VYSOKÉ UČENÍ TECHNICKÉ V BRNĚ

FAKULTA ELEKTROTECHNIKY A KOMUNIKAČNÍCH TECHNOLOGIÍ

ÚSTAV TELEKOMUNIKACÍ

Ing. Zoltán Galáž

ANALYSIS OF MOTOR AND NON-MOTOR DEFICITS IN PATIENTS WITH PARKINSON'S DISEASE BASED ON THE ACOUSTIC ANALYSIS OF DYSARTHRIC SPEECH

ANALÝZA MOTORICKÝCH A NEMOTORICKÝCH DEFICITŮ U PACIENTŮ S PARKINSONOVOU NEMOCÍ NA ZÁKLADĚ AKUSTICKÉ ANALÝZY DYSARTRICKÉ ŘEČI

SHORTENED VERSION OF PH.D. THESIS

Specialization:	Teleinformatics
Supervisor:	Ing. Jiří Mekyska, Ph.D.
Oponents:	
Date of Defense:	

KEYWORDS

Parkinson's disease, hypokinetic dysarthria, acoustic analysis, diagnosis, freezing of gait, machine learning, motor symptoms, non-motor symptoms, objective assessment, quantitative analysis, statistical processing.

KLÍČOVÁ SLOVA

Parkinsonova nemoc, hypokinetická dysartrie, akustická analýza, diagnóza, zárazy v chůzi, strojové učení, motorické příznaky, nemotorické příznaky, objektivní hodnocení, kvantitativní analýza, statistické zpracování.

ARCHIVED IN

Dissertation is available at the Science Department of Dean's Office FEEC, Brno University of Technology, Technická 10, Brno, 616 00.

MÍSTO ULOŽENÍ PRÁCE

Disertační práce je k dispozici na Vědeckém oddělení děkanátu FEKT VUT v Brně, Technická 10, Brno, 616 00.

ISBN 80-214-
ISSN 1213-4198

CONTENTS

INTRODUCTION

Nowadays, we observe two main phenomena in the genesis of Parkinson's disease (PD). Namely the progressive degeneration of dopaminergic neurons in the *substancia nigra pars compacta* of the midbrain, and/or development of α-synuclein-containing Lewy bodies within the surviving neurons. It is known that the associated motor symptoms such as tremor at rest, progressive bradykinesia, muscular rigidity, postural instability, gait freezing, voice/speech disorders, etc., and non-motor symptoms such as behavioural alternations, reduction of cognitive abilities, sleep disturbances, anxiety, depression, etc. have a detrimental impact on patients' health, physical and mental condition, social life, independence, and quality of life in general. Typically, PD is rare in young population and its prevalence rate grows with the advancement of a person's age. That's why it is mostly diagnosed in persons aged over 60 years. But, before the conclusive clinical diagnosis is finally made, there is a long period of the development of the underlying neurodegenerative process behind the disease, slowly but surely worsening the severity of its symptoms.

At some point, the cardinal motor symptoms are the ones that first bring patients to a hospital searching for help, and even though the disease gets finally diagnosed, at this stage, most of the dopaminergic neurons have already been damaged, unfortunately. As one can imagine, the conventional clinical diagnosis of PD is therefore based on the presence of the above-mentioned cardinal motor symptom. Nevertheless, the presence of these symptoms is still not enough, and other criteria such as the short-term positive response to dopaminergic (anti-parkinsonian) medication, and many others have to be met. It is therefore obvious that the diagnosis of PD is not an easy task. In fact, even today, an objective diagnostic test which allows a definitive 100 % accurate diagnosis of this disease has not been developed. Thus, clinicians are forced to use a battery of tests, heuristics, biomarkers, and inclusion/exclusion criteria to make the diagnosis as accurate as possible. Another drawback of the current state of affairs is that this set of examinations has to be taken in the medical environment under the supervision of skilled clinician/s, which is logistically demanding, costly and time-consuming. Not to mention the fact that the disease does not have to be diagnosed at the first trial. It is often the case that prior the diagnosis elderly people have to visit the hospital several times, which makes this whole process even more problematic.

Today, we are living in the era of modern technologies, smart devices, internet of things, etc. Even though older population might not be adopted to such a technological advancement, younger people essentially grow up surrounded by it. Nowadays, smart phones, smart watches and other devices can be easily used to record a large variety of biological signals such voice/speech, movements of hands, gait, heart rate,

and many more. With the previously mentioned facts in mind, it seems that one of the major obstacles of PD diagnosis is the lack of data available for the clinicians. Therefore, these modern devices could be potentially used to collect a large amount of data without necessity of the patient's presence at the clinic or any specialized supervision. Such data could be securely transmitted and stored on cloud, where only authorized persons could be allowed to access them. With this approach, clinicians would be provided with an additional information about the medical condition of their patients that could definitely help with their decision making that is related to diagnosis, assessment, treatment and/or monitoring of the disease. Imagine a system that would be able to access and process all clinical data (data acquired by a doctor as well as those acquired by a variety of specialized devices such as those discussed above) available for a patient. The large scale of data that would be available could provide such a system with the power to use advanced signal processing techniques to quantify and describe properties of the acquired biological signals that might even lay beyond human perception. Next, modern machine learning algorithms, statistical analyses and visualization methods could be applied to provide clinicians with powerful reports about the current state of biomarkers and their evolution in time, and so on and so forth. It is evident that not only doctors, but also patients themselves would benefit from such information. However, to reach that point, relationship between properties of these biological signals and other clinical symptoms of PD needs to be investigated and fully understood.

Speech is the most natural way of communication. In most cases, people use it without problems. However, when a disorder such as PD comes into play, speech disorder named hypokinetic dysarthria (HD) gets involved. The associated voice/speech deviations in the early stages of the disease are very hard to be clearly perceived. In addition to that, patients themselves are in most cases not aware of their handicap, and the perception of the changes in their voice and speech is different than the one reported by their family and relatives. But in general, and depending on the stage of the disease, at some point, speech communication difficulties will eventually come. In fact, HD is one of the most disabling symptoms of PD that occurs in most of the patients suffering from it, and therefore, even though HD has a detrimental impact on the patient's quality of life, it might be used as a rich source of information for its diagnosis, assessment and monitoring.

1 HYPOTHESES AND GOALS

1.1 Hypotheses

Taking the previously mentioned facts into account, it is hypothesized that quantitative acoustic analysis of voice/speech signals can be used to robustly and complexly describe and identify HD in PD, and to indirectly assess other non-speech symptoms of PD. Specifically, it is assumed that parametrization of voice/speech deficits in HD and application of statistical analysis and/or modern machine learning techniques is capable of estimating the values of clinical rating scales that are conventionally used to assess motor and non-motor symptoms of PD at the baseline, as well as in the horizon of two years.

1.2 Goals and objectives

The main goal of this doctoral thesis is to *investigate possibilities of using quantitative objective evaluation of HD, employing modern clinically interpretable speech parametrization, statistical analysis and machine learning techniques, in direction of PD identification and assessment.* More specifically, this thesis has five main objectives that can be briefly summarized as follows:

1. *Robust computerized quantification of HD manifestations in PD* – to use modern clinically interpretable speech parameterization techniques to quantify manifestations of HD in the area of phonation, articulation, prosody and speech fluency that are known to occur with idiopathic PD.
2. *Complex analysis and identification of dysprosody in HD* – to study dysprosody in HD and to investigate an influence of prosodic demands such as precise control of speech melody variability during recitation or modulation of stress in speech, on computerized identification of HD.
3. *Assessment of non-speech symptoms of PD at the baseline* – to analyse the possibilities of using acoustic analysis of HD to estimate the scores of a variety of clinical rating scales that are nowadays being commonly used to assess motor and non-motor symptoms of PD at the baseline.
4. *Assessment of gait freezing in PD in the horizon of two years* – to analyse the possibilities of using acoustic analysis of HD at the baseline for predicting the change in the severity of gait freezing in PD in the horizon of two years.
5. *Analyse pathological mechanism shared by HD and gait freezing in PD* – to investigate if there are any pathological mechanisms shared by voice/speech disorders in HD and freezing of gait in PD.

2 IDENTIFICATION OF DYSPROSODY

Many researchers have studied speech disorders associated with dysprosody in HD [2, 22, 29, 41, 42, 45, 46]. The literature have demonstrated the presence of reduced variability in pitch and intensity of speech. However, the results on speech rate/pausing disturbances remain to be inconsistent [1]. This study proposes a novel approach for accurate identification of dysprosody in HD. So far, there has been no work dealing with HD analysis and identification using a poem recitation and/or a comparison between neutral, stress-modified and rhymed speech. Furthermore, most works do not perform the gender-differentiation, which neglects the information about gender-specific patterns of dysprosody in HD.

2.1 Methodology

2.1.1 Description of the dataset

For the purpose of this study, 149 Czech native speakers were examined: 98 patients with idiopathic PD (59 males and 39 females), and 51 healthy speakers (25 males and 26 females). All PD patients were examined on their regular dopaminergic medication approximately 1 hour after the L-dopa [28] dose. For more information about demographical and clinical characteristics of the used cohort, especially for the group of male and female participants, see Table 2.1 and Figure 2.1.

With respect to the speech task used to quantify prosodic disorders present in HD, the following three speech tasks comprising two reading tasks (emotionally-neutral and stress-modified) and a poem recitation task were considered:

1. Reading a short paragraph with neutral emotion. In Czech (original) – *I na tom, že člověk si opatří psa, aby nebyl sám, je mnoho pravdy. Pes opravdu nechce být sám.*; In English – *Even the fact that a man gets a dog to not be alone is pretty much true. A dog really don't want to be alone.*
2. Stress-modified reading. In Czech (original) – *Teď musíš být chvíli trpělivý, než to dokončme. Už mě to nebaví, dej mi už konečně pokoj! Tak co, jak to dopadlo?*; In English – *Now, you have to be patient until we finish it. I'm tired of it already, leave me alone! So, how did it go?*
3. Poem recitation task. In Czech (original) – *Chcete vidět velký lov? Budu lovit v džungli slov. Osedlám si Pegasa, chytím báseň do lasa!*, ; In English – *Would you like to see a big hunt? I will be hunting in a jungle of words. I will saddle the Pegasus, I will catch a poem into a lasso.*

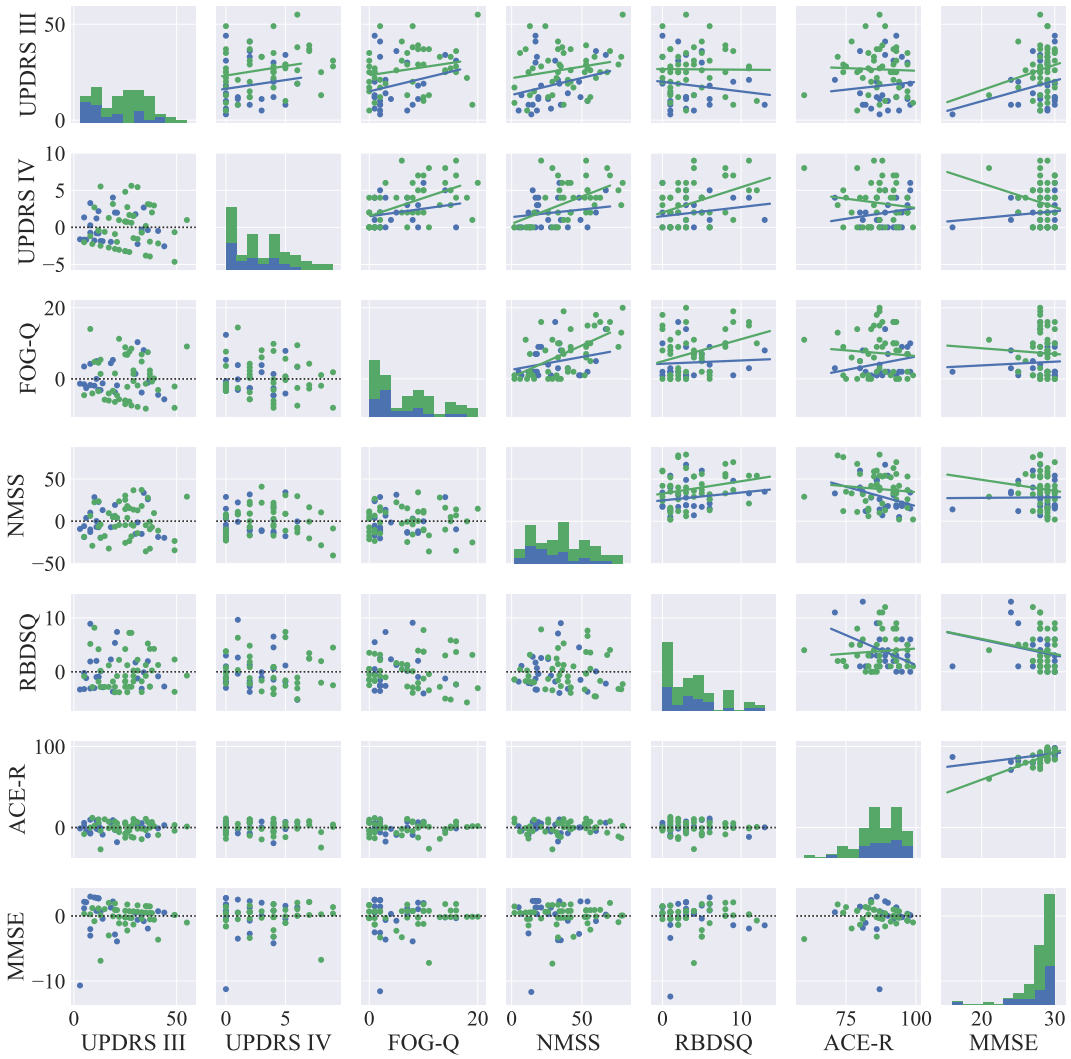

Fig. 2.1: Descriptive statistical graphs of clinical characteristics of PD patients. The structure of the graph: main diagonal – histograms are visualized; upper triangular part – scatter plots with the fitted lines of the robust linear regression models; and lower triangular part – residuals plots. Colour notation: blue colour – female speakers, and green colour – male speakers.

2.1.2 Feature extraction

To describe dysprosody in HD, clinically interpretable acoustic features [11] were used. To quantify variability of speech intonation, features derived from fundamental frequency (F0) [10] were computed: standard deviation of F0 (FOSD); relative standard deviation of F0 (relF0SD); variation range of F0 (FOVR); and relative variation range of F0 (relF0VR). To quantify variability of speech intensity: features derived from squared energy operator (SEO) and Teager-Kaiser energy operator (TEO) were computed: standard deviation of SEO/TEO (SEOSD/TEOSD); relative stan-

Tab. 2.1: Demographic and clinical characteristics of the participants.

characteristics	PD (females)	PD (males)	HC (females)	HC (males)
Number of speakers	44	53	22	29
Age (years)	68.48 ± 7.64	66.21 ± 8.78	62.25 ± 9.83	65.40 ± 9.04
PD duration (years)	7.61 ± 4.85	7.83 ± 4.39	-	-
UPDRS III	22.06 ± 13.73	26.85 ± 10.22	-	-
UPDRS IV	2.72 ± 3.01	3.15 ± 2.59	-	-
RBDSQ	3.42 ± 3.48	3.85 ± 2.99	-	-
FOG	6.94 ± 5.72	6.67 ± 5.57	-	-
NMS	36.03 ± 26.72	38.19 ± 19.72	-	-
BDI	18.57 ± 23.94	9.69 ± 6.23	-	-
MMSE	27.38 ± 3.63	28.56 ± 1.05	-	-
LED (mg/day)	862.44 ± 508.30	1087.00 ± 557.47	-	-

Table notation: UPDRS III – Unified Parkinson's disease rating scale, part III: evaluation of motor function [17]; UPDRS IV – Unified Parkinson's disease rating scale, part IV: evaluation of complications of therapy (Hoehn and Yahr scale, staging of severity of Parkinson's disease) [17]; RBDSQ – The REM sleep behavior disorder screening questionnaire [47]; FOG-Q – Freezing of gait questionnaire [21]; NMSS – Non-motor symptoms scale [14]; BDI – Beck depression inventory [7, 8]; MMSE – Mini-mental state examination [18]; LED – L-dopa equivalent daily dose (mg/day) [28].

dard deviation of SEO/TEO (relSEOSD/relTEOSD); variation range of SEO/TEO (SEOVR/TEOVR); relative variation range of SEO/TEO (relSEOVR/relTEOVR). Finally, to quantify speech rate/pausing: 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 ms)), articulation rate (AR), and speech index of rhythmicity (SPIR) were computed.

2.1.3 Analytical setup

To measure the strength of a monotonic relationship between feature vectors and the associated response variable [43], Spearman's correlation coefficient (ρ) was used. Next, mutual information was computed [50]. Mutual information is a measure of the amount of the information shared by two random variables. It is defined as:

$$I(X;Y) = \int_X \int_Y f(x,y) \log_2 \left(\frac{f(x,y)}{f_X(x) f_Y(y)} \right), \tag{2.1}$$

where X and Y are both random variables with the associated joint probability density function $f(x,y)$, marginal density functions $f_X(x)$ and $f_Y(y)$ respectively. For the purpose of this study, marginal entropies $H(X)$ and $H(Y)$, and joint entropy $H(X,Y)$ were used to compute MI. With this approach, MI is defined as:

$$I(X;Y) = H(X) + H(Y) - H(X,Y). \tag{2.2}$$

Moreover, Mann-Whitney U test was used to compare the distribution of the prosodic features between HC and patients with PD. The Mann-Whitney U test is a non-parametric statistical test that is used to assess whether two independent groups of variables are significantly different from each other [9]. It is defined as:

$$U = R_1 - \frac{n_1(n_1+1)}{2}, \tag{2.3}$$

where n_1 is the sample size for sample 1, and R_1 is the sum of the ranks in sample 1. Note that it is not specified which sample is considered sample 1, and therefore and equally valid statement can be made using sample 2 (n_2 instead of n_1 and R_2 instead of R_1, respectively).

To evaluate an individual power of each of the acoustic features to discriminate healthy and dysarthric speech, every feature was used separately as an input to the random forest (RF) classifier (univariate models). Next, to build models capable of HD discrimination based on the combination of the acoustic features, multivariate models were built as well. However, to select only the relevant set of features [24], a sequential floating forward selection (SFFS) [39] algorithm was applied. To evaluate the performance of the models, Matthew's correlation coefficient [25, 31] (MCC), accuracy (ACC), sensitivity (SEN), and specificity (SPE) were computed. MCC was also used as a measure for assessing the classification performance of the models during a feature selection process.

To evaluate the statistical power of the predictions made by the classifier [15], permutation test was used [38]. In this study, the significance level (α) of 0.01 was selected. Tested classification models with p-values bellow α were consider sufficiently high above chance level. Matthew's correlation coefficient was chosen as a test statistic for the permutation test as it is the measure used to assess the classification performance of the models during a feature selection step. The number of permutations was selected to be equal to 1000 and the classifier validation was conducted using stratified 10-fold cross-validation with 20 repetitions [15, 35].

2.2 Experimental evaluation

Results of the univariate analysis are summarized in Table 2.2. As can be seen, the best classification performance in terms of the classification accuracy computed for the univariate models can be summarized as follows: a) poem recitation task – ACC = 64.2 % (female participants), ACC = 64.6 % (male participants), and ACC = 68.5 % (all participants); b) reading with neutral emotion – ACC = 62.7 % (female participants), ACC = 69.1 % (male participants), and ACC = 58.4 % (all

Tab. 2.2: Statistical analysis of the prosodic features.

gender	features	disorder	ρ	MI	p	ACC	SEN	SPE
			Poem recitation task					
	relSEOSD	monoloudness	0.10	0.97		64.2	72.5	51.9
females	NST	speech rate	-0.23	0.88		62.7	67.5	55.6
	NSR	speech rate	0.23	0.88		61.2	65.0	55.6
	F0VR	monopitch	-0.12	0.90		64.6	64.3	65.4
males	TEOVR	monoloudness	-0.18	0.90		62.2	62.5	61.5
	rellF0SD	monopitch	-0.21	0.90		61.0	67.9	46.2
	TPT	speech rate	-0.17	0.94	*	68.5	70.8	64.2
all	NST	speech rate	-0.11	0.79		63.1	69.8	50.9
	NSR	speech rate	0.11	0.79		61.8	69.8	47.2
			Reading with neutral emotion					
	F0VR	monopitch	-0.07	0.97		62.7	67.5	55.6
females	TPT (50 ms)	speech rate	0.05	0.94		61.2	60.0	63.1
	AR	speech rate	-0.05	0.94		61.2	60.0	63.0
	relSEOVR	monoloudness	-0.06	0.90		64.6	71.4	50.0
males	relTEOSD	monoloudness	0.27	0.90	*	62.2	64.3	57.7
	relTEOVR	monoloudness	0.28	0.90	**	61.1	66.1	50.0
	relSEOSD	monoloudness	0.04	0.94		58.4	60.4	54.7
all	relSEOVR	monoloudness	-0.01	0.94		57.7	60.4	52.8
	TPT (50 ms)	speech rate	0.03	0.82		54.4	58.3	47.2
			Stress-modified reading task					
	relTEOVR	monoloudness	-0.36	0.97	**	68.7	70.0	66.7
females	F0SD	monopitch	-0.22	0.97		64.2	65.0	63.0
	relTEOSD	monoloudness	-0.38	0.97	**	59.7	55.0	66.7
	TPT (50 ms)	speech rate	-0.29	0.83	*	67.1	71.4	57.7
males	AR	speech rate	0.29	0.83	*	67.1	71.4	57.7
	relTEOVR	monoloudness	0.03	0.90		59.8	67.9	42.3
	TPT	speech rate	-0.26	0.93	**	59.7	65.6	49.1
all	F0SD	monopitch	-0.26	0.94	**	57.8	61.5	51.0
	TEOVR	monoloudness	-0.15	0.94		57.7	62.5	49.1

Table notation: ρ – Spearman's rank correlation coefficient; MI – mutual information; p – p-values of Mann-Whitney U test (* means $p < 0.05$; ** means $p < 0.01$); ACC – classification accuracy; SEN – classification sensitivity; SPE – classification specificity. ACC, SEN, SPE: expressed in %.

participants); and c) stress-modified reading – ACC = 68.7 % (female participants), ACC = 67.1 % (male participants), and ACC = 59.7 % (all participants).

Regarding the Mann-Whitney U test, there are few statistically significant differences, specifically: a) poem recitation task – $p < 0.01$ for TPT (all participants); b)

reading with neutral emotion – $p < 0.05$ for relTEOSD (males), and $p < 0.01$ for relTEOVR (males); and c) stress-modified reading – $p < 0.05$ for TPT (50 ms) (males), AR (males), and $p < 0.01$ for relTEOVR (females), relTEOSD (females), TPT (all participants), and F0SD (all participants). Next, a comparison of the features for monopitch (F0SD), monoloudness (SEOSD), and speech rate abnormalities (NSR) between PD patients and HC can be seen in Table 2.3 and Figures 2.2, 2.3, and 2.4.

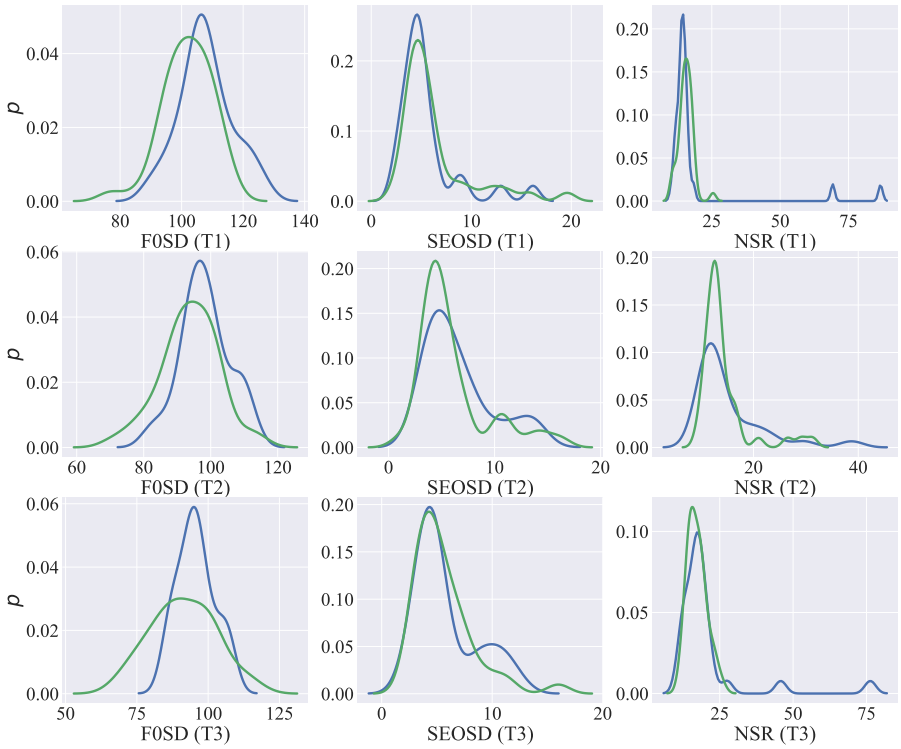

Fig. 2.2: Density estimation plots for female speakers only: T1 – poem recitation; T2 – emotionally-neutral reading; and T3 – stress-modified reading. Colour notation: blue colour (HC), and green colour (PD patients).

From the perspective of the monopitch, reduced variation in F0 can be observed in 8 out of the total number of 9 scenarios. The only exception occurs in the case of male speakers reading a passage with neutral emotion, which in general does not require that much variation in speech intonation or stress, so that this particular deviation is quiet acceptable. Regarding the monoloudness, interestingly there are 7 cases in which PD patients show more variation in speech intensity than HC, which is in contradiction with the original assumption of lowered variation in speech intensity in patients with PD in comparison with HC. The only two exceptions

Tab. 2.3: Comparison of acoustic features between PD speakers and HC.

gender	features	disorder	PD	HC	diff [%]
			Poem recitation		
females	F0SD	monopitch	101.87 ± 8.01	107.42 ± 8.91	PD ¡ HC (5.17)
	SEOSD	monoloudness	6.30 ± 3.67	5.61 ± 3.17	PD ¿ HC (12.30)
	NSR	speech rate	15.53 ± 2.66	19.24 ± 17.94	PD ¡ HC (19.28)
males	F0SD	monopitch	75.55 ± 14.06	80.90 ± 11.18	PD ¡ HC (6.61)
	SEOSD	monoloudness	6.44 ± 3.48	7.02 ± 3.67	PD ¡ HC (8.26)
	NSR	speech rate	18.09 ± 12.71	16.14 ± 2.95	PD ¿ HC (17.10)
all	F0SD	monopitch	86.49 ± 17.60	94.06 ± 16.64	PD ¡ HC (8.05)
	SEOSD	monoloudness	6.42 ± 3.54	6.28 ± 3.48	PD ¿ HC (2.23)
	NSR	speech rate	17.00 ± 9.82	17.65 ± 12.71	PD ¡ HC (3.68)
			Reading with neutral emotion		
females	F0SD	monopitch	93.69 ± 8.58	98.27 ± 7.58	PD ¡ HC (4.66)
	SEOSD	monoloudness	5.97 ± 3.37	6.03 ± 2.67	PD ¡ HC (1.00)
	NSR	speech rate	14.24 ± 4.78	13.91 ± 4.59	PD ¿ HC (2.37)
males	F0SD	monopitch	71.81 ± 12.36	70.51 ± 11.33	PD ¿ HC (1.84)
	SEOSD	monoloudness	6.67 ± 3.03	5.42 ± 2.15	PD ¿ HC (23.06)
	NSR	speech rate	13.49 ± 3.29	12.86 ± 1.86	PD ¿ HC (4.90)
all	F0SD	monopitch	80.86 ± 15.40	84.26 ± 16.90	PD ¡ HC (4.04)
	SEOSD	monoloudness	6.39 ± 3.16	5.68 ± 2.42	PD ¿ HC (12.50)
	NSR	speech rate	13.77 ± 3.97	13.42 ± 3.46	PD ¿ HC (2.61)
			Stress-modified reading		
females	F0SD	monopitch	91.27 ± 11.13	95.72 ± 6.71	PD ¡ HC (4.65)
	SEOSD	monoloudness	5.57 ± 2.71	5.55 ± 2.53	PD ¿ HC (0.36)
	NSR	speech rate	17.20 ± 3.22	19.44 ± 12.40	PD ¡ HC (11.52)
males	F0SD	monopitch	68.87 ± 10.11	74.86 ± 13.44	PD ¡ HC (8.00)
	SEOSD	monoloudness	6.53 ± 3.27	5.57 ± 2.62	PD ¿ HC (17.24)
	NSR	speech rate	19.35 ± 8.95	18.37 ± 4.45	PD ¿ HC (5.33)
all	F0SD	monopitch	78.05 ± 15.42	85.50 ± 14.44	PD ¡ HC (8.71)
	SEOSD	monoloudness	6.16 ± 3.07	5.48 ± 2.52	PD ¿ HC (12.41)
	NSR	speech rate	18.49 ± 7.15	18.82 ± 9.17	PD ¡ HC (1.75)

Table notation: diff [%] – difference between the mean values for patients with PD and HC. All prosodic features for PD patients and HC are represented as mean ± sd.

lies in the neutral reading task, and the poem recitation task performed by male speakers. However, in contrast to that, female patients did show significantly lower speech intensity when compared to HC while performing the poem recitation. Thus, the results suggest a presence of a gender-related pattern of parkinsonian speech intensity variation and control deterioration. Finally, in the case of speech rate

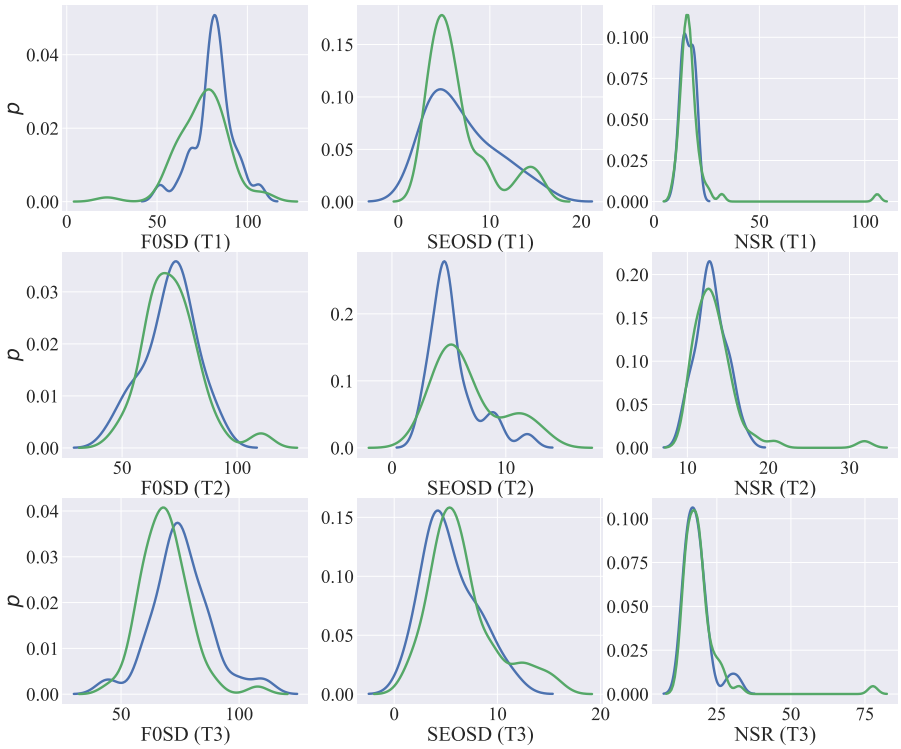

Fig. 2.3: Density estimation plots for male speakers only: T1 – poem recitation; T2 – emotionally-neutral reading; and T3 – stress-modified reading. Colour notation: blue colour (HC), and green colour (PD patients).

abnormalities, PD patients seem to have lower speech rate that HC when performing a task that requires stress (stress-modified reading) or changes in the melody of speech (poem recitation). In the case of male participants, PD patients seem to have higher speech rate when compared to HC. And finally, in the case of female participants, the same phenomenon can be observed for all participants.

As can be seen in Table 2.2, when the extracted prosodic features are taken individually, the resulting classification performance of the trained models does not reach satisfactory level of accuracy. However, this is somewhat expected since dysprosody in HD is rarely expressed as manifestation in a single prosodic domain. It is rather a combination of monopitch, monoloudness and abnormalities in speech rate and pausing. And moreover, HD is also known to be manifested slightly differently from patient to patient, which makes the prediction task even more difficult. Nevertheless, the univariate models can at least provide an indication about the con-

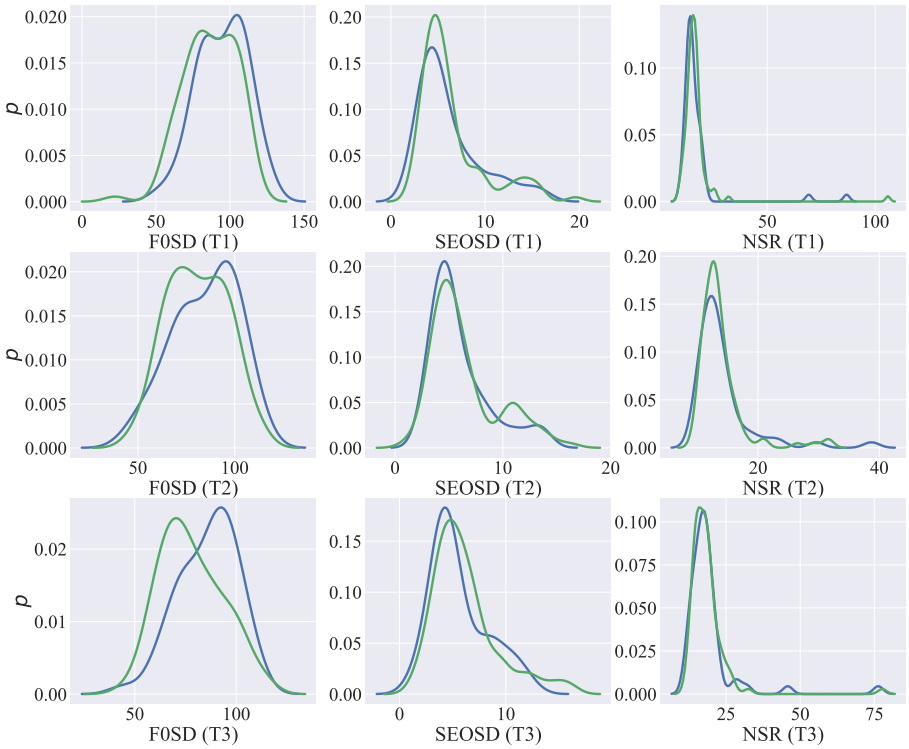

Fig. 2.4: Density estimation plots for all of the speakers: T1 – poem recitation; T2 – emotionally-neutral reading; and T3 – stress-modified reading. Colour notation: blue colour (HC), green colour (PD patients).

tribution of each of the selected acoustic features to discrimination of dysarthric and healthy speech. So, taking the previously mentioned facts into account, a feature selection procedure was applied as the next step towards obtaining a parsimonious, information-rich subsets of features, which provide maximum clinical information about the underlying prosodic pathology in patients with PD. Subsequently, the multivariate models were built using the selected features. The classification performance of these models can be seen in Table 2.4, and Table 2.5, respectively.

Consequently, t-distributed stochastic neighbourhood embedding (t-SNE) [30] algorithm was used to visualize the multi-dimensional space of prosodic features in the two-dimensional one. For this purpose, all the extracted acoustic features were used. The visualization was performed for all the three speech tasks separately to show the clusters of healthy and dysarthric speakers when speech prosody is quantified in a robust way (i. e. monopitch, monoloudness, and speech rate/pausing

abnormalities are described altogether). This method was also applied for female speakers, male speakers, and all speakers (both genders combined). The results of this method are presented in Figure 2.5. As can be seen, using the prosodic description it is not strong enough to conclusively and definitely identify HD in patients with PD. It is important to stress the fact that the results are strongly related to the dataset used in this study.

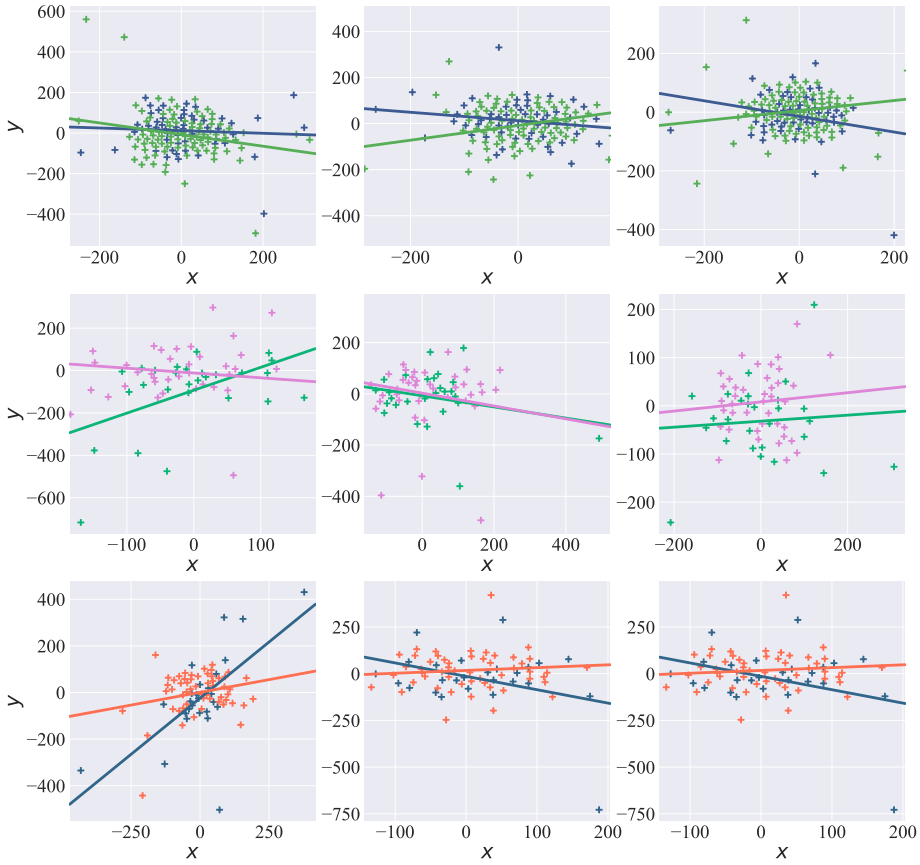

Fig. 2.5: Visualization of t-distributed stochastic neighbourhood embedding (lines fitted using robust linear regression). Graph grid notation: 1. row – all speakers, 2. row – female speakers, 3. row – male speakers; 1. column – poem recitation task, 2. column – emotionally-neutral reading, 3. column – stress-modified reading. Colour notation: all speakers – dark blue colour (HC), and dark green colour (PD patients); female speakers – medium green colour (HC), and purple colour (PD patients); and male speakers – medium blue colour (HC), and orange colour (PD patients).

Tab. 2.4: Classification results for groups of acoustic features.

feat.	gender	MCC	ACC	SEN	SPE	p	No.
			Poem recitation				
F1	females	0.14 ± 0.38	58.04 ± 17.71	63.00 ± 25.87	50.33 ± 32.73	0.0910	3
	males	0.19 ± 0.39	58.15 ± 19.01	57.06 ± 24.90	61.00 ± 33.94	0.0840	1
	all	0.17 ± 0.26	59.33 ± 12.43	61.35 ± 14.68	55.53 ± 21.15	0.1690	1
F2	females	0.24 ± 0.40	61.28 ± 18.47	59.50 ± 27.14	63.66 ± 28.70	0.2590	8
	males	0.26 ± 0.41	65.93 ± 17.78	72.06 ± 20.38	53.33 ± 34.17	0.0090	6
	all	0.23 ± 0.27	62.78 ± 12.58	66.24 ± 16.70	56.40 ± 25.64	0.0020	3
F3	females	0.19 ± 0.43	59.28 ± 20.25	60.00 ± 27.66	58.66 ± 30.53	0.0790	2
	males	0.29 ± 0.42	63.49 ± 20.05	61.73 ± 24.04	67.66 ± 32.54	0.0130	1
	all	0.27 ± 0.24	64.02 ± 11.35	65.80 ± 19.08	61.00 ± 24.55	0.0010	1
			Reading with neutral emotion				
F1	females	0.13 ± 0.42	54.85 ± 19.35	49.50 ± 28.34	63.33 ± 31.22	0.1450	3
	males	0.11 ± 0.33	59.39 ± 13.56	65.60 ± 17.45	45.33 ± 32.12	0.1340	2
	all	0.08 ± 0.24	54.67 ± 11.19	56.68 ± 17.85	50.86 ± 24.79	0.5910	3
F2	females	0.16 ± 0.44	58.19 ± 20.29	59.50 ± 28.07	56.00 ± 31.90	0.1590	3
	males	0.37 ± 0.42	70.53 ± 19.62	74.60 ± 20.34	62.33 ± 30.64	0.0080	4
	all	0.19 ± 0.28	60.90 ± 13.00	63.97 ± 15.55	55.33 ± 23.27	0.1420	3
F3	females	0.30 ± 0.37	64.04 ± 17.63	65.00 ± 25.25	62.66 ± 29.84	0.0510	2
	males	0.20 ± 0.31	60.31 ± 15.40	61.73 ± 20.89	58.00 ± 26.98	0.0710	1
	all	0.12 ± 0.30	56.17 ± 13.89	56.53 ± 15.02	55.80 ± 23.82	0.3260	4
			Stress-modified reading				
F1	females	0.31 ± 0.35	64.38 ± 15.63	67.50 ± 25.87	60.33 ± 30.84	0.0990	2
	males	0.21 ± 0.38	61.74 ± 17.71	63.80 ± 21.02	57.66 ± 31.26	0.1280	2
	all	0.15 ± 0.26	58.37 ± 13.13	61.60 ± 17.21	52.73 ± 19.60	0.1560	2
F2	females	0.40 ± 0.26	69.66 ± 12.01	72.00 ± 21.21	65.66 ± 26.60	0.0240	3
	males	0.24 ± 0.41	63.30 ± 18.24	64.73 ± 20.68	59.66 ± 33.51	0.0360	4
	all	0.20 ± 0.21	61.95 ± 9.15	66.35 ± 13.55	53.86 ± 21.57	0.1150	2
F3	females	0.15 ± 0.39	58.90 ± 15.39	63.50 ± 19.69	51.66 ± 35.99	0.4570	2
	males	0.24 ± 0.32	64.44 ± 14.19	68.73 ± 18.64	55.66 ± 30.41	0.0650	1
	all	0.13 ± 0.25	58.15 ± 12.36	62.02 ± 17.18	51.13 ± 21.88	0.6840	2

Table notation: F1 – monopitch features; F2 – monoloudness features; F3 – speech rate features; F4 – general prosodic features; MCC – Matthew's correlation coefficient (dimensionless) [31]; ACC – classification accuracy (expressed in %); SEN – classification sensitivity (expressed in %); SPE – classification specificity (expressed in %); No. – number of selected features; p – p-values of classification calculated by permutation test (1000 permutations).

To specify the results presented in these two tables: Table 2.4 shows the results of the multivariate classification analysis employed on the subsets of the prosodic features. Specifically, models for monopitch (F1), monoloudness (F2), and speech

Tab. 2.5: Classification results for all acoustic features.

feat.	gender	MCC	ACC	SEN	SPE	p	No.
T1	females	0.36 ± 0.42	66.57 ± 19.80	66.00 ± 25.13	68.33 ± 30.90	0.0070	5
	males	0.35 ± 0.34	67.84 ± 17.33	68.93 ± 22.54	66.00 ± 27.34	0.0020	3
	all	0.33 ± 0.16	67.30 ± 08.42	68.84 ± 14.18	64.66 ± 14.98	0.0040	1
T2	females	0.37 ± 0.40	68.47 ± 18.64	72.00 ± 26.06	63.33 ± 31.94	0.2110	3
	males	0.38 ± 0.29	69.52 ± 14.02	70.40 ± 18.93	68.00 ± 26.90	0.0050	8
	all	0.16 ± 0.32	59.62 ± 15.80	64.93 ± 18.77	50.13 ± 21.36	0.0350	4
T3	females	0.42 ± 0.35	70.71 ± 16.24	71.00 ± 19.13	70.33 ± 30.91	0.0001	1
	males	0.37 ± 0.34	70.03 ± 16.05	73.53 ± 19.28	63.00 ± 29.41	0.0120	5
	all	0.25 ± 0.26	63.20 ± 12.44	65.06 ± 14.92	60.00 ± 21.07	0.0130	3

Table notation: T1 – poem recitation task; T2 – reading with neutral emotion; T3 – stress-modified reading task; MCC – Matthew's correlation coefficient (dimensionless) [31]; ACC – classification accuracy (expressed in %); SEN – classification sensitivity (expressed in %); SPE – classification specificity (expressed in %); No. – number of selected features; p – p-values of classification calculated by permutation test (1000 permutations).

rate abnormalities (F3) were built. The assumption behind this approach was that despite insufficiency of the univariate models, investigation of the classification performance of each of the prosodic manifestations in HD can improve the performance of the models when more features are being used (it needs to be pointed out that these features do however describe the same phenomenon so that they are quiet correlated. But as mentioned previously, RF classifier is robust in dealing with high-dimensional and highly correlated data). Table 2.5 shows the results of the multivariate classification analysis employed on all of the prosodic features (F4).

The best classification performance in terms of classification accuracy achieved using the prosodic features for each speech task separately can be summarized as follows: a) poem recitation task – ACC = 67.84 % the model was trained using only 3 features based on the analysis of general prosodic impairment (TPT, TEOSD, SEOSD) computed for male participants; b) reading with neutral emotion – ACC = 69.52 %, the model was trained using 8 features based based on the analysis of general prosodic impairment (SEOSD, relF0SD, TST, TPT, relSEOSD, TPT (50 ms), relSEOVR, NSR) computed for male participants; and finally c) stress-modified reading – ACC = 70.71 %, the model was trained using just a single feature based on the analysis of monoloudness (relTEOVR) computed for female participants. It is worth noting that some of these models did not achieve sufficiently low p-values of the permutation test (strict significance level of 0.01 was chosen) that is needed to reject the null hypothesis. This may indicated that more data are required in order to get significant results [23].

3 ASSESSMENT OF PARKINSON'S DISEASE

So far, relatively small number of studies evaluating possibilities of computerized estimation of PD severity based on the quantitative acoustic analysis of dysarthric speech have been employed [4, 6, 16, 19, 32, 33, 37, 40, 48–50]. Nevertheless additional studies are needed to evaluate these results and to show whether speech prosody assessment might serve as a good biomarker for predicting a malignant course of the disease. Therefore, this study builds upon the previous finding and applies robust analysis of dysprosody in HD to indirectly estimate degree of PD severity assessed by a large number of well-known and widely-used clinical rating scales that are nowadays being commonly used to evaluate motor and non-motor symptoms in patients with PD.

3.1 Methodology

3.1.1 Description of the dataset

In the frame of this study, robust analysis and estimation of motor and non-motor symptoms of idiopathic PD using the acoustic analysis of dysarthric speech were employed. These symptoms were evaluated by skilled neurologists and clinical psychologists who examined and rated each PD patient participating in this study according to a variety of widely used and recognized clinical rating scales such as: UPDRS III (evaluation of motor functions) [17]; UPDRS IV (evaluation of complications of therapy; Hoehn and Yahr scale, staging of severity of PD) [17]; FOG-Q (evaluation of freezing and other gait-related deficits) [21]; NMSS (evaluation of non-motor deficits) [14]; RBDSQ (evaluation of sleep disorders, especially in the REM sleep) [47]; ACE-R (evaluation of cognitive dysfunctions) [27]; MMSE (evaluation of cognitive dysfunctions) [18]; and BDI (evaluation of depression) [7, 8]. These scales are nowadays commonly used in the clinical practice to assess and rate the severity of motor and non-motor manifestations associated with PD. Other clinical rating scales exist as well. However, in the frame of this study, this exact subset of the rating scales is considered exclusively.

To follow the results summarized in the previous chapter, the same speech tasks, recording setup, dataset, etc. (database) were used. For more information see, Chapter 2 (Section 2.1.1). However, in this particular study, only patients with PD are considered. Moreover, a subset of the patients was needed to be selected. The reason for that is the necessity of having the complex clinical information about each of the patients. More specifically, not every patient did undergo all examinations so that for some patients full clinical status is not available. Hence, to ensure that no

missing or corrupted data will be present in the dataset, a subset of 72 PD patients (47 males and a group of 25 females) were selected. For more information about demographical and clinical characteristics of the used cohort, see Table 3.1, and Figure 3.1.

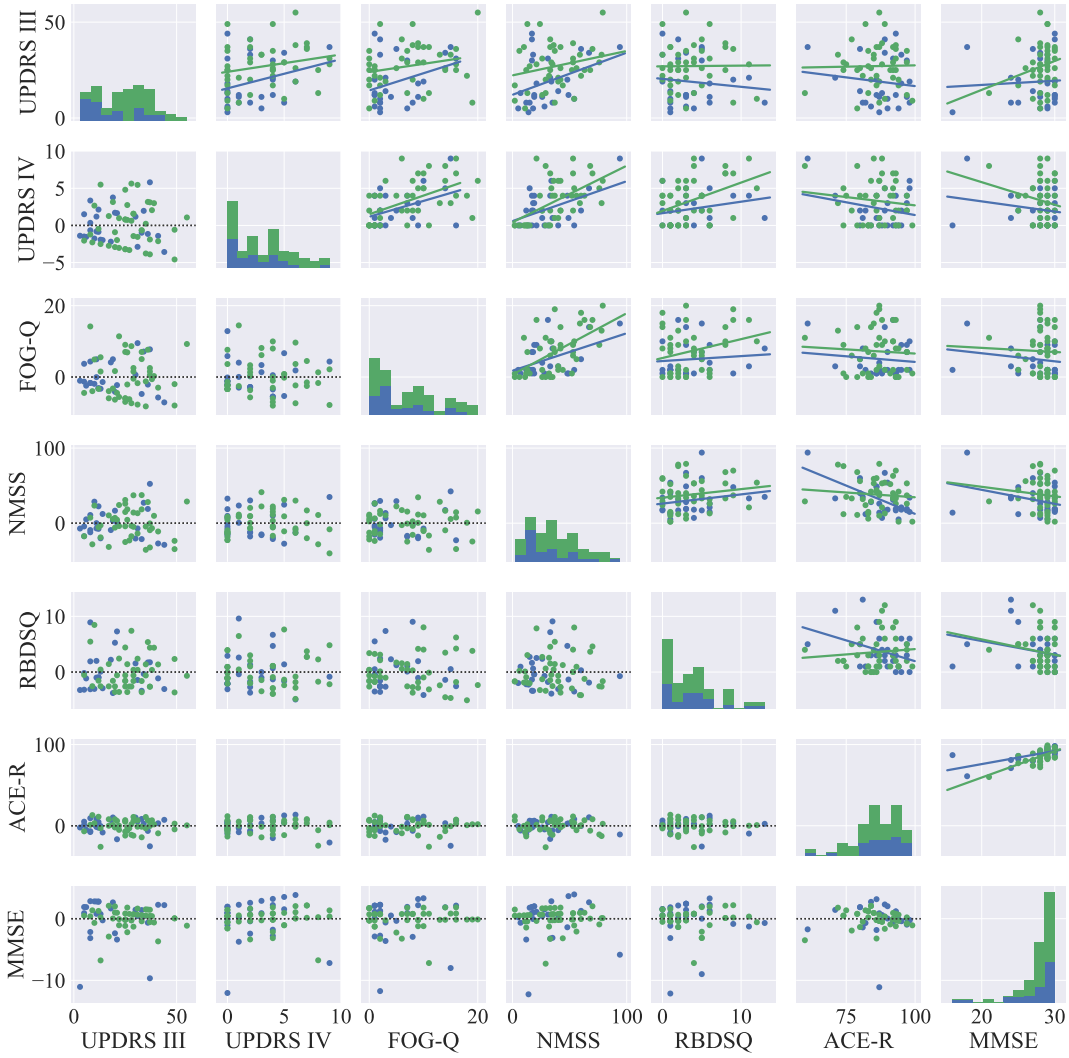

Fig. 3.1: Descriptive statistical graphs of clinical characteristics of PD patients. The structure of the graph: main diagonal – histograms are visualized; upper triangular part – scatter plots with the fitted lines of the robust linear regression models; and lower triangular part – residuals plots. Colour notation: blue colour – female speakers, and green colour – male speakers.

3.1.2 Feature extraction

As in the case of study presented in Chapter 2, to describe dysprosody in HD, clinically interpretable acoustic features [11] were used. To quantify variability of speech

Tab. 3.1: Clinical characteristics of the patients.

charact.	mean	std	min	Q1	Q2	Q3	max	r (d)	r (s)
LED (mg/day)	995.10	566.28	0.00	600.00	825.00	1325.50	2275.00	2275	∞
UPDRS III	24.06	12.22	3.00	12.75	25.00	33.25	55.00	52	108
UPDRS IV	2.94	2.68	0.00	0.00	2.00	5.00	9.00	9	23
FOG-Q	6.46	5.63	0.00	1.00	5.50	10.00	20.00	20	24
NMSS	35.23	20.75	2.00	17.75	33.00	52.25	94.00	94	360
RBDSQ	3.67	3.13	0.00	1.00	3.00	5.00	13.00	13	13
ACE-R	87.33	8.02	60.00	82.75	87.00	93.00	99.00	39	100
MMSE	27.88	2.54	16.00	28.00	28.50	29.00	30.00	14	30
BDI	10.46	6.14	0.00	6.00	9.00	13.50	27.00	27	63

[1] Table notation: charact. – characteristics (clinical); Qx – x-th quartile (Q1 [first], Q2 [second], Q3 [third]); r (d) – range (max − min) computed from the values actually present in the dataset; r (s) – range of the values in the scale; LED – L-dopa equivalent daily dose (mg/day) [28]; UPDRS III – Unified Parkinson's disease rating scale, part III: evaluation of motor function [17]; UPDRS IV – Unified Parkinson's disease rating scale, part IV: evaluation of complications of therapy (Hoehn and Yahr scale, staging of severity of Parkinson's disease) [17]; FOG-Q – Freezing of gait questionnaire [21]; NMSS – Non-motor symptoms scale [14]; RBDSQ – The REM sleep behavior disorder screening questionnaire [47]; ACE-R – Addenbrooke's cognitive examination-revised [27]; MMSE – Mini-mental state examination [18]; BDI – Beck depression inventory [7, 8].

intonation, features derived from fundamental frequency (F0) [10] were computed: standard deviation of F0 (FOSD); relative standard deviation of F0 (relF0SD); variation range of F0 (FOVR); and relative variation range of F0 (relF0VR). To quantify variability of speech intensity: features derived from squared energy operator (SEO) and Teager-Kaiser energy operator (TEO) were computed: standard deviation of SEO/TEO (SEOSD/TEOSD); relative standard deviation of SEO/TEO (relSEOSD/relTEOSD); variation range of SEO/TEO (SEOVR/TEOVR); and relative variation range of SEO/TEO (relSEOVR/relTEOVR). Finally, to quantify speech rate/pausing: 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 ms)), articulation rate (AR), and speech index of rhythmicity (SPIR) were computed.

3.1.3 Analytical setup

To the relationship between the values of the prosodic features and other motor and non-motor symptoms (assessed by the selected clinical rating scales) of PD, Spearman's correlation coefficient (ρ) was used (short description of this method can be found in the previous chapter). The significance level of correlation (p) of 0.05 was selected. Due to the limited number of samples and the exploratory character of the study, the correction for multiple comparisons was not performed.

Consequently, multivariate regression models were built (10-fold cross-validation with 20 repetitions). To reduce the number of features and create the regression models with low dimensionality, and better clinical interpretability [24], a modified version of sequential floating forward selection [39] algorithm was applied. For regression, classification and regression trees (CART) [12] were used.

To measure the prediction performance of the trained models, several conventional and widely-used regression metrics such as mean absolute error (MAE), root mean squared error (RMSE) were employed. Moreover, a novel regression metric named estimation error rate (EER) was computed to express the prediction error in percentage, which is particularly useful for easy and fast interpretation.

3.2 Experimental evaluation

Results of the correlation analysis are summarized in Table 3.2. The table shows top three acoustic features sorted according to their significance level of correlation expressed by Spearman's correlation coefficient. The following results showing the correlations for the specific prosodic areas under the focus were achieved (* means $p < 0.05$; ** means $p < 0.01$; T1 – poem recitation task; T2 – reading with neutral emotion; and T3 – stress-modified reading):

1. UPDRS III – reduced variation in pitch and intensity of speech (T1–T3). T1: -0.38^{**} (F0VR), -0.30^{**} (TEOSD); T2: -0.28^{*} (F0SD), 0.28^{*} (SEOSD); and T3: -0.37^{**} (F0VR), -0.30^{*} (TEOVR).
2. UPDRS IV – reduced variation in intonation (T1, T2), intensity of speech (T3), and speech rate abnormalities (T1, T2). T1: 0.29^{*} (relF0SD), 0.21^{*} (TPT (50 ms)); T2: 0.32^{**} (relF0SD), 0.31^{**} (TPT (50 ms)); and T3: -0.21^{*} (TEOSD), -0.21^{*} (F0VR).
3. FOG-Q – speech rate abnormalities (T1–T3), and reduced variation in intensity of speech (T1). T1: -0.35^{**} (NST), 0.28^{*} (relF0SD); T2: 0.42^{**} (TPT (50 ms)); and T3: -0.27^{*} (NST).
4. NMSS – reduced variation in intonation and intensity of speech (T1–T3). T1: -0.29^{*} (TEOSD), 0.26^{*} (relF0SD); T2: 0.36^{*} (relF0SD), 0.31^{*} (relTEOVR); and T3: -0.36^{*} (TEOVR), -0.29^{*} (F0VR).
5. RBDSQ – reduced variation in intensity of speech (T1, T2). T1: 0.20^{*} (TEOSD); and T2: -0.23^{*} (SEOSD).
6. ACE-R – speech rate abnormalities (T1), and reduced variation in intensity of speech (T3). T1: -0.43^{*} (TPT); T3: 0.23^{*} (TEOSD).
7. MMSE – speech rate abnormalities (T1): -0.26^{*} (TPT).
8. BDI – none of the features showed statistically significant correlation.

Tab. 3.2: Correlation analysis of the prosodic features.

	T1			T2			T3		
scale	features	ρ	p	features	ρ	p	features	ρ	p
	F0VR	-0.38	**	F0SD	-0.28	*	F0VR	-0.37	**
UPDRS III	TEOSD	-0.30	**	SEOSD	0.28	*	F0SD	-0.34	**
	TEOVR	-0.25	*	SEOVR	0.28	*	TEOVR	-0.30	*
	relF0SD	0.29	*	relF0SD	0.32	**	TEOSD	-0.21	*
UPDRS IV	TPT (50 ms)	0.21	*	TPT (50 ms)	0.31	**	F0VR	-0.21	*
	relF0VR	0.20		AR	-0.31	**	TEOVR	-0.18	
	NST	-0.35	**	TPT (50 ms)	0.42	**	NST	-0.27	*
FOG-Q	NSR	0.35	**	AR	-0.42	**	NSR	0.27	*
	relF0SD	0.28	*	NST	-0.41	**	TST	-0.26	*
	TEOSD	-0.29	*	relF0SD	0.36	**	TEOVR	-0.36	**
NMSS	relF0SD	0.26	*	relTEOVR	0.31	**	TEOSD	-0.35	**
	TPT	0.20		relF0VR	0.26	*	F0VR	-0.29	*
	TEOSD	0.20	*	SEOSD	-0.23	*	SPIR	-0.18	
RBDSQ	TEOVR	0.20		relSEOVR	0.10		TEOSD	0.17	
	SEOSD	-0.17		SEOVR	-0.09		relF0SD	0.17	
	TPT	-0.43	**	TEOSD	0.18		TEOSD	0.23	*
ACE-R	TST	-0.33	**	relSEOVR	0.16		TEOVR	0.20	
	TSR	0.33	**	relF0VR	-0.16		SPIR	0.20	
	TPT	-0.26	*	relTEOVR	-0.17		F0SD	-0.18	
MMSE	relSEOSD	-0.17		NST	0.16		SEOVR	-0.13	
	TST	-0.16		TPT	-0.16		TPT (50 ms)	-0.12	
	relTEOVR	0.18		SEOSD	0.21		relTEOSD	0.18	
BDI	SEOSD	0.16		relTEOVR	0.15		relSEOSD	-0.18	
	relTEOSD	0.15		relSEOVR	-0.14		relSEOVR	-0.16	

[1] Table notation: T1 – poem recitation task; T2 – emotionally-neutral reading task; T3 – stress-modified reading task; ρ – Spearman's correlation coefficient; p – significance level of correlation (* means $p < 0.05$; ** means $p < 0.01$); UPDRS III – Unified Parkinson's disease rating scale, part III: evaluation of motor function [17]; UPDRS IV – Unified Parkinson's disease rating scale, part IV: evaluation of complications of therapy (Hoehn and Yahr scale, staging of severity of Parkinson's disease) [17]; FOG-Q – Freezing of gait questionnaire [21]; NMSS – Non-motor symptoms scale [14]; RBDSQ – The REM sleep behavior disorder screening questionnaire [47]; ACE-R – Addenbrooke's cognitive examination-revised [27]; MMSE – Mini-mental state examination [18]; BDI – Beck depression inventory [7, 8].

Moreover, three specific clinical rating scales were selected: UPDRS III (evaluation of motor deficits), FOG-Q (evaluation of gait freezing), ACE-R (evaluation of cognitive deficits). For these three scales, regression plots can be seen in Figure 3.2. The figure provides a visual impression about the strength of a linear relationship between the most correlated acoustic features and the values of the selected clinical rating scales (a single feature is chosen for each scenario).

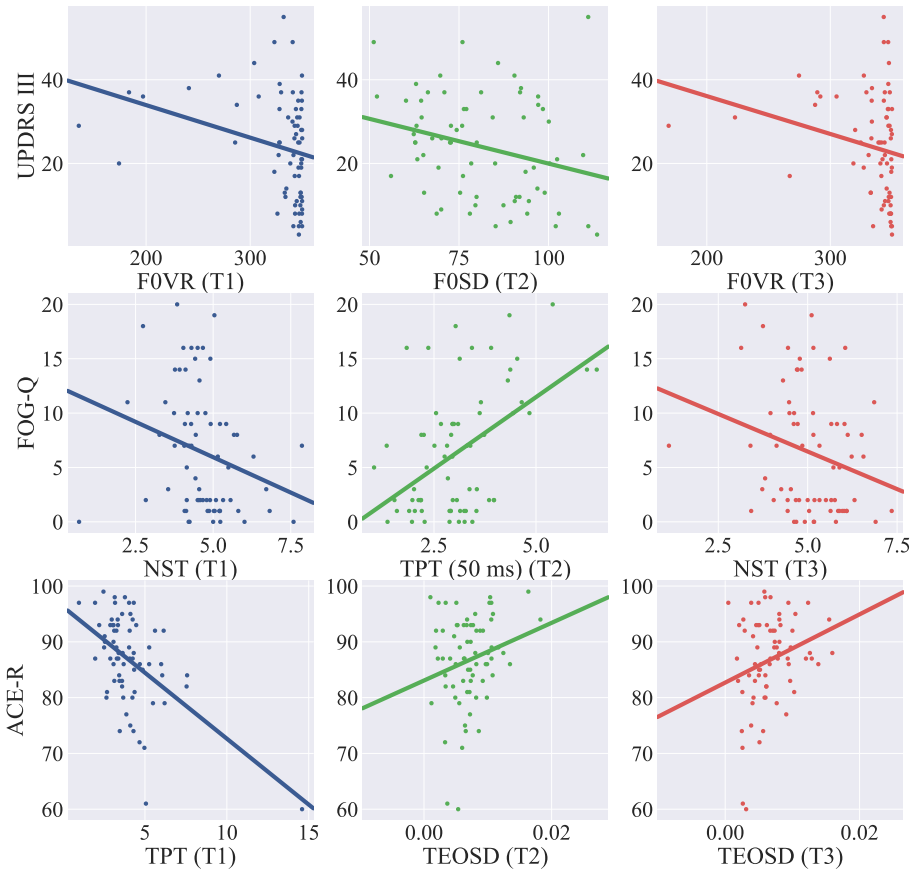

Fig. 3.2: Regression plots of the selected acoustic features/rating scales (UPDRS III, FOG-Q, ACE-R) for all three speech tasks performed by all PD patients: T1 – poem recitation task; T2 – emotionally-neutral reading; and T3 – stress-modified reading.

As can be seen, the strong relationship between reduced variation in pitch and UPDRS III (in the case of all the three speech tasks) is evident. Specifically, the flatter the intonation, the more severe motor disability assessed by UPDRS III can be observed. Next, the strong relationship between speech rate/pausing abnormalities and FOG-Q (in the case of all the three speech tasks) is present as well. Specifically, the faster the speech during poem recitation, larger number of pauses (longer that 50 ms) during emotionally-neutral reading, and faster the speech during stress-modified reading, the more severe gait freezing episodes assessed by UPDRS III can be observed. And finally, the strong relationship between speech rate/pausing abnormalities and ACE-R in the case of poem recitation can be seen. For the other two

tasks, the association is much weaker (less statistically significant as well). Specifically, the faster speech (less time spent on pausing) during poem recitation, and larger deviation of speech intensity during reading (emotionally-neutral, and stress-modified), the more severe cognitive deficits assessed by ACE-R can be observed. These observations emphasize the fact that poem recitation task is a great candidate to emphasize monopitch in HD, but also some cognitive deficits that are probably related to worse control of speech tempo (patients try to compensate it by occasional rushes of speech, etc.).

Next, for UPDRS III, FOG-Q, and ACE-R, multivariate regression models using the features selected by the feature selection algorithm, were built and visualized (visualization of the approximation of decision making performed by the regression tree) using the three graphs, see Figure 3.3, Figure 3.4, and Figure 3.5, respectively. Moreover, the results of the multivariate regression analysis are summarized in Table 3.3, and Table 3.4.

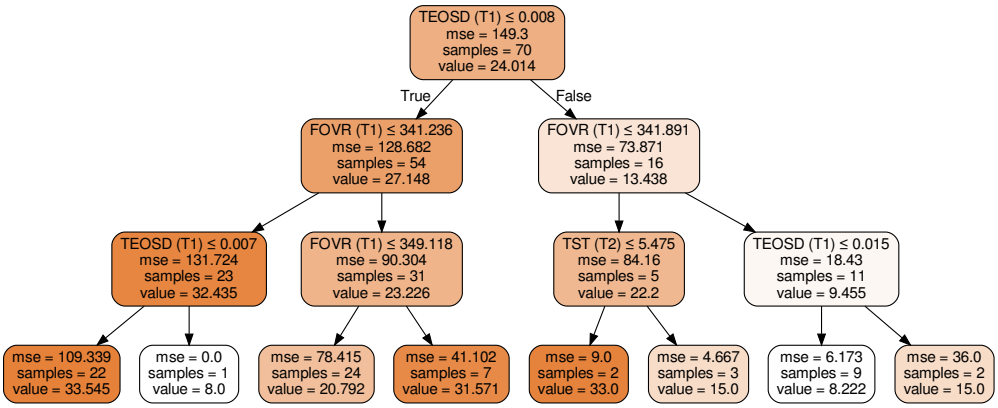

Fig. 3.3: Visualization of the regression tree built to estimate UPDRS III. The tree was trained using a single training run applied on all data in the dataset/selected features (hence the decision making of the tree is an approximation of the behavior responsible for the results summarized in Table 3.4). In the case of this tree: TST (T2), F0VR (T1), and TEOSD (T1) were used. For explanation of the speech task and acoustic feature notation, see Section 3.1.1, and Section 3.1.2, respectively.

With respect to the separate analysis (analysis of the speech tasks separately in direction of evaluating their sufficiency to assess severity of PD by estimating the clinical rating scales that are used to assess motor and non-motor deficits occurring with this disease), the following results were achieved: a) T1 – most of the selected acoustic features are based on the description of reduced variability in intonation and intensity of speech. The lowest estimation error rate was obtained in the case

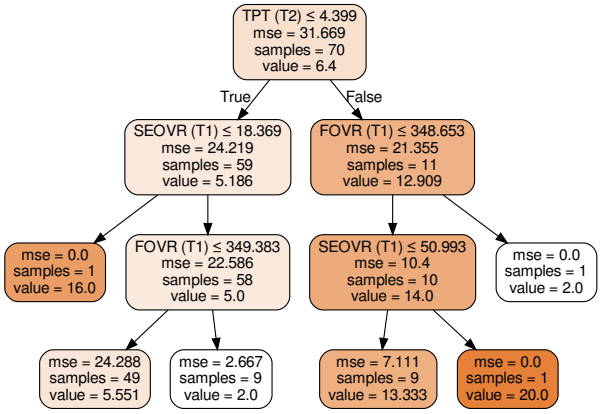

Fig. 3.4: Visualization of the regression tree built to estimate FOG-Q. The tree was trained using a single training run applied on all data in the dataset/selected features (hence the decision making of the tree is an approximation of the behavior responsible for the results summarized in Table 3.4). In the case of this tree: TPT (T2), F0VR (T1), and SEOVR (T1) were used. For explanation of the speech task and acoustic feature notation, see Section 3.1.1, and Section 3.1.2, respectively.

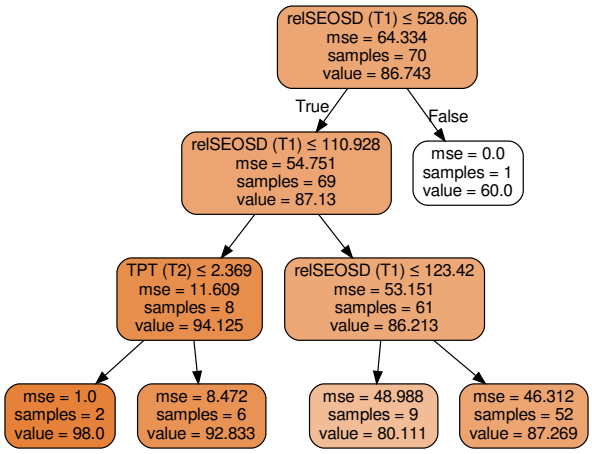

Fig. 3.5: Visualization of the regression tree built to estimate ACE-R. The tree was trained using a single training run applied on all data in the dataset/selected features (hence the decision making of the tree is an approximation of the behavior responsible for the results summarized in Table 3.4). In the case of this tree: TPT (T2), and relSEOSD (T1) were used. For explanation of the speech task and acoustic feature notation, see Section 3.1.1, and Section 3.1.2, respectively.

Tab. 3.3: Results of the regression analysis for individual speech tasks.

scale	MAE	RMSE	EER	No.	selected features
Poem recitation task					
UPDRS III	9.41 ± 2.73	11.45 ± 3.13	18.11 ± 5.26	2	F0VR, TEOSD
UPDRS IV	2.12 ± 0.65	2.69 ± 0.72	23.65 ± 7.26	1	relTEOVR
FOG-Q	4.52 ± 1.57	5.83 ± 1.92	22.61 ± 7.89	3	TEOVR, F0VR, TST
NMSS	18.55 ± 4.48	21.88 ± 4.83	20.16 ± 4.87	1	relTEOVR
RBDSQ	2.86 ± 0.80	3.50 ± 0.94	22.06 ± 6.17	1	relSEOVR
ACE-R	6.18 ± 1.84	7.67 ± 2.23	15.86 ± 4.72	1	relSEOSD
MMSE	1.83 ± 0.77	2.52 ± 1.26	13.13 ± 5.50	2	SEOVR, TPT
BDI	5.65 ± 1.52	6.68 ± 1.80	20.94 ± 5.63	2	TPT, relSEOVR
Reading with neutral emotion					
UPDRS III	10.44 ± 2.46	12.12 ± 2.63	20.08 ± 4.74	1	TPT
UPDRS IV	2.31 ± 0.50	2.66 ± 0.55	25.76 ± 5.63	1	TPT
FOG-Q	3.86 ± 1.22	4.80 ± 1.42	19.31 ± 6.14	3	relF0VR, TPT, NSR
NMSS	14.51 ± 4.29	17.78 ± 4.98	15.78 ± 4.66	3	TPT, relF0SD, TEOSD
RBDSQ	2.51 ± 0.68	3.04 ± 0.89	19.38 ± 5.24	1	TPT
ACE-R	6.80 ± 2.12	8.34 ± 2.66	17.44 ± 5.45	1	TPT
MMSE	1.61 ± 0.68	2.25 ± 1.24	11.50 ± 4.87	1	TPT
BDI	4.92 ± 1.39	6.00 ± 1.66	18.22 ± 5.18	1	TPT
Stress-modified reading					
UPDRS III	10.50 ± 3.23	13.10 ± 4.01	20.20 ± 6.22	2	relTEOSD, F0SD
UPDRS IV	2.45 ± 0.64	2.99 ± 0.67	27.27 ± 7.16	4	SEOVR, F0VR, AR, TPT
FOG-Q	4.90 ± 1.29	5.78 ± 1.44	24.50 ± 6.47	2	TSR, TST
NMSS	17.29 ± 4.88	20.91 ± 5.67	18.80 ± 5.31	1	TEOVR
RBDSQ	2.64 ± 0.71	3.23 ± 0.80	20.33 ± 5.49	2	TEOSD, F0VR
ACE-R	6.18 ± 1.84	7.67 ± 2.23	15.86 ± 4.72	3	TST, TSR, relSEOSD
MMSE	1.76 ± 0.67	2.36 ± 1.12	12.62 ± 4.83	1	TST
BDI	5.44 ± 1.63	6.72 ± 1.81	20.15 ± 6.04	3	TEOVR, F0VR, SEOSD

[1] Table notation: MAE – mean absolute error; RMSE – root mean squared error; EER – relative estimation error rate (MAE divided by the range of actual values of clinical rating scale present in the dataset; expressed in %); No. – number of selected features; UPDRS III – Unified Parkinson's disease rating scale, part III: evaluation of motor function [17]; UPDRS IV – Unified Parkinson's disease rating scale, part IV: evaluation of complications of therapy (Hoehn and Yahr scale, staging of severity of Parkinson's disease) [17]; FOG-Q – Freezing of gait questionnaire [21]; NMSS – Non-motor symptoms scale [14]; RBDSQ – The REM sleep behavior disorder screening questionnaire [47]; ACE-R – Addenbrooke's cognitive examination-revised [27]; MMSE – Mini-mental state examination [18]; BDI – Beck depression inventory [7, 8].

of MMSE (SEOVR, TPT): EER = 13.13 ± 5.50 %, closely followed by ACE-R (relSEOSD): EER = 15.86 ± 4.72, and UPDRS III (F0VR, TEOSD): EER = 18.11 ± 5.26 %; b) T2 – in 6 out of the total number of 8 analysed clinical rating scales, the feature selection found only a single feature based on the description of

Tab. 3.4: Results of the regression analysis for a combination of speech tasks.

scale	MAE	RMSE	EER	No.	selected features
UPDRS III	9.10 ± 2.93	11.27 ± 3.46	17.52 ± 5.64	3	TST[2], F0VR[1], TEOSD[1]
UPDRS IV	2.31 ± 0.50	2.65 ± 0.56	25.75 ± 5.64	1	TPT[2]
FOG-Q	3.45 ± 1.28	4.54 ± 1.61	17.28 ± 6.42	3	TPT[2], F0VR[1], SEOVR[1]
NMSS	17.03 ± 4.35	20.50 ± 4.86	18.52 ± 4.73	1	TPT[2]
RBDSQ	2.26 ± 0.83	2.88 ± 1.03	17.44 ± 6.40	3	F0SD[1], SEOSD[1], TPT[2]
ACE-R	6.20 ± 1.85	7.68 ± 2.22	15.72 ± 4.75	2	TPT[2], relSEOSD[1]
MMSE	1.60 ± 0.68	2.25 ± 1.24	11.49 ± 4.92	1	TPT[2]
BDI	4.91 ± 1.40	6.00 ± 1.66	18.21 ± 5.21	1	TPT[2]

[1] Table notation: [1] – poem recitation task; [2] – reading with neutral emotion; [3] – stress-modified reading task; MAE – mean absolute error; RMSE – root mean squared error; EER – relative estimation error rate (mean absolute error divided by the range of actual values of clinical rating scale present in the dataset; expressed in %); No. – number of selected features; UPDRS III – Unified Parkinson's disease rating scale, part III: evaluation of motor function [17]; UPDRS IV – Unified Parkinson's disease rating scale, part IV: evaluation of complications of therapy (Hoehn and Yahr scale, staging of severity of Parkinson's disease) [17]; FOG-Q – Freezing of gait questionnaire [21]; NMSS – Non-motor symptoms scale [14]; RBDSQ – The REM sleep behavior disorder screening questionnaire [47]; ACE-R – Addenbrooke's cognitive examination-revised [27]; MMSE – Mini-mental state examination [18]; BDI – Beck depression inventory [7, 8].

speech rate and pausing abnormalities (TPT) to be sufficient enough to describe the relationship between dysprosody in HD and severity of PD. The lowest estimation error rate was obtained in the case of MMSE (TPT): EER = 11.50 ± 4.87 %; and c) T3 – features based on the description of reduced variability of intensity of speech and speech rate abnormalities dominated most of the models. The lowest estimation error rate was obtained in the case of MMSE (TST): EER = 12.62 ± 4.83 %, closely followed by ACE-R (TST, TSR, relSEOSD): EER = 15.86 ± 4.72 %, and NMSS (TEOVR): EER = 18.80 ± 5.31 %.

Regarding the combined analysis (analysis of the combination of these speech tasks in direction of evaluating the power of the combined model to robustly and complexly assess severity of PD), combination of the speech tasks resulted into lower estimation error rates in most of the cases in which more than a single acoustic feature was selected. The prediction power of the regression models was slightly improved in the following clinical rating scales (improvements are expressed in %): UPDRS III = 0.59 %, FOG-Q = 2.04 %, RBDSQ = 1.94 %. However, as can be seen, in most of the cases, a single prosodic feature seems to be sufficiently describing a relationship between dysprosody in HD and other non-speech symptoms occurring with PD. Hypothetically, the prediction power of these models could be increased when taking other HD manifestations into account. Nevertheless, the results show that dysprosody is related with other motor (as assessed by UPDRS III, or FOG-Q) and non-motor (as assessed by MMSE or ACE-R) symptoms in PD.

4 ASSESSMENT OF FREEZING OF GAIT

Currently, there are only a few works addressing a relationship between FOG and speech disorders associated with PD [5, 13, 20, 34, 36]. To address this issue, this study is focused on investigation of pathological mechanisms shared by HD and FOG in patients with PD using a partial correlation analysis. Moreover, this study also provides an investigation of the possibilities of using quantitative acoustic analysis of dysarthric speech at the baseline for assessing the severity of FOG at the baseline (i. e. at time of the examination) as well as for assessing its progress in the horizon of two years (i. e. at the time of the follow-up examination).

4.1 Methodology

4.1.1 Description of the dataset

For the purpose of this study, 75 patients with idiopathic PD (48 males and 27 females) were enrolled at the First Department of Neurology, St. Anne's University Hospital in Brno, Czech Republic. All the patients were Czech native speakers. After two years, 41 of these patients (27 males and 14 females) were re-examined. For more information about demographical and clinical characteristics of the used cohort, see Table 4.1 and Figure 4.1. The patients were examined on their regular dopaminergic medication approximately 1 hour after the L-dopa [28] dose.

To quantify FOG, every patient was examined by a trained movement disorders specialist who rated the gait-related difficulties according to a specialized six-item Likert-scale (5-point scale where a score of 0 indicates absence of the symptom, while a score of 4 indicates the most severe stage; therefore the total score ranges from 0–24): Freezing of gait questionnaire [21]. The scale can be theoretically divided into two parts: 1st part (question 1–question 2) assesses walking and gait-related difficulties affecting patient's daily activities and independence; 2nd part (question 3–question 6) assesses gait freezing specifically. There is also a total score (T) computed as a sum of the two sub-scores (T_1 for Q1–Q2, and T_2 for Q3–Q6) summarizing the two parts ($T = T_1 + T_2$, where $T_1 = Q1 + Q2$, and $T_2 = Q3 + Q4 + Q5 + Q6$). This study was focused on gait freezing exclusively, therefore only the second part of the questionnaire and its total score are considered.

Furthermore, to provide more insight into the evolution of gait-related deficits (specifically Q6–Q6 score (sum of Q3–Q6) and the total score (sum of Q1–Q6)) between the two examinations (session 1, and session 2), box plots are presented as well. These graphs can be seen in Figure 4.2.

Fig. 4.1: Descriptive statistical graphs of clinical characteristics of PD patients (data for δ session (session 2 − session 1)). The structure of the graph: main diagonal – histograms are visualized; upper triangular part – scatter plots with the fitted lines of the robust linear regression models; and lower triangular part – residuals plots. Colour notation: blue colour – female speakers, and green colour – male speakers.

It is important to notice that, only participants with no missing data for the selected clinical rating scales were chosen. The same group of speakers were used later to built the regression models. With this approach, consistency of the dataset was ensured (even though the number of samples must have been reduced). After the filtration, 32 speakers (11 females, and 21 males) were left for the analysis.

Regarding the speech task used to quantify HD, a complex set of tasks was used to robustly quantify voice/speech disorders occurring with this disease. The speech acquisition protocol was actually derived from the standardized 3F Dysarthria Pro-

Tab. 4.1: Clinical characteristics of the patients (session 1, 2).

charact.	mean	std	min	Q1	Q2	Q3	max
Session 1 (48 males/27 females)							
PD duration (years)	7.48	4.15	4.00	1.00	11.00	7.00	21.00
UPDRS III	23.89	12.05	13.00	3.00	33.00	25.00	55.00
LED (mg/day)	997.26	554.05	610.00	0.00	1324.00	870.00	2275.00
NMSS	35.60	20.58	18.00	2.00	53.00	33.00	94.00
RBDSQ	3.76	3.22	1.00	0.00	5.00	3.00	13.00
MMSE	27.97	2.49	28.00	16.00	29.00	29.00	30.00
ACE-R	87.11	7.98	83.00	60.00	93.00	88.00	99.00
BDI	10.51	6.08	6.00	0.00	15.00	9.00	27.00
FOG-Q (Q3)	1.49	1.55	0.00	0.00	3.00	1.00	4.00
FOG-Q (Q4)	1.09	1.30	0.00	0.00	2.00	1.00	4.00
FOG-Q (Q5)	0.92	1.19	0.00	0.00	2.00	0.00	4.00
FOG-Q (Q6)	0.75	1.03	0.00	0.00	1.00	0.00	4.00
FOG-Q (total)	4.25	4.57	1.00	0.00	10.00	3.00	16.00
Session 2 (27 males/14 females)							
PD duration (years)	9.68	4.69	6.50	4.00	12.00	9.00	24.00
UPDRS III	28.15	12.93	20.00	5.00	36.00	29.00	61.00
LED (mg/day)	1128.67	469.20	767.50	375.00	1357.00	1070.00	2852.00
NMSS	55.54	33.72	29.00	2.00	70.50	57.00	138.00
RBDSQ	3.61	2.29	2.00	0.00	5.00	3.00	10.00
MMSE	28.02	2.08	27.00	22.00	30.00	29.00	30.00
ACE-R	84.88	9.68	79.50	51.00	92.50	87.00	97.00
BDI	10.76	5.12	6.50	2.00	15.00	10.00	25.00
FOG-Q (Q3)	1.71	1.50	0.00	0.00	3.00	2.00	4.00
FOG-Q (Q4)	1.22	1.31	0.00	0.00	2.00	1.00	4.00
FOG-Q (Q5)	1.24	1.20	0.00	0.00	2.00	1.00	4.00
FOG-Q (Q6)	1.05	1.16	0.00	0.00	2.00	1.00	4.00
FOG-Q (total)	5.22	4.76	2.00	0.00	13.50	6.00	16.00

[1] Table notation: charact. – characteristics (clinical); Qx – x-th quartile (Q1 [first], Q2 [second], Q3 [third]); UPDRS III – Unified Parkinson's disease rating scale, part III: evaluation of motor function [17]; LED – L-dopa equivalent daily dose [28]; NMSS – Non-motor symptoms scale [14]; RBDSQ – The REM sleep behavior disorder screening questionnaire [47]; MMSE – Mini-mental state examination [18]; ACE-R – Addenbrooke's cognitive examination-revised [27]; BDI – Beck depression inventory [7, 8]; FOG-Q – Freezing of gait questionnaire [21].

file [26] and included fourteen speech tasks, specifically: monologue, expiration with closed/open lips, sustained phonation (/a/, /i/), diadochokinesis, rhythmical units, basic intonation/stress templates, and reading with different/no emotions.

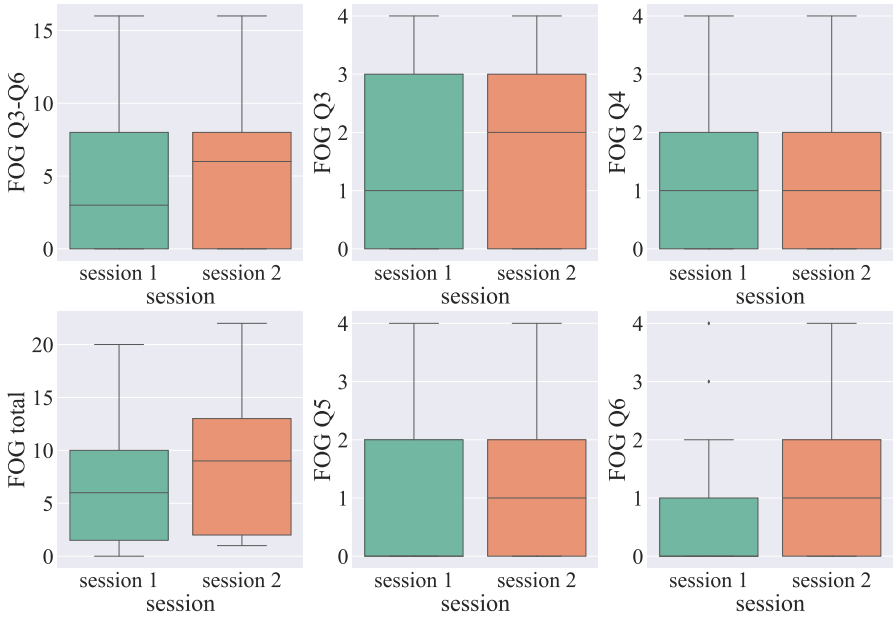

Fig. 4.2: Box plots visualizing the evolution of gait-specific deficits assessed by FOG-Q, specifically Q3–Q6 score and the total score (sum of Q1–Q6). Colour notation: green colour (session 1, e. e. baseline examination), and blue colour (session 2, i. e. follow-up examination).

4.1.2 Feature extraction

To quantify voice/speech disorders in the PD patients, a set of acoustic features based on a recommendation given in recent review on acoustic analysis if voice/speech signals in patients suffering from HD [11] was computed. It specifically covers the area of phonation, articulation, and prosody. To provide better insight into ability of these features to describe HD, a short description per HD area is presented.

In terms of phonation, the acoustic features describing airflow insufficiency (MPT) during expiration with closed (T2) or opened (T3) lips, irregular pitch fluctuations (relF0SD) during phonation of the vowel /a/ (T4), microperturbations in frequency (jitter) and amplitude (shimmer) during phonation of the vowel /a/ (T4), tremor of jaw (F1SD, F2SD) during phonation of the vowel /a/ (T4), increased acoustic noise (mean HNR) during phonation of the vowel /a/ (T4), and aperiodicity of voice (DUV) during phonation of the vowel /a/ (T4) were computed.

With respect to articulation, the acoustic features describing rigidity of tongue and jaw (F1IR, F2IR, F1SD, F2SD) during monologue (T1), rhythmical reading (T6), basic intonation templates (T7–9), paragraph reading (T10), and reading

with different emotions (T11–14), slow alternating motion rate (DDK rate) during diadochokinetic task (T5), and irregular alternating motion rate (DDK reg) during diadochokinetic task (T5) were computed.

Finally, regarding the acoustic features describing monopitch (relF0SD) and monoloudness (relSEOSD) during monologue (T1), rhythmical reading (T6), basic intonation templates (T7–9), paragraph reading (T10), and reading with different emotions (T11–14), inappropriate silences (SPIR) during paragraph reading (T10), unnatural speech rate (TSR, NSR) during basic intonation templates (T7–9), paragraph reading (T10), and reading with different emotions (T11–14) were computed.

4.1.3 Analytical setup

To assessed the strength of a relationship between the patients' clinical data and the selected items of FOG-Q in both sessions (session 1, session 2), Pearson's correlation with the significance level 0.05 was used. With this approach, it was possible to identify those clinical measures (PD duration, UPDRS III, LED (mg/day), NMSS, RBDSQ, MMSE, ACE-R, BDI) that are significantly correlated with the specific symptoms of gait freezing in PD, which is a very valuable information because it shows which clinical aspects of PD tend to be associated with FOG in the baseline and in the follow-up (after 2 years). Using the δ session (session $2 -$ session 1), it is even possible to see if the evolution of other clinical aspects of PD is related with the evolution of the associated gait problems.

Next, to assess the strength of a relationship between voice/speech disorders in HD and freezing of gait in patients with PD, Pearson's (linear relationship) and Spearman's (monotonic relationship) partial correlation coefficients between the acoustic features and the values of FOG-Q were computed. The significance level of correlation in this case was set to 0.05 as well. During the computation of partial correlations, the factors such as patients' age and gender [3, 44], dopaminergic medication [28] and a variety of associated motor and non-motor symptoms assessed by UPDRS III [17], BDI [7, 8], and ACE-R [27] were controlled for. As in the previous case, the aim was to identify those acoustic features that are significantly correlated with the specific symptoms of gait freezing in PD.

Finally, to evaluate the power of the acoustic features (in session 1; baseline) in predicting the change of the severity of gait freezing in PD (Δ FOG-Q), multivariate regression analysis was employed. For this purpose, we employed classification and regression trees (CART) in the supervised machine learning setup using stratified 10-fold cross-validation with 100 repetitions) [12]. As previously, see Chapters 2 and 3, feature selection process was applied to obtain the feature sets with the maximum clinical interpretability and also the power to predict FOG-related deficits

in patients with PD. For this purpose, a modified version of sequential floating forward selection [39] algorithm was used. To evaluate the prediction performance of the trained models, mean absolute error (MAE), root mean squared error (RMSE), and estimation error rate (EER) were computed. For more information about these metrics, see Chapter 3.

4.2 Results

Regarding the classical correlation analysis, the values of Pearson's correlation coefficients computed between clinical data (e. g. scores of the clinical rating scales assessing motor and non-motor symptoms of PD) and selected items of FOG-Q (i. e. Q3–Q6, and the total score) can be found in Table 4.2. This type of correlation was computed for all three sessions: session 1 (baseline examination), session 2 (follow-up examination), and δ session (description of the change in the particular item of the rating scale) The results are discussed bellow.

In both sessions, significant correlations of all FOG-Q items with duration of PD and UPDRS III (except Q5 in session 1) were identified. Next, LED was found significantly correlated with all FOG-Q items in session 1, but not in the session 2. Next, NMSS and all items of FOG-Q were found significantly correlated in session 1, however in the session 2 only few significant correlations were found. Regarding RBDSQ, no specific pattern can be observed. FOG-Q items correlated variably with RBDSQ, however, significant correlation for FOG-Q (total score) was found in both sessions. The scales assessing cognitive functions (MMSE, ACE-R) were not find significantly correlated with the items of FOG-Q. And finally, BDI score was not found significantly correlated with FOG-Q in session 1. Nevertheless, significant correlations can be observed in session 2. Regarding the correlations between Δ (Q3–6, total score) and Δ of the clinical scores, significant correlations between the changes in FOG-Q items and changes in LED, NMSS and MMSE were identified. Next, the results for partial correlation analysis are summarized in Table 4.3. The associated regression plots can be seen in Figure 4.3.

With respect to the partial correlation analysis, the correlations among acoustic features quantifying impaired phonation, articulation and prosody, and selected items of FOG-Q (Q3–Q6, and the total score) were computed. It is important to point out that the partial correlation analysis was performed for session 1 only to focus on the investigation of the relationship between FOG and HD in the baseline. For a better overview, only the acoustic features with significant correlation in both Pearson's, and Spearman's correlations were selected. Regarding Q3 (assessment of occurrence of freezing), this item was found correlated mostly with the

Tab. 4.2: Correlations among patients' FOG-Q items and clinical description.

FOG-Q	ρ (Q3)	p	ρ (Q4)	p	ρ (Q5)	p	ρ (Q6)	p	ρ (total)	p
Session 1										
PD dur. (years)	0.47	**	0.35	**	0.35	**	0.39	**	0.44	**
UPDRS III	0.24	*	0.24	*	0.24		0.23	*	0.25	*
LED (mg/day)	0.36	**	0.33	**	0.37	**	0.24	*	0.37	**
NMSS	0.45	**	0.43	**	0.39	**	0.52	**	0.49	**
RBDSQ	0.25	*	0.28	*	0.14		0.28	*	0.27	*
MMSE	-0.01		-0.07		0.06		-0.06		-0.02	
ACE-R	-0.06		-0.15		0.04		-0.13		-0.08	
BDI	0.05		0.06		0.09		0.13		0.09	
Session 2										
PD dur. (years)	0.41	**	0.41	**	0.38	**	0.42	**	0.44	**
UPDRS III	0.36	*	0.45	**	0.35	*	0.39	*	0.42	**
LED (mg/day)	0.28		0.03		0.17		0.15		0.18	
NMSS	0.39	*	0.30		0.26		0.36	*	0.36	*
RBDSQ	0.34	*	0.28		0.38	*	0.41	**	0.38	*
MMSE	-0.26		-0.14		-0.14		-0.07		-0.17	
ACE-R	-0.25		-0.18		-0.18		-0.13		-0.20	
BDI	0.36	*	0.36	*	0.38	*	0.38	*	0.40	**
Δ (Session 2 − Session 1)										
PD dur. (years)	-0.22		0.06		0.18		0.25		0.04	
UPDRS III	0.03		0.17		0.17		0.10		0.16	
LED (mg/day)	-0.28		-0.33		-0.35		-0.18		-0.40	*
NMSS	0.20		0.04		0.20		0.42	*	0.28	
RBDSQ	0.09		0.24		0.06		0.24		0.21	
MMSE	-0.35	*	-0.26		-0.29		-0.12		-0.36	*
ACE-R	-0.17		-0.25		-0.24		-0.06		-0.25	
BDI	-0.26		-0.10		-0.06		0.02		-0.16	

[1] Table notation: ρ – Spearman's correlation coefficient; p – significance level of correlation (* means $p < 0.05$; ** means $p < 0.01$); UPDRS III – Unified Parkinson's disease rating scale, part III: evaluation of motor function [17]; LED – L-dopa equivalent daily dose [28]; NMSS – Non-motor symptoms scale [14]; RBDSQ – The REM sleep behavior disorder screening questionnaire [47]; MMSE – Mini-mental state examination [18]; ACE-R – Addenbrooke's cognitive examination-revised [27]; BDI – Beck depression inventory [7, 8]; FOG-Q – Freezing of gait questionnaire [21] (Q1–Q6, and T – total score), for more details, see Section 4.1.

interpercentile range of the first formant, and with net speech rate (extracted from reading of a short sentence). The strongest correlation can be seen in the case of NSR extracted from the short imperative sentence reading ($\rho\,(\mathrm{P}) = -0.40$, $p < 0.01$, and $\rho\,(\mathrm{S}) = -0.44$, $p < 0.01$). In the case of Q4 (assessment of the duration of the longest freezing episode), 2 significant negative correlations were identified for the interpercentile range of the first formant (extracted from paragraph reading, and

Tab. 4.3: Partial correlations among features and FOG-Q (session 1) items.

HD area	specific disorder	features	ρ (P)	p (P)	ρ (S)	p (S)
		FOG (Q3)				
prosody	unnatural speech rate	NSR (T8)	0.41	**	0.34	*
articulation	rigidity of tongue and jaw	F1IR (T10)	-0.40	**	-0.44	**
articulation	rigidity of tongue and jaw	F1IR (T9)	-0.34	*	-0.38	**
articulation	rigidity of tongue and jaw	F1IR (T13)	-0.32	*	-0.39	**
articulation	rigidity of tongue and jaw	F1IR (T14)	-0.30	*	-0.30	*
articulation	rigidity of tongue and jaw	F1IR (T7)	-0.30	*	-0.35	*
		FOG (Q4)				
articulation	rigidity of tongue and jaw	F1IR (T10)	-0.40	**	-0.43	**
articulation	rigidity of tongue and jaw	F1IR (T9)	-0.35	*	-0.40	**
		FOG (Q5)				
articulation	rigidity of tongue and jaw	F1IR (T14)	-0.47	**	-0.49	**
prosody	unnatural speech rate	NSR (T8)	0.36	*	0.42	**
articulation	rigidity of tongue and jaw	F1IR (T10)	-0.36	*	-0.40	**
articulation	rigidity of tongue and jaw	F1IR (T13)	-0.29	*	-0.36	*
		FOG (Q6)				
prosody	unnatural speech rate	TSR (T11)	0.33	*	0.33	*
prosody	unnatural speech rate	NSR (T8)	0.33	*	0.33	*
prosody	unnatural speech rate	NSR (T11)	0.32	*	0.32	*
		FOG (total score)				
articulation	rigidity of tongue and jaw	F1IR (T10)	-0.40	**	-0.45	**
articulation	rigidity of tongue and jaw	F1IR (T14)	-0.38	**	-0.38	**
articulation	unnatural speech rate	NSR (T8)	0.36	*	0.38	**

[1] Table notation: ρ (P) – Pearson's correlation coefficient; p (S) – significance level of correlation according to ρ (P); ρ (S) – Spearman's correlation coefficient; p (S) – significance level of correlation according to ρ (S) (* means $p < 0.05$; ** means $p < 0.01$); FOG-Q – Freezing of gait questionnaire [21].

short declarative sentence reading). The strongest correlation can be seen in the case of paragraph reading (ρ (P) $= -0.40$, $p < 0.01$, and ρ (S) $= -0.43$, $p < 0.01$). With respect to Q5 (assessment of the duration of the typical start hesitation), the interpercentile range of the first formant (extracted from paragraph reading, reading of 9 words in a bored manner, reading of 5 words excitedly), and with the net speech rate (extracted from short imperative sentence reading) were found significantly correlated with this particular item of the questionnaire. The strongest correlation can be seen in the case of F1IR extracted from the reading of 5 words in an excited manner (ρ (P) $= -0.47$, $p < 0.01$, and ρ (S) $= -0.49$, $p < 0.01$). In the case of Q6

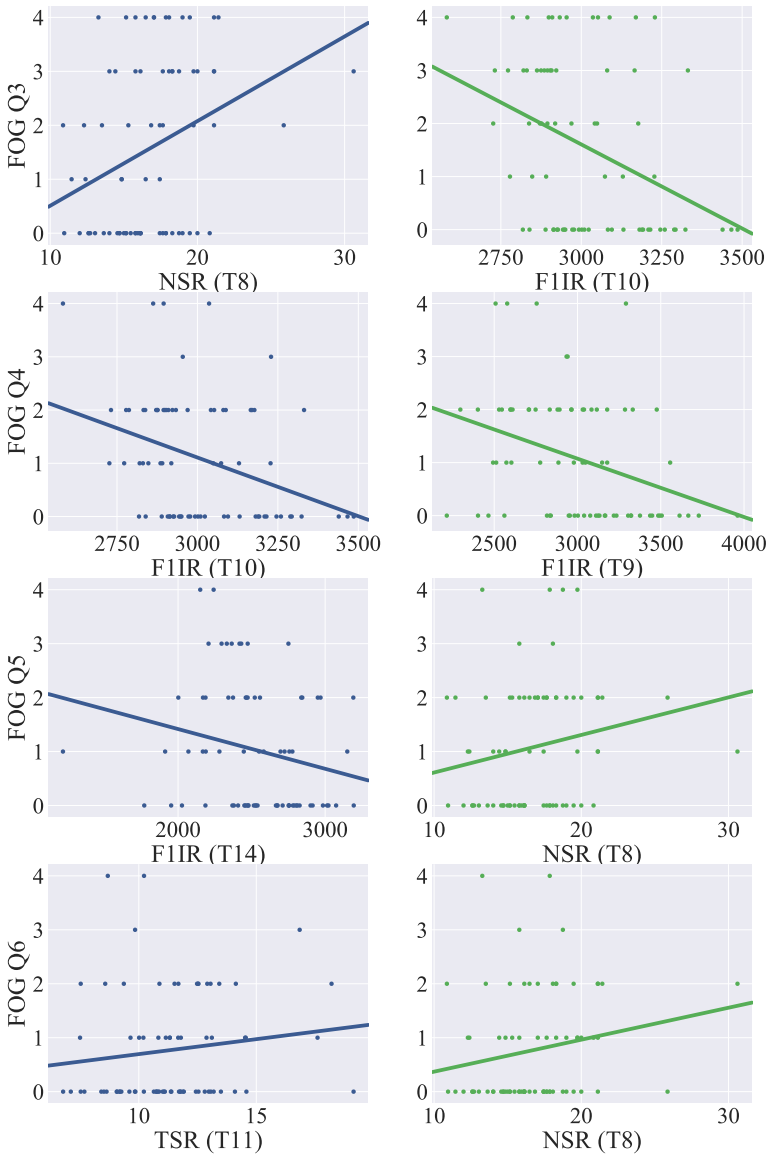

Fig. 4.3: Regression plots (scatter plots with the fitted line of the robust linear regression estimator) of the most correlated acoustic features (partial correlation) for Q3–Q6, see Table 4.3. Colour notation: blue colour (the most correlated feature); and green colour (the second most correlated feature).

(assessment of the duration of the typical turning hesitation), significant correlations were found for total speech rate (extracted from reading of a sentence of 8 words in a neutral manner) and net speech rate (extracted from short imperative sentence reading and reading of a sentence of 8 words in a neutral manner). The strongest correlation can be seen in the case of TSR extracted from the reading of a sentence of 8 words in a neutral manner ($\rho\,(\mathrm{P}) = 0.33$, $p < 0.05$, and $\rho\,(\mathrm{S}) = 0.33$, $p < 0.05$). And finally, with respect to the total score (Q3–Q6), interpercentile range of the first formant (extracted from paragraph reading and reading of 5 words excitedly), and net speech rate (extracted from short imperative sentence reading) were found significantly correlated with this item. The strongest correlation can be seen in the case of F1IR extracted from the paragraph reading ($\rho\,(\mathrm{P}) = -0.40$, $p < 0.05$, and $\rho\,(\mathrm{S}) = -0.45$, $p < 0.05$). Next, the results of the multivariate regression analysis can be seen in Table 4.4. Moreover, the models for FOG-G (Q5), and FOG (Q6) are visualized (visualization of the approximation of decision making performed by the regression tree) using the three graphs, see Figure 4.4, and Figure 4.5, respectively.

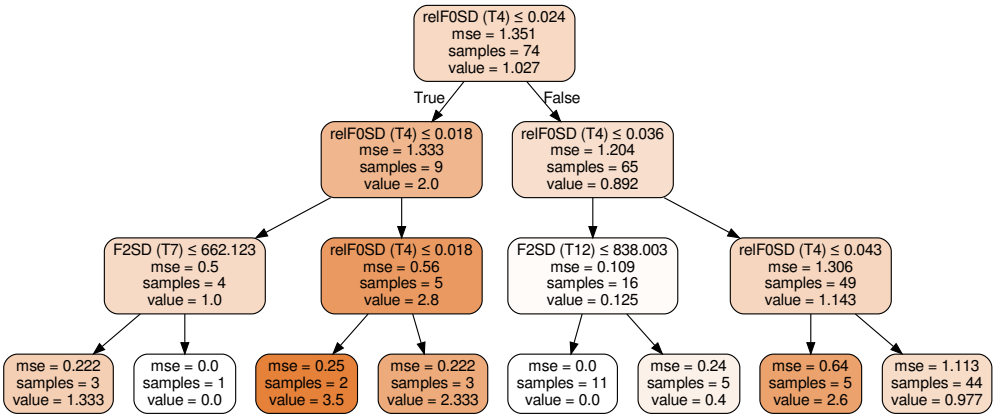

Fig. 4.4: Visualization of the regression tree built to estimate FOG-Q (Q5). The tree was trained using a single run applied on all data (all speech tasks and all acoustic features) in the dataset for the features selected by the feature selection algorithm (hence the decision making of the tree is an approximation of the behavior responsible for the results summarized in Table 4.4). In the case of this tree, three acoustic features are used: F2SD (T7), relF0SD (T4), and F2SD (T12).

Regarding the multivariate regression analysis, the results can be seen in Table 4.4. The table contains the results related to the prediction of the change in FOG severity in a two-year horizon. When considering the three HD dimensions separately, the following results were achieved. The change in Q3 was predicted with the estimation error of 20.96 % using 3 prosodic features. Specifically, TSR

Tab. 4.4: FOG deficits prediction using classification and regression trees.

FOG-Q	MAE	RMSE	EER	No.	selected features
			Articulation		
Q3	0.86 ± 0.26	1.03 ± 0.28	20.96 ± 6.38	1	$F1SD^{11}$
Q4	0.76 ± 0.28	0.89 ± 0.31	20.78 ± 7.73	2	$F1IR^{6}, F2SD^{14}$
Q5	0.49 ± 0.22	0.64 ± 0.34	10.52 ± 4.67	4	$F2SD^{7}, F1IR^{11}, F1SD^{12}, DDKr^{5}$
Q6	0.60 ± 0.28	0.77 ± 0.43	13.85 ± 6.41	1	$F1SD^{6}$
T	2.15 ± 0.63	2.53 ± 0.73	21.89 ± 6.44	1	$F1SD^{9}$
			Phonation		
Q3	1.11 ± 0.30	1.29 ± 0.33	27.11 ± 7.33	1	$jitter^{4}$
Q4	0.94 ± 0.28	1.14 ± 0.31	25.84 ± 7.57	2	$shimmer^{4}, jitter^{4}$
Q5	0.62 ± 0.24	0.81 ± 0.33	13.42 ± 5.18	1	MPT^{3}
Q6	0.60 ± 0.24	0.79 ± 0.34	13.95 ± 5.57	1	MPT^{2}
T	2.32 ± 0.75	2.91 ± 0.90	23.64 ± 7.63	1	$relF0SD^{4}$
			Prosody		
Q3	0.85 ± 0.33	1.04 ± 0.39	20.90 ± 8.01	3	$TSR^{11}, TSR^{10}, relF0SD^{11}$
Q4	0.80 ± 0.24	0.96 ± 0.27	21.88 ± 6.69	1	TSR^{7}
Q5	0.56 ± 0.22	0.71 ± 0.31	12.09 ± 4.77	3	$relSEOSD^{9}, SPIR^{10}, relF0SD^{1}$
Q6	0.55 ± 0.20	0.71 ± 0.26	12.75 ± 4.54	2	TSR^{7}, NSR^{14}
T	2.07 ± 0.71	2.59 ± 0.88	21.10 ± 7.20	4	$TSR^{11}, TSR^{10}, TSR^{9}, NSR^{8}$
			Combination		
Q3	0.83 ± 0.27	1.01 ± 0.31	20.40 ± 6.73	3	$F1SD^{11}, relF0SD^{6}, F2IR^{1}$
Q4	0.76 ± 0.28	0.89 ± 0.31	20.78 ± 7.73	2	$F1IR^{6}, F2SD^{14}$
Q5	0.51 ± 0.21	0.66 ± 0.32	11.03 ± 4.59	3	$F2SD^{7}, relF0SD^{4}, F2SD^{12}$
Q6	0.50 ± 0.21	0.65 ± 0.29	11.73 ± 4.93	4	$TSR^{7}, HNRm^{4}, F2SD^{7}, TSR^{11}$
T	2.00 ± 0.69	2.48 ± 0.82	20.35 ± 7.08	3	$F1SD^{9}, TSR^{11}, F1IR^{6}$

[1] Table notation: MAE – mean absolute error; RMSE – root mean squared error; EER – relative estimation error rate (mean absolute error divided by the range of actual values of clinical rating scale present in the dataset; expressed in %); No. – number of selected features; $feature^{x}$ – acoustic feature and the label of the speech task (x), see Section 4.1; FOG-Q – Freezing of gait questionnaire [21] (Q3–Q6, and T – total score), for more details, see Section 4.1 as well.

(reading of a sentence of 8 words in a neutral manner), TSR (paragraph reading), and relF0SD (reading of a sentence of 8 words in a neutral manner). The change in Q4 was predicted with the estimation error of 20.78 % using 2 articulatory features. Specifically, F1IR (rhythmical reading), and F2SD (reading of 5 words in an excited manner). The change in Q5 was predicted with the estimation error of 10.52 % using 4 articulatory features. Specifically, F2SD (short interrogative sentence reading), F1IR (reading of a sentence of 8 words in a neutral manner), F1SD (reading of a sentence of 6 words angrily), and DDKr (diadochokinetic task). The change in Q6

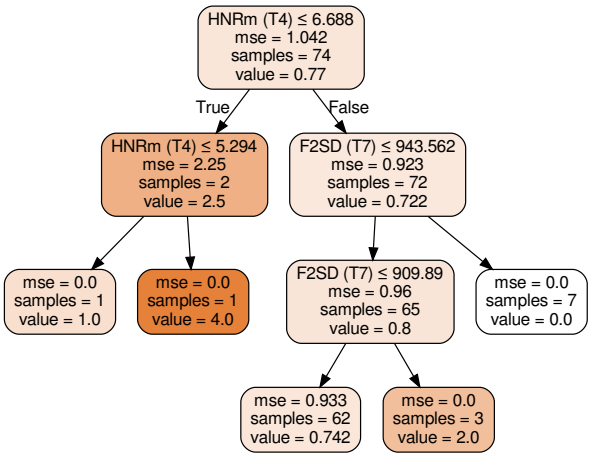

Fig. 4.5: Visualization of the regression tree built to estimate FOG-Q (Q6). The tree was trained using a single run applied on all data (all speech tasks and all acoustic features) in the dataset for the features selected by the feature selection algorithm (hence the decision making of the tree is an approximation of the behavior responsible for the results summarized in Table 4.4). In the case of this tree, three acoustic features are used: TSR (T7), HNRm (T4), F2SD (T7), and TSR (T11).

was predicted with the estimation error of 12.75 % using 2 prosodic features. Specifically, TSR (short interrogative sentence reading), and NSR (reading of 5 words in an excited manner). The change in total score (Q3–Q6) was predicted with the estimation error of 21.10 % using 4 prosodic features. Specifically, TSR (reading of a sentence of 8 words in a neutral manner), TSR (paragraph reading), TSR (short declarative sentence reading), and NSR (short imperative sentence reading). And finally, when considering a combination of the features, the prediction was improved in the case of Q3, Q6, and total score (the difference, i. e. improvement is shown [in percentage]): Q3 (0.56), Q6 (1.02), and total score (0.75). However, as can be seen, the improvements are not that significant, which shows a strong relationship between separate HD areas and specific FOG deficits.

5 CONCLUDING SUMMARY

This doctoral thesis deals with quantitative acoustic analysis of dysarthric speech applied in the field of objective non-invasive computerized diagnosis and assessment of idiopathic PD. The first study presented in this thesis is focused on robust quantification, description and identification of monopitch, monoloudness and speech rate/pausing abnormalities in patients with PD. In the frame of this study, speech recordings acquired from 98 PD patients and 51 healthy speakers were investigated. For this purpose, three specifically-designed speech tasks were recorded to quantify variability of speech melody, speech-stress control and naturalness of speech rate and pausing. With respect to the analyses, a complex comparison between HC and patients with PD in terms of gender-related distinctions occurring with parkinsonian dysprosody, and a unique investigation of the possibilities of HD identification using specific prosodic scenarios was performed. In addition, permutation test was applied to evaluate the statistical power of the predictions made by the multivariate classification models trained to discriminate healthy and dysarthric speech.

The second study presented in this thesis is focused on computerized and objective assessment of motor and non-motor symptoms of PD based on the quantitative acoustic analysis of dysarthric speech at the baseline. In the frame of this study, speech recordings and clinical data acquired from 72 PD patients were investigated. For this purpose, the same speech tasks as well as the acoustic features as in the case of the previous study was used. As opposed to the previous study, the correlation analysis aiming at investigating the relationship between dysprosody in HD and other non-speech symptoms of PD was employed. In addition to that, multivariate regression models capable of precise assessment of PD severity were built. These regression models used only the information about prosodic deficits of the patients at the baseline to predict the scores of a variety of clinical rating scales that are nowadays being commonly used to assess severity of motor and non-motor symptoms of PD.

The third study presented in this thesis is focused on computerized and objective assessment of freezing of gait in PD in the horizon of two years based on the quantitative acoustic analysis of dysarthric speech at the baseline. In the frame of this study, a robust set of acoustic features and speech task quantifying phonation, articulation, prosody, and speech fluency were used. For this purpose, speech recordings and clinical data acquired from 75 and 41 PD patients at the baseline and at the follow-up examination were investigated, respectively. In this study, multivariate regression models capable of predicting the change of gait-related deficits in the horizon of two years based on the information about severity of HD at the baseline are built. Furthermore, partial correlation analysis was performed in direction of

investigating pathological mechanisms shared by HD and freezing of gait in PD.

The main goal of this doctoral thesis that was to *investigate possibilities of using quantitative objective evaluation of HD, employing speech parametrization, statistical analyses and machine learning techniques, in direction of identification and assessment of PD*, as well as all its objectives were successfully accomplished. Specifically, the following goals were achieved:

1. *Robust computerized quantification of HD manifestations in PD* was performed. In the area of phonation, microperturbations in frequency and amplitude, irregular pitch fluctuations, tremor of jaw, increased acoustic noise, insufficient breath support and aperiodicity of voice were quantified. In the area of articulation, rigidity of tongue and jaw, slow alternating motion rate during diadochokinesis and irregular alternating motion rate during diadochokinesis were quantified. In the area of speech prosody, monopitch and monoloudness were quantified. And finally, in the area of speech fluency, inappropriate silences and unnatural speech rate were quantified. These acoustic features provided a basis for complex computerized description of HD in PD.
2. *Complex analysis and identification of dysprosody in HD* was employed. To quantify dysprosody in HD, conventional prosodic features, quantifying monopitch, monoloudness and speech rate/pausing abnormalities, were computed from the recordings of three specialized speech tasks: a) poem recitation task (description of flat speech melody), b) stress-modified reading (description of insufficient stress-control), and c) emotionally-neutral reading (description of speech rate/pausing abnormalities). Next, a comparison between dysarthric and healthy speech was performed. Additionally, multivariate classification models were built to discriminate between PD patients and HC. All of the analyses were employed in the gender-specific setup. Finally, each dimension of dysprosody was evaluated separately as well.
3. *Assessment of non-speech symptoms of PD at the baseline* was employed. To follow and build on top of the findings and conclusions of the previous study focused on identification of dysprosody in HD, the same acquisition and parameterization setup was used. Here, correlation analysis between prosodic features and values (scores) of a variety of clinical rating scales assessing motor and non-motor symptoms of PD was performed. Moreover, the computed prosodic features were used to train and evaluate multivariate regression models that were proved to be capable of assessing the values of these rating scales based solely on the information about the severity of HD at the baseline.
4. *Assessment of gait freezing in PD in the horizon of two years* was employed. To robustly describe HD in PD, a large variety of speech tasks such as sustained phonation, expiration, reading, free speech (monologue), diadochokinesis, etc.

and acoustic features quantifying all dimensions of speech production were studied. These features were consequently used to to train and evaluate multivariate regression models that were proved to be capable of predicting the change in the freezing of gait occurring with PD in the horizon of two years based solely on the information about the severity of HD at the baseline.

5. *Analysis of pathological mechanism shared by HD and gait freezing in PD* was employed. To investigate if there are pathological mechanisms shared by HD and freezing of gait in PD. Partial correlation analysis, controlling for the effect of other confounding factors such as age, gender, dopaminergic medication, etc., between the acoustic features and values of the specialized clinical rating scale assessing gait-related deficits in PD was performed. This analysis pointed out to some interesting facts about the relationship between HD and gait freezing in patients with PD.

Regarding the future direction of the research described in this thesis, application of the presented methodology for assessing of other common parkinsonian symptoms such as depression or cognitive deficits at the baseline as well as in the direction of two years is considered. Moreover, investigation of pathological mechanisms shared by HD and other symptoms of PD besides freezing of gait is considered. Next, application of quantitative acoustic analysis of dysarthric speech in direction of tuning the parameters of novel perspective PD treatment methods such as rTMS is considered as well. And finally, the ultimate goal behind this research is the fusion of clinical and paraclinical data in order to develop and evaluate a decision support system that would help clinicians with diagnosis, assessment, and monitoring of PD.

BIBLIOGRAPHY

[1] Ackermann, H.; Konczak, J.; Hertrich, I.: The temporal control of repetitive articulatory movements in Parkinson's disease. *Brain Lang.*, vol. 56, no. 2, 1997: pp. 312–9.

[2] Adams, S. G.; Dykstra, A.; McNeil, M.: Hypokinetic dysarthria. *Clinical Management of Sensorimotor Speech Disorders*, 2009: pp. 166–186.

[3] Arias-Vergara, T.; Vásquez-Correa, J. C.; Orozco-Arroyave, J. R.: Parkinson's Disease and Aging: Analysis of Their Effect in Phonation and Articulation of Speech. *Cogn Comput*, vol. 9, no. 6, 2017: pp. 731–48.

[4] Asgari, M.; Shafran, I.: Predicting severity of Parkinson's disease from speech. In *Engineering in Medicine and Biology Society (EMBC), 2010 Annual International Conference of the IEEE*, Aug 2010, ISSN 1557-170X, pp. 5201–5204.

[5] Bartels, A.; Balash, Y.; Gurevich, T.; Schaafsma, J.; Hausdorff, J.; Giladi, N.: Relationship between freezing of gait (FOG) and other features of Parkinson's: FOG is not correlated with bradykinesia. *J Clin Neurosci*, vol. 10, no. 5, 2003: pp. 584–8.

[6] Bayestehtashk, A.; Asgari, M.; Shafran, I.; McNames, J.: Fully automated assessment of the severity of Parkinson's disease from speech. *Comput Speech Lang*, vol. 29, no. 1, 2015: pp. 172–185.

[7] Beck, A. T.; Steer, D. J., R. A. Rissmiller: Use of the Beck Depression Inventory-II with depressed geriatric inpatients. *Behav. Res. Ther.*, vol. 38, no. 3, 2000: pp. 311–318.

[8] Beck, A. T.; Ward, C.; Mendelson, M.; Mock, J.; Erbaugh, J.: An inventory for measuring depression. *Archives of General Psychiatry*, vol. 4, no. 6, 1961: pp. 561–571.

[9] Birnbaum, Z.; et al.: On a use of the Mann-Whitney statistic. In *Proceedings of the Third Berkeley Symposium on Mathematical Statistics and Probability, Volume 1: Contributions to the Theory of Statistics*, The Regents of the University of California, 1956.

[10] Boersma, P.; Weenink, D.: Praat: Doing Phonetics by Computer (Version 5.3.10). 2012. URL `http://www.fon.hum.uva.nl/praat/`

[11] Brabenec, L.; Mekyska, J.; Galaz, Z.; Rektorova, I.: Speech disorders in Parkinson's disease: early diagnostics and effects of medication and brain stimulation. *J Neural Transm*, vol. 124, no. 3, 2017: pp. 303–34.

[12] Breiman, L.; Friedman, J. H.; Olshen, R. A.; Stone, C. J.: *Classification and Regression Trees*. Monterey, CA: Wadsworth and Brooks, 1984, ISBN 0-534-98053-8.

[13] Cantiniaux, S.; Vaugoyeau, M.; Robert, D.; Horrelou-Pitek, C.; Mancini, J.; Witjas, T.; Azulay, J.-P.: Comparative analysis of gait and speech in Parkinson's disease: hypokinetic or dysrhythmic disorders? *J Neurol Neurosurg Psychiatry*, vol. 81, no. 2, 2010: pp. 177–84.

[14] Chaudhuri, K. R.; Martinez-Martin, P.; Brown, R. G.; Sethi, K.; Stocchi, F.; Odin, P.; Ondo, W.; Abe, K.; MacPhee, G.; MacMahon, D.; et al.: The metric properties of a novel non-motor symptoms scale for Parkinson's disease: Results from an international pilot study. *Mov. disord.*, vol. 22, no. 13, 2007: pp. 1901–1911.

[15] Combrisson, E.; Jerbi, K.: Exceeding chance level by chance: The caveat of theoretical chance levels in brain signal classification and statistical assessment of decoding accuracy. *J. Neurosci. Methods*, vol. 250, jul 2015: pp. 126–136, doi: 10.1016/j.jneumeth.2015.01.010.

[16] Eskidere, Ö.; Ertaş, F.; Hanilçi, C.: A comparison of regression methods for remote tracking of Parkinson's disease progression. *Expert Systems with Applications*, vol. 39, no. 5, 2012: pp. 5523–5528, ISSN 0957-4174, doi: http://dx.doi.org/10.1016/j.eswa.2011.11.067.

[17] Fahn, S.; Elton, R. L.: *UPDRS Development Committee (1987) Unified Parkinson's Disease Rating Scale*. Recent developments in Parkinson's Disease. Macmillan, Florham Park, 1987.

[18] Folstein, M.; Folstein, S. E.; McHugh, P. R.: Mini-mental state: A practical method for grading the cognitive state of patients for the clinician. *J Psychiatr Res*, vol. 12, no. 3, 1975: pp. 189–98.

[19] Galaz, Z.; Mekyska, J.; Kiska, T.; Zvoncak, V.; Mucha, J.; Harar, P.; Smekal, Z.; Eliasova, I.; Mrackova, M.; Kostalova, M.; et al.: Monitoring Progress of Parkinson's Disease Based on Changes in Phonation: a Pilot Study. In *Telecommunications and Signal Processing (TSP), 2018 41st International Conference on*, IEEE, 2018, pp. 218–221.

[20] Giladi, N.; McDermott, M. P.; Fahn, S.; Przedborski, S.; Jankovic, J.; Stern, M.; Tanner, C.: Freezing of gait in PD: prospective assessment in the DATATOP cohort. *Neurology*, vol. 56, no. 12, 2001: pp. 1712–21.

[21] Giladi, N.; Shabtai, H.; Simon, E. S.; Biran, S.; Tal, J.; Korczyn, S. D.: Construction of freezing of gait questionnaire for patients with Parkinsonism. *Parkinsonism Relat Disord*, vol. 6, no. 3, 2000: pp. 165–70.

[22] Goberman, A. M.: Correlation between acoustic speech characteristics and non-speech motor performance in Parkinson's disease. *Med Sci Monit*, vol. 11, no. 3, 2005: pp. CR109–116.

[23] Golland, P.; Liang, F.; Mukherjee, S.; Panchenko, D.: Permutation Tests for Classification. In *Learn. Theory*, 2005, pp. 501–515, doi: 10.1007/11503415_34.

[24] Guyon, I.; Gunn, S.; Nikravesh, M.; Zadeh, L. A.; Eds: *Feature Extraction: Foundations and Applications*. New York: Springer, 2006.

[25] Jurman, G.; Riccadonna, S.; Furlanello, C.: A Comparison of MCC and CEN Error Measures in Multi-Class Prediction. *PLoS ONE*, vol. 7, no. 8, aug 2012: pp. e41882, doi: 10.1371/journal.pone.0041882.

[26] Kostalova, M.; Mrackova, M.; Marecek, R.; Berankova, D.; Eliasova, I.; Janousova, E.; Roubickova, J.; Bednarik, J.; Rektorova, I.: The 3F Test Dysarthric Profile – Normative Speach Values in Czech. *Cesk Slov Neurol Neurochir*, vol. 76, no. 5, 2013: pp. 614–18.

[27] Larner, A. J.: Addenbrooke's cognitive examination-revised (ACE-R) in day-to-day clinical practice. *Age Ageing*, vol. 36, no. 6, 2007: pp. 685–686.

[28] Lee, J. Y.; Kim, J. W.; Lee, W. Y.; Kim, J. M.; Ahn, T. B.; Kim, H. J.; Cho, J.; Jeon, B. S.: Daily dose of dopaminergic medications in Parkinson's disease: clinical correlates and a posteriori equation. *Neurol Asia*, vol. 15, no. 2, 2010: pp. 137–43.

[29] Ma, J. K.-Y.; Schneider, C. B.; Hoffmann, R.; Storch, A.: Speech Prosody Across Stimulus Types for Individuals with Parkinson's disease. *J. Parkinson's Dis.*, February 2015, ISSN 1877-7171, doi: 10.3233/jpd-140451.

[30] Maaten, L. v. d.; Hinton, G.: Visualizing data using t-SNE. *Journal of machine learning research*, vol. 9, no. Nov, 2008: pp. 2579–2605.

[31] Matthews, B. W.: Comparison of the predicted and observed secondary structure of T4 phage lysozyme. *Biochim. Biophys. Acta (BBA)*, vol. 405, no. 2, 1975: pp. 442–51.

[32] Mekyska, J.; Galaz, Z.; Mzourek, Z.; Smekal, Z.; Rektorova, I.: Assessing Progress of Parkinson's Using Acoustic Analysis of Phonation. In *Bioinspired Intelligence (IWOBI), 2015 4th International Work Conference on*, IEEE, June 2015, pp. 111–118, doi: 10.1109/IWOBI.2015.7160153.

[33] Mekyska, J.; Janousova, E.; Gomez-Vilda, P.; Smekal, Z.; Rektorova, I.; Eliasova, I.; Kostalova, M.; Mrackova, M.; Alonso-Hernandez, J. B.; Faundez-Zanuy, M.; et al.: Robust and complex approach of pathological speech signal analysis. *Neurocomputing*, vol. 167, 2015: pp. 94–111.

[34] Moreau, C.; Ozsancak, C.; Blatt, J.-L.; Destee, A.; Defebvre, L.: Oral festination in Parkinson's disease: biomechanical analysis and correlation with festination and freezing of gait. *Mov Disord*, vol. 22, no. 10, 2007: pp. 1503–6.

[35] Ojala, M.; Garriga, G.: Permutation Tests for Studying Classifier Performance. In *Data Mining, 2009. ICDM '09. Ninth IEEE International Conference on*, Dec 2009, ISSN 1550-4786, pp. 908–913, doi: 10.1109/ICDM.2009.108.

[36] Park, H. K.; Yoo, J. Y.; Kwon, M.; Lee, J.-H.; Lee, S. J.; Kim, S. R.; Kim, M. J.; Lee, M. C.; Lee, S. M.; Chung, S. J.: Gait freezing and speech disturbance in Parkinson's disease. *Neurol Sci*, vol. 35, no. 3, 2014: pp. 357–63.

[37] Peterek, T.; Dohnálek, P.; Gajdoš, P.; Šmondrk, M.: Performance evaluation of Random Forest regression model in tracking Parkinson's disease progress. In *Hybrid Intelligent Systems (HIS), 2013 13th International Conference on*, IEEE, 2013, pp. 83–87.

[38] Phipson, B.; Smyth, G. K.: Permutation P-values Should Never Be Zero: Calculating Exact P-values When Permutations Are Randomly Drawn. *Stat. Appl. Genet. Mol. Biol.*, vol. 9, no. 1, jan 2010, doi: 10.2202/1544-6115.1585.

[39] Pohjalainen, J.; Rasanen, O.; Kadioglu, S.: Feature Selection Methods and Their Combinations in High-Dimensional Classification of Speaker Likability, Intelligibility and Personality Traits. *Comput. Speech Lang.*, 2014.

[40] Rektorova, I.; Mekyska, J.; Janousova, E.; Kostalova, M.; Eliasova, I.; Mrackova, M.; Berankova, D.; Necasova, T.; Smekal, Z.; Marecek, R.: Speech prosody impairment predicts cognitive decline in Parkinson's disease. *Parkinsonism Relat Disord*, vol. 29, 2016: pp. 90–5.

[41] Rusz, J.; Cmejla, R.; Ruzickova, H.; Klempir, J.; Majerova, V.; Picmausova, J.; Roth, J.; Ruzicka, E.: Acoustic assessment of voice and speech disorders in Parkinson's disease through quick vocal test. *Mov. Disord.*, vol. 26, no. 10, 2011: pp. 1951–1952.

[42] 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*, vol. 129, no. 1, 2011: pp. 350–367.

[43] Sheskin, D. J.: *Handbook of Parametric and Nonparametric Statistical Procedures*. Chapman & Hall/CRC, 2007.

[44] Skodda, S.; Flasskamp, A.; Schlegel, U.: Instability of syllable repetition as a marker of disease progression in Parkinson's disease: a longitudinal study. *Mov Disord*, vol. 26, no. 1, 2011: pp. 59–64.

[45] Skodda, S.; Schlegel, U.: Speech rate and rhythm in Parkinson's disease. *Mov. Disord.*, vol. 23, no. 7, 2008: pp. 985–92, ISSN 1531-8257.

[46] Skodda, S.; Visser, W.; Schlegel, U.: Gender-related patterns of dysprosody in Parkinson's disease and correlation between speech variables and motor symptoms. *J. Voice*, vol. 25, no. 1, 2011: pp. 76–82.

[47] Stiasny-Kolster, K.; Mayer, G.; Schafer, S.; Muller, J. C.; Heinzel-Gutenbrunner, M.; Oertel, W. H.: The REM sleep behavior disorder screening questionnaire – A new diagnostic instrument. *Mov Disord*, vol. 22, no. 16, 2007: pp. 2386–93.

[48] Tsanas, A.: New nonlinear markers and insights into speech signal degradation for effective tracking of Parkinson's disease symptom severity. *Age (years)*, vol. 64, no. 8.1, 2010: pp. 63–6.

[49] Tsanas, A.; Little, M.; McSharry, P.; Ramig, L.: Accurate telemonitoring of Parkinson's Disease progression by noninvasive speech tests. *IEEE Trans. Bio-Med. Eng.*, vol. 57, no. 4, 2010: pp. 884–893.

[50] Tsanas, A.; Little, M. A.; McSharry, P. E.; Ramig, L. O.: Nonlinear speech analysis algorithms mapped to a standard metric achieve clinically useful quantification of average Parkinson's disease symptom severity. *J. R. Soc. Interface*, vol. 8, no. 59, 2010: pp. 842–855.

Curriculum Vitæ

Zoltán Galáž

Education

2017–∗	Masaryk University, Faculty of Medicine (Neurosciences), Kamenice 5, 625 00 Brno, degree: Ph.D. (expected in: 2021).
2014–2018	Brno University of Technology, Faculty of Electrical Engineering and Communication, Technická 3058/10, 616 00 Brno, degree: Ph.D. (expected in: 2018).
2011–2014	Brno University of Technology, Faculty of Electrical Engineering and Communication, Technická 3058/10, 616 00 Brno, degree: M.Sc (achieved in: 2014).
2008–2011	Brno University of Technology, Faculty of Electrical Engineering and Communication, Technická 3058/10, 616 00 Brno, degree: B.Sc (achieved in: 2011).
2004–2008	Gymnasium of Jan Botto, Trnava, Slovak Republic.

Internships

2018	University of Arizona Health Sciences (Neurology), University of Arizona, Department of Neurology, 1501 N. Campbell Ave, Tucson, Arizona, USA.
2016	Instituto para el Desarrollo Tecnológico y la Innovación en Comunicaciones (IDeTIC), Universidad de Las Palmas de Gran Canaria, 35001 Las Palmas de Gran Canaria, Spain.
2015	Facultad de Informática, Universidad Politécnica de Madrid (UPM), Campus de Montegancedo, s/n, 28660 Boadilla del Monte, Madrid, Spain.

Teaching

2014–2015	Assistant lecturer, Signals and systems analysis.
2015–2016	Assistant lecturer, Digital Signal Processing.
2017–2018	Assistant lecturer, Speech Processing.

Awards

2016	Publication in Computer Methods and Programs in Biomedicine (IF: 2.199) in the editor's choice (first author).
2015	Top 10 pedagogue at Brno University of Technology (anonymous student poll evaluating the quality of education).
2015	Fist place at "Conference of Faculty of Electrical Engineering and Communication" (EEICT 2015).
2014	Brno University of Technology dean's prize for the master thesis "Analysis of hand-written text in patients with neurological disorders".

Employment history

2017–∗	*researcher*: Applied Neuroscience Research Group, Central European Institute of Technology, Masaryk University, Kamenice 5, 625 00 Brno, Czech Republic.
2017–∗	*editor*: Elektrorevue journal, International Society for Science and Engineering, o.s., Klíčova 1261/2c, 618 00 Brno.
2017–∗	*software developer and data scientist*: Inventurist LLC, 585 Broadway Redwood City, CA, 94063, USA.
2015–∗	*research team leader*: Brain Disease Analysis Laboratory (BDALab) Department of Telecommunications, Faculty of Electrical Engineering and Communication, Brno University of Technology, Technická 12, 616 00 Brno, Czech Republic.
2015–∗	*researcher*: Signal Processing Laboratory (SPLab) Department of Telecommunications, Faculty of Electrical Engineering and Communication, Brno University of Technology, Technická 12, 616 00 Brno, Czech Republic.

Participation in projects

2018–2020	The Czech Science Foundation (18-16835S): *Research of advanced developmental dysgraphia diagnosis and rating methods based on quantitative analysis of online handwriting/drawing.*

2017–2021	H2020 Marie Sklodowska-Curie Research and Innovation Staff Exchange (H2020-MSCA-RISE-2016 734718): *Novel Network-Based Approaches for Studying Cognitive Dysfunction in Behavioral Neurology.*
2017–2020	Ministry of the Interior of Czech Republic (VI20172020078): *System for centralized supervision of complex and large objects of state's critical infrastructure.*
2016–2019	Ministry of Health of Czech Republic (NV16-30805A): *Effects of non-invasive brain stimulation on hypokinetic dysarthria, micrographia, and brain plasticity in patients with Parkinson's disease.*
2015–2017	Technology Agency of Czech Republic (TA04031666): *Intelligent Telematics Information System of Public Transportation II.*
2015–2016	European Cooperation in Science & Technology (LD14091): *De-Identification for Privacy Protection in Multimedia Content.*
2012–2015	Ministry of Health of Czech Republic (NT13499): *Speech, its disorders and cognitive function in Parkinson's disease.*

Invited Lectures

2016	*Statistical methods used in the field of objective analysis of Parkinson's Disease*, University of Defence, Kounicova 65 662 10 Brno, Invited by doc. RNDr. Jaroslav Michálek, CSc.
2015	*The power of Parkinson's disease*, TEDx Trencin (2015). For more information, see: video, Invited by the organizers.

Research activity

- Publications in journals with impact factor: 7
- Publications in journals without impact factor: 5
- Publications in books: 1
- Publications in conference proceedings: 13
- Software/tools: 29
- Publications indexed by WoS: 11
- Publications indexed by Scopus: 12
- H-index according to WoS: 3
- H-index according to Scopus: 4

ABSTRACT

Hypokinetic dysarthria (HD) is a speech disorder occurring in up to 90 % of patients suffering from idiopathic Parkinson's disease (PD) that significantly contributes to unnaturalness and incomprehensibility of speech of these patients. The main aim of this doctoral thesis is to investigate possibilities of using quantitative para-clinical analysis of HD, employing speech parametrization, statistical analyses, and machine learning techniques, for diagnosis and remote objective assessment of PD. This thesis demonstrates that it is possible to use computerized acoustic analysis to sufficiently describe HD, especially dysprosody, which is characterized by flat speech melody and unnatural speech rate. Moreover, it demonstrates it is also possible to use robust clinically interpretable acoustic parameters quantifying various manifestations of HD, such as phonation, articulation and prosody, to assess the severity of motor and non-motor symptoms of PD. Next, it presents the investigation of pathophysiological mechanisms shared by HD and freezing of gait in PD. And finally, it proves it is also possible to accurately estimate the change in gait-related deficits in the horizon of two years using acoustic analysis at the baseline.

ABSTRAKT

Hypokinetická dysartrie (HD) je častým symptomem vyskytujícím se až u 90 % pacientů trpících idiopatickou Parkinsonovou nemocí (PN), která výrazně přispívá k nepřirozenosti a nesrozumitelnosti řeči těchto pacientů. Hlavním cílem této disertační práce je prozkoumat možnosti použití kvantitativní paraklinické analýzy HD, s použitím parametrizace řeči, statistického zpracování a strojového učení, za účelem diagnózy a objektivního hodnocení PN. Tato práce dokazuje, že počítačová akustická analýza je schopná dostatečne popsat HD, speciálně tzv. dysprozodii, která se projevuje nedokonalou intonací a nepřirozeným tempem řeči. Navíc také dokazuje, že použití klinicky interpretovatelných akustických parametrů kvantifikujících různé aspekty HD, jako jsou fonace, artikulace a prozodie, může být použito k objektivnímu posouzení závažnosti motorických a nemotorických symptomů vyskytujících se u pacientů s PN. Dále tato práce prezentuje výzkum společných patofyziologických mechanizmů stojících za HD a zárazy v chůzi při PN. Nakonec tato práce dokazuje, že akustická analýza HD může být použita pro odhad progrese zárazů v chůzi v horizontu dvou let.