyska@feec.vutbr.cz

Abstract: This paper deals with quantitative analysis of prosodic impairment in idiopathic Parkinson's disease (PD). Experimental dataset consisted of 97 PD patients and 55 healthy speakers. The prosodic features expressing monopitch, monoloudness and speech rate deficits are extracted from stress-modified reading task. Classification accuracies of 70.71 % for females, 70.03 % for males, and 63.20 % for a mixture of both gender were achieved. According to permutation test (1000 permutations, $\alpha = 0.01$), the models were shown statistically significant. Promising potential of prosodic features to identify HD was confirmed.

Keywords: Parkinson's disease, hypokinetic dysarthria, dysprosody, speech processing

1 INTRODUCTION

Parkinson's disease (PD) is a chronic neurodegenerative disorder [1] affecting approximately 1.5 % of people aged over 65 years [1]. It is associated with a substantial reduction of dopaminergic neurons especially in *substancia nigra pars compacta* [2] resulting into variety of motor and non-motor symptoms. According to the previous studies [3], approximately 60–90 % of PD patients suffer from perceptually distinctive motor speech disorder referred to as hypokinetic dysarthria (HD) [4].

HD affects many aspects of speech, namely the area of phonation, articulation, prosody, speech fluency and faciokinesis [5, 6, 7, 8]. Detrimental impact of HD on human verbal communication and daily social life has been observed [9]. Prosodic impairment (dysprosody) is a common speech flaw in HD. It is mainly characterized by variable speech rate, reduced variations in pitch (monopitch) and intensity (monoloudness) [10] leading to unnatural and unintelligible speech. Dysprosody has been observed even in the early stages of PD [11].

At present, PD can not be definitely cured. Therefore, the current medicine is focused on treating its cardinal motor symptoms. However, effectiveness of the treatment depends upon the stage of the disease during which it is initiated [6]. Therefore, the accurate and early diagnosis is necessary for the clinicians to efficiently treat the patients. Nevertheless, the early diagnosis of PD requires a complex understanding of its manifestations on human body. Acoustic analysis of dysprosody is non-invasive, paraclinical method that has a great potential to provide clinicians an objective assessment of motor symptoms of PD.

The aim of this work is to evaluate methods of quantitative analysis of dysprosody that can identify HD and indirectly presence of PD. For this purpose, a variety of prosodic speech features are calculated. These features assess monopitch, monoloudness and speech rate deficits in HD. The features are extracted from the stress-modified reading task specially designed to capture dysprosody. Consequently, the features are used to train the classification model capable of HD identification. The rest of this paper is organized as follows. Section 2 presents the dataset and the methodology. Experimental results are presented in section 3, and section 4 provides discussion and some conclusions.

2 MATERIALS AND METHODS

2.1 DATA ACQUISITION

In this work, we investigated the speech recording acquired from 152 Czech native speakers. The speakers comprised 97 PD patients (53 men/44 women; mean age 67.52 $\pm$ 8.29 years; mean disease duration 7.80 $\pm$ 4.42 years) and 55 healthy speakers (29 men/22 women; mean age 63.96 $\pm$ 9.21 years). The participants were enrolled at the First Department of Neurology, St. Anne's University Hospital in Brno, Czech Republic. The healthy controls (HC) had no history or presence of speech disorders, brain diseases, including neurological and psychiatric illnesses. All patients were examined on their regular dopaminergic medication approximately 1 hour after the L-dopa dose. All patients signed an informed consent form that had been approved by the Ethics Committee of St. Anne's University Hospital in Brno.

The speech task was composed of 3 sentences (indicative, imperative, and interrogative). It comprised 22 words (87 characters). Regarding speech prosody, the task requires the stress-control in order to emphasize the emotions according to the sentences. The participants were asker to read the following sentences: in Czech – *Teď musíš být chvíli trpělivý, než to dokončíme. Už mě to nebaví, dej mi už konečně pokoj! Tak co, jak to dopadlo?*, English translation – *Now, you have to be patient until we finish it. I'm tired of it already, leave me alone! So, how did it go?* The speech signals were sampled with sampling frequency $f_s = 48\,\text{kHz}$ and subsequently downsampled to 16 kHz.

2.2 SPEECH FEATURES

Following our recent research of parkinsonian dysprosody [10], we extracted prosodic speech features that quantify: a) monopitch; b) monoloudness; c) speech rate abnormalities. Completely, we computed Standard Deviation of fundamental frequency ($F0_S$), Relative Standard Deviation of F0 ($F0_{r.\,S}$), Variation Range of F0 ($F0_R$), Relative Variation Range of F0 ($F0_{r.\,R}$), Standard Deviation of Squared Energy Operator/Teager-Kaiser Energy Operator (SEO_S/TEO_S), Relative Standard Deviation of SEO/TEO ($SEO_{r.\,S}$/$TEO_{r.\,S}$), Variation Range of SEO/TEO (SEO_R/TEO_R), Relative Variation Range of SEO/TEO ($SEO_{r.\,R}$/$TEO_{r.\,R}$), Total Speech Time (TST), Net Speech Time (NST), Total Pause Time (TPT), Total Speech Rate (TSR), Net Speech Rate (NSR), Total Pause Time (pauses longer than 50 ms) (TPT 50), Articulation Rate (AR) and SPeech Index of Rhytmicity (SPIR). These features are conventional in the field of dysarthric speech analysis [13].

To extract F0 contour, Praat acoustic analysis software [14] was used. To extract the rest of speech features, Neurological Disorder Analysis Tool (NDAT) [6, 12] written in MATLAB and developed at the Brno University of Technology was used.

2.3 STATISTICAL ANALYSIS

The features were normalized before the analysis on a per-feature basis to have 0 mean and a standard deviation of 1. Random Forests (RF) [15] classifier was used to investigate the power of prosodic features to discriminate healthy and dysarthric speech. Forward selection approach (modified version of Sequential Floating Forward Selection [16] algorithm) was employed to find a non-redundant combination of the features with the maximum clinical information about deterioration of speech prosody in HD. Matthew's correlation coefficient [17] (MCC $= TP \times TN + FP \times FN / \sqrt{N}$) was computed as a criterion of the feature selection. $N = (TP+FP)(TP+FN)(TN+FP)(TN+FN)$, TP (true positive) and FP (false positive) represents the number of correctly identified PD subjects and a number of subjects identified as PD, but being healthy. Similarly, TN (true negative) and FN (false negative) represent the total number of correctly identified HC, and PD patients identified as HC. Additionally, classification accuracy (ACC), sensitivity (SEN) and specificity (SPE) were computed.

Table 1: Statistical analysis of the prosodic features computed for the analysed speech task.

Feat.	G	MCC	ACC [%]	SEN [%]	SPE [%]	p	No.
F1	F	0.3064 ± 0.3507	64.3810 ± 15.6312	67.5000 ± 25.8775	60.3333 ± 30.8405	0.0990	2
	M	0.2129 ± 0.3776	61.7460 ± 17.7139	63.8000 ± 21.0281	57.6667 ± 31.2640	0.1280	2
	A	0.1458 ± 0.2642	58.3714 ± 13.1348	61.6000 ± 17.2150	52.7333 ± 19.6071	0.1560	2
F2	F	0.3995 ± 0.2622	69.6667 ± 12.0142	72.0000 ± 21.2132	65.6667 ± 26.6007	0.0240	3
	M	0.2412 ± 0.4092	63.3016 ± 18.2413	64.7333 ± 20.6810	59.6667 ± 33.5182	0.0360	4
	A	0.2004 ± 0.2135	61.9512 ± 9.1573	66.3556 ± 13.5546	53.8667 ± 21.5747	0.1150	2
F3	F	0.1492 ± 0.3861	58.9048 ± 15.3989	63.5000 ± 19.6980	51.6667 ± 35.9910	0.4570	2
	M	0.2430 ± 0.3193	64.4444 ± 14.1973	68.7333 ± 18.6493	55.6667 ± 30.4185	0.0650	1
	A	0.1300 ± 0.2503	58.1595 ± 12.3655	62.0222 ± 17.1806	51.1333 ± 21.8800	0.6840	2
F4	F	0.4229 ± 0.3514	70.7143 ± 16.2431	71.0000 ± 19.1397	70.3333 ± 30.9139	0.0001	1
	M	0.3708 ± 0.3446	70.0317 ± 16.0522	73.5333 ± 19.2891	63.0000 ± 29.4103	0.0120	5
	A	0.2497 ± 0.2586	63.2024 ± 12.4402	65.0667 ± 14.9209	60.0000 ± 21.0711	0.0130	3

Table notation: G – gender ([F] females, [M] males, [A] combination of both genders); F1 – monopitch features; F2 – monoloudness features; F3 – speech rate features; F4 – general prosodic features; MCC – Matthew's correlation coefficient; ACC – classification accuracy; SEN – classification sensitivity; SPE – classification specificity; No. – Number of selected features; p – p-values of classification calculated by permutation test (1000 permutations).

In this work, the classifier validation was conducted using 10-fold validation approach with 5 repetitions. Furthermore, to determine if the classification results were obtained by chance or by the actual relationship between the class labels and the data values, non-parametric statistical test referred to as permutation test was used [18]. To achieve sufficient statistical relevancy, 1000 permutations was performed. Significance level α was chosen equal to 0.01. Classification models with p values smaller than α was considered statistically significant.

3 RESULTS

In this work, we analysed dysprosody in patients with idiopathic PD. For this purpose, a set of prosodic speech features quantifying monopitch, monoloudness, and speech rate disturbances was extracted [10]. Non-relevant features were excluded by forward selection algorithm. For determining the power of the features to distinguish between healthy and speech with prosodic impairment, random forests classifier in a supervised learning setup (10-fold validation/5 repetitions) was used. Additionally, 1000 permutations were performed during the evaluation of the models. The results of classification process are summarized in Table 1.

As can be seen from the table, the best classification performance in terms of ACC was achieved by $\mathrm{F4-ACC} = 70.71\,\%$ using a single feature ($\mathrm{TEO_{r.\,R}}$) extracted from the reading of female participants. This model was also proven to be statistically relevant ($p < \alpha$). Regarding male speakers, almost identical classification performance was achieved $\mathrm{F4-ACC} = 70.00\,\%$, however the model was trained with 5 features ($\mathrm{F0_S}$, TPT, $\mathrm{F0_{r.\,R}}$, $\mathrm{SEO_S}$, $\mathrm{SEO_{r.\,R}}$) and also less statistically significant. Combining both genders in one group resulted into classification model of the following performance $\mathrm{F4-ACC} = 63.20\,\%$ ($\mathrm{TEO_R}$, TST, TSR). Of note is the fact that all prosodic submodels (F1 – F3) did not achieve sufficient statistical significance (see Table 1 $p > \alpha$). On the contrary, the model based on general prosodic features (monopitch, monoloudness and speech rate) was shown to be statistically significant.

The results of this paper show the promising potential of prosodic features for quantification of HD in PD. However, subsequent research is warranted to fully understand HD and its manifestation on human speech prosody. Moreover, despite relatively low classification accuracy of the prosodic submodels, further investigation of monopitch, monoloudness and speech rate deficits can lead to development of more robust speech features and therefore more precise HD identification.

4 CONCLUSION

In this paper, we performed a quantitative analysis of dysprosody in a set of 97 patients with idiopathic PD and 55 HC using the speech task based on the stress-modified reading of indicative, interrogative and imperative sentences. The presence of prosodic impairment in patients with PD was observed, which is in accordance with the previous research [19, 13, 20]. To emphasize the possible gender-related discrepancies in prosodic deterioration in HD, the analysis was conducted for both gender separately and also for a mixture of both. Based on the results of this study, we conclude that dysprosody in PD is likely to be gender-differentiated. This observation was considered reasonable taking into account the anatomic differences (e.g. length of the vocal tract, etc.) between the speech production system of men and women.

We showed that acoustic analysis of dysprosody can discriminate healthy and dysarthric speech with the classification accuracy over 70 %. Important thing to note is the simplicity and clinical interpretability of selected speech features. In our opinion, these facts compensate the lack of classification performance compared to more sophisticated methods of vocal pathology identification, such as the non-linear features used to quantify dysphonia [21]. The future investigation in this field of science should lead to a fusion of the two approaches to comprehensively describe the manifestation of PD on human speech production.

In our future study we will focus on deeper statistical analysis of dysprosody in patients with PD. We aim to estimate PD severity using the analysis of prosodic features. Furthermore, a longitudinal study of prosodic impairment in HD is planned. Deeper understanding of dysprosody in PD may lead to development of more robust features capable of assessing motor and non-motor aspects of PD and consequently result into precise, early diagnosis, disease tracking and efficiency of treatment monitoring.

ACKNOWLEDGEMENT

Research described in this paper was financed by the National Sustainability Program under grant LO1401 and by project COST IC1206. For the research, infrastructure of the SIX Center was used.

REFERENCES

[1] M. C. de Rijk, et al., 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 1 (1997) 10–5.

[2] O. Hornykiewicz, Biochemical Aspects of Parkinson's Disease, Neurology 51 (1998) S2–S9.

[3] L. O. Ramig, C. Fox, S. Sapir, Speech treatment for Parkinson's disease, Expert Rev. Neurother. 8 (2008) 297–309.

[4] F. L. Darley, A. E. Aronson, J. R. Brown, Differential Diagnostic Patterns of Dysarthria, J. Speech Hear. Res. 12 (1969) 246–269.

[5] J. Mekyska, I. Rektorova, Z. Smekal, Selection of Optimal Parameters for Automatic Analysis of Speech Disorders in Parkinson's Disease, in: Telecommunications and Signal Processing (TSP), 2011 34th International Conference on, pp. 408–412.

[6] I. Eliasova, J. Mekyska, M. Kostalova, R. Marecek, Z. Smekal, I. Rektorova, Acoustic evaluation of short-term effects of repetitive transcranial magnetic stimulation on motor aspects of speech in Parkinson's disease, J. Neural Transm. 120 (2013) 597–605.

[7] Mekyska, J.; Galaz, Z.; Mzourek, Z.; Smekal, Z.; Rektorova, I.; Eliasova, I.; Kostalova, M.; Mrackova, M.; Berankova, D.; Faundez-Zanuy, M.; Lopez-de-Ipina, K.; Alonso-Hernandez, J.B., "Assessing progress of Parkinson's disease using acoustic analysis of phonation," in Bioinspired Intelligence (IWOBI), 2015 4th International Work Conference on , vol., no., pp. 111-118, 10-12 June 2015.

[8] N. Elfmarkova, M. Gajdos, M. Mrackova, J. Mekyska, M. Mikl, I. Rektorova, Impact of Parkinson's disease and levodopa on resting state functional connectivity related to speech prosody control, Parkinsonism. Relat. Disord. 22 Suppl 1 (2016) 52–5.

[9] M. D. Pell, H. S. Cheang, C. L. Leonard, The impact of parkinson's disease on vocal-prosodic communication from the perspective of listeners, Brain Lang. 97 (2006) 123–134.

[10] Z. Galaz, J. Mekyska, Z. Mzourek, Z. Smekal, I. Rektorova, I. Eliasova, M. Kostalova, M. Mrackova, D. Berankova, Prosodic analysis of neutral, stress-modified and rhymed speech in patients with parkinson's disease, Comput. Methods. Programs. Biomed. (2016) (in print).

[11] J. Rusz, R. Cmejla, H. Ruzickova, 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 129 (2011) 350–367.

[12] J. Mekyska, E. Janousova, P. Gomez-Vilda, Z. Smekal, I. Rektorova, I. Eliasova, M. Kostalova, M. Mrackova, J. B. Alonso-Hernandez, M. Faundez-Zanuy, K. L. Ipiña, Robust and complex approach of pathological speech signal analysis, Neurocomputing 167 (2015) 94–111.

[13] S. Skodda, H. Rinsche, U. Schlegel, Progression of dysprosody in Parkinson's disease over time–a longitudinal study, Mov. Disord. 24 (2009) 716–722.

[14] P. Boersma, D. Weenink, Praat: Doing Phonetics by Computer (Version 5.3.10), 2012.

[15] L. Breiman, Random forests, Mach. Learn. 45 (2001) 5–32.

[16] J. Pohjalainen, O. Rasanen, S. Kadioglu, Feature selection methods and their combinations in high-dimensional classification of speaker likability, intelligibility and personality traits, Computer Speech and Language (2014).

[17] B. W. Matthews, Comparison of the predicted and observed secondary structure of t4 phage lysozyme, Biochim Biophys Acta (BBA) 405 (1975) 442–51.

[18] B. Phipson, G. K. Smyth, Permutation p-values should never be zero: Calculating exact p-values when permutations are randomly drawn, Statistical Applications in Genetics and Molecular Biology 9 (2010).

[19] S. Skodda, U. Schlegel, Speech rate and rhythm in Parkinson's Disease, Mov Disord 23 (2008) 985–92.

[20] S. Skodda, W. Gronheit, U. Schlegel, Intonation and speech rate in Parkinson's Disease: General and dynamic aspects and responsiveness to levodopa admission, J Voice 25 (2011) 199–205.

[21] A. Tsanas, M. A. Little, P. E. McSharry, L. O. Ramig, Nonlinear speech analysis algorithms mapped to a standard metric achieve clinically useful quanti?cation of average Parkinson's disease symptom severity, J. R. Soc. Interface 8 (2010) 842–855.