Relationship between LTAS-based spectral moments and acoustic parameters of hypokinetic dysarthria in Parkinson’s disease

Jan Svihlík\(^1,2\), Vojtech Illner\(^1\), Petr Kryza\(^1\), Mario Sousa\(^3\), Paul Krack\(^3\), Elina Tripoliti\(^1\), Robert Jech\(^5\), Jan Ruz\(z\)\(^1\)

\(^1\)Czech Technical University in Prague, Czech Republic
\(^2\)University of Chemistry and Technology, Prague, Czech Republic
\(^3\)University Hospital of Bern, Switzerland
\(^4\)University College London Hospitals NHS Foundation Trust, United Kingdom
\(^5\)General University Hospital, Prague, Czech Republic

svihlj1@fel.cvut.cz, vojtech.illner@fel.cvut.cz, kryzepet@fel.cvut.cz, mario.sousa@insel.ch, paul.krack@insel.ch, e.tripoliti@ucl.ac.uk, jech@cesnet.cz, rusz mz@gmail.com

Abstract

Although long-term averaged spectrum (LTAS) descriptors can detect the change in dysarthria of patients with Parkinson’s disease (PD) due to subthalamic nucleus deep brain stimulation (STN-DBS), the relationship between LTAS variables with measures that relate to laryngeal physiology remain unknown. We aimed to find connections between LTAS-based moments and the main acoustic characteristics of hypokinetic dysarthria in PD as the response to STN-DBS stimulation changes. We analyzed reading passages of 23 PD patients in ON and OFF STN-DBS states compared to 23 healthy controls. We found a relation between the stimulation-induced change in several spectral moments and acoustic parameters representing voice quality, articulatory decay, articulation rate, and mean fundamental frequency. While the difference between PD and controls was significant across most acoustic descriptors, only the spectral mean and fundamental frequency variability could differentiate between ON and OFF conditions.

Index Terms: Parkinson’s disease, STN-DBS, dysarthria, speech disorder, long-term averaged spectrum, spectral moments, acoustic parameters

1. Introduction

Parkinson’s disease (PD) is a neurodegenerative disorder characterized by the progressive loss of dopaminergic neurons in the substantia nigra. It affects approximately 1% to 2% of individuals aged 60 years and above [1]. The primary motor symptoms of PD encompass rest tremor, bradykinesia, rigidity, and postural instability. While there are ongoing efforts to develop neuroprotective treatments, currently, there is no available therapy that can halt or slow down the progression of the disease. Pharmacotherapy aims to mitigate the symptomatic motor manifestations of PD, with levodopa (L-DOPA) being the most potent medication. In cases where long-term motor complications arise from L-DOPA therapy, deep brain stimulation of the subthalamic nucleus (STN-DBS), a neurosurgical approach, is employed. [2]. It is considered the most significant advance in the treatment of PD since the introduction of L-DOPA. [3].

According to Duffy (2013), speech, being the most complex motor skill, is highly vulnerable to negative changes in the neural structures that control motor abilities [4]. Hypokinetic dysarthria, a collective term for speech disorders, is observed in up to 90% of patients with PD [5]. The main characteristics of hypokinetic dysarthria include monopitch, reduced stress, imprecise consonant articulation, inappropriate silences, harsh voice, and overall deficits in speech timing [6]. The impact of STN-DBS on speech remains uncertain. While certain aspects of speech may improve with STN-DBS, the most common side effect is stimulation-induced dysarthria [7]. In fact, the deterioration of speech can offset the motor benefits of STN-DBS in terms of overall improvement in quality of life. Previous research indicates that 78% of STN-DBS patients experienced impaired speech intelligibility 1 year after the DBS implantation [8], with a greater decline observed in DBS-implanted patients compared to the control group of medically treated patients (-16.9% vs. -4.5% reduction in intelligibility in the medication ON state compared to baseline).

One potential method to reliably evaluate the severity of dysarthria is through the utilization of Long-term Averaged Spectrum (LTAS) and its spectral moments. This approach involves employing the discrete Fourier transform on segmented parts of a speech signal, followed by averaging the resulting spectra. By examining four key measures - spectral mean, standard deviation, skewness, and kurtosis - important insights can be gained regarding the concentration, variation, tilt, and peakedness of the LTAS energy distribution [9]. Individuals with PD typically exhibit lower spectral mean and standard deviation values, while displaying higher spectral skewness and kurtosis compared to control subjects [10], [11]. Previous research has identified correlations between perceived dysarthria severity and the spectral mean and kurtosis [11], suggesting that LTAS moments may serve as an acoustic proxy for assessing dysarthria severity. Notably, a recent study has demonstrated the ability of LTAS moments to detect changes between the ON and OFF states of STN-DBS [12]. This highlights the potential of LTAS as a surrogate descriptor for monitoring dysarthria severity induced by stimulation.

From a physiological standpoint, the spectral mean is hypothesized to reflect changes in the fundamental frequency of the voice [13]. Spectral skewness, on the other hand, is believed to be linked to glottal closure during phonation, where reduced skewness indicates hyperadduction and increased skewness indicates hypoadduction [14]. However, a direct relationship between LTAS and pathophysiology of speech production is still poorly understood. Therefore, this paper investigates correlations between response on DBS stimulation change of LTAS-based moments and main characteristics of hypokinetic dysarthria in PD captured by acoustic analysis. An additional aim was to investigate whether representative acoustic param-
eters of hypokinetic dysarthria are used to distinguish between DBS ON, DBS OFF states, and HC.

2. Method

2.1. Participants

A total of 23 individuals with PD participated in the study, including four females. The mean age of the PD group was 61.7 years (SD = 5.0, range: 53–72). These individuals had received bilateral STN-DBS in combination with dopaminergic medication. The assessment of the PD group was conducted under two conditions: the STN-DBS switched OFF condition (referred to as DBS OFF) and the STN-DBS switched ON condition (referred to as DBS ON). For more specific details regarding the participants’ characteristics and experimental conditions, please refer to [12].

A control group of 23 healthy controls (HC) was also included in the study. The HC group consisted of individuals matched in terms of age and sex, with 4 female participants. The mean age of the HC group was 61.5 years (SD = 5.6, range: 52–72). None of the control group participants had a history of neurological or communication disorders. All participants were native Czech speakers. The study was conducted in accordance with the principles of the Helsinki Declaration and was approved by the Ethics Committee of the General University Hospital in Prague, Czech Republic. Each participant provided written informed consent.

2.2. Speech Examination

The audio recordings took place in a quiet room with minimal background noise. A head-mounted condenser microphone (specifically, the BeyerDynamic Opus 55) was used, positioned at a distance of approximately 5 cm from the subject’s mouth [15]. The speech data were sampled at a frequency of 48 kHz with a resolution of 16 bits. Each participant underwent a single recording session, which was conducted by a speech specialist. During the recording session, the participants were instructed to perform phonetically balanced reading passage tasks. These tasks involved reading a standardized text consisting of 313 words. The text was carefully chosen to include familiar vocabulary and grammatical structures that are relevant and up-to-date.

2.3. Long-Term Average Spectrum

To process the speech signal, pause intervals (>30 ms) were identified and eliminated using an automatic segmentation tool designed for connected speech [16]. Only the segments containing speech were retained for further analysis, including the estimation of the LTAS and its derived moments. The speech signal was then divided into sections of 50 milliseconds in length with a 50% overlap [17]. The Hamming window, as described in [18], was applied to each section. Power spectra were computed using the fast Fourier transform for every section, and these spectra were then averaged to obtain the LTAS estimation. This specific technique is referred to as Welch’s method [19]. For each patient with PD, the spectral moments (spectral mean and standard deviation) were calculated, following the approach described in [20]. Additionally, derived measures such as skewness and kurtosis were computed. These calculations were performed for both the DBS ON and DBS OFF stimulation states of the PD patients, as well as for the HC subjects.

The first spectral moment (spectral mean) is given by equation 1:

\[ M_1^s = \frac{1}{W} \sum_{i=1}^{N} F(i) \text{PSD}(i), \]

where \( W = \sum \text{PSD}(i) \), \( F \) represents a frequency vector and \( \text{PSD} \) denotes a power spectral density. The second spectral moment (spectral standard deviation) \( SD^s \) is given by equation 2:

\[ SD^s = \sqrt{\frac{1}{W} \sum_{i=1}^{N} F^2(i) \text{PSD}(i) - (M_1^s)^2}. \]

The third spectral moment represented by spectral skewness \( M_3^s \) is given by equation 3:

\[ M_3^s = \frac{1}{W (SD^s)^3} \sum_{i=1}^{N} (F(i) - M_1^s)^3 \text{PSD}(i). \]

The spectral skewness is a metric used to assess the asymmetry of the frequency distribution of energy. A negative skewness value indicates that the left tail of the energy distribution is longer, and vice versa. The fourth spectral moment based on the spectral kurtosis \( M_4^s \) is given by equation 4:

\[ M_4^s = \frac{1}{W (SD^s)^4} \sum_{i=1}^{N} (F(i) - M_1^s)^4 \text{PSD}(i). \]

Spectral kurtosis is a measure that quantifies the presence of outliers in the energy distribution. A higher kurtosis value indicates a distribution with heavy tails.

2.4. Acoustic parameters of hypokinetic dysarthria

We selected the following 7 acoustic parameters that allow quantitative acoustic assessment that corresponds to the perceptual description of hypokinetic dysarthria by the landmark study of Darley et al [21]. The standard deviation of power (stdPWR), defined as the standard deviation of speech intensity contour extracted from voiced segments, captures the abnormal variation of loudness reflecting typically poor respiratory-phonatory coordination and control. The standard deviation of fundamental frequency (stdF0), defined as the standard deviation of fundamental frequency contour converted to semitone scale, a perceptual feature of decreased voice melody variation (monopitch), is associated with the hypooadduction of the vocal folds. The mean value of fundamental frequency (F0) was also used as a descriptor due to relationships between LTAS and mean F0 change found in previous study [13]. Cepstral peak prominence (CPP), defined as the measure of cepstral peak amplitude normalized for overall amplitude, is an acoustic measure of voice quality that is related to hoarseness and breathiness. Resonant frequency attenuation (RFA), defined as the differences between the maxima of the second formant region and minima of the local valley region called antiformant, represents decreased spectral energy as a result of decayed articulatory movements [22]. Since difficulties in initiating speech have a considerable effect on pause duration, Duration of pause intervals (DPI), defined as the median length of pause intervals, was used to reflect hypokinesis of the movements involved in initiating speech and pause production. Net speech rate (NSR), defined as the total number of words divided by the total duration of speech after the removal of pauses, is a standard measure of dysfunctional speech.
frequently manifesting as slow speech, not only due to the slowness of individual movements but also as a compensatory mechanism for increasing the intelligibility of speech. The comprehensive algorithmic description of acoustic measurements used can be found in a previous study [23].

2.5. Statistical analysis
Since the spectral moments and the acoustic parameters were not normally distributed (Kolmogorov–Smirnov test), correlations between DBS ON and DBS OFF (Δ = OFF - ON) state differences of LTAS spectral moments and ON and OFF state differences of acoustic parameters were computed using the Spearman approach. To evaluate the statistical differences between the groups under examination (HC, DBS ON, and DBS OFF), a nonparametric Friedman test was employed. Following the Friedman test, post hoc analysis was conducted using the Fisher least significant difference adjustment method. The significance level was set at \( p < 0.05 \).

3. Results
3.1. LTAS correlation with acoustic parameters
The statistically significant correlation between acoustic parameters and LTAS moments can be seen in Figure 1.

3.2. Group differences
We found statistically significant differences for the DPI, RFA, stdF0, and between the HC and both DBS ON, and DBS OFF groups, while differences between DBS OFF and DBS ON were found only for the stdF0 (Figure 2). The statistically significant differences for the NSR and stdPWR were found only between HC and DBS OFF group. In addition, we found statistically significant differences in all the spectral features between the HC and both DBS ON, and DBS OFF groups, whereas the differences between DBS OFF and DBS ON were found only for 1st spectral moment (Figure 3).

4. Discussion
In this study, we examined the relationship between LTAS variables with acoustic measures that relate to main aspects of parkinsonian dysarthria pathophysiology. The change in DBS conditions highlighted mutual changes in LTAS variables and voice quality, articulatory decay, articulation rate, and mean fundamental frequency. Due to the effect of STN-DBS, first and second LTAS coefficients were increased while third and fourth were decreased from OFF to ON states.

The increase in mean F0 was reflected by an increase in the third and fourth LTAS coefficients. Although we did not find a direct correlation between mean F0 and the first or second coefficient, the reversed direction of change between the first and second two LTAS coefficients can be expected due to DBS, suggesting that mean F0 decline has also an impact on lowering spectral mean. This would be in good agreement with previous studies suggesting that the spectral mean probably reflects the changes in the fundamental frequency of the voice [13].

The highest correlations were observed between an increase in LTAS mean (as well as a decrease in skewness and kurtosis) and an increase of voice quality via CPP from OFF to ON states. This finding is consistent with a previous study [11], which observed differences in LTAS between groups of speakers with normal and disordered speech, despite similar sentence

![Figure 1: Results of correlation analysis between acoustic parameters and LTAS moments for differences (Δ = OFF - ON). Correlation between mean F0 ΔF0 and (a) spectral skewness ΔM2 and (b) spectral kurtosis ΔM4. Correlation between resonant frequency attenuation ΔRFA and (c) spectral standard deviation ΔSD and (d) spectral kurtosis ΔM4. Correlation between net speech rate ΔNSR and (e) spectral mean ΔM2 and (f) spectral standard deviation ΔSD. Correlation between cepstral peak prominence ΔCPP and (g) spectral mean ΔM2, (h) spectral skewness ΔM2 and (i) spectral kurtosis ΔM4.](image-url)
intelligibility. The study suggested that the observed differences in LTAS were primarily attributed to reduced vocal cord movements rather than articulatory dysfunction. However, one would expect that the resulting high-frequency noise in the spectrum may shift the spectral mean higher and reduce skewness, which is opposite to our experimental findings. However, the values of CPP were not generally different between PD and healthy controls in the current study, which is contrary to previous studies in PD without STN-DBS [24], [25], and thus this finding needs further validation in future research.

For the articulation subsystem, an increase of spectral mean was associated with increase in articulatory clarity via RFA. Our results thus confirm that changes in LTAS descriptors for speakers with dysarthria may be attributed to both phonatory and articulatory impairments, although dysphonia seems to play a bigger role in the estimation of resulting spectral coefficients. Indeed, the LTAS is a composite signal reflecting the spectrum of the glottal source (i.e., fundamental frequency) as well as the resonant characteristics of the vocal tract (i.e., formants).

Last but not least, our study found a relationship between an increase in spectral mean as well as standard deviation and a decrease in articulation rate. To the best of our knowledge, this is the first time study to report that LTAS characteristics are influenced by articulation rate. PD typically manifest oral festination and short “rushes” of speech that are closely connected with changes in speaking rate. However, discrepant findings about speaking rate in PD exist in the literature [26]. PD patients were assumed to produce speech at a faster rate because of articulatory difficulties, in which the contrast between different speech sounds might be blurred, causing an increased speech rate [27]. Also, self-timing impairment for motor movements might play a role in increased articulation rate [28]. On the other hand, a number of previous studies have demonstrated slower speaking rates in some patients with PD [26]. These extreme speech rate disturbances in both directions could be expected due to STN-DBS and might thus influence resulting LTAS coefficients as observed here.

5. Conclusions

We found that the majority of selected acoustic features of hypokinetic dysarthria were able to differentiate between HC and STN-DBS ON or OFF states. However, only spectral mean and stdF0 were able to distinguish also between STN-DBS ON and OFF states. The spectral mean not only captures changes in F0 but also encompasses other important aspects of hypokinetic dysarthria, including articulatory precision and speech rate. Therefore, measuring the LTAS from speech signals presents a promising technological approach for the automated assessment of stimulation-induced dysarthria. This becomes particularly relevant considering the potential future advancements in adaptive DBS for PD patients. Adaptive DBS holds the promise of significantly improving the quality of life for individuals with PD [29]. In addition, LTAS may serve in a future as a predictive index of speech deterioration from pre-operative to post-operative STN-DBS [30].

6. Acknowledgements

This study was supported by the Czech Science Foundation, grant no. GF21-14216L, Czech Ministry of Education, Youth and Sports: Programme EXCELES, ID Project No. LK22NPO5107, CTU in Prague grant no. SGS23/170/OHK3/3T/13, and SNF 32003BL_197709.
7. References


