Long-Term Effect of Repetitive Transcranial Magnetic Stimulation on Parkinson’s Disease Patients with Different Severity of Hypokinetic Dysarthria

Krystof Novotny
Department of Telecommunications
Brno University of Technology
Brno, Czech Republic
ORCID: 0009-0005-7232-0841

Abstract—The prevalence of Parkinson’s disease (PD), severe neurodegenerative disorder, has steadily increased. Among the symptoms of PD is hypokinetic dysarthria (HD), a motor speech disorder, characterised by respiratory, articulatory, prosodic and phonatory impairments. It has been demonstrated that both motor and non-motor symptoms of PD can be improved using the repetitive Transcranial Magnetic Stimulation (rTMS). This study analyses acoustic speech characteristics of 19 participants diagnosed with PD before (one pre-stimulus) and after (four post-stimulus) evaluation sessions of rTMS treatment. The participants were divided into two groups – receiving either rTMS or sham stimulation (1:1 randomization). Based on the pre-stimulus subresults of the Test 3F, participants were stratified into two cohorts, according to their possible HD severity level. Speech recordings were also taken during each evaluation session. The outcome of the follow-up acoustic analysis resulted in 16 parameters for each of those sessions. Their evaluation demonstrated the dependence of the effect of rTMS treatment on the severity level. The actively stimulated group of the first cohort showed consistent improvement in articulation and prosody (sham did not) while the actively stimulated group of the second stratified cohort showed consistent improvement in phonation (sham did not). The study provides early preliminary insights into the benefits of rTMS for the alleviation of HD manifestations (symptomatic treatment of PD). In addition, it provides new insights into the possible relationship between the effectiveness of rTMS and the degree of severity in HD.

Index Terms—Parkinson’s disease, hypokinetic dysarthria, repetitive transcranial magnetic stimulation, acoustic analysis, digital speech and voice biomarkers

I. INTRODUCTION

Parkinson’s disease (PD) is the second most common neurodegenerative disorder worldwide with an age-standardised prevalence rate of 106.28 per 100,000 cases in 2019 (94.09 for Central Europe and 126.01 for Western Europe). Moreover, the number is following an increasing trend in most parts of the world [1]. The prevalence rates are then significantly higher for the older population [2]. The disease is caused by a malfunction of the motor loop of the basal ganglia and is manifested by tremors, muscle stiffness and bradykinesia [3]. Thus, it is a motor disorder. This also has a significant impact on the speech and voice of patients. Up to 90% of them have a motor speech disorder known as hypokinetic dysarthria (HD) [4]. Its manifestations include monotony of the fundamental pitch and loudness of speech, a blurred hoarse voice, unnatural and inconsistent intonation, incorrect phrasing, inappropriate pauses and sudden changes in speech rate [5]. These impairments have an undeniable impact on patients’ quality of life.

Apart from pharmacological and surgical treatments [6], one of the possible solutions is the repetitive Transcranial Magnetic Stimulation (rTMS). Studies on patients with PD have shown that the superior temporal gyrus (STG) is the part of the brain that is responsible for projecting the motor aspects of speech onto the voice during its production. It also implements prosody and manages auditory feedback processing [7]. Excitation of neurons in this area by rapid magnetic field changes (the principle of functioning of rTMS based on magnetic induction) can bring improvements in the named areas of speech [8]. The advantage of this approach is that it is one of the non-invasive brain stimulation methods – it does not involve surgical procedures and has relatively mild side effects. Its effectiveness has already been demonstrated by multiple studies on the primary symptoms of PD [9], [10]. There are fewer relevant studies in the context of dysarthric speech [11].

The short-term effects of rTMS applied to the primary motor cortex on speech were investigated by Dias et al. The results of the speech task of the sustained vowel [a:] showed an improvement in both the fundamental pitch and the intensity of voice [12]. Excitation of the same brain region was also investigated by Hartelius et al. but the studied features quantifying sustained fricative, prolonged vowel phonation, diadochokinetic rate and sentence intelligibility did not yield clear results. In addition, a strong placebo effect was also mapped [13]. More positive results were obtained by Eliasova et al. when rTMS was performed on an adjacent region of the cerebral cortex – the primary sensorimotor cortex. The best results here were observed for the five sustained vowel
tasks. There was a general improvement in voice quality and intensity, speech rate and tongue movement. This was also a study focused on short-term effects [14]. Brabenec et al. were the first to provide evidence of improved articulation after STG excitation as part of a short-term effect. Significant improvement was achieved in parameters monitoring speech formants characteristic for tongue and jaw movements [15].

Even fewer studies have addressed the long-term effects of rTMS on speech. Moreover, it should be noted that they are all based on the same data collection (the excited region was STG). Brabenec et al. conducted the first of these. Speech is quantified based on the Test 3F. The study focuses mainly on the phonation region, which is the only one for which they observed a noticeable effect of rTMS. The improvement over baseline is long-lasting in the active group. However, so is the significant placebo effect in the sham group [16]. Gomez-Rodellar et al. involve multiple speech biomarkers and observe long-term improvement in jitter, cepstral peak prominence and selected features quantifying the tremor of vocal cord tension (divided into frequency bands corresponding to those standardized in electroencephalography) [17]. Next, Brabenec et al. report noticeable difference in the active group (compared to sham) for parameters describing the left anterior arcuate fasciculus. This is the region connecting the auditory feedback area with the motor regions involved in articulation. Furthermore, the values correlate with the time evolution of the phonetic subscore of the Test 3F [18]. Subsequently, Gomez-Rodellar et al. expand knowledge in the relationship between biomechanical correlates of the vocal cord tension and rTMS therapy [19]. However, all longitudinal studies have yielded mixed results. Improvement is not uniform across the active group and the sham group often shows a placebo effect.

It is hypothesized that the effectiveness of rTMS therapy might be related to the varying severity of HD. Steurer et al. reported different behaviors of dysarthric speech symptoms in a cohort of 83 PD patients. Three different cohorts showing distinct speech characteristics were examined: a group with no HD, a group with mild HD, and a group with moderate HD. Furthermore, within these, they also documented correlations of the digital speech and voice biomarkers with the results of other clinical tests and screenings [20].

The aim of this paper is to stratify PD patients based on the Test 3F into different cohorts possibly representing the severity level of HD. Subsequently, digital speech and voice biomarkers should be calculated from the recordings of each session of all patients. Using one session completed by healthy controls (HC), the relationship of the calculated parameters to normative values should be determined. The goal is to investigate the evolution of each feature over time after active treatment (or sham stimulation), taking into account the severity level of HD (the outcome of the stratification).

As a result, this study aims to be the first ever to explore the possible influence of severity level of HD on rTMS therapy outcomes. This could shed light on new findings in the treatment of HD.

II. MATERIALS AND METHODS

A. Database

The input data for this work were audio recordings from the HIDI database, which contains a total of 19 PD patients (14 females/5 males, mean age 71.38 ± 7.43). The group was divided into two parts. The first group (10 participants, mean age 71.83 ± 6.69) received active stimulation (labeled "STG") according to the protocol described in [16] during ten sessions spread over two weeks at the Central European Institute of Technology, Masaryk University. The second group (9 participants, mean age 70.87 ± 8.57), underwent a similar process using the same device, but here no magnetic field was applied (labelled "SHAM"). Audio recordings of each participant were acquired during five sessions. Pre-stimulus T0 before undergoing the actual stimulation, post-stimulus T1 two weeks after, post-stimulus T2 six weeks after, post-stimulus T3 10 weeks after and post-stimulus T4 14 weeks after the stimulation. Among other examinations, patients also completed the Test 3F at each session. The database was created under the European Union’s Horizon 2020 research and innovation programme under the Marie Skłodowska-Curie grant agreement no. 734718 (CoBeN) and under the grant of the Czech Ministry of Health no. 16-30805A. All patients were recorded in the ON state.

At the same time, a dataset of HC was used to enable the calculation of normative values. It contains a total of 31 individuals (15 females/16 males, mean age 67.10 ± 6.05) who completed the same speech tasks as the representatives from the patient database within one session. The recordings were obtained within the framework of the Ministry of Health project no. NU20-04-00294. The studies mentioned above were approved by the local Ethics Committee and all participants signed informed consent.

B. Speech processing

Two speech tasks were investigated – sustained phonation of the vowel [a:] (as long as possible), and a free speech monologue (at least 90 s long). Maximum phonation time (MPT, the total length of the phonation) was calculated from the first task. The following parameters were also calculated from the first task, but after its adjustment – only the section of the recording starting at 2 s of its time and ending at 4 s was treated. Such parameters are: relative standard deviation of the first (relF1SD) and second (relF2SD) formant, relative standard deviation of the fundamental frequency (relF0SD), jitter (PPQ), shimmer (APQ) and harmonics-to-noise ratio (HNR).

For the second mentioned task, the relative standard deviation of the first (relF1SD) and second (relF2SD) formant and the relative standard deviation of the fundamental frequency (relF0SD) were extracted. The following features were also calculated from the second task, but after its modification by removing the silent regions that exceeded 250 ms. From the pauses removed, the median of their duration (DurMED) and the mean absolute deviation of their duration (DurMAD)
were obtained, and the ratio of their number to the duration of the whole task (SPIR) was also computed. Similarly, after removing regions of silence exceeding 50 ms, the parameters relative standard deviation of speech energy (relSEOSD), energy evolution of speech (EEVOL) and percentual pause ratio (PPR) were extracted.

The values of all calculated parameters were adjusted so that the worsening of the manifestation of a particular symptom was tracked by increasing the parameter value. Parameters with the opposite manifestation (an increase in value meant an improvement) were adjusted by multiplying by -1.

In the next step, the values of all parameters (dimension a) across all sessions of all participants (dimension b) were collectively processed, including the results of one session of HC. The data prepared in this way were adjusted using linear regression. The effect of the age of the speakers at the time of each session and the effect of gender were removed.

From the set of values of HC, norms were established for each parameter – the median \( MED_{HC} \) and the value corresponding to the 95th percentile \( PR_{95,HC} \) were determined. Using these values, it was then possible to relate to the norms all calculated values of PD patients from all sessions for each parameter according to the following relationship:

\[
DIST_{PD,k,T} = \frac{VAL_{PD,k,T} - MED_{HC}}{PR_{95,HC} - MED_{HC}}.
\]

The result measures the relative distance \( DIST_{PD} \) of the investigated feature value \( VAL_{PD} \) for a particular session \( T \) of the selected patient \( k \) from the norm determined by the HC.

C. Stratification

Based on the hypothesis of different effects of rTMS on HD according to its severity, participants were stratified. Input values for the k-means cluster analysis were the three Test 3F subscores obtained in T0. Using the silhouette method, it was decided that there were most likely to be two distinct cohorts in the dataset. Thus, the k-means algorithm identified 8 participants (4 STG and 4 SHAM) as the first group (labeled "0") and 11 participants (6 STG and 5 SHAM) as the second group (labeled "1"). The groups were then compared using the Mann–Whitney U (MWU) test applied to the individual Test 3F subscores. In general, group 0 values exceeded group 1 in all subscores. In addition, the test evaluated the differences as significant in all three cases (see Fig. 1).

D. Evaluation

The main part of the analysis relied on plotting the stacked grouped bar graphs. For each calculated parameter, one graph was plotted showing four clusters of results. Each cluster belongs to one cohort (STG 0, STG 1, SHAM 0, SHAM 1). Four results are then displayed within each cluster. These are percentages of cases where the cohort improved/deteriorated for that parameter compared to T0. Each column then considers the more strictly selected values compared to the previous one. Column one – improvement over T0 was achieved at least in one post-stimulus session, column two – improvement was achieved at least in two post-stimulus sessions, column three – improvement was achieved at least in three post-stimulus sessions and column four – improvement over T0 was achieved in all post-stimulus sessions. The examined values are the data generated by equation (1) for adjustment with respect to norms.

Additionally, the median percentage improvement is also observed for each case. It is calculated from the set of all values where there was an improvement in relation to the norm compared to T0, for all patients meeting the minimum number of improvements condition.

III. RESULTS

The best result for stratified group 0 was yielded by relF1SD computed from the monologue task. In STG 0, there was an improvement in 50 % of cases in all four post-stimulus sessions and in another 25 % in three sessions (see Fig. 2). In SHAM 0, then, there was no improvement in relF1SD of the monologue in two or more sessions (only 75 % in one session). Additional results of improvement (and discrimination from SHAM) for this group were then provided by the parameters PPR, SPIR and relF2SD calculated from the monologue task.

The HNR parameter delivered the best results for stratification group 1, with STG 1 showing improvement in all four post-stimulus sessions in 67 % of cases and in two sessions in the remaining 33 %. In contrast, the improvement among all sessions in SHAM group 1 was not achieved at all (only 20 % in three sessions and another 20 % in one session). Similar results with slight variations were obtained within group 1 for PPQ, APQ, MPT and relF0SD calculated from extended phonation.

IV. DISCUSSION

There is not much significant difference in the distribution of values of individual parameters in prestimulus sessions within the stratified groups. However, longitudinal analysis showed a difference in the effect of rTMS on a particular cohort.

The first two parameters for which the most consistent results in improvement were observed for group 0 both deal with the range of values of the selected formant in terms of the longer monologue. A larger range of values (expressed
in relative standard deviation) indicates a higher articulation ability. For the relF1SD parameter, this indirectly refers to the size of the pharyngeal cavity and thus to the mobility of the tongue root. For the relF2SD parameter (which does not yield as clear results as relF1SD, but similar nonetheless), on the other hand, it refers to the oral cavity, i.e. the openness of the lips and the position of the tongue in the mouth. The other two parameters (PPR and SPIR) with consistent improvement results in group 0 both relate to pauses in speech. In all four cases, for STG 0, there was a steady improvement in the parameter values in the majority of cases, whereas for SHAM 0 the improvements were rather sporadic. Both STG 1 and SHAM 1 gave inconsistent results in these cases.

The HNR parameter (monitoring the increase of the noise (irregularity of vocal fold vibration). All four of these features show similar patterns to HNR, with improvements for STG 1, rather sporadic improvements for SHAM 1, and mixed results for all of group 0.

Looking at the relationship between the selected parameters, a possible link emerges. The parameters describing the difference between STG and SHAM group 0 can be described as articulatory and prosodic, and the parameters describing the differences between STG and SHAM group 1 can all be seen as members of the phonation category. Thus, this longitudinal study suggests that the first stratified group can be viewed as a cohort of people in whom the influence of rTMS can lead to an improvement in articulation and prosody abilities, and the second stratified group as a cohort in which rTMS has a positive effect on phonation abilities.

Both short-term improvements in voice quality and tongue movement after the application of rTMS were observed by Eliasova et al. in [14]. Therefore, this result is consistent with the behaviour of both our groups 0 and 1. However, it should be mentioned here that in their research a different part of the brain was stimulated. Brabenec et al. then provided positive results regarding the short-term effect on speech formants in [15]. This matches the behavior of our group 0. In their case, the STG area was stimulated (this is a dataset partially identical to the one used here). Further research then provides consistent evidence from longitudinal studies of a positive effect of rTMS in the phonation domain (voice quality, vocal cord tremor...) [16], [17], [19]. This corresponds to the behaviour of our group 1.

A significant phenomenon present in this work is the placebo effect. A consistent long-term improvement is observed in a number of parameters (relSEOSD, EEVOL) not only for STG but also for SHAM. The hypothesis of a likely placebo effect is supported by the aforementioned studies that also dealt with it. In addition, an alternative explanation may be the phenomenon described in [21] and [22]. The studies focused on comparing patients’ general communicative speech with speech where patients focused on being understood as clearly as possible. Noticeable differences were observed between the two styles of speech, indicating that PD patients may be able to produce more intelligible speech with sufficient concentration and thus minimize the classic features of dysarthric speech. Therefore, the results of our research may also be influenced by fluctuations in momentary concentration and patients’ efforts to achieve higher intelligibility across different sessions.
However, the objectively biggest weakness of this work is the limited database. Moreover, when dealing with stratification and further subdividing the dataset into smaller cohorts, the low number of research participants is a major drawback.

V. CONCLUSION

This work had three main objectives. The first aim was to stratify the dataset of PD patients based on their speech. This resulted in two cohorts differing in the subscores of the Test 3F at T0.

The second aim was to obtain values of the digital biomarkers of speech and voice that would allow the results of the speech tasks of individual patients to be compared with each other. Thus, 16 different features were calculated for each session of each research participant. In addition, these values were then related to the norms obtained from the HC.

The last and the major aim of the whole study was to investigate whether the effect of rTMS on the different stratified groups differs in any way. The findings suggest that rTMS may have varying effects on speech impairments. Depending on the severity level of HD, there was an improvement in a specific group of parameters characterizing a particular speech area (articulation and prosody versus phonation).

The research offers novel observations regarding the advantages of using rTMS to ease the speech impairments in PD patients. Furthermore, it sheds light on the potential correlation between the effectiveness of rTMS and various HD severity levels in PD patients.

The database of research participants offers even more unexplored data in relation to the effects of rTMS on HD according to its severity. Thus, follow-up research is likely. For example, studies observing behavior within each stratified group in search of relationships between changes in digital speech and voice biomarkers and the development of other clinical testing outcomes are suggested.

ACKNOWLEDGMENT

This study was supported by EU – Next Generation EU (project no. LX22NPO5107 (MEYS)). I would like to express my sincere gratitude to my supervisor Jiri Mekyska for his guidance, inspirational suggestions and knowledge sharing throughout this research project.

REFERENCES


