Measurement of Tremor in the Voices of Speakers with Parkinson's Disease

Markus Brückl, Alain Ghio, François Viallet

Measurement of Tremor in the Voices of Speakers with Parkinson's Disease

Markus Brückl
Fachgebiet Kommunikationswissenschaft
Technische Universität Berlin
Berlin, Germany
Email: markus.brueckl@tu-berlin.de

Alain Ghio
Laboratoire Parole & Langage
Université d'Aix-Marseille
Marseille, France
Email: alain.ghio@lpl-aix.fr

François Viallet
Laboratoire Parole & Langage
Université d'Aix-Marseille
Marseille, France
Email: francois.viallet@lpl-aix.fr

***Abstract*—A study is presented analyzing tremor in the voice of speakers that were diagnosed with Parkinson's disease (PD). The examined sounds are sustained /a/s, originating from a large dysarthric speech corpus. Six measures of vocal tremor are extracted from these vowels by applying a self-developed algorithm that is based on autocorrelation of contours and implemented as a script of an open-source speech analysis program. Univariate analyses of covariance reveal significantly raised tremor magnitudes (tremor intensity indices and tremor power indices) in PD speakers off medication as compared to a control group as well as within PD speakers in off medication condition as compared to on medication. No significant differences are found between the control group and PD speakers on medication as well as for tremor frequencies. However, the greater part of variance in tremor measures is always accounted for the speakers' age.**

I. Introduction

Parkinson's disease (PD), also idiopathic parkinsonism, is a neurodegenerative disorder of the central nervous system, which is mainly destroying the substantia nigra. This seriously impairs the secretion of the neurotransmitter dopamine and that in turn affects emotion, cognition as well as autonomous and motor neuronal activity. Relative motor inactivity or increased latency leads to one of the main symptoms of PD, if not the most formative one: tremor (commonly shake, tremble), an unintentional muscular control deficit that results in cyclic movement deviations. It was not by accident that James Parkinson himself named "his" disease "shaking palsy".

Functionally speaking, all tremor causing phenomena can be seen as disturbances of or latencies in the neuronal regulation of a muscular process, e.g. the production of speech. The production of speech, especially the process of phonation, probably is the fastest and most complex motor activity that humans are capable of. Thus, if there is a neuronal deficit that is generally causing tremor, then it should affect phonation. Vocal tremor is often defined as an *unintentional* low-frequency modulation of the vocal fold vibration. If intentionally used in singing, such modulations are known as vibrato. And an acoustic speech signal may also show further "tremulous" components that are e.g. due to articulatorily motivated jaw movements. Thus, for a reliable measurement of vocal *tremor* in natural voices, a vowel (e.g. /a/) phonation that should be sustained as constantly as possible is to be preferred.

Though, unlike other tremors the acoustic representation of vocal tremor channels into two components: A frequency and an amplitude tremor. And probably all of the neuronal disturbances or latencies of voice production are interweaved in both tremor types. That entails the fact that vocal tremor also can be observed as a symptom of other diseases as well as in healthy people, e.g. as a consequence of aging (as far as there is healthy aging at all – neurotransmitters are dramatically reduced by aging). Thus, diagnosing vocal tremor alone does not allow to refer unambiguously to any underlying cause. But on the other hand this comprises also the power of vocal tremor analysis as an additional tool for the determination or diagnosis of a wide variety of phenomena and diseases and especially PD [1].

However, as a look into recent literature reveals, the effect of PD on vocal tremor is still not too clearly understood. Speech pathologists [2, p. 41, our translation] state: "Indeed, this phenomenon [tremor] should not constitute an outstanding feature of hypokinetic dysarthria [...] and moreover should be bound to advanced stadia of disease [...]." And a little further on, special tremor frequencies around 9 Hz are suspected to indicate PD. Support for the relevance of tremor frequencies comes from a very recently published comprehensive study [3] involving 30 PD speakers and a control group that finds the frequency of amplitude tremor to be the only acoustic tremor measure that differs significantly between PD speakers and a control group – but within the PD group the mean value of amplitude tremor frequency lies below 5 Hz. In addition the author discovers that "acoustic voice tremor did not relate in any significant way to PD disability or phenotype." But she also finds that PD speakers "were more likely to show greater auditory perceived [...] magnitude[s] of frequency and amplitude tremor in comparison to controls, however without statistical significance" (sic!). Other speech researchers [4] have found acoustic tremor magnitude measures depending rather sensitively on PD, but not tremor frequency. Yet others [5] have tested 132 acoustic dysphonia measures with different classification algorithms in order to predict PD. They reached at best 98.6% accuracy (with 10 remaining features) – but without any (direct) tremor measure being involved.

II. Data – the AHN corpus

The examined data are a subset of the data referred to as Aix Hospital Neurology (AHN) corpus [6]. Although this corpus also comprises other pathologies than PD and other data, e.g. absolute SPL and airflow, this study concentrates only on acoustic signals, more specifically on sustained /a/-vowels from 363 speakers, recorded in mono with a resolution

of 16 Bit at a sampling frequency of 25 kHz. 239 of these speakers (83 females, 156 males) are diagnosed with PD and 124 (73 females, 51 males) are control speakers without any pathology. Within the PD group most (228) individuals were recorded under two conditions, on and off medication (L-DOPA, respectively Levodopa) that is administered in order to compensate for the decrease of the dopamine level.

For reasons of comparability we decided to take only quasi-stationary parts of equal duration (3 s) from the sustained /a/s under examination. A duration of 3 s seems appropriate, because reliable detection of (tremor) frequencies requires a few (at least 2-4) cycles – and we aim to measure frequencies as low as 1.5 Hz. Thus, 70 of the 591 vowels were too short and had to be excluded. 234 PD speakers (182 under both medication conditions), aged 66.66 a in average, SD = 9.81 a, and 105 control speakers, with a mean age of 62.29 a, SD = 10.85 a, remained.

The actual number of objects in the statistical analyses again is reduced, since tremor can not be detected in every vowel (see subsection III-A) and this leads to "undefined" measurements and thus missing values.

III. Acoustic measurement of vocal tremor

The tremor extraction algorithm is based on autocorrelations of the F_0 contour and of the amplitude contour and implemented in the script language of the speech-processing program Praat [7]. This script, named TREMOR.PRAAT, version 2.06, can be downloaded from [8]. Although the algorithm has evolved since its first publication [9], additional information and further description is found there. Nonetheless, the algorithm's essential ideas are outlined below.

A. tremor.praat's output values

The operational definitions for the tremor measures are (in the most general sense) adopted from the likely most commonly used tremor extracting instrument MDVP [10]. Accordingly the first outputted value, the frequency tremor frequency (FTrF) is defined as the frequency of the strongest low-frequency modulation of the fundamental frequency (F_0). The second, the amplitude tremor frequency (ATrF) analogously is the frequency of the strongest low-frequency modulation of the amplitude (intensity).

These two parameters are referred to as "tremor frequencies". The two tremor intensity indices and the two tremor power indices, outlined below, are subsumed under the term "tremor magnitudes".

Using TREMOR.PRAAT, the tremor frequencies are determined by auto-correlating the contours (see the upper subfigures of Fig. 1). So, if the highest autocorrelation coefficient that can be detected in the contour is smaller than the threshold (that can be set individually, see [8]; standard value: 0.15), it is assumed that there is no tremor and therefore no tremor frequency nor intensity nor power – and the output will be "undefined".

The frequency tremor intensity index (FTrI) is defined as the intensity/magnitude of the strongest low-frequency modulation of F_0, the amplitude tremor intensity index (ATrI) as the intensity/magnitude of the strongest low-frequency modulation of the sound's (intensity-per-period-) amplitude (A). These intensities/magnitudes are expressed relative to the mean F_0 ($\overline{F_0}$), respectively the mean A ($\overline{A}$), in the analyzed sound. Thus, they do not have a physical unit. TREMOR.PRAAT operationalizes these definitions by relativizing the contours at first, see 1 and 2:

$$relative\ F_0 = \frac{F_0 - \overline{F_0}}{\overline{F_0}} \tag{1}$$

$$relative\ A = \frac{A - \overline{A}}{\overline{A}} \tag{2}$$

Hereon and after de-declining the whole contours by subtracting their linear fit in order to compensate for natural declinations, TREMOR.PRAAT computes the intensity indices by applying the following equation:

$$(F,A)TrI = \left(\frac{\sum_{i=1}^{m} |max_i|}{m} + \frac{\sum_{j=1}^{n} |min_j|}{n} \right) \div 2 \tag{3}$$

where n and m denote the number of local minima resp. maxima. The time marks of the extrema are found with Praat's built-in function "To PointProcess (peaks)", once the tremor frequencies are known. This step is visualized in the middle and bottom graphs of Fig. 1: The dotted vertical lines mark the times of found extrema. The searched magnitudes (min and max) are the ordinates that are assigned to these times.

Additionally to these four common measures of vocal tremor, two new ones are introduced: the indices of tremor power (FTrP and ATrP). These measures result from weighting the intensity indices with factors that are depending on tremor frequencies. These factors are defined smaller for lower frequencies and therefore a lower power index would emerge if the same tremor intensity was found at a lower tremor frequency.

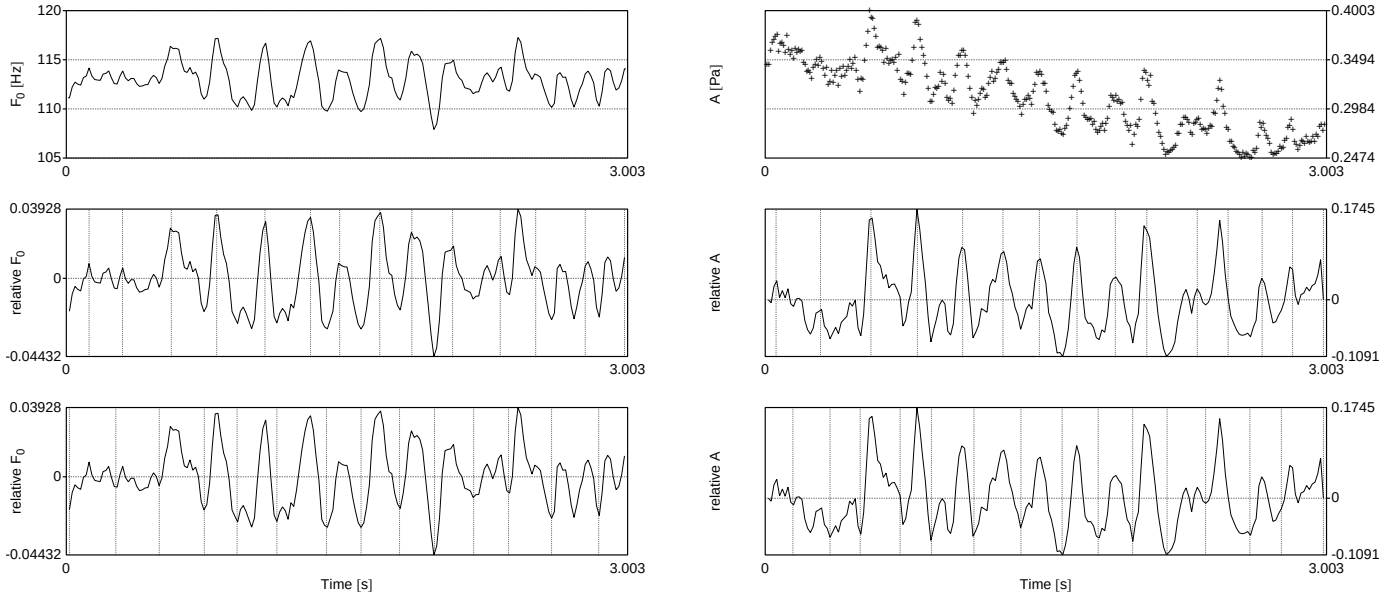
$$FTrP = FTrI \cdot \frac{FTrF}{FTrF + 1} \tag{4}$$

$$ATrP = ATrI \cdot \frac{ATrF}{ATrF + 1} \tag{5}$$

These power indices are thought to be biologically and psychologically more significant for the concept "tremor level" or "tremor magnitude" than the known intensity indices, since the (perceived) effect of events (e.g. deviations, tremors) of the same size (e.g. intensity) is bigger, if they occur more often per time unit (i.e. with a higher rate).

B. Changes since version 1

Besides the removal of two minor bugs (see TREMOR.PRAAT [8] for details) the development of the tremor algorithm since its first publicized version [9] comprises the modularization of the one script into procedures, accompanied by the supply of a start form in order to set the most important arguments. This hopefully facilitates the usage. The second invention pertains to a further method to extract amplitudes per period (besides Praat's built-in function "To AmplitudeTier (period)"): Now the calculation of the root mean square per period can be chosen alternatively. Thirdly, the resampling of these amplitudes at a constant rate, which is needed for a proper autocorrelation of the amplitude contour, is re-formulated.



(a) Frequency tremor analysis

(b) Amplitude tremor analysis

Fig. 1. Steps of tremor analysis: The upper sub-figures show the F_0 and the amplitude (A) contour, the middle sub-figures depict the normalized and de-declined contours with time marks at the tremor maxima. The lower sub-figures show time marks for the tremor minima. Fig. (a) shows the frequency and Fig. (b) the amplitude tremor analysis of a 3 seconds lasting quasi-stationary part of the sustained vowel /a/ from a male speaker (id 19) of the AHN corpus. At the recording time he was 74.24 years old, PD patient, and off medication. FTrF=4.75Hz, FTrI=2.019%, FTrP=1.668, ATrF=4.64Hz, ATrI=6.462%, ATrP=5.316.

C. Used settings

The presented analyses were done using the standard values for required arguments, except for three changes: Firstly, for female speakers the range of the initial pitch analysis was adjusted to 120-450 Hz. Secondly, in order to analyze vowels (within the gender groups) with equally set arguments and to speed up the processing a lot, "Mode" was set to "Run mode" in TREMOR.PRAAT's start form, see [8] – by abandoning the possibility to adjust the algorithm to extraordinary input signals (see also section VI). And thirdly, for amplitude tremor analysis the analysis time step was reduced to 10 ms – just because the re-sampling of the amplitude values at a constant rate requires a little more precision as the frequency tremor analysis, where an initial extraction per period is not needed.

IV. STATISTICAL METHODS

Primarily, we aim to test the influence of PD on vocal tremor. Thus, the main objective is to compare (means of) tremor measures between the group of PD speakers and the control group. A second aim is to test these measures within PD speakers in relation to the presence or absence of medication that is attenuating the first cause of PD, the loss of dopamine. Hence, two different types of analyses are needed, one for independent samples and one for the dependent ones. The effects of the speakers' age and gender shall also be considered. Thus, analyses of covariance (ANCOVAs) are the pertinent statistical procedures, where gender must be considered an observed and fixed factor and age as covariate. In the comparison of measures within PD speakers, medication is a manipulated and fixed within-factor and comparing the PD and the control group, pathology is an observed and fixed grouping-factor.

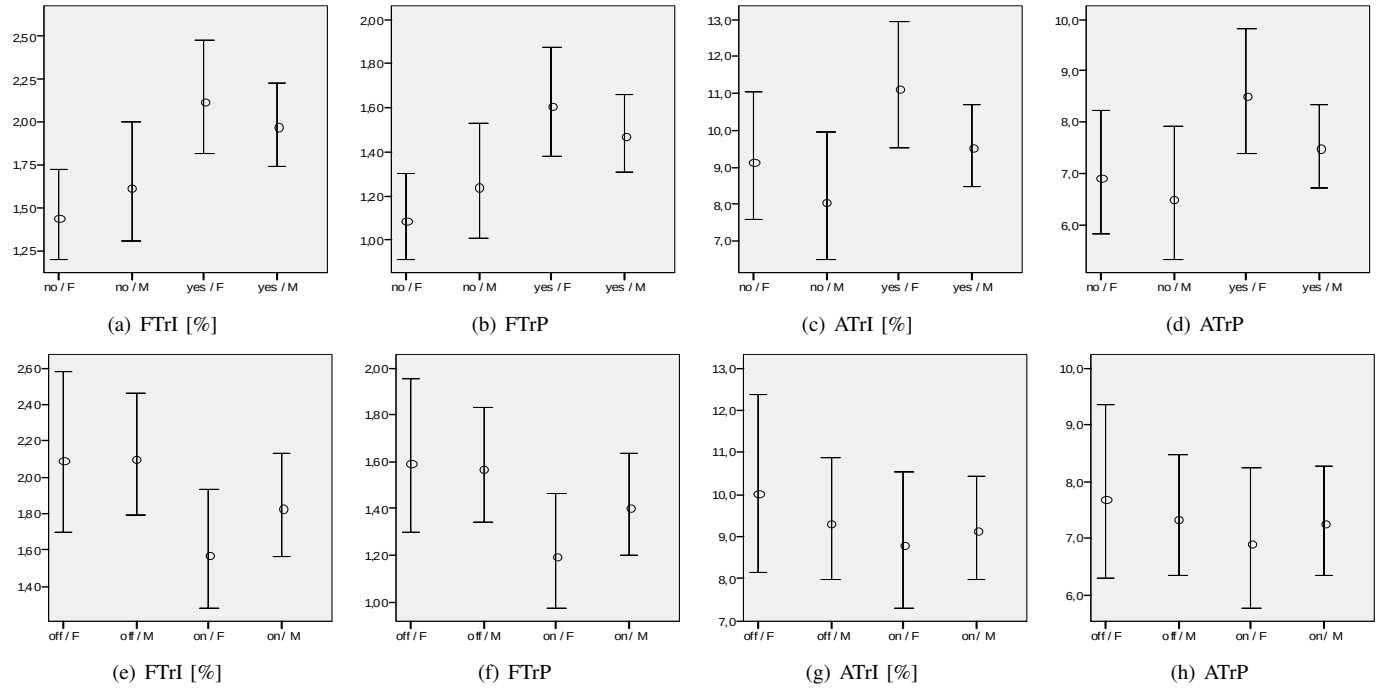
Since it could be argued that six measures of *one* construct "vocal tremor" are to be tested, a multivariate analysis was considered but dismissed: It is not needed, because the measures are sufficiently theoretically independent for separate analyses [11, p. 472]. And it would be counter-productive, since missing values in single measures would result in a total exclusion of cases (vowels), finally resulting in a big loss of data.

Thus, in total 18 univariate ANCOVAs are resulting from six measures (dependent variables) in three types of analyses: (1) control vs. PD group on medication, (2) control vs. PD group off medication, and (3) within PD speakers, off vs. on medication. Since all designs are unbalanced, type 3 square sums are indicated. Only saturated models are used. The tested hypotheses are directional for tremor magnitude measures (and non-directional for tremor frequencies):

H_0: There is no difference in tremor magnitude between PD speakers and the control group respectively between the off and the on medication condition or even lowered tremor values are found in the PD group respectively in off medication condition.

H_1: Raised tremor magnitude values are found in PD speakers respectively in the off medication condition.

Because this is the first attempt to find differences due to PD for measures obtained with TREMOR.PRAAT, a type I error level of α=0.05 seems to be suitable. Since the tremor data are positively skewed, a logarithmic transformation is needed to (better) fit the assumption of normally distributed measures of parametric statistical analysis.



(a) FTrI [%] (b) FTrP (c) ATrI [%] (d) ATrP

(e) FTrI [%] (f) FTrP (g) ATrI [%] (h) ATrP

Fig. 2. Estimated marginal means and their (two-tailed) 95% confidence intervals. The values are re-transformed from logarithmic to original scale. Sub-figures (a)-(d) show results of the analysis between PD speakers (yes) off medication and the control group (no) outlined in Section V-A, Sub-figures (e)-(h) depict results from the analysis within PD speakers, off vs. on medication, see Section V-B. Gender is classified by F and M.

V. Results

Both tremor frequencies FTrF and ATrF are neither influenced by any of the factors nor related to the covariate speaker age in any of the analyses.

Also, a significant impact of gender as well as of all interactions including gender can not be detected. However, as can be gathered from Fig. 2, an effect of gender on the tremor magnitude measures may exist: Female speakers show, by trend, greater differences.

The analyses of the PD speakers under medication and the control group exhibit probabilities for a correct H_0 around chance level for tremor magnitudes due to the factor pathology. But frequency tremor magnitudes correlate most highly significantly ($p < 0.05\%$) and amplitude tremor magnitudes still highly significantly ($p = 0.6\%$) to the covariate speaker age.

A. PD speakers off medication vs. the control group

In contrast, the comparisons of the tremor magnitude measures (intensity and power indices) between the PD speakers off medication and the control group reveal significantly increased values for PD speakers, see Table I and Fig. 2 (a-d). The greater effects are found for frequency tremor measures as compared to amplitude tremor measures and for the power indices as compared for the intensity indices. Though, advanced age as well raises the tremor magnitude measures. Moreover, the effects that can be attributed to the speakers' age roughly attain twice the size of those pertaining to pathology. The effect of PD on vocal tremor even would seem to be greater (e.g. for FTrI it would achieve an $\eta_P^2 = \eta_G^2 = 6.27\%$), if speaker age would not be considered as a covariate. But this can be attributed to PD speakers being slightly older than the control group in combination with the bigger effect of age.

TABLE I. Results of the ANCOVAs comparing the tremor measures between PD speakers off medication and the control group. p denotes the one-tailed probability of committing a type I error, η_P^2 the partial, and η_G^2 the generalized effect size, see [12], [13].

Measure	Effect	p [%]	η_P^2 [%]	η_G^2 [%]
ln(FTrI)	Age	< 0.05	7.34	7.04
	PD	0.05	4.04	3.73
ln(FTrP)	Age	< 0.05	8.18	7.84
	PD	0.05	3.94	3.61
ln(ATrI)	Age	0.15	4.18	4.05
	PD	2.0	1.99	1.89
ln(ATrP)	Age	0.1	4.46	4.34
	PD	1.7	2.12	2.02

B. Within PD speakers, differed by medication

A similar picture emerges from the comparison within PD speakers, differed by medication, see Table II and Fig. 2 (e-h) – at least as long as only the frequency tremor measures are considered, although these differences, too, are very slightly smaller than those seen in the above comparison. In contrast, the amplitude tremor magnitude measures do not seem to be affected by medication. This results in only the frequency tremor measures differing significantly between the off and on medication conditions. "The dramatic decrease in the effect size estimated [from η_P^2 to η_G^2] is a function of counteracting the correlation between the observations." [12]

TABLE II. RESULTS OF THE ANCOVAS COMPARING THE TREMOR MEASURES WITHIN PD SPEAKERS, ON AND OFF MEDICATION.

Measure	Effect	p [%]	η_P^2 [%]	η_G^2 [%]
ln(FTrI)	Age	<0.05	14.11	10.66
	Dopa	0.1	7.27	1.84
ln(FTrP)	Age	<0.05	15.85	12.33
	Dopa	0.1	7.33	1.68
ln(ATrI)	Age	1.5	4.97	3.50
	Dopa	12.75	1.39	0.41
ln(ATrP)	Age	1.3	5.21	3.76
	Dopa	17.6	0.93	0.26

VI. DISCUSSION

So, although the amplitude tremor magnitude measures do differ between the control group and PD speakers off medication, they do not vary within PD speakers due to medication.

The effect sizes can (at least roughly) be interpreted as explained variance (just as R^2 in regression analysis). If one likes to interpret them in an absolute manner, Bakeman [13] can be followed: "Cohen (1988, pp. 413-414), who did not consider repeated measures designs explicitly, defined an η^2 [...] of .02 as small, one of .13 as medium, and one of .26 as large. It seems appropriate to apply the same guidelines to η_G^2 as well." So, the effect of PD on the tremor magnitude measures can be regarded as "small". That the greater part of the variance in tremor magnitude measures can be explained by the speakers' age, seems to be quite remarkable, since age does not vary extensively in the AHN corpus. But on the other hand this is no surprise, because detecting age was the purpose that the used tremor algorithm was initially developed [14] for.

Nevertheless, based on our findings we have to reject the statements of Nebel & Deuschl [2, p. 41] as well as the results from Gillivan-Murphy [3] and confirm the findings of Cnockaert et al. [4] as well as Gillivan-Murphy's speculations based on (insignificant results on) auditory perceived scales: The magnitude measures (the intensity and power indices) of vocal tremor indeed are observable features of PD and may serve to diagnose the disease even better and also in early stadia, if they are combined with other dysphonia measures (see Tsanas et al. [5]) – and if they are measured properly.

Please note that TREMOR.PRAAT is strongly depending on its first step, the pitch analysis, to work accurately. This requires, among other things, appropriate settings of the minimal and the maximal pitch arguments. Also, suitable values for the "tremor octave cost" arguments are needed. So, although the analyses for this study were adjusted to fit the two (depending on gender) most probable signal properties, it is very likely that more valid results could be achieved, if these arguments were set individually for each analyzed sound in the "analysis mode", see [8].

VII. CONCLUSIONS

Amplitude and frequency tremor magnitudes (intensity and power indices) are increased in sustained vowels that are produced by people that are diagnosed with PD and off medication. Frequency tremor magnitudes also differ between the on and off medication conditions. Hence, these measures probably can be used – together with other (vocal) measures and as long as the speakers' age is controlled – to diagnose PD, maybe even in early stadia. However, increased vocal tremor magnitudes, above all, seem to display a lowered neurotransmitter level.

ACKNOWLEDGMENTS

We would like to thank Paul Boersma and David Weenink for inventing and constantly developing PRAAT.

REFERENCES

[1] P. Gillivan-Murphy, N. Miller, "Voice tremor: what we know and what we do not know", in *Current Opinion in Otolaryngology & Head & Neck Surgery,* vol. 19 no. 3, pp. 155-159, 2011.

[2] A. Nebel, G. Deuschl, *Dysarthrie und Dysphagie bei Morbus Parkinson.* Stuttgart: Thieme, 2008.

[3] P. Gillivan-Murphy, *Voice tremor in Parkinson's disease (PD); Identification, characterisation and relationship with speech, voice, and disease variables.* Dissertation, University of Newcastle, retrieved 15 June 2015 from: https://theses.ncl.ac.uk/dspace/bitstream/10443/2170/1/Gillivan-Murphy 13.pdf.

[4] L. Cnockaert et al., "Effect of Intensive Voice Therapy on Vocal Tremor for Parkinson Speakers", in *INTERSPEECH 2007 – 8th Annual Conference of the International Speech Communication Association, August 27-31, Antwerpen, Proceedings,* 2007.

[5] A. Tsanas et al., "Novel speech signal processing algorithms for high-accuracy classification of Parkinson's disease", in *IEEE Transactions on Biomedical Engineering,* vol. 59 no. 5, pp. 1264-1271, 2012.

[6] A. Ghio et al., "How to manage sound, physiological and clinical data of 2500 dysphonicand dysarthric speakers?", *Speech Communication,* vol. 54, no. 5, pp. 664-679, 2012.

[7] P. Boersma, D. Weenink. (2014, Dec.), Praat: doing phonetics by computer. [Computer program, Version 5.4.02, Online]. Available: http://www.praat.org.

[8] M. Brückl, "tremor2.06. A Praat-script to measure vocal tremor [Computer program]" Version 2.06, retrieved 15 June 2015 from http://www.brykl.de/tremor2.06.zip.

[9] M. Brückl, "Vocal tremor measurement based on autocorrelation of contours" in *INTERSPEECH 2012 – 13th Annual Conference of the International Speech Communication Association, September 9-13, Portland (OR), Proceedings,* 2012.

[10] Kay Elemetrics Corp., *Multi-Dimensional Voice Program (MDVP), Model 5105.* [Computer program, Version 2.6.2], 1993/2003.

[11] J. Bortz, C. Schuster, *Statistik für Human- und Sozialwissenschaftler.* Berlin: Springer, 2010.

[12] S. Olejnik, J. Algina, "Generalized Eta and Omega Squared Statistics: Measures of Effect Size for Some Common Research Designs", in *Psychological Methods,* vol. 8 no. 4, pp. 434-447, 2003.

[13] R. Bakeman, "Recommended effect size statistics for repeated measures designs", in *Behavior Research Methods,* vol. 37 no. 3, pp. 379-384, 2005.

[14] M. Brückl, *Altersbedingte Veränderungen der Stimme und Sprechweise von Frauen*. Berlin: Logos, 2011.