VĚDECKÉ SPISY VYSOKÉHO UČENÍ TECHNICKÉHO V BRNĚ
Edice Habilitační a inaugurační spisy, sv. 801
ISSN 1213-418X

Jiří Mekyska

UTILIZING DIGITAL SPEECH BIOMARKERS IN THE ASSESSMENT AND MONITORING OF PARKINSON'S DISEASE

BRNO UNIVERSITY OF TECHNOLOGY
FACULTY OF ELECTRICAL ENGINEERING AND COMMUNICATION
DEPARTMENT OF TELECOMMUNICATIONS

doc. Ing. Jiří Mekyska, Ph.D.

UTILIZING DIGITAL SPEECH BIOMARKERS IN THE ASSESSMENT AND MONITORING OF PARKINSON'S DISEASE

VYUŽITÍ DIGITÁLNÍCH ŘEČOVÝCH BIOMARKERŮ PŘI HODNOCENÍ A MONITOROVÁNÍ PARKINSONOVY NEMOCI

TEZE PŘEDNÁŠKY
K PROFESORSKÉMU JMENOVACÍMU ŘÍZENÍ

VYSOKÉ UČENÍ TECHNICKÉ V BRNĚ
NAKLADATELSTVÍ VUTIUM

BRNO 2024

KEYWORDS

Parkinson's disease, hypokinetic dysarthria, acoustic analysis, digital speech biomarkers, monitoring, assessment

KLÍČOVÁ SLOVA

Parkinsonova nemoc, hypokinetická dysartrie, akustická analýza, digitální řečové biomarkery, monitorování, hodnocení

ARCHIVED AT

Department of Science and International Relations
Faculty of Electrical Engineering and Communication
Brno University of Technology, Technicka 10, 616 00 Brno, Czech Republic

ISBN 978-80-214-6281-6
ISSN 1213-418X

AUTHOR'S DECLARATION

I have used ChatGPT (version 4) to summarize research articles and proofread English texts, enhancing clarity and fluency. I take full responsibility for the final content and confirm that these tools were used in accordance with the guidelines for generative AI tools issued by the Brno University of Technology.

Contents

Author bio

Jiří Mekyska is an associate professor at the Faculty of Electrical Engineering and Communication, Brno University of Technology. His research focuses on interpretable and trustworthy digital biomarkers that assist in the diagnosis, assessment, and monitoring of various disorders, including Parkinson's disease, Alzheimer's disease, Lewy body dementia, and neurodevelopmental dysgraphia. He serves as both a researcher and the head of the Brain Diseases Analysis Laboratory (BDALab) at the Department of Telecommunications. In this role, he teaches, supervises undergraduate and doctoral students, manages the BDALab, and acts as a principal or co-investigator in research projects, with results published in top-tier journals.

From a pedagogical point of view, he currently teaches four courses related to digital signal processing, one for ERASMUS+ students and another as part of a double-degree program. He has developed numerous teaching resources, including presentations, online lectures, and interactive demos. Jiří Mekyska has supervised four doctoral students, 23 master's students, and 15 bachelor's students, all of whom successfully defended their theses. He currently leads four doctoral students and has mentored five internships for international students. He has given a number of invited lectures at prestigious institutes and universities (e.g. at the Massachusetts Institute of Technology, Johns Hopkins University or the University of Southern California). For two academic years, he was ranked among the top 10 teachers at his faculty.

In terms of science, he has authored over 110 publications indexed in the Web of Science (with 1 565 citations, excluding self-citations, and an h-index of 23). He is the author of a number of articles published in Q1 journals (some of them are in the first decile). He is/was the principal investigator/co-investigator of 5 research projects (Czech Science Foundation, Ministry of Health of the Czech Republic, Technology Agency of the Czech Republic, Ministry of Education, Youth and Sports of the Czech Republic). He also participated in another eighteen projects. His research team (BDALab), consisting of seven members, has earned global recognition in the fields of biomedical signal processing, neuroscience, and psychology. He collaborates with numerous national and international universities, research institutes, and hospitals, including St. Anne's University Hospital in Brno, the Central European Institute of Technology, the Czech Academy of Sciences, Pompeu Fabra University, the University of Arizona, the University of Edinburgh, the University of Haifa, and Johns Hopkins University.

He is also active in several expert panels and groups. For example, he currently serves on the Science and Innovation Panel of the Innovative Health Initiative and previously chaired Panel P103 (Cybernetics, Artificial Intelligence, and Information Processing) for the Czech Science Foundation.

Introduction

Parkinson's disease (PD) is a chronic idiopathic disorder characterized by the progressive degeneration of dopaminergic neurons in the substantia nigra pars compacta [1], [2]. This degeneration is caused by the abnormal accumulation of the alpha-synuclein protein, which forms Lewy bodies (protein aggregates) that disrupt cellular function and are implicated in neuron death [3]. The loss of dopaminergic neurons leads to dopamine deficiency, primarily affecting the basal ganglia, a brain region responsible for movement control [4].

Globally, PD is the second most common neurodegenerative disorder after Alzheimer's disease, with prevalence increasing with age (and more prevalent in males). It affects approximately 1 % of individuals over 60 and about 2 % of those over 65 [5]. Environmental risk factors for PD include prolonged exposure to pesticides, heavy metals, and certain industrial chemicals [5], [6]. While motor symptoms like bradykinesia, muscular rigidity, resting tremor, and postural instability characterize PD, non-motor symptoms such as cognitive impairment, mood disorders, and sleep disturbances are also prevalent and significantly impact patients' quality of life [5]–[7].

Clinically, PD diagnosis is typically based on the presence of the above-mentioned motor symptoms, although prodromal signs (symptoms that precede the classical motor signs) may offer early diagnostic clues. These signs can include REM sleep behavior disorder, loss of smell (anosmia), constipation, and mood changes, which may precede the onset of motor symptoms by several years [5], [6]. Advances in biomarkers, particularly those derived from imaging studies, have enabled more accurate identification of individuals at risk for PD and facilitated research into early intervention strategies [5], [6].

Therapeutic options for PD are primarily symptomatic. Levodopa, a precursor to dopamine, remains the cornerstone of treatment for motor symptoms, though long-term use can lead to complications like motor fluctuations and dyskinesias [8], [9]. Deep brain stimulation (DBS) of the subthalamic nucleus has shown efficacy for some patients, particularly in controlling tremor and reducing medication requirements [8].

Due to the underlying basal ganglia pathology, up to 90 % of PD patients also experience a motor speech disorder called hypokinetic dysarthria (HD) [8], [10]. HD manifests mainly in these dimensions of speech:

- Respiration – Individuals with HD often experience reduced respiratory support, leading to decreased air pressure available for speech. This may result in a lower vocal intensity and shorter phrases due to insufficient breath support. The reduced control over respiratory muscles can limit the effectiveness of coordinated speech breathing [11].
- Phonation – Phonatory issues in HD primarily involve reduced vocal loudness, often perceived as a soft, breathy, or weak voice. Phonatory irregularities, such as increased jitter, shimmer, and reduced harmonic-to-noise ratio, are common. These reflect impairments in the vocal fold adduction and control, often contributing to a strained or effortful phonatory quality [8], [11]–[13].

- Resonance – Resonance abnormalities include a tendency toward hypernasality, though this is usually less pronounced compared to other dysarthria types. The altered resonance stems from the lack of control over the velopharyngeal mechanism, possibly due to generalized muscle rigidity and decreased range of movement in the articulatory organs [11].
- Articulation – Articulatory precision is notably reduced in HD, often resulting in slurred or imprecise consonant articulation. This manifests as a result of diminished range and force of articulatory movements. Consonants may appear "blurred" due to the inability to fully achieve the necessary articulatory positions, which is compounded by overall bradykinesia and rigidity of the facial and lingual muscles [8], [11], [13].
- Prosody – Prosodic deficits are significant in HD, typically resulting in monopitch, monoloudness, and inappropriate pauses. Speech rate abnormalities are common, with some individuals displaying a rapid, staccato-like rate, while others may experience sudden bursts of speed interspersed with prolonged pauses. These prosodic disturbances severely impact speech intelligibility and the natural rhythm of speech, contributing to the overall flat and unexpressive speech typical of HD [8], [11], [13], [14].

These combined effects of HD on speech components contribute to reduced intelligibility and communication effectiveness, posing substantial challenges in everyday interactions for those with PD.

Since PD manifests in speech, this modality can be advantageously leveraged for supportive diagnosis, assessment, or monitoring. For this purpose, digital speech biomarkers are particularly useful. The U.S. Food and Drug Administration (FDA) defines a digital biomarker as "a characteristic or set of characteristics, collected from digital health technologies, that is measured as an indicator of normal biological processes, pathogenic processes, or responses to an exposure or intervention, including therapeutic interventions" [15]. In the case of digital speech biomarkers, the digital health technology may consist of, for example, a smartphone with a built-in microphone or an external microphone connected to a laptop. An example of a characteristic could be the variability in the fundamental frequency (F0) of speech. Finally, in our context, the pathogenic process is represented by PD.

The aim of this work is to introduce the concept of acoustic speech analysis in patients with PD in depth and to highlight its benefits through two key examples: 1) research of a new treatment approach for HD based on repetitive transcranial magnetic stimulation (rTMS), and 2) predicting cognitive decline in PD patients.

1 Concept of acoustic analysis of speech in patients with PD

The general concept of acoustic analysis of speech in patients with PD is visualised in Fig. 1.1. The process usually consists of these steps:

1. Data acquisition – Depending on the intended application, speech data can be recorded under controlled conditions, such as in a clinical setting [16], or in less controlled environments, such as a patient's home [17]. To ensure reliable results and prevent misinterpretations, certain requirements must be met concerning the microphone, environment, sampling process, etc. Guidelines have been published to support this process [18], [19]. In addition, to capture all possible manifestations of HD, speech and voice are typically recorded using a comprehensive protocol that includes tasks such as free speech (e.g., a patient talking about her/his hobbies), reading words, sentences, or paragraphs; sustained phonation of vowels; diadochokinetic (DDK) tasks (rapid repetition of specific syllables or sound sequences, such as "pa–ta–ka," as quickly and accurately as possible); expiration tasks, and more [18], [20].
2. Labeling – For task-specific analysis, such as measuring the rate of DDK, when multiple tasks are present within a single recording, then segmentation is required . This means isolating the part of interest from the entire recording, which can be done automatically or manually. In the manual approach (generally more accurate), a person listens to the recording, observes the spectrogram, and identifies specific time points, typically marking the beginning and end of the prompt. These time points are then associated with labels that provide textual information about the prompt's content.

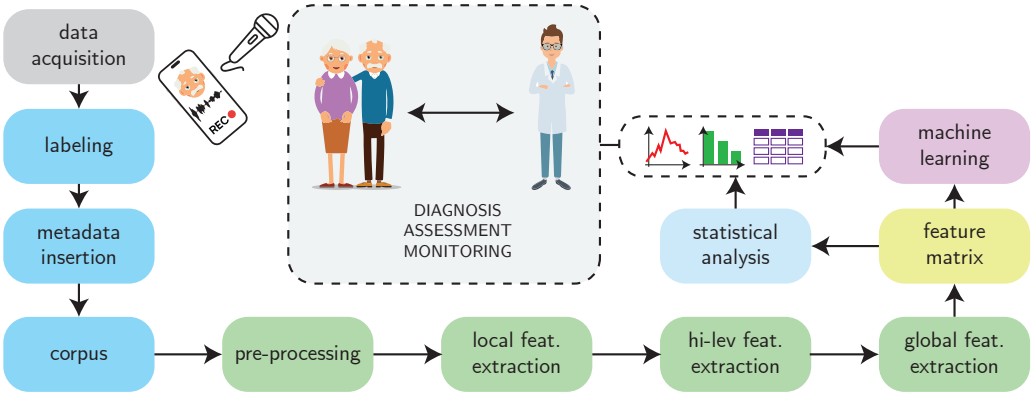

Fig. 1.1: Concept of acoustic analysis of speech in patients with PD.

3. Metadata insertion – Once the recordings are segmented, the acoustic data are associated with metadata essential for subsequent analysis and for providing insights into the

patient's profile. Typically, this includes demographic information (e.g., age and sex) and clinical information, such as scores from assessments like the UPDRS III (Unified Parkinson's Disease Rating Scale, Part III: Motor Examination), RBDSQ (REM Sleep Behavior Disorder Screening Questionnaire), FOG (Freezing of Gait Questionnaire), NMSS (Non-Motor Symptoms Scale), BDI (Beck Depression Inventory), ACE-R (Addenbrooke's Cognitive Examination – Revised), and medication details via LED (L-Dopa Equivalent Daily Dose) [12], [20]. After this step, we obtain a database of annotated speech recordings, commonly referred to as a corpus.

4. Pre-processing – Before proceeding to acoustic analysis and feature extraction, additional pre-processing of the signals may be necessary. For instance, we can perform resampling, normalization (standardizing signal amplitude to a consistent level), filtering with a pre-emphasis filter (a simple high-pass filter that accentuates formants), etc. [21]

5. Local feature extraction – Although the field of acoustic speech analysis is highly dynamic, with complex models playing an increasingly important role, the analysis of speech and voice pathology requires clinical interpretability, which often necessitates the use of so-called handcrafted features, i.e. features designed by humans based on heuristic insights. Examples of such features include fundamental frequency or speech rate. Generally, a speech feature is a metric extracted from a speech signal. In this work, a digital speech biomarker is considered a speech feature associated with a pathogenic process. The use of digital speech biomarkers in the field of PD is further discussed in Section 1.1.

6. High-level feature extraction – Features represented by tensors (e.g., segmental features) are typically transformed into scalar values using statistical measures such as the median, relative standard deviation, 95th percentile, or slope (using the Theil–Sen estimator), among others [22].

7. Global feature extraction – Extracting certain features requires input from multiple speech tasks. This is the case, for example, with the vowel articulation index, which is calculated based on formant frequencies extracted from several vowels. In this work, features calculated from multiple tasks will be considered as global features. The output of this step (or the previous one) is typically a feature matrix that undergoes further processing.

8. Statistical analysis – For an initial examination of the data, we typically use kernel density estimations, violin plots, or correlation matrices. When necessary, we control for confounding variables (e.g., age or medication levels) through regression. Since most features do not follow a normal distribution (as verified by tests like the Kolmogorov-Smirnov test), we apply non-parametric methods such as the Mann-Whitney U test, Wilcoxon signed-rank test, and/or Pearson's correlation with a significance level of $\alpha = 0.05$ during exploratory analysis. For analyses involving numerous features, we also apply false discovery rate correction.

9. Machine learning – Depending on the specific application, we model the feature space using approaches such as logistic regression, classification and regression trees, bagging and gradient boosting methods, or artificial neural networks. To avoid overfitting and ensure robust results, we typically employ a cross-validation strategy with multiple repetitions. Hyperparameters are optimized through random search or Bayesian optimization. Classifier performance is generally assessed by sensitivity, specificity, balanced accuracy, and the Matthews correlation coefficient. For regressors, we evaluate performance using metrics like mean absolute error, mean squared error, root mean squared error, and estimation error rate. To gain insight into model robustness, we conduct permutation testing. Finally, model interpretation is achieved using feature importance scores or SHAP (SHapley Additive exPlanations) [14]. In some cases, we visualize model performance with ROC (Receiver Operating Characteristic) curves.

1.1 Digital speech biomarkers associated with HD

In cases of pathological voice, tension on the vocal folds can vary greatly, causing the signal to become aperiodic and noise-like, making it challenging to identify patterns within the acoustic signal. Sub-harmonics and chaotic elements often appear, which can undermine the effectiveness of traditional speech signal analysis techniques. This challenge affects not only the voice but also speech, as articulation issues can lead to unintelligibility. In response, researchers have developed a range of advanced, handcrafted digital biomarkers to aid in identifying these pathologies. This group of digital speech biomarkers includes, e.g., those based on empirical mode decomposition, correlation dimension, fractal dimension, Hurst exponent, largest Lyapunov exponent, approximate entropy, sample entropy, correlation entropy, recurrence probability density entropy, detrended fluctuation analysis, pitch period entropy, normalized noise energy, segmental signal-to-dysperiodicity ratio, modulation spectra, and/or bicepstra [8], [22], [23].

However, the limitation of most of these biomarkers is that they are typically very challenging to clinically interpret. To facilitate the adoption of this technology by clinicians (e.g., neurologists or speech-language pathologists), we must avoid "black boxes" and clearly explain how each biomarker functions and how it is linked to physiological processes. In recent years, researchers have attempted to clinically explain some commonly used features, such as Mel-frequency cepstral coefficients (MFCC) [24] and cepstral peak prominence [25]. Nevertheless, the majority of works focused on clinical practice continue to implement simple acoustic measures [18].

Based on the manifestations of HD [11], a comprehensive review of scientific papers [8], [13], [14], [18], [20], [23], [26]–[29], discussions with experts, and over fifteen years of experience in the acoustic analysis of dysarthria, we have prepared a set of the most commonly used digital speech biomarkers linked to specific pathologies of HD. These biomarkers are listed in Tables 1.1–1.3. The utilization of these biomarkers is outlined in Section 2.

Tab. 1.1: Respiratory and phonatory digital biomarkers commonly used in HD analysis.

Respiration			
Specific manifestation of HD	**Tasks**	**Biomarker**	**Definition**
Airflow insufficiency	SPL	MPT	Maximum phonation time, the aerodynamic efficiency of the vocal tract, measured as the maximum duration of a sustained vowel.
Phonation			
Specific manifestation of HD	**Tasks**	**Biomarker**	**Definition**
Irregular pitch fluctuations	SP	relF0SD	Standard deviation of the fundamental frequency relative to its mean; variation in the frequency of vocal fold vibration.
Microperturbations in frequency	SP	PPQ	Frequency perturbation, indicating the extent of variation in the voice range. Jitter is defined as the variability of the F0 of speech from one cycle to the next.
Microperturbations in amplitude	SP	APQ	Amplitude perturbation, representing roughness in speech. Shimmer is defined as the sequence of maximum signal amplitude extent within each vocal cycle.
Increased noise	SP	HNR	Harmonics-to-noise ratio, the amount of noise in the speech signal due to incomplete vocal fold closure and/or turbulences in the vocal tract. HNR is defined as the ratio of harmonic (periodic) components to noise (non-periodic) components in a signal.
Aperiodicity	SP	DUV	Degree of unvoiced segments, the fraction of pitch frames marked as unvoiced.
Tremor of jaw	SP	relF1SD, relF2SD	Standard deviation of the first/second formant relative to its mean. Formants are related to resonances of the oro-naso-pharyngeal tract and are modified by the position of the tongue and jaw.
Increased hoarseness	SP, MO, RE	CPP	Cepstral peak prominence, defined as the difference between the cepstral peak representing the fundamental frequency and the linear regression line calculated from the magnitude-quefrency cepstra.

[1] MO – monologue; RE – reading; SP – sustained phonation of a vowel at a comfortable pitch and loudness; SPL – sustained phonation of a vowel at a comfortable pitch and loudness, held as constant and long as possible.

Tab. 1.2: Articulatory digital biomarkers commonly used in HD analysis.

Specific manifestation of HD	**Tasks**	**Biomarker**	**Definition**
Decreased tongue movement (imprecise vowels)	MO, RE	VAI	Vowel articulation index, based on formant centralization, defined as VAI = (F1a + F2i)/(F1i + F1u + F2a + F2u), where FXy is the Xth formant extracted from vowel y.
Rigidity of tongue and jaw	MO, RE	relF1SD, relF2SD	Standard deviation of the first/second formant relative to its mean.
Reduced intelligibility	RE	#lndmrk	The number of speech landmarks relative to total speech time, representing moments of different abrupt acoustic changes related to consonant production.
Slow alternating motion rate	DDK	PR	Pace rate, representing the number of syllable vocalizations per second, considering the first 30 syllables.
Instability of diadochokinetic pace	DDK	COV	Coefficient of variation, defined as the ratio of the standard deviation of the duration of the fourth to tenth DDK cycles to the average duration of the first three cycles.
Instability of diadochokinetic pace	DDK	RI	Rhythm instability, defined as the sum of absolute deviations from a regression line modeling each DDK cycle duration, weighted by total DDK performance time.
Acceleration of diadochokinetic pace	DDK	PA	Pace acceleration, defined as PA = 100 × (avCycDur4_6 – avCycDur7_9) / avCycDur1_3, where avCycDurX_Y is the average duration of cycles X_Y.
Acceleration of diadochokinetic pace	DDK	RA	Rhythm acceleration, defined as the gradient of the regression line modeling DDK cycle durations (positive values indicate acceleration).

[1] MO – monologue; RE – reading; DDK – diadochokinetic task.

Tab. 1.3: Prosodic digital biomarkers commonly used in HD analysis.

Specific manifestation of HD	**Tasks**	**Biomarker**	**Definition**
Monoloudness	MO, RE	relSEOSD	Speech loudness variation, defined as the standard deviation of the intensity contour relative to its mean after removing silences exceeding 50 ms.
Unstable mean loudness	MO, RE	EEVOL	Energy evolution, defined as the slope of intensity.
Monopitch	MO, RE	relF0SD	Pitch variation, defined as the standard deviation of the F0 contour relative to its mean.
Inappropriate silences	RE	SPIR	Number of pauses (longer than 50 ms) relative to total speech time.
Higher proportion of silence time	RE	PPR	Percentual pause ratio, defined as the total duration of silences (longer than 50 ms) divided by the total duration of speech.
Longer duration of silences	RE	DurMED	Median duration of silences longer than 50 ms.
Higher variability of silence duration	RE	DurMAD	Median absolute deviation of silence duration (longer than 50 ms).
Unnatural speech rate	RE	AR	Number of speech sounds produced per second after removing pauses longer than 50 ms.

[1] MO – monologue; RE – reading.

2 Examples

2.1 Research of a new treatment approach for HD based on rTMS

In the first example, we demonstrate how digital speech biomarkers could be applied in researching a new treatment method for HD. Although Levodopa is the most commonly used pharmacological treatment for PD, its effect on HD varies significantly among individuals and may depend on the progression of PD or the specific speech dimensions in which HD is most prominent [8], [20], [30]. Another treatment approach for PD is DBS; however, its parameters (such as the precise intracranial electrode targets) are primarily tailored based on DBS effects on limb control rather than on HD. Consequently, DBS may even have a negative impact on speech [8], [31].

Based on the findings above, we collaborated with neuroscientists from the Central European Institute of Technology, St. Anne's University Hospital, and Masaryk University to explore and propose a new treatment approach using rTMS [32]. This non-invasive technique employs magnetic fields to modulate neuronal excitability in specific and interconnected brain regions. Although rTMS has shown promise for various PD symptoms [33], there are still significant knowledge gaps regarding its effects on HD. Thus, the aim of this study was to investigate whether rTMS could positively impact HD and to determine the optimal rTMS settings for the most beneficial effect.

2.1.1 Participants and methods

We enrolled 16 patients with clinically confirmed PD, all presenting mild to moderate HD, as evaluated perceptively via the 3F Test [34]. Demographic and clinical details are provided in Table 2.1. None of the participants had a history or current symptoms of hallucinations, psychosis, depression, or dementia. Each participant underwent MRI scanning before and immediately following each rTMS session, with speech recorded inside the scanner using a previously described fMRI protocol [35]. All participants were tested in an ON-medication state without dyskinesias and had not received speech therapy during the study. All participants were native Czech speakers who provided informed consent as approved by the local ethics committee.

A cross-over design was employed, with each participant undergoing five sessions of rTMS over three different brain sites: the left orofacial motor area (OFM1), the right superior temporal gyrus (STG), and a control site at the vertex (V). Each participant received both 1 Hz (low frequency) and 10 Hz (high frequency; except for OFM1) rTMS, randomized across sessions, with at least a one-day interval between sessions. The stimulation was applied using a figure-8 coil positioned over the designated areas with frameless stereotaxy 2.1. The 1 Hz

Tab. 2.1: Demographic and clinical information (except for gender, values are reported as mean ± std.

Clinical/demographic information	Value
Gender (Female/Male)	5/11
Age [years]	67.21 ± 6.18
Duration of PD [years]	6.81 ± 5.00
LED [mg]	758.25 ± 489
UPDRS III	18.6 ± 7.33
ACE–R	91.37 ± 4.68
Beck Depression Inventory II	7.68 ± 3.58
3F Test Score	67.05 ± 8.87

stimulation consisted of 1800 pulses per session, and the 10 Hz stimulation involved 2250 pulses. Each stimulation session included an fMRI scan pre- and post-stimulation.

A 3T MR scanner was used for functional and anatomical imaging. T1-weighted high-resolution images facilitated navigation, while resting-state and task-related BOLD scans were used to examine connectivity and activation changes due to rTMS. Preprocessing included realignment, normalization, and spatial smoothing of the functional data. Analysis focused on detecting changes in connectivity and task-related activation in specific brain regions.

Each participant completed a distinct reading task within the MR scanner both before and immediately after each rTMS session. Speech data acquisition spanned 15 minutes and involved overtly reading short, emotionally neutral sentences or viewing a string of "X" as a baseline condition. The task comprised 48 sentence-reading trials and 24 baseline trials, presented in a pseudo-random order. Each stimulus was displayed for 5 seconds, with an 11-second black screen interval between stimuli.

Due to the MR scanner's noise affecting recording conditions, we focused on digital speech biomarkers that partly characterize speech prosody and articulation. Specifically, prosodic parameters included measurements of monopitch (relF0SD) and inappropriate silences (SPIR and TPT – total pause time). Articulation was quantified using formants (relF1SD, relF2SD), resonances of the oro-naso-pharyngeal tract influenced mainly by simultaneous tongue and jaw movements. The front-back (horizontal) tongue gesture affects the second formant, while the open-close gesture, primarily controlled by the jaw, impacts the first formant.

Effects of each stimulation condition on the relative change in digital speech biomarkers were analyzed using linear mixed models or nonparametric Friedman tests. Paired t-tests or Wilcoxon signed-rank tests were applied to compare these parameters before and after each stimulation condition.

More details regarding the methodology can be found in our original article [32].

Fig. 2.1: Frameless rTMS performed at the Central European Institute of Technology.

2.1.2 Results and discussion

The acoustic analysis revealed that 1 Hz rTMS applied to the STG led to a significant improvement in articulation. This improvement was specifically marked by an increase in relF2SD, a key digital speech biomarker reflecting movements of the tongue and jaw. The observed changes were significantly greater than those achieved by stimulating either the control site (V) or OFM1 with 10 Hz rTMS. Following this low-frequency STG stimulation, a clinical speech pathologist reported improved speech intelligibility in ten patients, while no change was perceived in six patients. Secondary analysis indicated that 1 Hz stimulation of the STG increased TPT for pauses longer than 50 ms, suggesting potential effects on speech fluency as well as articulation. High-frequency stimulation of the STG also led to a modest increase in the range of the first formant, though this effect was less prominent.

In the fMRI analysis, task-related changes following rTMS were evident in the functional connectivity of the STG. Specifically, 1 Hz stimulation of the STG increased functional connectivity between the STG and the right parahippocampal gyrus (PHG), a region implicated in auditory-motor feedback mechanisms. This connectivity change was positively correlated with the observed improvements in articulation, suggesting that rTMS-induced connectivity enhancements within the auditory feedback pathway could contribute to the observed gains in speech articulation. In addition, the high-frequency stimulation of the STG also increased resting-state functional connectivity between the STG and the right inferior parietal lobule (IPL).

This study was the first to demonstrate that low-frequency rTMS targeting the auditory

feedback area (specifically, the right STG) can induce significant acute effects on speech articulation in PD patients with HD. These findings suggest that modulation of the STG through rTMS can improve motor-speech control by enhancing the function of the brain's auditory feedback network. The STG plays a critical role in encoding complex auditory information during vocalization [36], and low-frequency stimulation may specifically strengthen its involvement in feedback control, thereby improving articulation in speech. In clinical terms, the study suggests that low-frequency rTMS over the right STG could be a promising non-invasive intervention for addressing HD in PD.

However, the study examined only the immediate effects of rTMS, which led us to explore its longitudinal effects as well. In the following two studies, we demonstrated that 10 sessions of 1 Hz rTMS over the STG had long-lasting effects, primarily in the field of phonetics [21], [37], thereby enhancing the therapeutic potential of this type of stimulation.

2.2 Prediction of cognitive decline in PD patients

Identifying PD patients who are at increased risk of developing dementia (PD–D) is essential for effective patient care management and for conducting clinical trials aimed at prevention. Major risk factors for PD–D include increased age, more advanced parkinsonism with features such as postural instability and gait challenges, and mild cognitive impairment (MCI) [38]. MCI affects around 40 % of PD patients, marked by both subjective and objective declines in cognitive function while maintaining typical social interactions and daily activities [39].

Based on these facts, we teamed up with neuroscientists, psychologists, and speech-language pathologists from the Central European Institute of Technology, St. Anne's University Hospital, University Hospital Brno, Masaryk University, and University Hospital Ostrava with the aim to explore whether digital speech biomarkers alone, or in combination with certain clinical scores, could predict cognitive decline in PD patients [40].

2.2.1 Participants and methods

In this longitudinal study, we enrolled 44 non-depressed PD patients, each examined twice with an approximate two-year interval. These patients, diagnosed with mild to moderate PD and free of other central nervous system disorders, were assessed using various scales, including UPDRS III, BDI, RBDSQ, NMSS30 (Non-Motor Symptoms Scale), and FOG (Freezing of Gait Questionnaire). Cognitive performance was evaluated with the ACE–R, which classified participants based on cognitive status: normal cognition (PD–NC), mild cognitive impairment (PD–MCI), and PD dementia (PD–D). These classifications were tracked over time to assess changes in cognitive status. All participants were native Czech speakers tested in an ON-medication state, and all provided informed consent, as approved by the local ethics committee.

The speech protocol was adapted from the 3F Test [34] and comprised five tasks focused on assessing prosody. These tasks included reading 135 words at a comfortable pitch and volume, producing interrogative, imperative, and declarative sentences, and reciting a poem with two

rhymes. Speech was recorded using a large capsule cardioid microphone. The analysis targeted 13 features, including the relative standard deviation (relF0SD) and relative variation range (relF0VR) of fundamental frequency, which reflect reduced melody variations (monopitch). To assess monoloudness, features such as the relative standard deviation (relSEOSD) of squared energy were extracted. Measures such as TST, AR, and SPIR were used to evaluate speech rate and pausing. In total, 65 parameters were derived from combinations of speech features and tasks. A correlation analysis was then conducted to examine relationships between these parameters and changes in ACE-R scores. Ultimately, 10 parameters showing significant Pearson correlation coefficients were selected for further analysis of their predictive power for changes in ACE-R scores or cognitive status.

For statistical analyses, data normality was verified using the Shapiro-Wilk test, and differences in speech and clinical characteristics between patients with and without cognitive decline were assessed using two-sample t-tests or Fisher's exact tests. Linear regression models were applied to analyze continuous changes in ACE-R scores, while logistic regression assessed categorical changes in cognitive status. Univariate models were initially developed to explore the independent effects of variables on ACE-R changes and cognitive status decline. Significant variables were then included in multivariate models, incorporating relevant clinical and demographic covariates. The best linear and logistic regression models were selected based on R-square and Nagelkerke R-square values, respectively.

More details regarding the methodology can be found in our original article [40].

2.2.2 Results and discussion

The study's results revealed that over a two-year follow-up, 25 % of PD patients showed cognitive decline, evidenced by an average reduction of 3.8 points in their ACE–R scores. Patients with cognitive deterioration also had longer disease duration, higher baseline scores on the RBDSQ, and lower SPIR values. These baseline characteristics significantly distinguished patients who experienced cognitive decline from those who did not. Multivariate analysis identified a model combining pitch variation and RBDSQ score, which explained 37.2 % of the variability in ACE–R score changes, suggesting that these factors could serve as promising predictors of cognitive worsening in PD.

In the logistic regression analysis, SPIR emerged as a predictor of cognitive status changes, with a baseline predictive accuracy of 73.2 %, rising to 80.5 % when combined with RBDSQ scores. SPIR, which measures the number of pauses relative to total speech time, highlighted a temporal-based impairment correlated with cognitive decline but independent of other clinical and demographic factors. This finding is significant because SPIR captures timing-related motor deficits that do not respond to dopaminergic treatments, similar to motor symptoms like freezing of gait, which also lack responsiveness to these treatments [41].

Research indicates that patients with idiopathic RBD (Rapid Eye Movement Sleep Behavior Disorder) are at a significantly higher risk for developing neurodegenerative synucleinopathies [42]. RBD is characterized by a lack of typical muscle atonia during REM sleep,

leading to motor activity aligned with dream content. Among PD patients, those with RBD exhibit a notably higher frequency of MCI and dementia compared to those without RBD [43]. Our findings showed that combining RBDSQ scores with either pitch variation or SPIR yielded high predictive accuracy for overall cognitive decline and changes in cognitive status, respectively. Our study supports the significant relationship between speech disorders and idiopathic RBD [44]. Moreover, it has been shown that automatic assessment of speech disorders in RBD patients can support the early diagnosis of PD [45].

In clinical terms, the study suggests that a simple acoustic analysis of speech could predict cognitive decline in PD patients, which is essential for enabling more personalized care and helping patients manage daily life more effectively. Early identification allows for timely interventions, such as cognitive therapies and support systems, which can delay or lessen cognitive symptoms and help maintain independence longer. It also guides clinicians in optimizing treatment plans and aids researchers in developing therapies targeted at preventing cognitive decline, all of which contribute to a higher quality of life for patients. Additionally, in a follow-up study, we demonstrated that mHealth technology (e.g., a smartphone) can be potentially used to assess cognitive decline in PD patients, allowing for easy and effective cognitive screening outside a clinical setting [17].

3 Conclusion

This thesis presents the concept of assessing and monitoring PD using digital speech biomarkers that quantify specific manifestations of HD, a motor speech disorder prevalent in this neurodegenerative disease. The work demonstrates two applications of these biomarkers. First, it illustrates how digital speech biomarkers were employed in designing a novel treatment approach using rTMS, which positively impacts articulation and enhances brain auditory feedback pathways. Second, the study explores the predictive capability of digital speech biomarkers for cognitive decline in PD patients, suggesting that speech-based assessments can aid in the early detection of cognitive risks.

Over the last decade, our team has made significant contributions to the field of rTMS, successfully advancing research projects, bridging knowledge gaps, and pushing beyond the current state of the art [21], [37], [46]. Research into non-invasive stimulation for speech in PD patients continues, with current efforts focused on transcranial direct current stimulation (tDCS), a more affordable method that patients can potentially administer themselves. For example, in a recent study, we demonstrated that anodal tDCS could influence the temporal characteristics of speech [16]. We are now investigating its effects in a longitudinal study.

Regarding cognitive decline, we have shown that digital speech biomarkers can effectively support the prediction of this non-motor feature of PD. However, cognitive decline is not the only aspect that can be predicted based on acoustic speech analysis. In another study, we focused on gait alterations. HD and freezing of gait (FOG) are both axial motor symptoms frequently observed in PD patients. FOG is characterized by sudden, brief episodes in which patients feel unable to move their feet, especially during transitions, turning, or in confined spaces. Both HD and FOG arise from motor circuit dysfunctions in the basal ganglia and are less responsive to dopaminergic treatments than other PD symptoms, such as tremors and rigidity. This limited responsiveness suggests that non-dopaminergic pathways also contribute to the pathology of HD and FOG. In [41], we found that FOG in PD patients is primarily associated with improper articulation, disturbed speech rate, and reduced intelligibility. Additionally, we demonstrated that baseline acoustic analysis of HD can serve as a predictor for the development of FOG deficits over a two-year follow-up period.

In addition to the above-mentioned, we have also demonstrated that digital speech biomarkers can be effectively used to understand phonation and articulation in PD [22], [47]–[50], investigate compensatory mechanisms [51], passively assess PD [17], explore the effects of pharmacological treatment [52], [53], enable multilingual diagnosis of PD [14], and examine speech subtypes in PD [20]. These studies underscore the potential of digital speech biomarkers as non-invasive, objective tools to support clinicians in diagnosing PD, managing its symptoms, monitoring progression, and tailoring patient care.

3.1 Future directions

Over the last decade, there has been a growing belief that PD could be diagnosed solely through the acoustic analysis of the sustained vowel sound [a] [23], [26]. Numerous studies were published supporting this approach, suggesting it as a valid diagnostic method. However, a few research teams, including ours, have clearly demonstrated that this approach is odd [54]. HD in PD can result in distinct speech subtypes, with the phonatory subtype occurring in only about 30 % of patients [20], [30]. Hopefully, this insight will be quickly accepted by the community, encouraging researchers to view HD as a multidimensional disorder that requires a comprehensive set of digital speech biomarkers for accurate assessment.

Neurologists are still seeking methods to facilitate PD screening. Digital speech biomarkers, as an easy-to-use technology, serve as strong candidates for this purpose. However, while HD is present in up to 90 % of PD patients, it cannot be the sole source of diagnostic information. Consequently, speech analysis is increasingly expected to be combined with other modalities, such as handwriting, sleep, gait, etc. [55] This multimodal approach can further enhance the sensitivity and specificity of PD screening.

Currently, most trustworthy and explainable machine learning models for the diagnosis and assessment of PD rely on shallow methods, such as logistic regression, random forests, and gradient boosting (also due to the limitations of small sample datasets in PD research). These methods provide stable performance and are easier to interpret in clinical settings. However, recent advancements have opened doors for applying deep learning architectures to small sample datasets while maintaining clinical relevance and explainability. Techniques like transfer learning allow models to leverage knowledge from larger datasets in related tasks, while data augmentation strategies expand small datasets through synthetic data generation (on the other hand, the generation of synthetic pathological samples remains highly challenging, and these methods must be applied with caution). Furthermore, explainability techniques, such as layer-wise relevance propagation (LRP) and SHAP, offer insights into model decisions by highlighting the most influential features, enhancing the interpretability of deep neural networks (DNNs) in a clinical context. These advancements are paving the way for deeper, more complex models to be used effectively in PD research, potentially improving diagnostic accuracy and patient assessment.

When diagnosing, assessing, or monitoring PD using digital speech biomarkers, mHealth systems, such as smartphone-based applications, offer significant advantages. However, these systems currently rely on active patient interaction with the device, e.g., via specific speech tasks or vocal exercises, which presents limitations. Patients may experience fatigue, may not consistently follow instructions, or may lack technical literacy, all of which can impact data quality and adherence over time. In the future, passive speech assessment through smartphones or smart assistants offers an alternative, capturing natural speech during daily interactions without requiring active input from patients [45]. This passive approach could improve convenience, adherence, and the frequency of monitoring, providing richer, real-world data for longitudinal analysis. However, passive assessment also poses challenges, including

privacy concerns, the need for sophisticated background noise filtering, and the potential for less controlled data variability. Balancing these factors will be essential in leveraging passive speech assessment for effective PD management.

Advances are being made in using digital speech biomarkers to predict how patients respond to treatments. Predictive models that analyze historical speech data can foresee treatment efficacy, helping clinicians identify the best therapeutic approach while reducing trial-and-error prescriptions. This level of personalization not only improves symptom control but also enhances patients' quality of life by aligning treatments more closely with their specific needs and progression patterns.

Bibliography

[1] D. W. Dickson, "Parkinson's disease and parkinsonism: Neuropathology", *Cold Spring Harbor Perspectives in Medicine*, vol. 2, no. 8, a009258, 2012.

[2] O. Hornykiewicz, "Biochemical aspects of Parkinson's disease", *Neurology*, vol. 51, no. 2, S2–S9, 1998.

[3] K. Wakabayashi, K. Tanji, S. Odagiri, Y. Miki, F. Mori, and H. Takahashi, "The Lewy body in Parkinson's disease and related neurodegenerative disorders", *Molecular Neurobiology*, vol. 47, pp. 495–508, 2013.

[4] P. Brodal, *The central nervous system*. Oxford university Press, 2010.

[5] D. Berg, R. B. Postuma, C. H. Adler, B. R. Bloem, P. Chan, B. Dubois, T. Gasser, C. G. Goetz, G. Halliday, L. Joseph, *et al.*, "MDS research criteria for prodromal Parkinson's disease", *Movement Disorders*, vol. 30, no. 12, pp. 1600–1611, 2015.

[6] S. Heinzel, D. Berg, T. Gasser, H. Chen, C. Yao, R. B. Postuma, and M. T. F. on the Definition of Parkinson's Disease, "Update of the MDS research criteria for prodromal Parkinson's disease", *Movement Disorders*, vol. 34, no. 10, pp. 1464–1470, 2019.

[7] R. B. Postuma, D. Berg, M. Stern, W. Poewe, C. W. Olanow, W. Oertel, J. Obeso, K. Marek, I. Litvan, A. E. Lang, *et al.*, "MDS clinical diagnostic criteria for Parkinson's disease", *Movement Disorders*, vol. 30, no. 12, pp. 1591–1601, 2015.

[8] L. Brabenec, J. Mekyska, Z. Galaz, and I. Rektorova, "Speech disorders in Parkinson's disease: Early diagnostics and effects of medication and brain stimulation", *Journal of Neural Transmission*, vol. 124, pp. 303–334, 2017.

[9] A. J. Espay, F. Morgante, A. Merola, A. Fasano, L. Marsili, S. H. Fox, E. Bezard, B. Picconi, P. Calabresi, and A. E. Lang, "Levodopa-induced dyskinesia in Parkinson disease: Current and evolving concepts", *Annals of Neurology*, vol. 84, no. 6, pp. 797–811, 2018.

[10] A. K. Ho, R. Iansek, C. Marigliani, J. L. Bradshaw, and S. Gates, "Speech impairment in a large sample of patients with Parkinson's disease", *Behavioural Neurology*, vol. 11, no. 3, pp. 131–137, 1999.

[11] J. R. Duffy, *Motor Speech Disorders E-Book: Motor Speech Disorders E-Book*. Elsevier Health Sciences, 2019.

[12] J. Mekyska, Z. Galaz, Z. Mzourek, Z. Smekal, I. Rektorova, I. Eliasova, M. Kostalova, M. Mrackova, D. Berankova, M. Faundez-Zanuy, *et al.*, "Assessing progress of Parkinson's disease using acoustic analysis of phonation", in *2015 4th international work conference on bioinspired intelligence (IWOBI)*, IEEE, 2015, pp. 111–118.

[13] L. Moro-Velazquez and N. Dehak, "A review of the use of prosodic aspects of speech for the automatic detection and assessment of Parkinson's disease", in *Automatic Assessment of Parkinsonian Speech: First Workshop, AAPS 2019, Cambridge, Massachussets, USA, September 20–21, 2019, Revised Selected Papers 1*, Springer, 2020, pp. 42–59.

[14] D. Kovac, J. Mekyska, V. Aharonson, P. Harar, Z. Galaz, S. Rapcsak, J. R. Orozco-Arroyave, L. Brabenec, and I. Rektorova, "Exploring digital speech biomarkers of hypokinetic dysarthria in a multilingual cohort", *Biomedical Signal Processing and Control*, vol. 88, p. 105 667, 2024.

[15] S. Vasudevan, A. Saha, M. E. Tarver, and B. Patel, "Digital biomarkers: Convergence of digital health technologies and biomarkers", *NPJ Digital Medicine*, vol. 5, no. 1, p. 36, 2022.

[16] L. Brabenec, D. Kovac, J. Mekyska, L. Rehulkova, V. Kabrtova, and I. Rektorova, "Short-term effects of transcranial direct current stimulation on motor speech in Parkinson's disease: A pilot study", *Journal of Neural Transmission*, pp. 1–7, 2024.

[17] D. Kovac, J. Mekyska, L. Brabenec, M. Kostalova, and I. Rektorova, "Research on passive assessment of Parkinson's disease utilising speech biomarkers", in *International Conference on Pervasive Computing Technologies for Healthcare*, Springer, 2022, pp. 259–273.

[18] J. Rusz, T. Tykalova, L. O. Ramig, and E. Tripoliti, "Guidelines for speech recording and acoustic analyses in dysarthrias of movement disorders", *Movement Disorders*, vol. 36, no. 4, pp. 803–814, 2021.

[19] J. G. Svec and S. Granqvist, "Tutorial and guidelines on measurement of sound pressure level in voice and speech", *Journal of Speech, Language, and Hearing Research*, vol. 61, no. 3, pp. 441–461, 2018.

[20] D. Kovac, K. Novotny, J. Mekyska, L. Brabenec, M. Kostalova, and I. Rektorova, "Exploring speech profiles in Parkinson's disease patients on long-term dopaminergic medication", in *2024 47th International Conference on Telecommunications and Signal Processing (TSP)*, IEEE, 2024, pp. 270–273.

[21] A. Gomez-Rodellar, J. Mekyska, P. Gomez-Vilda, L. Brabenec, P. Simko, and I. Rektorova, "A pilot study on the functional stability of phonation in EEG bands after repetitive transcranial magnetic stimulation in Parkinson's disease", *International Journal of Neural Systems*, vol. 33, no. 06, p. 2 350 028, 2023.

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

[23] Q. C. Ngo, M. A. Motin, N. D. Pah, P. Drotar, P. Kempster, and D. Kumar, "Computerized analysis of speech and voice for Parkinson's disease: A systematic review", *Computer Methods and Programs in Biomedicine*, vol. 226, p. 107 133, 2022.

[24] V. Illner, P. Kryze, J. Svihlik, M. Sousa, P. Krack, E. Tripoliti, R. Jech, and J. Rusz, "Which aspects of motor speech disorder are captured by Mel frequency cepstral coefficients? evidence from the change in STN-DBS conditions in Parkinson's disease", in *Proceedings of the Annual Conference of the International Speech Communication Association, INTERSPEECH 2023*, ISCA-International Speech Communication Association, 2023.

[25] M. Simek and J. Rusz, "Validation of cepstral peak prominence in assessing early voice changes of Parkinson's disease: Effect of speaking task and ambient noise", *The Journal of the Acoustical Society of America*, vol. 150, no. 6, pp. 4522–4533, 2021.

[26] L. Moro-Velazquez, J. A. Gomez-Garcia, J. D. Arias-Londoño, N. Dehak, and J. I. Godino-Llorente, "Advances in Parkinson's disease detection and assessment using voice and speech: A review of the articulatory and phonatory aspects", *Biomedical Signal Processing and Control*, vol. 66, p. 102 418, 2021.

[27] J. Rusz, R. Cmejla, H. Ruzickova, and E. Ruzicka, "Quantitative acoustic measurements for characterization of speech and voice disorders in early untreated Parkinson's disease", *The journal of the Acoustical Society of America*, vol. 129, no. 1, pp. 350–367, 2011.

[28] S. Boyce, H. J. Fell, and J. MacAuslan, "SpeechMark: Landmark detection tool for speech analysis.", in *INTERSPEECH*, 2012, pp. 1894–1897.

[29] J. Rusz, J. Hlavnicka, R. Cmejla, and E. Ruzicka, "Automatic evaluation of speech rhythm instability and acceleration in dysarthrias associated with basal ganglia dysfunction", *Frontiers in Bioengineering and Biotechnology*, vol. 3, p. 104, 2015.

[30] J. Rusz, T. Tykalova, M. Novotny, D. Zogala, K. Sonka, E. Ruzicka, and P. Dusek, "Defining speech subtypes in de novo Parkinson disease: Response to long-term levodopa therapy", *Neurology*, vol. 97, no. 21, e2124–e2135, 2021.

[31] R. Baudouin, J. R. Lechien, L. Carpentier, J.-M. Gurruchaga, Q. Lisan, and S. Hans, "Deep brain stimulation impact on voice and speech quality in Parkinson's disease: A systematic review", *Otolaryngology–Head and Neck Surgery*, vol. 168, no. 3, pp. 307–318, 2023.

[32] L. Brabenec, P. Klobusiakova, M. Barton, J. Mekyska, Z. Galaz, V. Zvoncak, T. Kiska, J. Mucha, Z. Smekal, M. Kostalova, *et al.*, "Non-invasive stimulation of the auditory feedback area for improved articulation in Parkinson's disease", *Parkinsonism & Related Disorders*, vol. 61, pp. 187–192, 2019.

[33] W. Zhang, B. Deng, F. Xie, H. Zhou, J.-F. Guo, H. Jiang, A. Sim, B. Tang, and Q. Wang, "Efficacy of repetitive transcranial magnetic stimulation in Parkinson's disease: A systematic review and meta-analysis of randomised controlled trials", *EClinicalMedicine*, vol. 52, 2022.

[34] M. Kostalova, M. Mrackova, R. Marecek, D. Berankova, I. Eliasova, E. Janousova, J. Roubickova, J. Bednarik, and I. Rektorova, "The 3F Test Dysarthric profile – normative speach values in Czech", *Czech and Slovak Neurology and Neurosurgery*, vol. 76, no. 109, p. 5, 2013.

[35] I. Rektorova, J. Barrett, M. Mikl, I. Rektor, and T. Paus, "Functional abnormalities in the primary orofacial sensorimotor cortex during speech in Parkinson's disease", *Movement Disorders*, vol. 22, no. 14, pp. 2043–2051, 2007.

[36] F. H. Guenther and G. Hickok, "Role of the auditory system in speech production", *Handbook of clinical neurology*, vol. 129, pp. 161–175, 2015.

[37] L. Brabenec, P. Klobusiakova, P. Simko, M. Kostalova, J. Mekyska, and I. Rektorova, "Non-invasive brain stimulation for speech in Parkinson's disease: A randomized controlled trial", *Brain Stimulation*, vol. 14, no. 3, pp. 571–578, 2021.

[38] M. Emre, D. Aarsland, R. Brown, D. J. Burn, C. Duyckaerts, Y. Mizuno, G. A. Broe, J. Cummings, D. W. Dickson, S. Gauthier, *et al.*, "Clinical diagnostic criteria for dementia associated with Parkinson's disease", *Movement Disorders*, vol. 22, no. 12, pp. 1689–1707, 2007.

[39] C. Baiano, P. Barone, L. Trojano, and G. Santangelo, "Prevalence and clinical aspects of mild cognitive impairment in Parkinson's disease: A meta-analysis", *Movement Disorders*, vol. 35, no. 1, pp. 45–54, 2020.

[40] I. Rektorova, J. Mekyska, E. Janousova, M. Kostalova, I. Eliasova, M. Mrackova, D. Berankova, T. Necasova, Z. Smekal, and R. Marecek, "Speech prosody impairment predicts cognitive decline in Parkinson's disease", *Parkinsonism & Related Disorders*, vol. 29, pp. 90–95, 2016.

[41] J. Mekyska, Z. Galaz, T. Kiska, V. Zvoncak, J. Mucha, Z. Smekal, I. Eliasova, M. Kostalova, M. Mrackova, D. Fiedorova, *et al.*, "Quantitative analysis of relationship between hypokinetic dysarthria and the freezing of gait in Parkinson's disease", *Cognitive Computation*, vol. 10, pp. 1006–1018, 2018.

[42] M. G. Miglis, C. H. Adler, E. Antelmi, D. Arnaldi, L. Baldelli, B. F. Boeve, M. Cesari, I. Dall'Antonia, N. J. Diederich, K. Doppler, *et al.*, "Biomarkers of conversion to α-synucleinopathy in isolated rapid-eye-movement sleep behaviour disorder", *The Lancet Neurology*, vol. 20, no. 8, pp. 671–684, 2021.

[43] M. Vendette, J.-F. Gagnon, A. Decary, J. Massicotte-Marquez, R. Postuma, J. Doyon, M. Panisset, and J. Montplaisir, "REM sleep behavior disorder predicts cognitive impairment in Parkinson disease without dementia", *Neurology*, vol. 69, no. 19, pp. 1843–1849, 2007.

[44] J. Rusz, J. Hlavnicka, M. Novotny, T. Tykalova, A. Pelletier, J. Montplaisir, J.-F. Gagnon, P. Dusek, A. Galbiati, S. Marelli, *et al.*, "Speech biomarkers in rapid eye movement sleep behavior disorder and Parkinson disease", *Annals of Neurology*, vol. 90, no. 1, pp. 62–75, 2021.

[45] V. Illner, M. Novotny, T. Kouba, T. Tykalova, M. Simek, P. Sovka, J. Svihlik, E. Ruzicka, K. Sonka, P. Dusek, *et al.*, "Smartphone voice calls provide early biomarkers of Parkinson in rapid eye movement sleep behavior disorder", *Movement Disorders*, 2024.

[46] I. Eliasova, J. Mekyska, M. Kostalova, R. Marecek, Z. Smekal, and I. Rektorova, "Acoustic evaluation of short-term effects of repetitive transcranial magnetic stimulation on motor aspects of speech in Parkinson's disease", *Journal of Neural Transmission*, vol. 120, pp. 597–605, 2013.

[47] P. Gomez-Vilda, A. Gomez-Rodellar, J. M. F. Vicente, J. Mekyska, D. Palacios-Alonso, V. Rodellar-Biarge, A. Alvarez-Marquina, I. Eliasova, M. Kostalova, and I. Rektorova, "Neuromechanical modelling of articulatory movements from surface electromyography and speech formants", *International Journal of Neural Systems*, vol. 29, no. 02, p. 1 850 039, 2019.

[48] P. Gomez, J. Mekyska, A. Gomez, D. Palacios, V. Rodellar, and A. Alvarez, "Characterization of Parkinson's disease dysarthria in terms of speech articulation kinematics", *Biomedical Signal Processing and Control*, vol. 52, pp. 312–320, 2019.

[49] A. Gomez-Rodellar, D. Palacios-Alonso, J. M. Ferrandez Vicente, J. Mekyska, A. Alvarez-Marquina, and P. Gomez-Vilda, "A methodology to differentiate Parkinson's disease and aging speech based on glottal flow acoustic analysis", *International Journal of Neural Systems*, vol. 30, no. 10, p. 2 050 058, 2020.

[50] P. Gomez-Vilda, A. Gomez-Rodellar, J. Mekyska, A. Alvarez-Marquina, and D. Palacios-Alonso, "Unfolding laryngeal neuromotor activity in Parkinson's disease by phonation inversion", in *International Work-Conference on the Interplay Between Natural and Artificial Computation*, Springer, 2024, pp. 300–309.

[51] P. Klobusiakova, J. Mekyska, L. Brabenec, Z. Galaz, V. Zvoncak, J. Mucha, S. Z. Rapcsak, and I. Rektorova, "Articulatory network reorganization in Parkinson's disease as assessed by multimodal MRI and acoustic measures", *Parkinsonism & Related Disorders*, vol. 84, pp. 122–128, 2021.

[52] N. Elfmarkova, M. Gajdos, M. Mrackova, J. Mekyska, M. Mikl, and I. Rektorova, "Impact of Parkinson's disease and levodopa on resting state functional connectivity related to speech prosody control", *Parkinsonism & Related Disorders*, vol. 22, S52–S55, 2016.

[53] M. Mravckova, R. Marecek, J. Mekyska, M. Kostalova, and I. Rektorova, "Levodopa may modulate specific speech impairment in Parkinson's disease: An fMRI study", *Journal of Neural Transmission*, vol. 131, no. 2, pp. 181–187, 2024.

[54] A. Favaro, T. Cao, T. Thebaud, J. Villalba, A. Butala, N. Dehak, and L. Moro-Velázquez, "Do phonatory features display robustness to characterize Parkinsonian speech across corpora?", in *Proceedings of the Annual Conference of the International Speech Communication Association, INTERSPEECH*, vol. 2023, 2023, pp. 2388–2392.

[55] J. C. Vasquez-Correa, T. Arias-Vergara, J. R. Orozco-Arroyave, B. Eskofier, J. Klucken, and E. Noth, "Multimodal assessment of Parkinson's disease: A deep learning approach", *IEEE Journal of Biomedical and Health Informatics*, vol. 23, no. 4, pp. 1618–1630, 2018.

ABSTRACT

Parkinson's disease (PD) is the second most common neurodegenerative disorder, characterized by a range of motor and non-motor symptoms. Up to 90 % of individuals with PD develop hypokinetic dysarthria (HD), a motor speech disorder that affects respiration, phonation, resonance, articulation, and prosody. Digital speech biomarkers (acoustic features related to underlying physiological processes) offer an objective means of quantifying specific manifestations of HD. These biomarkers can support neurologists, psychologists, and speech-language pathologists in the assessment and monitoring of PD. This work presents the concept of acoustic speech analysis in PD patients and highlights its benefits through two key examples: 1) the research of a new treatment approach for HD based on repetitive transcranial magnetic stimulation, and 2) the prediction of cognitive decline in PD patients. The work concludes with additional applications of digital speech biomarkers in PD and outlines future research directions in this evolving field.

ABSTRAKT

Parkinsonova nemoc (PN) je druhé nejčastější neurodegenerativní onemocnění, které se vyznačuje řadou motorických i nemotorických příznaků. Až u 90 % jedinců s PN se rozvine hypokinetická dysartrie (HD), porucha motorické realizace řeči, která se projevuje v oblasti respirace, fonace, rezonance, artikulace a prozodie. Digitální řečové biomarkery (akustické metriky související s fyziologickými procesy) nabízejí objektivní nástroj kvantifikace specifických projevů HD. Tyto biomarkery mohou podpořit neurology, psychology a klinické logopedy při hodnocení a monitorování PN. Tato práce představuje koncept akustické analýzy řeči u pacientů s PN a zdůrazňuje její přínosy prostřednictvím dvou klíčových příkladů: 1) výzkum nového přístupu léčby HD založeného na repetitivní transkraniální magnetické stimulaci a 2) predikce kognitivního deficitu u pacientů s PN. V závěru práce jsou zmíněny další aplikace digitálních řečových biomarkerů u PN a jsou nastíněny budoucí směry výzkumu v této oblasti.