Forum Acusticum 2011 - Aalborg - Denmark

Acoustic markers of speech degradation in early untreated Parkinson's disease

Jan Rusz and Roman Cmejla
Department of Circuit Theory, Faculty of Electrical Engineering, Czech Technical University in Prague, Technicka 2, 166 27, Prague 6, Czech Republic

Hana Ruzickova, Jiri Klempir, Veronika Majerova, Jana Picmausova, Jan Roth, and Evzen Ruzicka
Department of Neurology and Centre of Clinical Neuroscience, First Faculty of Medicine, Charles University in Prague, Katerinska 30, 120 00, Prague 2, Czech Republic

Summary
Parkinson's disease (PD) is a neurological illness associated with a variety of motor deficits and non-motor deficits involving areas such as speech, mood, behaviour, thinking, and sensation. The PD-related vocal impairment results from the involvement of various speech subsystems including respiration, phonation, articulation, and prosody. The abnormalities in these speech subsystems are traditionally assessed using several acoustic measurements. Signal processing algorithms, in turn, offer an objective method for precise evaluation of speech performance from recorded signals. In this paper, we study the disordered speech of people with early PD who have not undergone pharmacotherapy treatment. Here, we demonstrate that acoustic measurements can reveal subtle changes in speech and thus significantly separate healthy persons from patients with PD. The various speech data were recorded from 23 people with recently diagnosed PD and 23 healthy control (HC) speakers. PD patients were scored according to the Hoehn and Yahr stages and the Unified Parkinson's Disease Rating Scale. We have found 19 representative measurements that are able significantly to separate both groups of speakers. Compared to HC, the PD speakers show abnormalities in measures of jitter, shimmer, noise-to-harmonics ratios, diadochokinetic rate, clarity and accuracy of articulation, sound pressure level, intensity of speech, melody of speech, number of pauses, and ability to reproduce perceived rhythm. Our findings show that (a) 78% of PD subjects indicate symptoms of vocal impairment that differ from the speech performances of the wider norm of healthy speakers, (b) lowered loudness of speech are significantly correlated with overall severity of disease.

PACS no. 43.70.Dn

1. Introduction

Parkinson's disease (PD) is a chronic neurodegenerative disorder characterized by the progressive loss of dopaminergic neurons in the substantia nigra [1]. PD is the second most common neurodegenerative disorder after Alzheimer's disease [2], and affects 1-2% of people aged over 60 years [3].

Hypokinetic dysarthria in Parkinson's disease is a multidimensional impairment that affects all different aspects of speech such as respiration, phonation, articulation, and prosody [4]. Previous studies have found that approximately 70-90% of patients with PD show some form of vocal impairment [5, 6]. This alteration of speech may also be one of the earliest indicators of disease [7, 8]. Medical treatment, including neuro-pharmacological and neurosurgical methods, alleviates certain symptoms, but there is no causal cure now available, and early diagnosis is critical for maximizing the effect of treatment and improving the quality of patients' life [9, 10]. However, research has shown that medical interventions alone are not as effective for speech treatment as they are for motor symptoms, and the reduced ability to communicate is considered to be one of the most difficult aspects of the PD [11].

As the result, acoustic analysis can provide useful markers for the diagnosis of PD, for monitoring of deterioration or improvement of speech performances in the course of disease, and furthermore, for providing important feedback in voice and speech treatment [12].

The most salient features of PD-related vocal impairment are related to dysphonia, characterized by deficits in the production of vocal sound, and dysarthria, associated with motor speech disorder problems [6]. Nonetheless, patients with PD can manifest abnormalities related to each of the speech dimensions, including reduced loudness, monopitch, imprecise articulation, variability of speech rate, hoarseness, speech disfluencies, reduced stress, and others [13].

Although the investigation in the area of speech and voice disorders associated with PD has been examined at least for 50 years [14], there are only a few studies exploring the extent of vocal impairment in the early stages of PD, when the progression of PD-related speech symptoms is not affected by medication [15]. This weak evidence was our rationale for investigating the signs of PD-related speech degradation through a range of acoustic speech signal processing algorithms [16, 17]. In this study, we thus explore whether the voice and speech disorders are present from the early stages of PD before starting dopaminergic pharmacotherapy, and find the possible correlation of the voice parameters with respect to the global motor impairment (according to the Unified Parkinson's Disease Rating Scale [UPDRS III]), the stage of disease (Hoehn and Yahr [H&Y] Scale), and the duration of disease.

2. Methods

2.1. Participants and data

Data was used from the original study Rusz *et al.* [16], in which a grand total of 46 Czech native speakers were studied. Twenty-three individuals were diagnosed with an early stage of idiopathic PD. All PD patients were examined immediately after the diagnosis was made and before the symptomatic treatment was started. In addition, 23 healthy control (HC) speakers with no history of neurological or communication disorders were included. Table I summarizes the subject details.

The speech data was recorded in a quiet room with a low ambient noise level using an external condenser microphone placed at approximately 15 cm from the mouth. The voice signals were sampled at 48 kHz, with 16-bit resolution. The vocal tasks ranged from producing isolated vowels to reading short sentences and producing a spontaneous monolog about a given subject. In each vocal task, the best speech performances for every subject were retained. Table II details the speech data used.

Table I. Summary of the participants' data.

	PD (23 subjects)	HC (23 subjects)
Male	$n = 19$	$n = 16$
Female	$n = 4$	$n = 7$
Age (year)	61.74±12.60	58.08±12.91
Duration of PD (month)	30.22±22.21	n/a
H&Y stage	1-2	n/a
UPDRS III score	17.52±7.26	n/a

The values are given in the form mean ± standard deviation. Entries labelled 'n/a' are not applicable for HC speakers. H&Y stage represents disability scale comprised of stages 1 through 5, where 5 is most severe. UPDRS III score represents motor rating scaled from 0 to 108, where 108 represents severe motor impairment.

2.2. Measurement methods

In this section, we have selected the major part of traditional clinically used measurement methods for PD-related voice disorders assessment [13]. The selection of acoustical measurements is designed with attention paid to automatic feature extraction and individual subject differences (see Table III for a list of measurements used). The individual measurements used here can be subdivided according to the specific speech dimensions. We have mainly focused on three speech subsystems including phonation (the vibration of the vocal folds to create sound), articulation (the modification of the position and

Table II. Summary of the speech data.

Task no.	Speech data
/1/	Sustained phonation of /i/ on one breath at a comfortable pitch and loudness as constant and long as possible, at least 5-sec.
/2/	Rapid steady /pa/-/ta/-/ka/ syllables repetition on one breath as constant and long as possible, repeated at least 5-times.
/3/	Approximately 5-sec sustained vowels of /a/, /i/, /u/ on one breath at a comfortable pitch and loudness.
/4/	Reading the same standard text of 136 word.
/5/	Monologue, at least approx. 90-sec.
/6/	Reading the same text containing 8 variable sentences of 71 words with varied stress patterns on 10 indicated words.
/7/	Reading 10 sentences according specific emotions in a comfortable voice in response to an emotionally neutral sentence.
/8/	Rhythmically read text containing 8 rhymes of 34 words following the example set by the examinator.

(c) European Acoustics Association, ISBN: 978-84-694-1520-7, ISSN: 221-3767

Table III. Overview of the measurement methods used.

Determined from task	Acoustic measurements	Acoustic measurements description
1, 4-7	F0 SD	Variations of fundamental frequency, vibration rate of vocal folds.
1	Jitter	The average absolute difference between a period and the average of it and its four closest neighbours, divided by the average period.
1	Shimmer	Average absolute difference between the amplitudes of consecutive periods, divided by the average amplitude.
1	NHR	Noise-to-Harmonics-Ratio, the amplitude of noise relative to tonal components.
1	HNR	Harmonics-to-Noise-Ratio, the amplitude of tonal relative to noise components.
4, 5	Percent pause time	The percent change from the unedited sample length to the edited sample length.
4	Articulation rate	Te number of syllables produced per second, after removing silence period exceeding 60 ms.
4, 5	Number of pauses	The number of all pauses compared to total time duration, after removing silence period not lasting more than 60 ms.
4-6	Intensity SD	Variations of average squared amplitude within a predefined time segment ("energy") after removing silence period exceeding 60 ms.
2	DDK rate	The number of /pa/-/ta/-/ka/ syllable vocalizations per second.
2	DDK regularity	The degree of /pa-/ta/-/ka/ syllable vocalizations rate variations in the period.
3	Vowel space area	Quantitative measure which involves plotting the three corner vowels in F1/F2 plane.
8	Rhythm	Measurement of ability to reproduce perceived rhythm through dynamic time warping.
2	SPLD	Sound Pressure Level Decline, the robust linear regression of energy.
2	SDCV	Spectral Distance Change Variations, the variations of spectral distance changes in signal spectrum.
2	RFPC	Robust Formant Periodicity Correlations, the first autocorrelation coefficient of F2 contour.

shape of the speech organs, e.g. tongue, in the creation of sound), and prosody (the variation in loudness, pitch, and timing accompanying natural speech) [18].

2.2.1. The fundamental frequency

The fundamental frequency (F0) is the vibration rate of vocal folds. The measure of F0 variation (F0 SD) was extracted from vocal tasks of sustained phonation /1/ to demonstrate defects in phonation, and reading text /4/, monolog /5/, stress pronouncement /6/, emotions /7/ to demonstrate defects in prosody such as reduced melody of speech. The algorithm supplied in the software package PRAAT [19] and the direct-time domain pitch estimation algorithm [20] were used to obtain the F0 contour.

2.2.2. Perturbation measures and noise-to-harmonics ratios

The perturbation measures of jitter (the extent of variation of voice range) and shimmer (the extent of variation of expiratory flow), and noise-to-harmonics (NHR) and harmonics-to-noise (HNR) ratios (the amplitude of noise relative to tonal components in speech) were obtained using sustained phonation task /1/ to demonstrate defects in phonation. These measures were calculated using PRAAT [19].

2.2.3. Articulation rate and pause characteristics

The measurements of articulation rate, percent pause time, and number of pauses demonstrate defects in prosody, and were calculated for tasks of reading text /4/, while percent pause time and number of pauses were also calculated for the monolog /5/. These measurements were based on formulas and calculated using simple speech-pause detector presented in [16].

2.2.4. Intensity of voice

The measurements of intensity variations (Intensity SD) demonstrate defects in prosody and were determined using the tasks of reading text /4/, monolog /5/, and stress patterns /6/. The automatic intensity contours extraction is described in [16].

2.2.5. Diadochokinetic rate and regularity

The diadochokinetic (DDK) task is the measurement of subject's ability to repeat rapidly and steadily consonant-vowel (C-V) combination. The DDK rate is the number of syllable vocalizations per second. The DDK regularity assesses the ability to maintain a constant rate of C-V combinations. These two measurements demonstrate defects in articulation and were determined from repetition of three-syllable items of /pa/-/ta/-/ka/ task /2/. The automatic extraction of these measures is described in [16].

2.2.6. Vowel space area

The first (F1) and second (F2) formant frequencies were obtained from phonation of three corner vowels /3/ to demonstrate the defects in articulation. The vowel space area was calculated by plotting on an *xy* coordinate plane with F1 on

the *x*-axis and F2 on the *y*-axis. This total area is determined by measuring the entire triangle area. We used a robust formant tracker to obtain the formant sequences [21].

2.2.7. Rhythm

The ability to reproduce perceived rhythm was measured using a rhythmically read text /8/ to demonstrate deficits in prosody, computed as the similarity between subject performance and template recording on the basis of dynamic time warping algorithm. The automatic rhythm parameter extraction is described in [16].

2.2.8. New non-standard measurement methods of articulation

The three novel acoustic measurements of articulation were designed and performed using /pa/-/ta/-/ka/ syllable repetition /2/. As a consequence of rapid steady articulation, several deficits occur in PD-related dysarthria. A detailed description of these new measurements can be found in [16].

Sound pressure level decline (SPLD) measure the ability to maintain the intensity level; the signal envelope from a speech recording is constructed and normalized to the range [0, 100], and then the robust linear regression is applied to compute the intensity level decline.

Robust formant periodicity correlation (RFPC) measures the accuracy of articulation through measurement of the similarity of rate of tongue movement from a consonant into a vowel; e.g. similarity of F2 (second formant) slopes.

Spectral distance change variation (SDCV) quantifies the clarity of articulation using a Bayesian autoregressive change-point detector in the signal spectrum.

2.3. Statistics

Since not all the variables show Gaussian distribution, the non-parametric Wilcoxon signed rank-sum test was used for comparison between PD and HC groups.

To set the best classification performance, we discarded all measures with a statistically insignificant relationship between both groups. Subsequently, we apply the Wald task to the measurements' Gaussian kernel densities to assess separately the relevance of the each individual measure [22]. The Wald task enables making the decision to classify a subject as PD, HC, or "not sure" in case of indecisive situation. This last case occurs when the subject's speech performance matches the extent of speech performance of the wider norm of heatlhy people. In the case that observation matches the performance of the PD group (is classified as PD), the subject is rated by "1" positive point. In the opposite case, where the subject's speech performance matches the intact speech performance, the subject is rated by "-1" negative point. In the case of an indecisive situation, the subject's speech performance is matched by "0" points. To obtain the final results, we calculate the sum of points for each subject. A higher quantity of positive points predicts the greater vocal impairment, while the number of negative points corresponds to the performance of healthy speech production. A comprehensive description of how to applied Wald task and its motivation can be found in [16].

The search for correlations between the acoustic features and duration and severity of PD (according to UPDRS III score and H&Y stage) was performed using the Spearman rank-sum test.

3. Results

Table IV summarizes the means, standard deviations, and statistical significances between PD and HC group for all measurements.

Statistical significances between both groups were found in all measurements of phonation except pitch variations (F0 SD). We can consider that abnormalities in phonation captured by measurements of jitter, shimmer, NHR, and HNR can be clinically interpreted as hoarseness, hypophony, and tremolo.

From articulatory measurements, all measures except DDK regularity and vowel space area show statistical significant differences between PD and HC speakers. Figure 1 shows the result of calculating SPLD, RFPC, and SDCV values for a selected speech signal. As can be seen, the PD speech signal shows a lower ability to maintain the intensity level, which can be caused by weakness in the production of stable airflow from the lungs. The rate and similarity of tongue movement are well represented by the RFPC measure. The higher number peaks in SDCV represent a greater clarity of articulation.

In measures of prosody, the patients with PD show lower melody intonation in all F0 SD measurements and also decreased intensity variations in all intensity SD measurements. This observation can be caused by altered laryngeal tension, decreased breath support, and decreased range of motions. From pause characteristics, only the measurement of the number of pauses shows

(c) European Acoustics Association, ISBN: 978-84-694-1520-7, ISSN: 221-3767

Table IV. List of results of all measurements with mean values, standard deviations, and statistical significances for PD and HC groups.

Measurement	Subjects				Statistics
	PD		HC		
	Mean	SD	Mean	SD	
Phonation					
/1/ *Sustained phonation*					
F0 SD (semitones)	0.46	0.49	0.35	0.23	p = 0.29
Jitter (%)	0.83	0.75	0.32	0.32	**p < 0.01**
Shimmer (%)	7.51	4.97	2.72	2.27	**p < 0.001**
NHR (-)	0.16	0.27	0.02	0.04	**p < 0.01**
HNR (dB)	16.01	7.36	24.02	5.61	**p < 0.001**
Articulation					
/2/ *DDK task*					
DDK rate (syll/s)	6.22	0.63	7.16	0.73	**p < 0.001**
DDK regularity (-)	0.54	0.58	0.67	0.36	p = 0.49
SPLD (1/s)	5.68	2.99	3.85	3.01	**p < 0.05**
RFPC (-)	0.46	0.17	0.60	0.09	**p < 0.01**
SDCV (-)	0.14	0.03	0.18	0.03	**p < 0.001**
/3/ Sustained vowels					
Vowel area (semit.^2)	94.19	29.24	95.10	25.84	p = 0.66
Prosody					
/4/ *Reading text*					
F0 SD (semitones)	1.71	0.66	2.48	0.56	**p < 0.001**
Intensity SD (dB)	5.93	1.05	7.55	1.62	**p < 0.001**
Percent pause time (%)	0.30	0.02	0.29	0.02	p = 0.30
Articulation rate (syll/s)	6.09	0.78	6.09	0.84	p = 0.58
No. of pauses (pauses/s)	3.29	0.67	3.98	0.51	**p < 0.01**
/5/ *Monolog*					
F0 SD (semitones)	1.53	0.32	2.44	0.65	**p < 0.001**
Intensity SD (dB)	7.05	1.41	8.75	1.51	**p < 0.001**
Percent pause time (%)	0.32	0.03	0.31	0.03	p = 0.14
No. of pauses (pauses/s)	3.04	0.83	3.86	0.69	**p < 0.01**
/6/ *Stress patterns*					
F0 SD (semitones)	2.06	0.81	2.78	0.62	**p < 0.01**
Intensity SD (dB)	6.40	1.07	7.84	1.97	**p < 0.01**
/7/ *Emotional sentences*					
F0 SD (semitones)	2.59	0.74	3.82	0.56	**p < 0.001**
/8/ *Rhythmic text*					
Rhythm (-)	2.65	0.55	2.27	0.28	**p < 0.01**

significant differences between both groups. This situation can be indicated by breathiness and the starting time of tongue movement. The patients with PD also show a lower ability to reproduce perceived rhythm.

After pre-processing by removing statistically insignificant measurements and applying the Wald task to 19 remaining representative measures, we have found that 18 patients with PD (78%) indicate some form of vocal impairment that differs from the speech performance of the wider norm of healthy people. None of the HC speakers reached the dysarthric speech performance of people with PD.

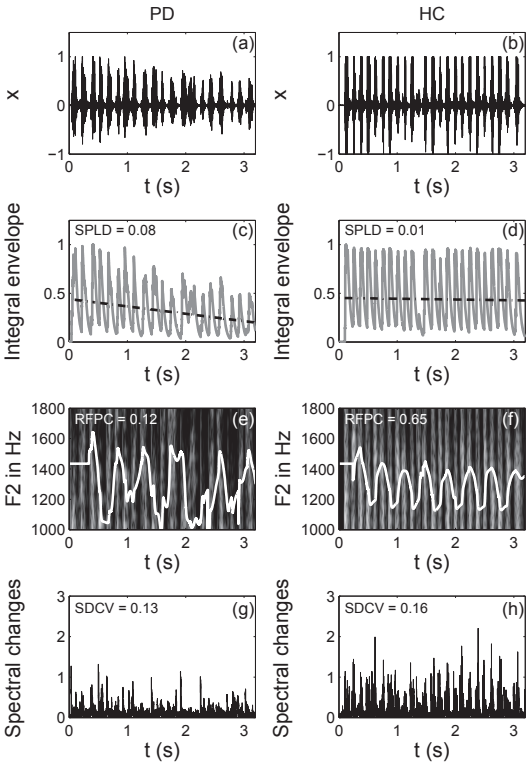

Figure 1. Details of novel articulation measures performed on rapid steady /pa/-/ta/-/ka/ syllables repetition. (a, b) Speech signals; (c, d) Light gray lines represent obtained signal envelopes, dashdot lines represent the SPLD; (e, f) RFPC; (f, g) SDCV. The left panel is for a person with PD, the right panel is for a HC subject.

We did find significant correlations between the measurements of intensity SD and: H&Y stage (for reading text: R = -46.80, $p < 0.05$; for stress patterns: R = -55.18, $p < 0.01$), UPDRS III score (for reading text: R = -57.98, $p < 0.01$; for stress patterns: R = -51.07, $p < 0.05$), and duration of disease (for reading text: R = -49.18, $p < 0.05$; for stress patterns: R = -45.96, $p < 0.05$). Subsequently, we have also found a correlation between measurement of rhythm and UPDRS III score (R = 48.12, $p < 0.05$). Therefore, on the basis of several correlations between reduced intensity variations and global motor impairment, we can assume that defects in reduced loudness of speech are related to the severity of the disease.

4. Conclusion

Our main finding is that 78% of recruited PD patients show some form of vocal impairment from early stages of the disease. Moreover, these PD-related voice and speech disorders can be reliably captured by designed acoustic measurements. Through two independent reading

tasks including reading text and stress patterns, we have noticed that reduced loudness of speech is correlated to global motor impairment, stage of disease, and duration of disease, indicating greater speech communication problems in the more disabled patients.

In addition, the knowledge of incomplete vocal fold closure, lack of lung pressure, and lower articulation clarity and accuracy as a consequence of difficult articulation of rapid syllables repetition has inspired us to design new articulation measures that are gaining significance in speech and voice disorders assessment.

The results of this paper can ease the clinical monitoring of voice and speech disorders progression as well as the effects of medication on speech production, and can serve as important feedback in voice treatment.

Acknowledgement

This project has been funded by the Czech Science Foundation, project GACR 102/08/H008, Czech Ministry of Health, project NT 11331-6/2010, Grant Agency of the Czech Technical University in Prague, project SGS 10/180/OHK3/2T/13, and Czech Ministry of Education, projects MSM 0021620849 and MSM 6840770012.

References

[1] O. Hornykiewicz: Biochemical aspects of Parkinson's disease. Neurology Suppl 2 51 (1998) S2-S9.

[2] M. C. de Rijk, L. J. Launer, K. Berger, M. M. B. Breteler, J.-F. Dartigues, M. Baldereschi, L. Fratiglioni, A. Lobo, J. Martinez-Lage, C. Trenkwalder, A. Hofman: Prelevance of Parkinson's disease in Europe: a collaborative study of population-based cohorts. Neurology 54 (2000) 21-23.

[3] C. D. Marsden: Parkinson's disease. Journal of Neurology, Neurosurgery, and Psychiatry 57 (1994) 672-681.

[4] F. L. Darley, A. E. Aronson, J. R. Brown: Differential diagnostic patterns of dysarthria. Journal of Speech and Hearing Research 12 (1969) 246-269.

[5] A. K. Ho, R. Iansek, C. Marigliani, J. Bradshaw, S. Gates: Speech impairment in large sample of patients with Parkinson's disease. Behavioural Neurology 11 (1998) 131-137.

[6] J. A. Logemann, H. B. Fisher, B. Boshes, E. R. Blonsky: Frequency and coocurence of vocal tract dysfunction in the speech of a large sample of Parkinson patients. Journal of Speech and Hearing Disorders 43 (1978) 47-57.

[7] J. R. Duffy: Motor Speech Disorders: substrates, differential diagnosis and management. New York: Mosby, 2nd ed., 2005.

[8] B. Harel, M. Cannizzaro, P. J. Snyder: Variability in fundamental frequency during speech in prodromal and incipient Parkinson's disease: A longitudinal case study. Brain Cognition 56 (2004) 24-29.

[9] N. Singh, V. Pillay, Y. E. Choonara: Advances in the treatment of Parkinson's disease. Progress in Neurobiology 81 (2007) 29-44.

[10] E. Tolosa, C. Craig, J. Santamaria, Y. Compta: Diagnosis and the premotor phase of Parkinson disease. Neurology 72 (2009) 12-20.

[11] M. Trail, C. Fox, L. O. Ramig, S. Sapir, J. Howard, E. C. Lai: Speech treatment for Parkinson's disease. Neurorehabilitation 20 (2005) 205-221.

[12] B. T. Harel, M. S. Cannizaro, H. Cohen, N. Reilly, P. J Snyder: Acoustic characteristic of Parkinsonian speech: a potential biomarker of early disease progression and treatment. Journal of Neurolinguistics 17 (2004) 439-453.

[13] A. M. Goberman, C. Coelho: Acoustic analysis of Parkinsonian speech I: Speech characteristics and L-Dopa therapy. Neurorehabilitation 17 (2002) 237-246.

[14] G. J. Canter: Speech characteristics of patients with Parkinson's disease: I. Intensity, pitch, and duration. Journal of Speech and Hearing Disorders 28 (1963) 221-229.

[15] F. J. Jimenez-Jimenez, J. Gamboa, A. Nieto, J. Guerrero, M. Orti-Pareja, J. A. Molina, E. Garcia-Albea, I. Cobeta: Acoustic voice analysis in untreated patients with Parkinson's disease. Parkinsonism & Related Disorders 3 (1997) 111-116.

[16] J. Rusz, R. Cmejla, H. Ruzickova, E. Ruzicka: Quantitative acoustic measurements for characterization of voice and speech disorders in early untreated Parkinson's disease. Journal of the Acoustical Society of America 129 (2011) 350-367.

[17] J. Rusz, R. Cmejla, H. Ruzickova, J. Klempir, V. Majerova, J. Picmausova, J. Roth, E. Ruzicka: Acoustic assessment of voice and speech disorders in Parkinson's disease through quick vocal test. Movement Disorders (2011) in press.

[18] A. M. Goberman: Correlation between acoustic speech characteristics and non-speech motor performance in Parkinson disease. Medical Science Monitor 11 (2005) 109-116.

[19] P. Boersma, D. Weenink: PRAAT, a system for doing phonetics by computer. Glot International 5 (2001) 341-345.

[20] H. Boril, P. Pollak: Direct time domain fundamental estimation of speech in noisy conditions. In: EUSIPCO Proceedings, Vienna, Austria, 2004.

[21] K. Mustafa, I. C. Bruce: Robust formant tracking for continuous speech with speaker variability. IEEE Transactions on Audio, Speech, and Language Processing 14 (2006) 435-444.

[22] A. Wald: Sequential analysis. New York: Wiley, 1947.