

Description of PD Phonation in Terms of EEG-Related Frequency Bands

Pedro Gómez¹^a, Jiří Mekyska²^b, Luboš Brabenec³^c, Patrik Šimko³^d, Irena Rektorová⁴^e,
Andrés Gómez⁵^f and Victoria Rodellar¹^g

¹*NeuSpeLab, CTB, Universidad Politécnica de Madrid, 28220 Pozuelo de Alarcón, Madrid, Spain*

²*Department of Telecommunications, Brno University of Technology, Brno, Czech Republic*

³*Applied Neuroscience Research Group, Central European Institute of Technology - CEITEC,
Masaryk University, Brno, Czech Republic*

⁴*First Department of Neurology, Faculty of Medicine and St. Anne's University Hospital,
Masaryk University, Brno, Czech Republic*

⁵*Usher Institute, Faculty of Medicine, University of Edinburgh, Edinburgh, U.K.*

Keywords: Parkinson's Disease, Functional Assessment of Phonation, Neuromotor EEG Activity Monitoring, Repetitive Transcranial Magnetic Stimulation.

Abstract: Parkinson's Disease (PD) is an increasing prevalence neurodegenerative condition affecting the life quality of people suffering from its neuromotor and cognitive performance. PD symptoms include vocalization and speech alterations, known as hypokinetic dysarthria (HD). One of the manifestations of HD is unstable phonation. Repetitive Transcranial Magnetic Stimulation (rTMS) is a non-invasive method that may improve some motor and non-motor symptoms of persons with PD (PwP). The present study concentrates on analyzing and comparing the phonation behavior of two cases before (pre-stimulus) and after (post-stimulus) ten sessions of rTMS treatment, to assess the extent of changes in their vocalization. Voice recordings of a sustained vowel [a:] taken immediately before and after the treatment, and at follow-up sessions (at six, ten, and fourteen weeks after the baseline assessment) were processed by inverse filtering to estimate a biomechanical correlate of vocal fold stiffness, which band-pass filtered into EEG-related frequency bands. Log-likelihood ratios between pre- and post-stimulus amplitude distributions of each frequency band, Mann-Whitney U-tests, and normalized difference scores showed significant improvements in the actively stimulated case, which were not observed in the sham case. Early preliminary insights into the capability of phonation quality assessment on monitoring neuromechanical activity from acoustic signals are shown.

1 INTRODUCTION

Parkinson's Disease (PD) is a neurodegenerative disorder with a prevalence of around 200 cases per 100,000 persons, at a growing incidence rate of 15 cases per 100,000 (Dorsey et al., 2007). It has a severe impact on the life quality of persons with PD (PwP)

affecting motor and non-motor symptoms (Duffy 2013). Perturbed respiration, phonation, articulation, and prosody are among motor symptoms hampering vocalization and speech, in what is known as hypokinetic dysarthria (HD), characterized by mono-pitch and mono loudness, imprecise articulation, impaired speech rate, and rhythm, and irregular pitch fluctuations. Repetitive transcranial magnetic

^a <https://orcid.org/0000-0003-3283-378X>

^b <https://orcid.org/0000-0002-6195-193X>

^c <https://orcid.org/0000-0002-8348-5757>

^d <https://orcid.org/0000-0002-0767-0918>

^e <https://orcid.org/0000-0002-5455-4573>

^f <https://orcid.org/0000-0003-3283-378X>

^g <https://orcid.org/0000-0001-9384-3290>

stimulation (rTMS) is a non-invasive method used to modulate neuronal excitability which has been proposed as a therapy to improve various symptoms of PD (Brabenec et al. 2019). The purpose of the present study is two-fold: on the one hand, to explore the relationship between neuromotor activity and phonation through the inversion of a cortico-muscular coupling (CMC) model (Branbilla 2021); on the other hand, to show how to characterize and monitor the efficiency of rTMS in two study cases using CMC model inversion.

The structure of the paper is as follows. Section 2 is devoted to explaining the conducting narrative capitalizing on the possibility of describing phonation-related biomechanical correlates as vocal fold stiffness in terms of EEG-related frequency bands based on the well-known relationship between EEG neuroelectrical activity in the premotor and supplementary motor cortex and the neuromuscular activity in the laryngeal nerves controlling voice production explained by CMC (McKeown et al, 2006). Section 3 describes the signal inversion methods to determine neuromotor activity in EEG-related frequency bands, and their statistical distributions in two cases of PwP submitted to active and sham rTMS. Section 4 compares the results from the two study cases, which are discussed in Section 5. Contributions, findings, and conclusions are disclosed in Section 6.

2 FUNDAMENTALS

The objective of the present study is to dive into the relationship between neuromotor and acoustical activity (neuroacoustical) involved in vocalization (phonation and articulation) with application to the characterization of PD hypokinetic dysarthria, as summarized in 0. The simplified neuroacoustical model (acoustical and neuromotor) of the system controlling vocalization (top-down view) is summarized in 0.a. The vocalization structure involves the lungs, larynx, and the oro-naso-pharyngeal cavities, considering the lips as the radiation place where the speech wave is projected to the surrounding media. Important muscular systems control each substructure, such as the diaphragm and intercostal (not shown) controlling lung pressure, the laryngeal muscles (thyroarytenoid, cricothyroid, and transverse and oblique arytenoid) of which the thyroarytenoid (*musculus vocalis*) is responsible for blocking and releasing airflow through the glottis and producing the basic vibration in voiced speech (phonation), the hypoglossus, extrinsic and intrinsic

glossal, controlling tongue movements, the mandibular, regulating jaw raising and lowering, and the orofacial, defining lip rounding gestures.

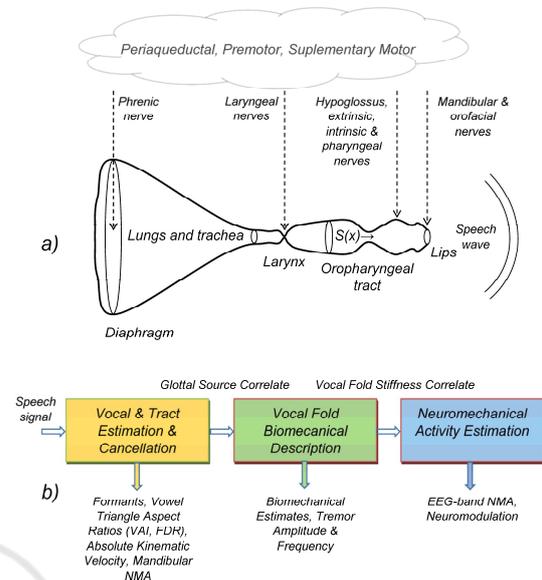


Figure 1: Acoustical and neuromotor pathways in vocalization: a) Top-down model from the neuromotor to the acoustical (speech wave); b) Bottom-up model from phonation to neuromechanical activity estimation.

The neuromotor activity (NMA) controlling each group of muscles is driven by common areas of activation for semantic processing, lexical selection, syntactic construction as well as oral articulation, involving mainly the periaqueductal, premotor, and supplementary motor brain areas, among others (Schulz et al., 2005). These areas project to the specific muscles by different neural pathways, such as the phrenic nerve (diaphragm), the superior and recurrent laryngeal (laryngeal muscles), the pharyngeal, hypoglossus, extrinsic, and intrinsic tongue nerves (oropharyngeal), and the trigeminal branch controlling mandibular and orofacial muscles. The inverse bottom-up model shown in 0.b provides a view of the model inversion methodology proposed in the present study to estimate and visualize NMA involved in phonation. The speech signal is processed by a linear predictive algorithm (Deller, Proakis, and Hansen, 1993) to produce a glottal source correlate. The inverse filter used in the processing is used to estimate acoustic-articulatory features, such as formants, vowel triangle aspect ratios, tongue-jaw kinematic movement, masseter NMA, etc. (Gómez et al., 2021). Among the mentioned larynx muscles, the thyroarytenoid is of special interest for the study because it is a very small mass low-inertial muscle, responding well to vibration frequencies beyond

1kHz, and is directly related to phonation. The study capitalizes on the estimation of the thyroarytenoid muscle stiffness on the glottal source correlate, which is used to produce estimations of the vocal fold biomechanical features. It is assumed the vocal fold stiffness to be controlled by the inferior laryngeal nerve NMA on the thyroarytenoid muscle, therefore, it may be used to estimate EEG-related frequency bands in the laryngeal nerve periaqueductal control. This inverse bottom-up chain could help in better monitoring phonation control in PwP (Rektorova et al., 2012). Eventual instabilities of the laryngeal NMA will appear as oscillations in the frequency bands (tremors).

3 MATERIALS AND METHODS

The present research includes results from two PwP participants (under active and sham rTMS) in a study devised to HD in PwPs (Brabenec et al., 2021). The NMA of both participants was monitored at several evaluation periods since stimulation using sustained emissions of the vowel [a:], and the results have been compared.

Table 1: Participants' demographic and clinical data; A: active stimulation; S: sham stimulation; M: Male Gender; Y: years. UPDRS-III: Unified Parkinson's Disease Rating Scale part III (motor).

PwP code (pre)	Active/Sham	Gender	Age (Y)	UPDRS -III
1400	A	M	64	10
1900	S	M	77	8

The speech processing was based on fragments of 4s long vowel emissions between the time instants at 2s and 6s from the vowel onset, sampled at 16 kHz and 16 bits. An inverse lattice-ladder filter (Deller, Proakis, and Hansen, 1993, Alku et al., 2019) evaluated the oropharyngeal tract model, estimating the residual prediction error, which once integrated, produced the glottal source correlate. The vocal fold stiffness (VFS) was estimated from the glottal source correlate adjusting its spectral power by a 2-mass model of the vocal fold biomechanics (Gómez et al. 2009) The VFS was de-biased and de-trended by a moving-average filter. See 0 for an example of unbiased VFS (UVFS) estimation from case 1400. The UVFS was band-pass filtered at the EEG-related frequency bands (δ : $f \leq 4$ Hz; θ : $4 \text{ Hz} < f \leq 8$ Hz; α : $8 \text{ Hz} \leq f \leq 16$ Hz; β : $16 \text{ Hz} < f \leq 32$ Hz; γ : $f > 32$ Hz; μ : $8 \text{ Hz} < f \leq 12$ Hz), producing a set of UVFS estimates

given by $\xi_{ij}^k(n)$, where $i=(0, \dots, I)$ is the session index ($I=5$), $j=a$ (active case) or $j=s$ (sham case) is the participant index, and $k=(1 \text{ for } \delta, \dots, 6 \text{ for } \mu)$ is the frequency band index of the six frequency bands defined above, and n is the time index. 0 shows the results of the band-pass separation in the time domain (left column) and their power spectrograms (right column). The distributions $p\{\xi\}$ of EEG-related frequency bands were estimated from normalized amplitude histograms as

$$p_{ia,s}^k(\xi) = h\{\xi_{ia,s}^k(n)\} \quad (1)$$

The distributions from post-stimulus recordings were compared on their overlap interval $\xi \in \Omega$ with pre-stimulus ones using log-likelihood ratios

$$\lambda_{a,s}^k(p_i|p_0) = \int_{\xi \in \Omega} \log\{p_{ia,s}^k(\xi)/p_{0a,s}^k(\xi)\} d\xi \quad (2)$$

as well as Mann-Whitney U-tests (T_{MW})

$$T_{a,s}^k = T_{MW}\{\xi_{ia,s}^k(n), \xi_{0a,s}^k(n)\} \quad (3)$$

and time-weighted scores (s_{ia}^k and s_{is}^k)

$$s_{ia,s}^k = \frac{\langle \xi_{ia,s}^k(n) \rangle - \langle \xi_{0a,s}^k(n) \rangle}{\langle \xi_{0a,s}^k(n) \rangle} w_i; \quad (4)$$

$$w_i = (d_i - d_0)/(d_i - d_0)$$

where the weight w_i is a normalizing factor to take into account the time interval between each post-stimulus date d_i and the corresponding pre-stimulus date d_0 normalized to the longest interval ($d_i - d_0$) in days, therefore, long-lasting beneficial effects were given larger importance than short-duration effects. The time intervals in days for the case 1400 were $T_1=15$, $T_2=47$, $T_3=74$, and $T_4=99$. The corresponding ones for case 1900 were $T_1=9$, $T_2=42$, $T_3=65$, and $T_4=96$. Finally, the VFS unbalance (the difference between two neighbor estimates of the total UVFS relative to their average) is added as a reference feature indirectly related to the jitter

$$u_{ia,s} = 2 \frac{\xi_{ia,s}(n) - \xi_{ia,s}(n-1)}{\xi_{ia,s}(n) + \xi_{ia,s}(n-1)} \quad (5)$$

4 RESULTS

Case 1400 corresponded with a participant who was submitted to active stimulation. The pre-stimulation recording of a sustained vowel [a:] produced an

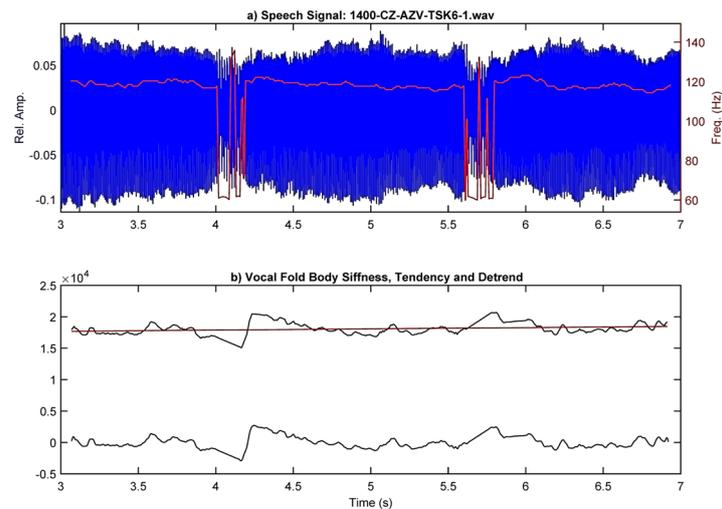


Figure 2: Example of the vocal fold body stiffness estimation from a segment of 4 s of phonation from case 1400, during the utterance of a sustained vowel [a:]: a) speech segment showing two events of phonation blocking at intervals 4.0 s - 4.2 s and 5.6 s - 5.8 s, where the f_0 drops down from 120 Hz to 60 Hz; b) vocal fold stiffness (black), its detrend tendency (red) and its unbiased (blue).

utterance with a duration of 11.163 s, during which five events of phonation blocking were observed. The segment selected for the analysis was extended between 3 s and 7 s, to include the first two blocking events, which are shown in 0.a, with an estimation of its fundamental frequency (f_0) profile superimposed in red. 0.b shows the estimation of the vocal fold stiffness (black), its detrend (red), and the unbiased detrended result (blue). In its turn, 0 depicts the results of splitting the unbiased vocal fold stiffness (UVFS) into the EEG-related frequency bands defined in Section 3 in the time domain (left column) and as spectrograms (right column), the first one (0.b) being given in a logarithmic representation to compensate different amplitude levels, and the remnant ones are shown in linear representation. The distributions of the pre-stimulus and post-stimulus EEG-related frequency bands are presented in 0. In Figure 2 an example of VFS estimation from the acoustic signal is shown, where two blocking events are appreciated as 4 s and 5.2 s. These events are marked by a descent in f_0 (red line in Figure 2.a) from 120 Hz to 60 Hz during approximately 200 ms, during which, the estimation of f_0 becomes erratic and unstable. This behavior is aligned with the UVFS profile shown in Figure 2.b, where f_0 instability is explained by a decay in the VFS at 4 s, followed by a correction action at 4.3 s to initiate a new slow decay to be corrected again at 5.6 s.

The graphical example showing the frequency band estimation protocol illustrated in Figure 3 adds new information to the evolution of these events. The activity on the δ -band in Figure 3.c shows the decay

starting really at 3.7 s, coming to a minimum at 4.1 s, and being strongly incremented immediately after, followed by a slow progressive decay and correction bursts between 5.0 s and 5.6 s, where a new pull-up is observed. The corrective actions are evident in all the bands, but the most interesting ones are β and γ , given in Figure 3.j and l. The β -band shows a strong activity burst between 22-30 Hz at 4.1-4.3 s, followed by a narrower one at 26 Hz and 4.4-4.7 s. The γ -band, on the contrary, shows a narrow burst at 38 Hz between 4.4-4.5 s, and a wider and stronger one at 42-46 Hz lasting from 5.4-5.7 s. According to recent research, the activity in these bands could be related to the neuromotor correction (β), and interaction of different brain areas (γ), including the striato-cortical, the cerebellum, and the temporal auditory areas (Ibarra-Lecue, Haegens, and Harris, 2022).

The information provided in Figure 4.a and b, corresponding to the activity on the β -band is also quite meaningful, as it shows the profile of the distributions for the five recordings corresponding to the active case 1400. Whereas the first evaluation (1400, pre-stimulus) is amplitude-widespread, all post-stimulus evaluations (T1-T4) concentrate in small amplitudes, all of them showing a χ^2 profile. The averages and standard deviations of the post-stimulus evaluations are sensibly smaller than the pre-stimulus one, pointing to a strong instability reduction after stimulation. By contrast, the five recordings corresponding to the sham case 1900 given in Figure 4.c and d, do not show any meaningful changes, except possibly in the second evaluation (T1, 1900).

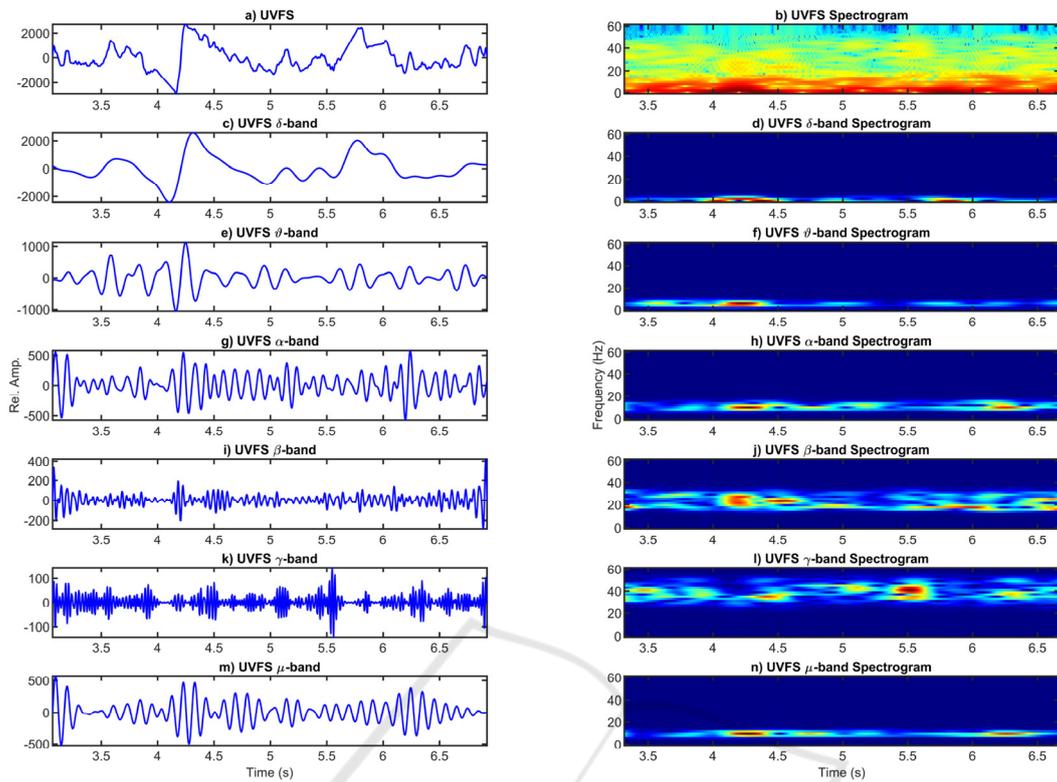


Figure 3: Spectral contents of the EEG-band description of 4 s of phonation from case 1400: a) unbiased vocal fold stiffness; b) spectrogram of the vocal fold stiffness; c) δ -band component; d) spectrogram of the δ -band; e-f) id. of the θ -band; g-h) id. of the α -band; i-j) id. of the β -band; k-l) id. of the γ -band; m-n) id. of the μ -band.

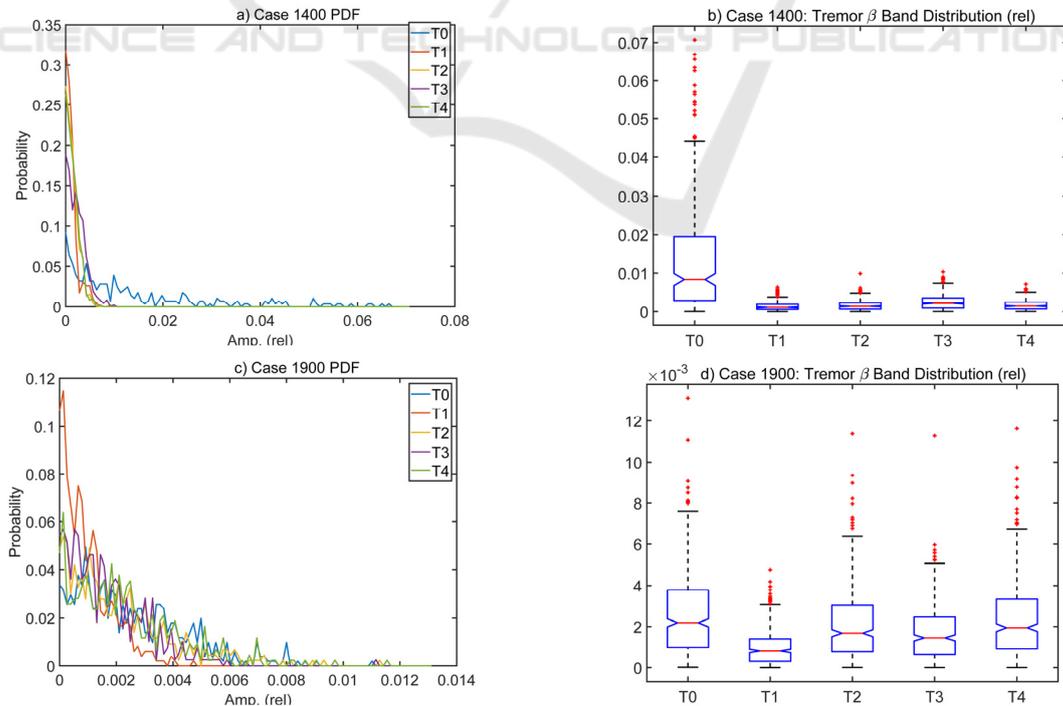


Figure 4: Probability density distributions of the EEG β -band activity from cases 1400 and 1900: a) pdfs from each evaluation of the active case (1400: T0-T4); b) their corresponding boxplots; c-d) id. from the sham case (1900: T0-T4).

The results of the comparisons between pre-stimulus and post-stimulus frequency distributions estimated following expressions (2)-(3) are given in Table 2 and Table 3. The results of evaluating the improvement scores following expression (4) are given in 0. 0 summarizes graphically the amplitude averages of the five non-overlapping frequency bands (δ , θ , α , β , and γ) in the five evaluation instants (T0-T4), and the VFS unbalance (as a reference) for the two cases being considered, their relative differences concerning T0, and the same differences weighted by the in-between-evaluation time intervals, following expression (4).

Table 2: LLR scores according to expression (2) on the β -band of post-stimulus evaluations (T1-T4) relative to pre-stimulus (T0).

PwP code	T1	T2	T3	T4
1400	0.675	0.607	0.469	0.593
1900	0.192	-0.005	0.083	-0.015

Table 3: p-values on the β -band from MW U-tests of post-stimulus evaluations (T1-T4) relative to pre-stimulus (T0) according to expression (3).

PwP code	T1	T2	T3	T4
1400	<0.001	<0.001	<0.001	<0.001
1900	<0.001	<0.001	<0.001	0.036

Table 4: Improvement scores according to expression (4) summarizing the β -band of post-stimulus evaluations (T1-T4) relative to pre-stimulus (T0).

PwP code	T1	T2	T3	T4
1400	-0.135	-0.418	-0.615	-0.878
1900	-0.058	-0.074	-0.233	-0.105

5 DISCUSSION

The aim of this study was focused on exploring the capability of EEG-related frequency bands to explain the activity on the neuromotor pathways related to phonation using acoustic signals analyzing sustained

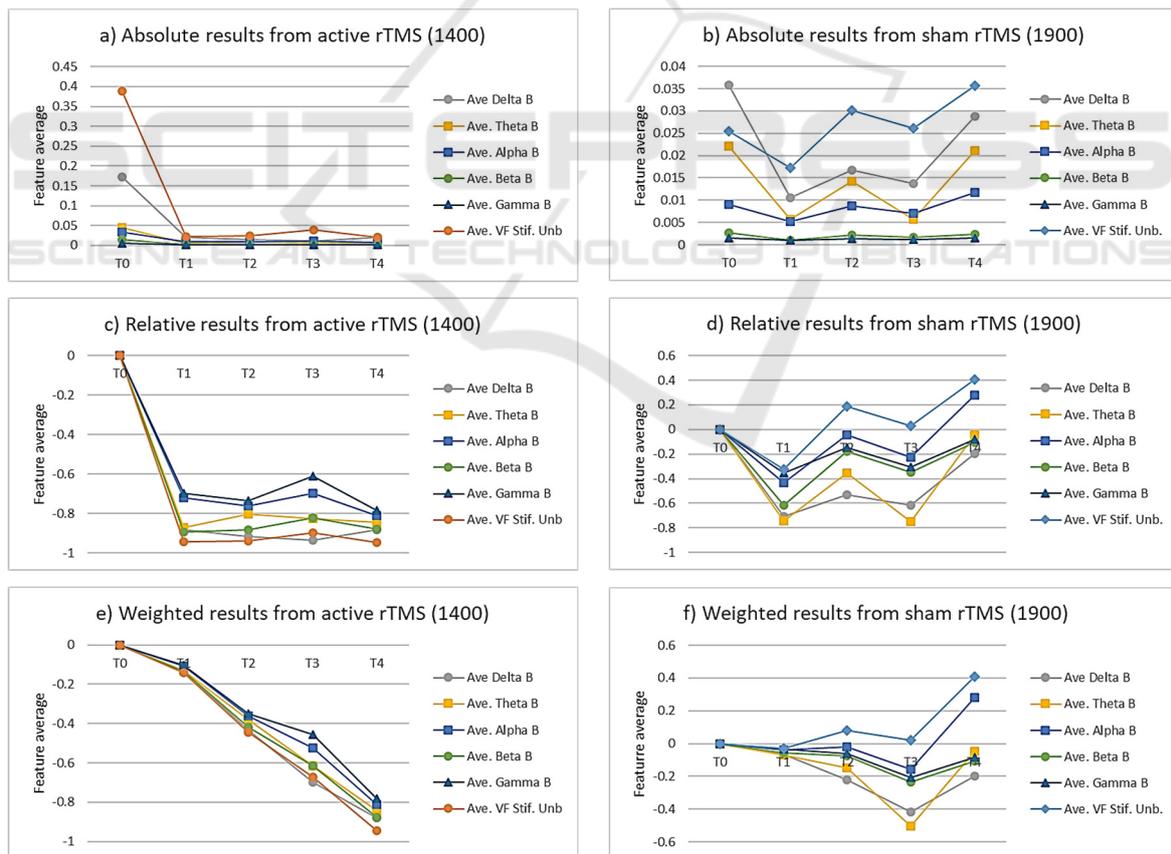


Figure 5: Normalized average values of each EEG-band (δ - γ) compared with the vocal fold stiffness unbalance from an active stimulation case (1400) and a sham case (1900): a-b) absolute values; c-d) relative post-stimulus values (T1-T4) compared with the pre-stimulus evaluation (T0); e-f) relative results weighted accordingly to the time interval between the pre-stimulus and each post-stimulus interval, as defined in expression (4).

vowel vocalizations from PwPs submitted to active and sham rTMS.

The comparison among pre- and post-stimulus estimations in terms of LLRs given in 0 confirms the observations on the β -band, pointing to strong improvements in the active case ($\lambda > 0$), whereas the sham case shows mixed behavior and moderate improvements in T1 and T3 which might be due to circumstantial or confounding factors. The p-values which are shown in 0 avail the estimations given in 0 for a significance level of 0.05 on the null hypothesis of equal medians.

After examining the global improvement scores on all the non-overlapping frequency bands given in 0, it may be concluded that taking the time interval between the pre-stimulus and each post-stimulus evaluation into account, the progress in the process induced by rTMS seems steady, at least for the observation time intervals considered. These findings may be better examined on the evolution templates given in 0.a and b, where the normalized amplitude average values of the frequency-band components are given, as well as the VFS unbalance regarding expression (5) which is added as a reference. The improvements of the phonation instability conditions for the active case 1400 are evident (0.a), whereas the evaluations from the sham case (0.b) do not show a clear tendency. When considering the difference between the pre-stimulus and each post-stimulus estimations by bands given in 0.c and d, the droppings observed in the active case (1400) become more evident when compared with the random behavior of the sham case (1900). This comparison is even more meaningful when comparing the same differences in all frequency bands weighted by the time intervals between each pre- and post-stimulus pair, as seen in 0.e and f. The monotonous descent observed in 0.e is indicative of the almost-permanent improvements observed in the active case during the period considered, contrasting with the quasi-erratic behavior of the sham case.

The character of this study is very specific, exploratory, and limited to the observations from the two cases considered, and further efforts would be required to generalize its potential application on a large database.

6 CONCLUSIONS

The present paper is intended to explore the possibilities of predicting the interactions on the EEG-related β - γ frequency bands of the NMA from the phonation acoustical signal. Albeit the specificity

of the cases studied is a limit to the findings observed, the methodology proposed to extract neuromotor activity from acoustical information to characterize PwP vocalization may provide new meaningful insights into the neuromotor activity related to phonation stability. The three scores used in the assessment of potential improvement behavior of PwP phonation after active rTMS are in full agreement, and can be used alternatively or combined. These facts may open new applications of signal processing in the field of speech neuromotor understanding, and neurodegenerative disease monitoring.

ACKNOWLEDGEMENTS

This research received funding from European Union's Horizon 2020 research and innovation program under the Marie Skłodowska-Curie grant agreement no. 734718 (CoBeN), a grant from the Czech Ministry of Health, 16-30805A, a grant from EU – Next Generation EU (project no. LX22NPO5107 (MEYS)), and grants TEC2016-77791-C4-4-R (Ministry of Economic Affairs and Competitiveness of Spain), and Teca-Park-MonParLoc FGCSIC-CENIE 0348-CIE-6-E (InterReg Programme). Andrés Gómez-Rodellar holds a scholarship from the Medical Research Council Doctoral Training Programme in the Usher's Institute (University of Edinburgh Medical School).

REFERENCES

- Alku, P., et al. (2019). OPENGLLOT-An open environment for the evaluation of glottal inverse filtering, *Speech Communication* 107 (2019) 38-47. <https://doi.org/10.1016/j.specom.2019.01.005>.
- Brabenec, L. et al. (2021) Non-invasive brain stimulation for speech in Parkinson's disease: A randomized controlled trial, *Brain Stimulation*, 14, 571-578. <https://doi.org/10.1016/j.brs.2021.03.010>.
- Brambilla, C. et al., (2021). Combined Use of EMG and EEG Techniques for Neuromotor Assessment in Rehabilitative Applications: A Systematic Review, *Sensors*, 21 7014. <https://doi.org/10.3390/s21217014>.
- Deller, J. R., Proakis, J. G., and Hansen, J. H. L. (1993) *Discrete-Time Processing of Speech Signals*, NewYork, Macmillan.
- Dorsey, E. R., et al (2007). Projected number of people with Parkinson's disease in the most populous nations, 2005 through 2030, *Neurology* 68(5) 384-386. <https://doi.org/10.1212/01.wnl.0000247740.47667.03>.

- Duffy, J. R. (2013). *Motor Speech Disorders: Substrates, Differential Diagnosis, and Management*, 3rd Ed., Elsevier.
- Gómez, P. et al. (2009). Glottal Source biometrical signature for voice pathology detection, *Speech Communication* 51(9) 759-781. <https://doi.org/10.1016/j.specom.2008.09.005>.
- Gómez, A., et al. (2021). Acoustic to kinematic projection in Parkinson's disease dysarthria. *Biomedical Signal Processing and Control* 66 Art. 102422. <https://doi.org/10.1016/j.bspc.2021.102422>.
- Ibarra-Lecue, I., Haegens, S., and Harris, A. Z. (2022) Breaking Down a Rhythm: Dissecting the Mechanisms Underlying Task-Related Neural Oscillations, *Front. Neural Circuits*, 16 846905. <https://doi.org/10.3389/fncir.2022.846905>.
- McKeown, M. J., et al. (2006). Cortical muscle coupling in Parkinson's disease (PD) bradykinesia. *Parkinson's Disease and Related Disorders*, P. Reiderer et al. (eds.), Springer, Vienna (2006) 31-40. https://doi.org/10.1007/978-3-211-45295-0_7.
- Mekyska, J., et al. (2015) Robust and complex approach of pathological speech signal analysis, *Neurocomputing* 167 94-111. <https://doi.org/10.1016/j.neucom.2015.02.085>
- Rektorova, I., et al. (2012). Functional neuroanatomy of vocalization in patients with Parkinson's disease, *J. Neu. Sci.* 313 7-12. <https://doi.org/10.1093/cercor/bhi06>.
- Schulz, G. M. et al. (2005). Functional neuroanatomy of human vocalization: an H215O PET study, *Cereb. Cortex* 15 1835-47. <https://doi.org/10.1016/j.jns.2011.10.020>.

SCIENCE AND TECHNOLOGY PUBLICATIONS