# LABORATORY TASKS FROM VOICE ANALYSIS IN THE STUDY OF BIOMEDICAL ENGINEERING USING MATLAB

*R. Čmejla[1], J. Rusz[1,2], L. Vokřál[2]*

[1] Czech Technical University in Prague, Faculty of Electrical Engineering
[2] Charles University in Prague, First Faculty of Medicine

**Abstract**

**Students of Biomedical Engineering of the Czech Technical University in Prague meet with the laboratory task "Voice Analysis" in the subject Biological Signals. They have the opportunity here to practically learn basics about fundamental glottal tone, formant frequencies, their parameters and detection, frequently used in voice pathology analysis.**

## 1 Introduction

The archival databases of anonymous speech and voice recordings are used in laboratory tasks, but students have the opportunity for analysis of own voice also.

Because the value of the frequency fluctuations of glottal tone (absolute jitter) reaches of the pathological values already in one-tenth of milliseconds, voice recording acquisition and analysis require a higher sampling frequency. The value of amplitude fluctuations (shimmer) is very sensitive to background noise also (for example fan noise presents in laboratory room when twenty students working on twenty sites). The level of background noise in standard computer laboratory does not allow measuring of those characteristics. Direct connection of a microphone or headset to the computer therefore does not allow acquisition of high-quality recording. But using the portable digital recording device, the quality recording is easily obtained in any office room, where there is no presence of more number of people and running computers (the required level of noise for these measurements is 45 dB). Digital Tape Recorder Edirol R09HR was chosen with regard to portability with high-quality uncompressed recording capabilities. Head microphone allows clear definition of the distance away from mouth. Frequency response of a condenser microphone Bayerdynamic Opus 55 MkII is linear in the low frequencies below 100 Hz therefore the microphone is fundamentally differed from conventional headsets. The digital audiometer Voltcraft SL-400 is used for a simple dynamic measurement of voice level. Audiometer measures the sound level from 30 dB to 130 dB with an accuracy of + / - 1,4 % and has a USB connection.

## 2 Methods

### 2.1 Voice analysis in time domain

The changes of vocal cords vibration are perceived as changes of pitch period or fundamental frequency in voice. The fundamental frequency is influenced by properties of the vocal cords (their elasticity, weight and length). Thus the fundamental frequency and consequently derived instability measurements provide information about the condition and functionality of the vocal cords. Abnormal characteristics of the fundamental frequency are found in a number of voice and speech disorders [1].

Pitch period is read from waveform in time domain in MATLAB (Fig.1). It is simple and easy detection of pitch period using available threshold. Subseqently, the mean of the fundamental frequency and pitch period is computed. Measurements are performed for all Czech long sustained vowels (at least five seconds).

Minimum and maximum length of pitch period $T_0$ and the fundamental frequency $f_0$ (frequency range) is determined from the scale of song vowel /a/.

Frequency and amplitude instability (jitter and shimmer, see Fig.1 and Fig.2) are derived from the instantaneous values of pitch periods $T_0$ [2].

Equation (1) gives the average absolute difference *Jita* between consecutive periods

$$Jita = \frac{1}{N-1}\left[\sum_{i=1}^{N-1}\left|T_i - T_{i+1}\right|\right]. \tag{1}$$

The value 83 μs of *Jita* is used as a threshold for pathology.

The average absolute difference *J%* between consecutive periods, and divided by the average period is given by equation (2).

$$J\% = \frac{\frac{1}{N-1}\left[\sum_{i=1}^{N-1}\left|T_i - T_{i+1}\right|\right]}{\frac{1}{N}\sum_{i=1}^{N}T_i}\cdot 100 \tag{2}$$

The value 1.040% of *J%* is a threshold for pathology.

The relative average perturbation *RAP*, the average absolute difference between a period and the average of two neighbours, divided by the average period.

$$RAP = \frac{\frac{1}{N-2}\left[\sum_{i=2}^{N-1}\frac{T_{i-1}+T_i+T_{i+1}}{3} - T_i\right]}{\frac{1}{N}\sum_{i=1}^{N}T_i}\cdot 100 \tag{3}$$

A threshold for pathology is *RAP* = 0.680%.

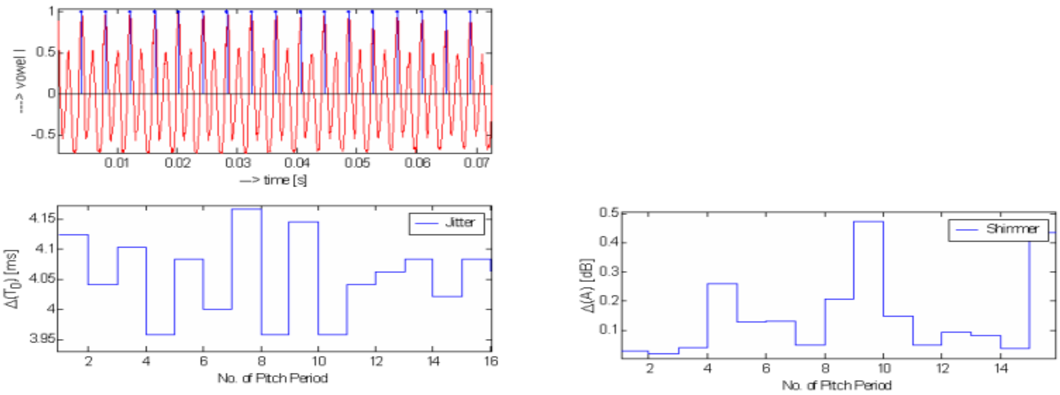


Figure 1: The red waveform in the top of the figure shows part of the sustained vowel and blue lines represent the glottal closures. The waveforms bellow represents frequency and amplitude instabilities at moments of glottal closures (jitter (left), shimmer (right)).

## 2.2 Voice representation in frequency domain

The view in the frequency domain is very important in speech analysis because pathological voices show a significant increase of energy in higher frequencies. Students will learn in this task with differences between broadband and narrowband spectrograms with regard to the fundamental frequency [3]. They will work with different parameters of the FFT and LPC spectral analysis.

## 2.3 Fundamentals frequency and formant detection alghoritms

Fundamental tone is analyzed using simple robust algorithms based on autocorrelation function and formant frequencies are obtained using AR modeling in MATLAB [4], see Fig.2.

Students also display and evaluate phonetogram (the dynamic range of the voice in terms of fundamental frequency and intensity).

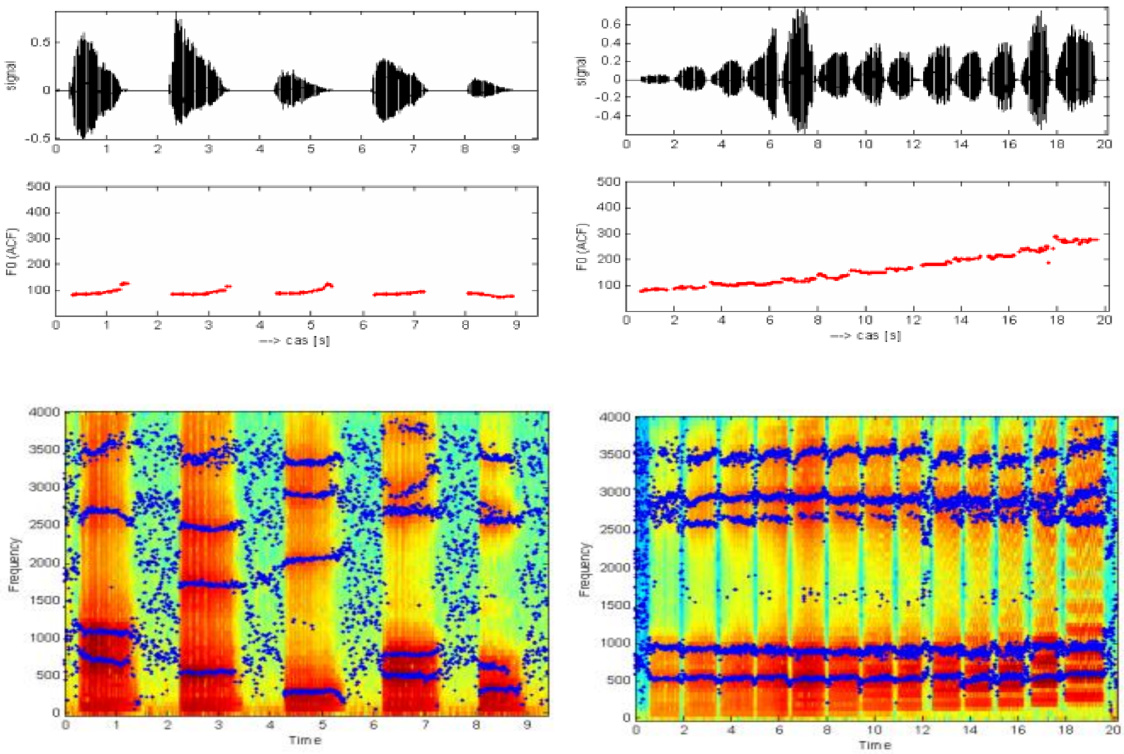


Figure 2: The left column displays the Czech vowels; in the right column there is the singing vowel /a/ (in a musical scale). The waveforms in time domain are in upper figures; the results of the fundamental frequency detection using algorithm based on autocorrelation function are in the middle figure and the results of formant detection using LPC are in the bottom.

## 2.4 Voice parameters in vowel synthesis

The signals used in the previous tasks are studied in terms of production and perception of vowels using Klatt cascade formant synthesizer. [5].

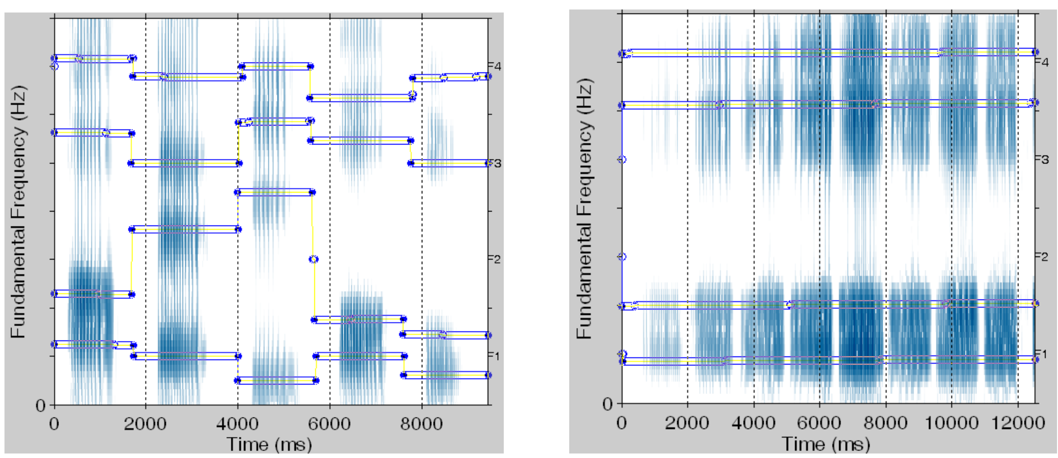


Figure 3: Left: formant tracks for Czech vowel synthesis; right: formant tracks for /a/ in musical scale

## 3 Conclusion

The measured dynamic and frequency values of voice and the amplitude and frequency instability (jitter and shimmer) are comparable with the results acquired using the professional and therefore much more expensive equipments. Cost of used hardware is approximately 12 thousand CZK.


## Acknowledgement

This work has been primarily supported by the grant FRV F3a 316 Audio Signal and Voice Analysis. Study has been also supported by the research programs MSM6840770012 Transdisciplinary Research in Biomedical Engineering, MSM 0021620806 Studies at the molecular and cellular levels in normal and in selected clinically relevant pathologic states, and grants GAČR102/08/H008 Analysis and modelling biomedical and speech signal, SGS10/180/OHK3/2T/13 Voice and Speech Disorders Evaluation


## References


[1] Rusz, J. - Čmejla, R.: *Speech Rate and Intensity Analysis in Parkinson's Disease*. Akustické listy. 2008, roč. 14, č. 2-4, s. 13-16. ISSN 1212-4702. (in Czech).

[2] Boersma, P.: *Should jitter be measured by peak picking or by waveform matching*? Folia Phoniatrica et Logopaedica Vol. 61, No. 5, 2009.

[3] Čmejla, R. - Sovka, P.: *Introduction in DSP*. Prague: CTU, 2005. 108 s. ISBN 80-01-03158-6. (in Czech).

[4] Uhlíř, J. - Sovka, P. - Pollák, P. - Hanžl, V. - Čmejla, R.: *Technology of Speech Communication*. Prague, CTU, 2007. 276 s. ISBN 978-80-01-03888-8. (in Czech).

[5] Assman, P.: http://www.utdallas.edu/~assmann/TRACKDRAW/trackdraw.html


---


Roman Čmejla
Czech Technical University in Prague, Faculty of Electrical Engineering, Department of Circuit Theory, Czech Republic
*e-mail*: cmejla@fel.cvut.cz

Jan Rusz
Czech Technical University in Prague, Faculty of Electrical Engineering, Department of Circuit Theory, Czech Republic
Charles University in Prague, First Faculty of Medicine, Department of Pathological Physiology, Laboratory of Biocybernetics, Czech Republic
*e-mail*: rusz.mz@gmail.com

Jan Vokřál
Charles University in Prague, First Faculty of Medicine, Department of Phoniatrics, Czech Republic
*e-mail*: vokral@LF1.cuni.cz