

ACOUSTIC ANALYSIS OF VOICE AND SPEECH CHARACTERISTICS IN EARLY UNTREATED PARKINSON'S DISEASE

J. Ruzs^{1,2}, R. Cmejla¹, H. Ruzickova², J. Klempir², V. Majerova², J. Picmausova², J. Roth², E. Ruzicka²

¹ Czech Technical University in Prague, Faculty of Electrical Engineering, Department of Circuit Theory, Prague, Czech Republic

² Charles University in Prague, First Faculty of Medicine, Department of Neurology and Centre of Clinical Neuroscience, Prague, Czech Republic

Abstract: Parkinson's disease (PD) is a neurological illness characterized by progressive loss of dopaminergic neurons, primarily in the substantia nigra pars compacta. Changes in speech associated with hypokinetic dysarthria are a common manifestation in patients with idiopathic PD. The aim of this study is to investigate the feasibility of automated acoustic measures for the identification of voice and speech disorders in PD. The speech data were collected from 46 Czech native speakers, 24 with early PD before receiving pharmacotherapy treatment. We have applied several traditional and non-standard measurements in combination with statistical decision-making strategy to assess the extent of vocal impairment of recruited speakers. Subsequently, we have applied support vector machine to find the best combination of measurements to differentiate PD from healthy subjects. This method leads to overall classification performance of 85%. Admittedly, we have found relationships between measures of phonation and articulation and bradykinesia and rigidity in PD. In conclusion, the acoustic analysis can ease the clinical assessment of voice and speech disorders, and serve as measures of clinical progression as well as in the monitoring of treatment effects.

Keywords: Parkinson's disease, speech disorders, hypokinetic dysarthria, acoustic analysis, biomedical application

I. INTRODUCTION

Following the recent findings on the pathogenesis of Parkinson's disease (PD), increased interest has been paid to the nonmotor symptoms indicating an early affection of the lower brainstem that may precede the accession of the main motor signs of PD [1].

As a part of the nonmotor symptoms, voice and speech disorders are still considered to occur inconstantly and to tend to be nonspecific, making them of little diagnostic usefulness in early disease. On the other hand, previous

research has shown that deficiencies in speech affect approximately 75-90% people with PD [2, 3]. The most salient features of PD speech impairment include deficits in the production of vocal sounds and motor involvement of articulation [3-5]. Moreover, it has been demonstrated that PD-related dysarthria can affect all different speech subsystems including *respiration*, *phonation*, *articulation*, and *prosody* [6, 7]. Patients with PD can manifest abnormalities related to all dimensions of speech including monoloudness, monopitch, imprecise articulation, variable speech rate, hoarseness, reduced stress, speech disfluencies, inappropriate silence, and others [7].

To clinically test voice and speech disorders, there are various vocal tests that have been proposed to assess the extent of these symptoms including sustained phonation, diadochokinetic (DDK) task (diadochokinesis connected with articulation), and variable reading of sentences or spontaneous speech [8, 9], that can be subsequently assessed with various traditional and novel acoustic measurements [10]. In our studies, we focus to characterize the speech and voice disorders in the early stages of PD, where the progression of speech symptoms is not affected by medication. In order to find PD-related speech features and separate patients with PD from healthy control (HC) persons, we use several traditional and novel acoustic measurement techniques as well as statistical learning or decision theory.

II. DATA

We used the database of PD speech recordings which we reported in [11]. From 2007 to 2009, a total of 46 Czech native participants were recruited for this research. 24 of these subjects (20 men and 4 women) fulfilling the diagnostic criteria for PD were examined immediately after the diagnosis was made and before the symptomatic treatment was started. As a control group, 22 persons (15 men and 7 women) with no history of neurological or communication disorders were included. None of the participants had been under voice therapy and all gave their consent to the vocal tasks and recording procedure.

The speech data were recorded in a quiet room with a low ambient noise level using an external condenser

microphone placed at approximately 15 cm from the mouth and coupled to a Panasonic NV-GS 180 video camera. The voice signals were sampled at 48 kHz, with 16-bit resolution; the video material was not used. All subjects were recorded at the time of a single session with a speech pathologist. Each participant was instructed to perform at least two times three vocal tasks including sustained phonation, diadochokinetic task, and running speech as a part of a larger protocol. Detailed description of the recording and data can be found in [11].

III. METHODS

We aimed to characterize the PD speech performance, applying the statistical decision-making theory to several acoustic measurements to explore how PD-related vocal symptoms differ from the speech performances of the wider norm of healthy speakers. Subsequently, we designed quick vocal test in order to reduce the time required for voice investigation, represent all speech subsystems, and create reliable assessment in practice, and tested performance of this test in separating PD subjects from HC participants. Finally, we search for a possible correlation of the voice parameters with respect to the duration and severity of disease.

A. Characteristics of voice and speech disorders

We have mainly focused on four speech subsystems including phonation, respiration, articulation, and prosody. For each of these subsystems, we computed several acoustic measures; all the algorithms are described in [10]. In examining phonation of PD speakers, we computed 4 features including *jitter* (the extent of variation of voice range), *shimmer* (the extent of variation of expiratory flow), *noise-to-harmonics* (NHR), and *harmonics-to-noise* (HNR) ratios (the amplitude of noise relative to tonal components in speech) [12]. In examining respiration, we used 1 feature of *Sound Pressure Level Decline* (SPLD – measure the ability to maintain intensity level). In examining articulation, we calculated 3 features including *DDK rate* (number of syllable vocalizations per second), *Robust Formant Periodicity Correlation* (RFPC – quantifies the accuracy of articulation), and *Spectral Distance Change Variation* (SDCV – quantifies the clarity of articulation) [10]. In examining prosody, we used 3 features including *fundamental frequency variation* (F0 SD), *intensity of voice variation* (Intensity SD), and *number of pauses* [10].

Two-sided Wilcoxon rank-sum test was performed to find differences between groups. To explore the extent of PD-related vocal impairment, we applied the *Wald task* decision-making theory to features' Gaussian kernel densities [13]. As a result, for all the features, the subjects were classified as PD (dysarthric speech performance),

HC (intact speech performance), or “not sure” (indecisive situation – performance of wider norm of healthy people). The higher quantity of classifications as PD is associated with progression of PD vocal impairment, the higher quantity of classifications as HC is associated with healthy speech performance.

B. Identification of voice and speech disorders

In order to create reliable assessment in clinical practice, there is a need to test and find the optimal combination of acoustic measurements that gain a useful amount of information for separating early PD from HC. Therefore, we constructed a feature vector with 8 representative measurements including jitter, shimmer, NHR, HNR, SPLD, RFPC, SDCV, and F0 SD. To reduce dimensionality of the data and find the combination of acoustic measurements with the best classification accuracy, the exhaustive search of all possible combinations of features was performed using the method from statistical learning theory called *support vector machine* (SVM) [14]. The SVM classifier with Gaussian radial basis kernel was applied because it allows smooth, curved decision boundaries. On the basis of the decision boundary, the SVM classifier enables to build a predictive model which decides whether a subject belong to the PD or HC group. The choice of optimal SVM parameters was determined by an exhaustive search over a range of values. Cross-validation with the leave-one-out method was used to validate reproducibility (for possible new outcome samples) of SVM classifier; the 50 iteration was used for validation of each combination.

C. Relationships between acoustic features and severity of disease

In addition to speech data, for each of the PD patients, we have administered the *duration of disease* prior to recording, stage of disease according to the *Hoehn & Yahr* (HY) scale (disability scale comprised of stages 1 through 5, where 5 is most severe), and global motor impairment according to the *Unified Parkinson's Disease Rating Scale* (UPDRS) III (motor rating scale from 0 to 108, where 108 represents severe motor impairment). UPDRS III contains 27 items, each scored from 0 (no disability) to 4 (severe disability). We have also administered the three UPDRS III composite subscores including *bradykinesia* (sum of the UPDRS III items 23, 24, 25, 26), *postural instability and gait disorders* (PIGD - sum of UPDRS III items 27, 28, 29, 30), and *rigidity* (UPDRS III item 22). Subsequently, the Person product-moment correlation was used to find relationships between acoustic features and HY stage, duration of PD, UPDRS III score, and UPDRS III composite subscores.

IV. RESULTS

Table 1 summarizes the comparison of subject parameters and speech parameters between Parkinsonian speakers and control group. The final data obtained were composed of 116 recordings (56 from the PD patients, and 60 from HC individuals). After applying the Wald task, we have found that 18/24 patients with PD indicate some form of vocal impairment that differs from the speech performance of the wider norm of healthy people. None of the HC speakers reached the specific dysarthric performance of patients with PD.

From all possible tested measurement combinations, 4 features including NHR, SPLD, RFPC, and F0 SD obtained the best classification score of 85.02%. The classification performance of the entire measurements subset was 81.67%. From individual measures, F0 SD obtained the best classification accuracy of 81.30%. The maximal correct overall classification accuracy was 76.40% using only sustained phonation, and 71.35% using only the DDK task.

In PD patients, there were no statistically significant correlations between the vocal parameters and the stage or duration of the disease. Accordingly, there were no statistically significant correlations between the vocal parameters and UPDRS III scores. However, the partial subscore of bradykinesia significantly correlated with the measure of articulation SDCV ($R = -0.44$, $P < 0.05$) and measures of phonation including jitter ($R = 0.42$, $P < 0.05$), NHR ($R = 0.43$, $P < 0.05$), and HNR ($R = -0.44$, $P < 0.05$). Admittedly, the subscore of rigidity correlated with HNR ($R = -0.43$, $P < 0.05$). There were no significant correlations between the vocal parameters and the UPDRS subscores of PIGD and speech.

V. DISCUSSION

Fig. 1 summarizes the procedure and results of the two-minute vocal test that was employed to evaluate voice and speech disorders in a group of patients with unmedicated PD in comparison to HC people. For the sake of acoustic analysis, the measurements were designed as robust as possible with respect to a possible real-time automatic evaluation in a common acoustic environment and with the presence of contradictory factors such as individual differences in voice and speech. The acoustic measures were used as features for classification of probands into the PD and HC groups. Despite the limited number of speech samples, the best classification accuracy gains performance of 85% was reached using the combination of four measures, each of them representing deficits in one of the speech subsystems related to PD.

According to our results, the deficits in speech prosody appear to contain the greatest amount of information in assessment of early PD-related vocal

Table 1: List of results of acoustic measures with mean±SD values and statistical comparisons between Parkinsonian and healthy groups.

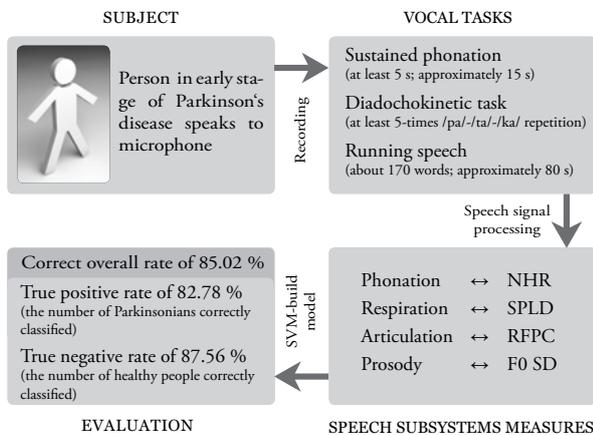
	Subjects		Difference between PD and HC
	PD (n = 24)	HC (n = 22)	
Subject parameters			
Age (year)	60.92±11.24	58.73±14.61	$P = .46$
Male	$n = 20$	$n = 15$	n/a
Female	$n = 4$	$n = 7$	n/a
Duration of PD (month)	31.29±22.25	n/a	n/a
HY stage	2.19±0.48	n/a	n/a
UPDRS III score	17.42±7.14	n/a	n/a
Sustained phonation			
<i>Phonation</i>			
Jitter (%)	0.91±0.68	0.33±0.21	$P < .001$
Shimmer (%)	8.57±4.60	3.25±1.57	$P < .001$
NHR (-)	0.22±0.25	0.04±0.03	$P < .001$
HNR (dB)	14.05±6.01	22.55±4.28	$P < .001$
DDK task			
<i>Respiration</i>			
SPLD (1/s)	5.68±2.99	3.85±3.01	$P < .05$
<i>Articulation</i>			
DDK rate (syll/s)	6.01±0.60	7.16±0.71	$P < .001$
RFPC (-)	0.43±0.14	0.58±0.10	$P < .001$
SDCV (-)	0.14±0.03	0.17±0.03	$P < .01$
Running speech			
<i>Prosody</i>			
F0 SD (semitones)	1.52±0.43	2.62±0.75	$P < .001$
Intensity SD (dB)	7.15±1.42	8.66±1.49	$P < .001$
No. pauses (pause/s)	3.24±0.85	3.83±0.70	$P < .01$

impairment. Similarly, reduced melody in running speech captured by F0 SD measurement was found in other studies in PD patients, both treated and untreated with dopaminergic drugs [7, 15, 16]. On the other hand, several previous studies suggested that the most salient features of PD speech are related to phonatory and articulatory impairment [3, 4]. Indeed, our findings of increased values in jitter, shimmer, and NHR/HNR that may be clinically interpreted as hypophonia, voice hoarseness, and tremolo are in agreement with a previous report on untreated patients with PD [15]. However, in PD patients treated by dopaminergic drugs, only the jitter values were increased compared to controls while shimmer values were similar to those of controls, and the NHR/HNR findings were controversial [16].

VI. CONCLUSION

In conclusion, our newly designed configuration of vocal tests appears suitable for identification of voice and speech disorders in early stages of PD where it can accurately differentiate PD patients from HC. It consists of vocal tasks commonly used in most of the research studies examining PD-related voice and speech disorders [7]. Furthermore, the measurement methods can be

Figure 1: Schematic diagram depicting the recording of the PD patient's speech signals through the vocal test. Signals are calculated using speech signal processing algorithms and evaluated using the SVM-based model.



performed automatically with sufficient accuracy and without assistance of speech pathologist. In the future, when precise automatic assessment of boundaries between vowels and consonants etc. become feasible, the quick vocal test can be extended using new measurement methods such as for example novel acoustic measure of formant centralization ratio [17] which was proposed to robustly differentiate dysarthric from healthy speech.

ACKNOWLEDGMENT

This research was partly supported by the Czech Science Foundation, project GACR 102/08/H008, Czech Ministry of Health, projects NT 11331-6/2010 and NT 12288-5/2011, Grant Agency of the Czech Technical University in Prague, project SGS 10/180/OHK3/2T/13, and Czech Ministry of Education, projects MSM 0021620849 and MSM 6840770012.

REFERENCES

[1] H. Braak K. Del Tredici, U. Rüb, R. A. de Vos, E. N. Jansen Steur, and E. Braak, "Staging of brain pathology related to sporadic Parkinson's disease," *Neurobiol. Aging.*, vol. 24, pp. 197-211, 2003.

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

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

[4] C. Ludlow, N. Connor, and C. Bassich, "Speech timing in Parkinson's and Huntington's Disease," *Brain. Lang.*, vol. 32, pp. 195-214, 1987.

[5] S. Skodda, and U. Schlegel, "Speech rate and rhythm in Parkinson's Disease," *Mov. Disord.*, vol. 23, pp. 958-992, 2008.

[6] F. L. Darley, A. E. Aronson, and J. R. Brown, "Differential diagnostic patterns of dysarthria," *J. Speech Hear. Res.*, vol. 12, pp. 246-269, 1969.

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

[8] P. H. Dejonckere, P. Bradley, P. Clemente, G. Cornut, L. Crevier-Buchman, G. Friedrich, P. Van De Heyning, M. Remacle, and V. Woisard, "A basic protocol for functional assessment of voice pathology, especially for investigating the efficacy of (phonosurgical) treatments and evaluating new assessment techniques. Guideline elaborated by the Committee on Phoniatrics of the European Laryngological Society (ELS)," *Eur. Arch. Otorhinolaryngol.*, vol. 258, pp. 77-82, 2001.

[9] S. Fletcher, "Time-by-count measurement of diadochokinetic syllable rate," *J. Speech Hear. Disord.*, vol. 15, pp. 757-762, 1972.

[10] J. Ruzs, R. Cmejla, H. Ruzickova, and E. Ruzicka, "Quantitative acoustic measurements for characterization of voice and speech disorders in early untreated Parkinson's disease," *J. Acoust. Soc. Am.*, vol. 129, pp. 350-367, 2011.

[11] J. Ruzs, R. Cmejla, H. Ruzickova, J. Klempir, V. Majerova, J. Picmausova, J. Roth, and E. Ruzicka, "Acoustic assessment of voice and speech disorders in Parkinson's disease through quick vocal test," *Mov. Disord.*, 2011, in press.

[12] P. Boersma, and D. Weenink, "PRAAT, a system for doing phonetics by computer," *Glott International*, vol. 5, pp. 341-345, 2001.

[13] A. Wald, "Sequential analysis," New York: Wiley, 1947.

[14] T. Hastie, R. Tibshirani, and J. H. Friedman, "The elements of statistical learning : data mining, inference, and prediction: with 200 full-colour illustrations," New York: Springer, 2001.

[15] F. J. Jimenez-Jimenez, J. Gamboa, A. Nieto, J. Guerrero, M. Orti-Pareja, J. A. Molina, E. Garcia-Albea, and I. Cobeta, "Acoustic voice analysis in untreated patients with Parkinson's disease," *Parkinsonism. Relat. D.*, vol. 3, pp. 111-116, 1997.

[16] J. Gamboa, F. J. Jimenez-Jimenez, A. Nieto, J. Montojo, M. Orti-Pareja, J. A. Molina, E. Garcia-Albea, and I. Cobeta, "Acoustic voice analysis in patients with Parkinson's disease treated with dopaminergic drugs," *J. Voice*, vol. 11, pp. 314-320, 1997.

[17] S. Sapir, L. O. Ramig, J. L. Spielman, and C. Fox, "Formant centralization ratio: a proposal for a new acoustic measure of dysarthric speech," *J. Speech Lang. Hear. Res.*, vol. 53, pp. 114-125, 2010.