

AKUSTICKÉ LISTY

České akustické společnosti
www.czakustika.cz

ročník 17, číslo 3

říjen 2011

Obsah

Pozvánka na Valnou hromadu	3
Production of a Train of Acoustic Pulses with a High Repetition Rate Generování posloupnosti akustických impulzů s vysokou opakovací frekvencí <i>Pavel Valoušek</i>	4
Měření dynamického rozsahu lidského hlasu Measurements of Dynamic Range of Human Voice <i>Hana Šrámková, Svante Granqvist, Tomáš Füst, Jan G. Švec</i>	11

Rada České akustické společnosti svolává ve smyslu stanov

VALNOU HROMADU,

kteřá se bude konat ve čtvrtek 26. ledna 2012 na Fakultě elektrotechnické ČVUT, Technická 2, Praha 6 – Dejvice.

Rámcový program:

13.00–13.45 Jednání v odborných skupinách. Rozpis místností pro jednání v odborných skupinách bude vyvěšen ve vstupním prostoru fakulty a na dveřích sekretariátu, místnost č. T2:B2-47.

13.45–14.15 Prezentace.

14.15–16.00 Plenární zasedání, místnost č. T2:C2-82.

Důležité upozornění: Člen společnosti, který se nebude moci Valné hromady osobně zúčastnit, pověří jiného člena, aby jej zastupoval. Jeden člen společnosti může zastupovat nejvýše tři členy. Formulář pověření je součástí tohoto čísla Akustických listů.

Production of a Train of Acoustic Pulses with a High Repetition Rate

Pavel Valoušek

ČVUT, Fakulta elektrotechnická, Technická 2, 166 27 Praha 6
e-mail: valousp@feld.cvut.cz

The goal of this paper is to find a reliable method for producing a sequence of isolated acoustic pulses. Pulses in a train represent bits of digital information for the purpose of signal processing or data transmission in the acoustic domain. The requirements for a high repetition rate and specific pulse shape are given by the sampling frequency of the common sources of the digital audio signal and specific digital modulation. Solitary pulses and pulse sequences suitable for application of a digital loudspeaker are described in this paper, as well as simulation and experimental methods of their production. Pulses are shaped by the compensation filter in order to eliminate the distortion introduced by the electro-acoustic transducer. Simulations and experiments were conducted for unipolar and derivative bipolar pulses. The length of the pulse period was found to be fully dependent on the frequency bandwidth of the electro-acoustic transducer and the shortest simulated pulse period was approximately 50 μs . Experimental results obtained with a dedicated digital processing unit show good correspondence with simulations and the fastest pulse repetition rate was about 12 kHz.

1. Introduction

The production of acoustic impulses has always been a task with only limited usefulness. Applications included currently outdated measurements in architectural acoustics, obtaining an impulse response for various acoustic systems etc. These direct methods were later replaced by more sophisticated ones, such as MLS, and were not developed much further.

1.1. Digital Loudspeaker

The research of transducers with direct digital-to-analog conversion [1] raised the need for definition of digital information in the acoustic domain. Such a definition is not historically or technically determined in any way. The term “digital loudspeaker” denotes a device performing digital-to-analog conversion of an audio signal usually from the electrical to the acoustical domain. Typically, it consists of a transducer array in which each transducer element is driven by a digital bit-stream with a specific binary weight, according to the driving scheme, digital modulation and binary coding [2, 3].

1.2. Digital Information in Acoustic Domain

Regardless of the principle, the transducers of the digital loudspeaker are usually driven by a square-shaped digital signal, and their output is more or less distorted. If a digital loudspeaker is to pass the digital information into the air for further processing in the acoustic domain, it has to convert electric energy quanta into acoustic ones [4].

2. Acoustic Pulse Approximation

The digital bit stream, generally observed as a sequence of energy pulses, retains this fundamental general nature in the acoustic domain as well. The single bit could be modeled as a quantum of acoustic energy taking the form of a short and well-controlled acoustic pulse.

An acoustic pulse representing single-bit information has to have a duration shorter than the sampling period in order to prevent energy overlap between adjacent sampling periods. The shape of acoustic pulses now becomes important. While a square-shaped pulse is a suitable approximation of digital information in the electric domain, the physical properties of fluids in the context of generation and propagation of acoustic waves do not allow such an approximation. Furthermore, it is necessary to consider the abilities of current electro-acoustic transducers.

2.1. Windowed Pulse

Let us assume a pulse that is created from one of the standard window functions [7,8]. To be suitable for pulse shaping applications, the function or pulse should have the following properties.

- Maximum attenuation of side lobes from the main lobe in the spectrum.
- A zero-ended and non-negative time waveform.
- The highest energy possible.

Window functions that generally comply with these frequency and time domain requirements are listed in Table 1.

Window function
Gaussian
Nuttall
Blackman-Harris
Blackman-Nuttall
Dolph-Chebyshev

Table 1: Window functions suitable for pulse shaping

The Gaussian function, as an example from all suitable window functions, can be produced by the time-based Gaussian filter

$$h(t) = \frac{\sqrt{\pi}}{a} e^{-\frac{\pi^2 t^2}{a^2}}. \quad (1)$$

It may be useful to produce a bipolar acoustic pulse with no DC component. This could be achieved by derivation of the Gaussian pulse

$$h(t) = -\frac{2\pi t}{a^3} e^{-\frac{\pi^2 t^2}{a^2}}. \quad (2)$$

The discrete-time plot of a Gaussian pulse and its derivation with a length of 65 samples, both including the spectrum, is shown in Figure 1 and 2 respectively.

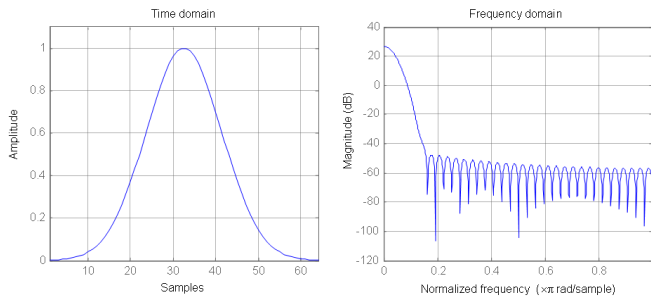


Figure 1: Time-plot (left) and spectrum (right) of the Gaussian pulse

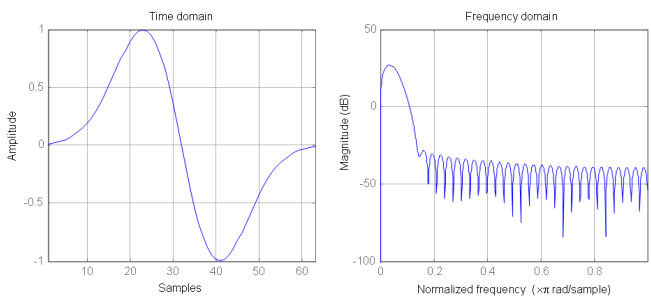


Figure 2: Time-plot (left) and spectrum (right) of derivation of the Gaussian pulse

2.2. Pulse Shaping

To achieve short acoustic pulses with a high repetition rate using standard moving-mass electro-acoustic trans-

ducers, it is necessary to compensate the driving signal for their impulse response. These techniques were already studied in the past with a view to the development of acoustic impulse methods [5, 6]. Basically, there are two main techniques: the step and the filtering techniques. The first technique can be used for transducers with dominant damped self-oscillations. This is not the case of common high-frequency transducers, and consequently the filtering technique is used here.

The main idea of the filtering method [5] is to place a compensating filter into the signal path, in order to compensate for the distortion introduced by the electro-acoustic transducer's impulse response. This filter can be theoretically placed before or after the transducer, but it would be naturally assumed to be placed in the electric domain, and implemented on the digital level.

The transducer's impulse response $h(t)$ is a response to the Dirac impulse

$$\delta(t) * h(t) = h(t). \quad (3)$$

The driving signal of the digital loudspeaker is a pulse representing the "logic one" level in one sampling period. When the sampling period is sufficiently long or the duty cycle is significantly reduced, then the driving signal can always be considered as a sequence of solitary Dirac impulses. The filter $g(t)$ that should compensate for $h(t)$ is again driven by a Dirac impulse

$$\delta(t) * g(t) * h(t) = \delta(t). \quad (4)$$

The Fourier transform turns the Dirac impulse into constant, and it comes out that $g(t)$ is just the inverse filter to $h(t)$. Unfortunately, the stability of $H(f)$ does not necessarily imply the stability of $G(f)$. When $H(f)$ is not a minimum-phase system, the de-convolution needed to achieve $g(t)$ is not solvable. Instead of a Dirac impulse usually masked as a short square pulse, it would be better to obtain the acoustic pressure with a pulse waveform that has favorable spectral properties. For the following considerations, let's again assume a Gaussian pulse as our example. Equation (4) turns into the term below, where $q(t)$ denotes a Gaussian pulse

$$\delta(t) * g(t) * h(t) = q(t). \quad (5)$$

The Fourier transform directly leads to an expression of $G(f)$

$$G(f) = \frac{Q(f)}{H(f)}. \quad (6)$$

For feasible de-convolution, the expression is further extended by complex conjugation $H^*(f)$ and the smallest possible constant k , to ensure stability and target precision of the expression

$$G(f) = \frac{Q(f)H^*(f)}{H(f)H^*(f) + k}. \quad (7)$$

Once $g(t)$ is determined, standard digital filtering techniques can be used for simulations, as well as for practical implementation.

3. Simulations

The purpose of these simulations is to obtain the shortest possible acoustic pulse with a suitable window shape, and consequently the fastest pulse repetition frequency, with good isolation between adjacent pulses. The calculated and optimized pulse lengths and repetition frequencies will be used later for experimental confirmation.

3.1. Transducer’s Impulse Response

To achieve high pulse repetition rates, electro-acoustic transducers with extended frequency range were selected and measured:

- o Vifa OX20SC00-04 textile dome type tweeter [9]
- o Dali hybrid tweeter module – textile dome transducer
- o Dali hybrid tweeter module – ribbon transducer

The impulse responses of all three transducers were measured using a B&K 1/4” microphone, a B&K microphone preamplifier and EASERA acoustic test software at sampling frequency 192 kHz. The microphone was located on the axis of the transducers and the entire system was placed in an acoustically damped chamber.

3.2. Spectrum Fitting

The method of estimating the optimal pulse length can be described as fitting the spectrum image of the pulse with the spectrum image of the transducer’s impulse response. Its low frequency part, where the transducers have significant attenuation, has to be compensated in order to match the spectrum of a short pulse. The length of the pulse influences the location of the spectral main lobe. The length of the pulse is varied from 50 samples down to 5 samples at a sampling frequency of 192 kHz. One period time duration is

$$T_{s192k} = 5.21 \mu s . \tag{8}$$

The pulse length along the time scale is then approximately between 260 μs and 26 μs. Figure 3 shows the frequency response of the Vifa tweeter, together with the spectrum of the Gaussian pulse of 50, 25, 10 and 5 samples

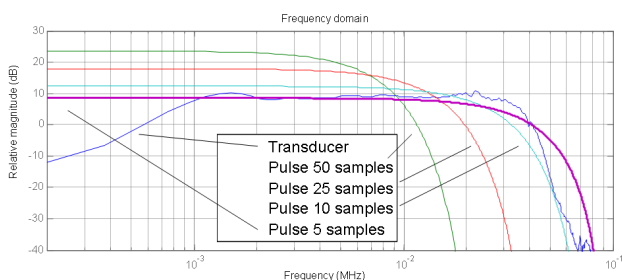


Figure 3: Spectrum of the Vifa tweeter and the Gaussian pulse

length. Figure 4 shows the situation for the Dali ribbon tweeter and derivation of the Gaussian pulse.

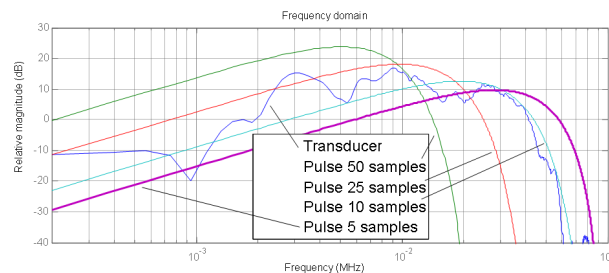


Figure 4: Spectrum of the Dali ribbon tweeter and derivation of the Gaussian pulse

Not all combinations of pulse shapes, lengths and transducer responses are introduced here, as general conclusions can be deduced from the spectral curves in Figures 3 and 4. The best spectral fit and transducer bandwidth utilizations are achieved by pulses with a length of around 10 samples, which corresponds to a time duration of approximately 50 μs.

When this single pulse duration is considered as one period of the digital driving signal-sampling frequency, then theoretically the achievable pulse repetition frequency reaches roughly 20 kHz.

3.3. Compensation Signal

The compensation signals are now calculated according to Eq. (7). These signals are expected to contain indispensable low frequency components. Therefore, both the impulse response of a transducer and the desired pulse are padded into a 1 kS long vector to obtain a compensation signal of the same length. The coefficient k was experimentally determined and kept for all following simulations equal to the value of 10⁻⁵ (-).

Figure 5 shows the spectrum of the Vifa tweeter, the Gaussian pulse with the duration of 10 samples and the compensation signal derived in order to produce such a Gaussian pulse with the Vifa tweeter.

The spectrum of the compensation signal has virtually perfect symmetry with the transducer frequency response

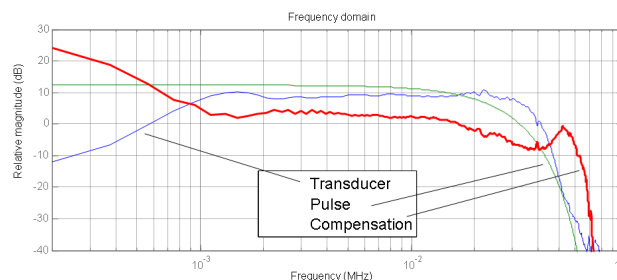


Figure 5: Spectrum of the Vifa tweeter, pulse and compensation signal

around the pulse spectrum curve. The compensation signal has 0 dB transfer at the crossing point of the transducer response and the pulse.

3.4. Single Pulse

The calculated compensation signal is then convolved with the transducer impulse response in order to verify the algorithm.

The pulse produced by convolution is shown in Figure 6, together with the pulse origin in the time and frequency domain. The time waveforms are shifted from each other by 20 samples for better visibility.

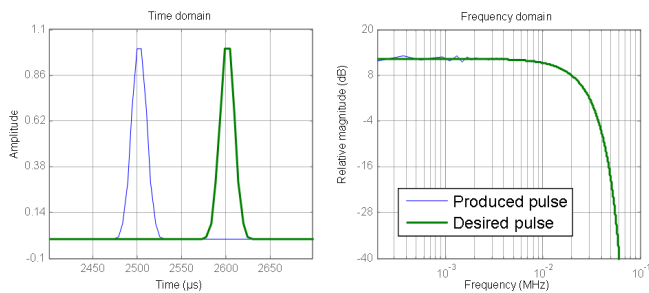


Figure 6: Time-plot and spectrum of the produced pulse and the desired one

The spectra in Figure 6 show an almost perfect match except for small variations in the lowest frequency region. Virtually identical simulation results can be achieved with all three evaluated transducers and all pulse types.

3.5. Pulse Sequence

A signal processing chain for the simulation of pulse repetition was designed in Simulink. Its scheme is shown in Figure 7 and it consists of one pulse generator and two filter stages. The first one has the response of the compensating signal derived above, and the second filter simulates the electro-acoustic transducer by its impulse response.

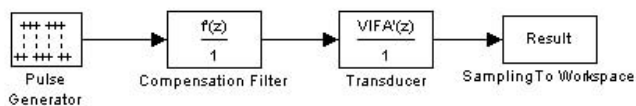


Figure 7: Simulation scheme for pulse repetition

The pulse generator has a period of 50 samples and 1 sample active pulse at the sampling frequency of 192 kHz. Signal waveforms after passing the compensation filter and the simulated transducer are shown in Figures 8 and 9 respectively. The output signal of the transducer with the pulse generator set to a period of 10 samples is shown in Figure 10.

As expected, the 10-sample pulse period can produce a train of pulses with a repeating frequency of 19.2 kHz

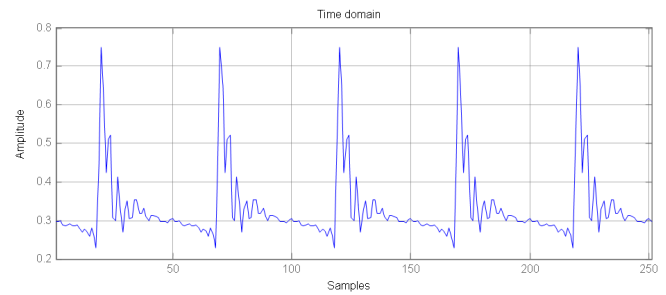


Figure 8: Signal waveform after the compensation filter, pulse period 50 samples

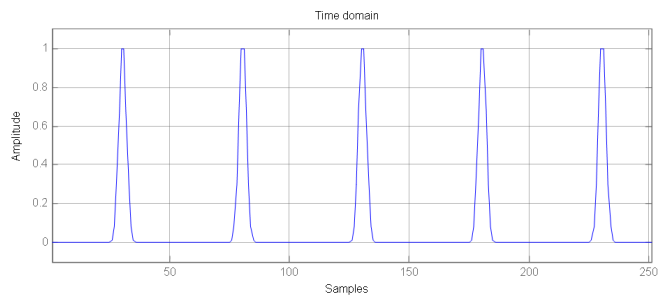


Figure 9: Signal waveform after the transducer, pulse period 50 samples

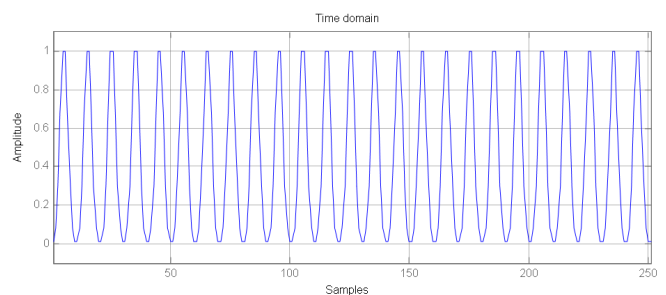


Figure 10: Signal waveform after the transducer, pulse period 10 samples

without visible distortion. On the digital logic level, we achieved a sequence of logic one bits, while the 50 samples pulse period can be regarded as a five-bit sequence with one bit at logic one level and four bits at zero level.

4. Measurements

The simulation results predict the feasible implementation of a digital loudspeaker model based on the idea of transferring complete digital information into the acoustic domain. A signal processing unit was designed and built in order to verify the key principles, such as pulse compensation and shaping.

The unit is based on a fully programmable signal processing FPGA core, two high-speed DACs for pulse shaping and power stage with analog bandwidth exceeding

1 MHz at full power. S/PDIF and USB interfaces provide digital audio input and remote control capability.

The FPGA design includes a 1024-coefficient FIR engine with re-loadable coefficients, a pulse generator and a pulse train modulator. All internal parameters such as DACs reference voltage, DC offset compensation, digital zero for bipolar signals, digital gain and others are controllable via USB interface. A photo of the assembled experimental unit is shown in Figure 11.

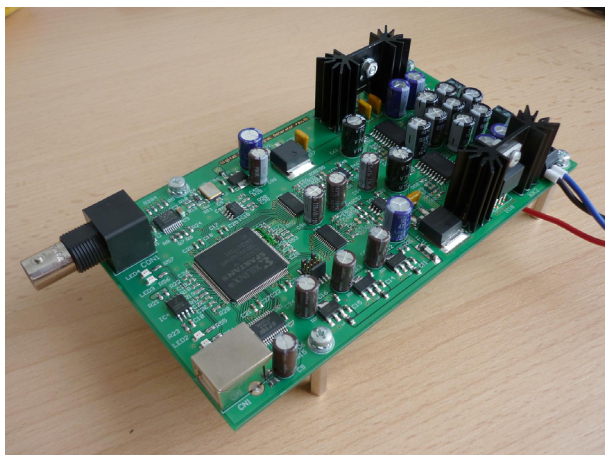


Figure 11: Photo of the experimental unit

The scheme of the experimental measurement setup is shown in Figure 12. The electro-acoustic transducer under measurement and the microphone are placed into an acoustically damped chamber. The driving signal for the transducer is switchable between the amplified line-out from the PC, and the direct output of the experimental unit. This made it possible to measure the transducer’s impulse response or the pre-compensated signal from the experimental unit.

All simulations and recordings were performed at a sampling rate of 192 kHz. The recorded waveforms were imported back to Matlab for analysis.

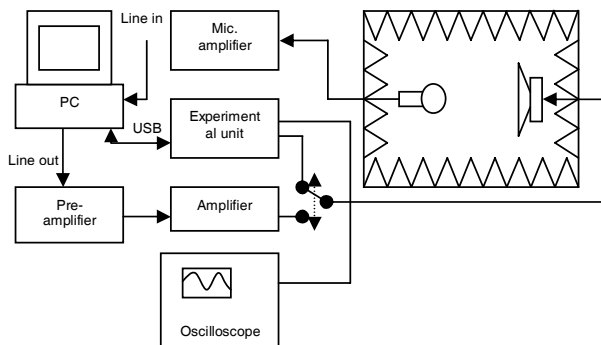


Figure 12: Block scheme of the measurement setup

4.1. Single Pulse

The compensation signal was calculated for the Vifa tweeter and for a Gaussian pulse with lengths of 50 and 20 samples. The recorded acoustic signal with its spectrum for each length is shown in Figures 13 and 14.

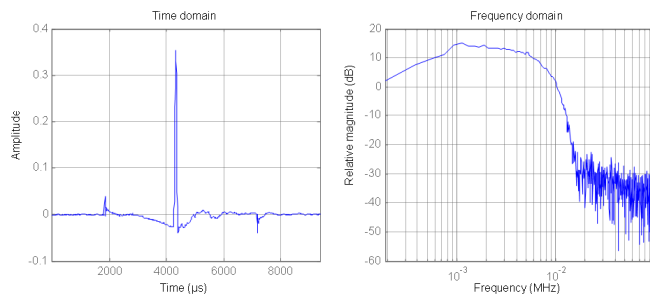


Figure 13: Recorded acoustic pulse, length 50 samples

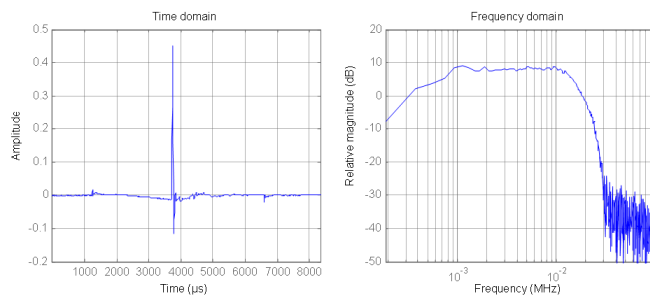


Figure 14: Recorded acoustic pulse, length 20 samples

The time plot of the 20-sample pulse reveals some undershoot at its end. The origin of this undershoot is probably caused by the uncertainty of the transducer’s impulse response measurement at very low and very high frequencies, and that an electro-acoustic transducer cannot effectively compensate the lowest frequency region of the pulse, requiring the presence of a DC component. Also, the mutual position of the microphone and the transducer, while obtaining the impulse response, was found to be crucial for the results.

The compensation signal was also calculated for the derivative pulse. Recorded signals for pulses with lengths of 50 and 13 samples are shown in Figures 15 and 16.

The time and spectral match of the derivative pulse with the transducer is subjectively better than for the unipolar pulse, apparently due to lack of the DC component in pulse spectral content.

4.2. Pulse Sequence

The purpose of the next set of measurements was to find the limits of the pulse repetition rate. The internal pulse train modulator of the experimental unit was used to generate a square modulated sequence of pulses with programmable distance. The recorded acoustic outputs are

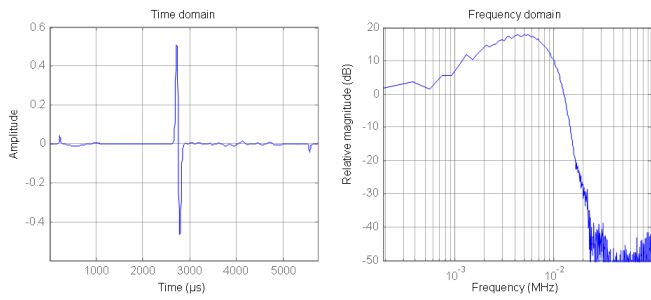


Figure 15: Recorded derivative pulse, length 50 samples

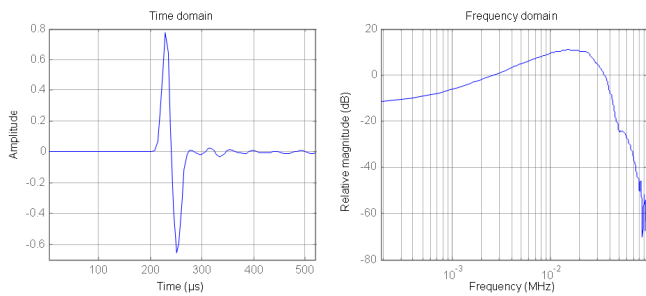


Figure 16: Recorded derivative pulse, length 13 samples

shown in Figure 17 with the driving signal set up for a pulse length of 10 samples, a distance of 32 samples between single pulses and 4 pulses in a square half-period, as an example of the fastest repetition rate achieved for unipolar pulses.

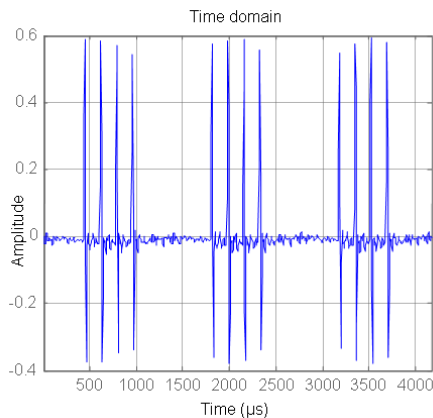


Figure 17: Recorded signal time plot, unipolar pulses

The pulse distance of 32 samples at the 192 kHz sample frequency can be interpreted as a 6 kHz repetition rate. Although it should theoretically be possible to get pulses even closer to each other, the driving signal has a significant DC component that allows for driving an electroacoustic transducer only for a very short time without transducer damage.

The fastest reliable distance between derivative pulses was experimentally found to be 16 samples at the 192 kHz

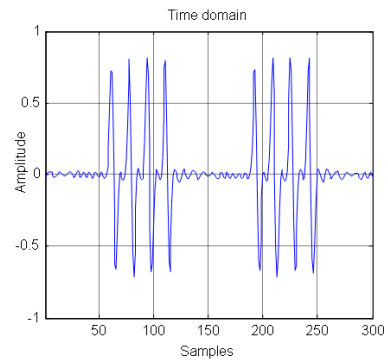


Figure 18: Recorded signal time plot, derivative pulses (4p)

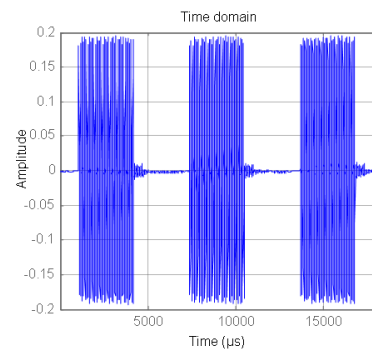


Figure 19: Recorded signal time plot, derivative pulses (32p)

sampling frequency. The effective pulse frequency is then 12 kHz. Two settings of the number of pulses in the square half-period were examined: first 4 pulses, and then 37 pulses in half-period. The acoustic output for the 4-pulse variant is shown in Figure 18.

The plot for the 37-pulse variant is shown in Figure 19. Figure 20 presents a detail of the beginning of the pulse train.

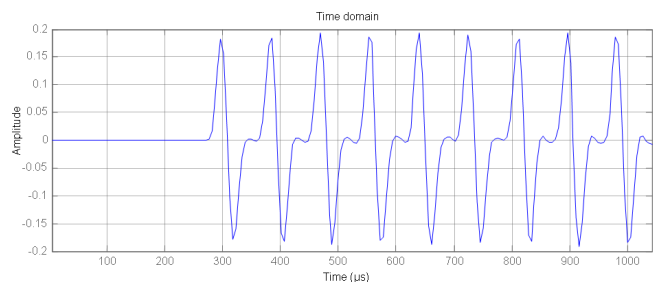


Figure 20: Recorded signal time plot, derivative pulses, detail

5. Conclusion

A solitary acoustic pulse of finite length with the shape of a certain class of window functions was chosen as the representation of a bit carrying the logic one value. The method of pre-filtering compensation was used to calculate the driving signal, in order to produce finite-length acoustic pulses using traditional electro-acoustic moving-mass transducers. This method was extensively simulated and experimentally verified.

The analysis of experimental recordings shows the advantage of using bipolar pulses over unipolar pulses in terms of achievable pulse length and, more importantly, pulse repetition rate. The second parameter appears to be a serious problem for electro-acoustic transducers when driven to produce a train of unipolar pulses. Initial experiments achieved the shortest unipolar pulse with duration of about 60 μs , with the possibility of a 6 kHz repetition rate. For bipolar pulses, a length of 70 μs was reached with the repetition rate as high as 12 kHz.

The results of this work can be used not only for applications of transducers with direct digital to analog conversion, but also to regular data transmission in the acoustic domain.

References

- [1] Flanagan, J. L.: Direct Digital-to-Analog Conversion of Acoustic Signals, *The Bell Systems Technical Journal*, Vol. 59, No. 9, 1980
- [2] Huang, Y. et al.: Distortion and Directivity in a Digital Transducer Array Loudspeaker, *J. Audio Eng. Soc.*, Vol. 49, No. 5, May 2001
- [3] Busbridge, S. C. et al.: Digital Loudspeaker Technology: Current State and Future Developments, *AES 112th convention*, Munich, Germany, May 2002
- [4] Valoušek, P.: Transferring Digital Information into Acoustic Domain, Doctoral thesis, CTU–FEE, Prague, February 2011
- [5] Marshall, I.: Production of Acoustic Pulses in Air, *Meas. Sci. Technol.*, Vol. 1, No. 5, pp. 413–418, 1990
- [6] Zipser, L. et al.: Generation of Solitary Acoustic Impulses, *IEEE Ultrasonics Symposium*, Vol. 1, pp. 555–558, San Antonio, 1996
- [7] National Instruments: Characteristics of Different Smoothing Windows, LabView 8.2 Help, 371361B-01, August 2006
- [8] Nuttall, A. H.: Some Windows with Very Good Side-lobe Behavior, *IEEE transactions on acoustics, speech and signal processing*, Vol. ASSP-29, No. 1, February 1981
- [9] Tymphany: Vifa OX20SC00-04 Product Datasheet, Online:
<http://www.tymphany.com/files/OX20SC00-04%20Rev%201.0.pdf>

Měření dynamického rozsahu lidského hlasu

Hana Šrámková^a, Svante Granqvist^b, Tomáš Füst^c, Jan G. Švec^a

^aKatedra biofyziky, Přírodovědecká fakulta Univerzity Palackého v Olomouci, Olomouc

^bRoyal Institute of Technology, Stockholm, Sweden

^cKatedra matematické analýzy a aplikací matematiky, Přírodovědecká fakulta Univerzity Palackého v Olomouci, Olomouc

e-mail: hana.sramkova@upol.cz, ssg@kth.se, Tomas.Furst@seznam.cz, jan.svec@upol.cz

The knowledge of the dynamic extremes of human voice is important for voice diagnostics and for scientific purposes. Numerous authors have published on the dynamic range of human voice but the methodology has been quite heterogenous and the exact values of the dynamic extremes are often unclear. There has been a need for a carefully documented study about measuring extreme limits of normal voice with easily reproducible methodology. We measured 37 subjects, 23 women and 14 men for this experiment. Two microphones were used (one head-mounted and one stand-mounted), which were calibrated at a standard distance 30 cm from the mouth. The head-mounted microphone was used for measuring the softest phonations and the stand-mounted for measuring the loudest ones. The loudest level was determined from calling the Czech words “Hej, Hou, Haló” as the peak sound level. The softest level was determined as a minimum of the fast time-weighted sound level from the softest sustained phonations on vowel [a:]. The standard weighting filters A, C or Z and a custom high-pass filter were used to decrease ambient noise. The level of the loudest phonation was found to be 127 dB at 30 cm from the mouth (male subject) and the softest phonation was found to be 31 dB (A-weighted) at 30 cm from mouth (male subject). These levels appear to be greater than those measured previously and can serve as a basis for specifying dynamic ranges of microphone intended for voice measurements.

Key words: Dynamic range, human voice, sound pressure level, normal healthy voice, measurement.

1. Úvod

Dynamický rozsah, tedy rozdíl mezi nejtisší a nejhlasitější možnou fonací (minimální a maximální hladinou akustického tlaku) je jedna z nezákladnějších charakteristik lidského hlasu. Může být ovlivněn mnoha faktory, jako je například pohlaví, věk, trénovanost či onemocnění hlasu. Znalost tohoto rozsahu, hodnot průměrných, ale i extrémních, je proto velmi důležitá pro diagnostické i pro výzkumné účely. Pro zjištění těchto hodnot a obecně pro měření dynamického rozsahu lidského hlasu je ovšem potřeba mít dostatečně kvalitní vybavení. Správně vybraný mikrofon umožňuje zaznamenat hlas v celém jeho dynamickém rozsahu a záznam tak nezhodnocovat ani nijak jinak neovlivňovat. Otázkou tedy je, jaký minimální dynamický rozsah by měl takový mikrofon mít, aby zaznamenal jak nejtisší, tak i nejhlasitější fonace jakékoli osoby? Podrobnější studium literatury věnované tomuto tématu prozrazuje značnou různorodost publikovaných údajů a nejednotnost metod měření a vyhodnocování dat [17]. V publikacích jsou většinou uváděny především průměrné zjištěné hodnoty a jejich směrodatné odchylky, často pouze ve formě grafů. Zjištěné extrémní naměřené hodnoty však obvykle uváděny nejsou, jelikož se velmi často jedná spíše o studii zaměřenou na ovlivnění hlasu některým z výše uvedených faktorů. K záznamu hlasu, analýze a vyhodnocování jsou tak často používány automatické komerční programy [10, 13, 14] či různé druhy fonetografů (Phog a další) [11], u kterých nebývá vždy jasné,

jak přesně a jakým algoritmem byly dané hodnoty vlastně určeny.

Cílem této práce je proto zjištění minimálních a maximálních hladin akustického tlaku znělého lidského hlasu u zdravých osob (nehledě na další faktory ovlivňující hlas) a porovnání výsledků s hodnotami uváděnými v literatuře. Je zde použita originální metoda experimentálního měření fonace lidského hlasu s využitím dvou mikrofonů v různých vzdálenostech, kalibrovaných na standardní vzdálenost 30 cm od úst. Navíc jsou zde uvedeny všechny naměřené hodnoty, jejich průměrné hodnoty a směrodatné odchylky, konkrétní použité výpočty a jsou zde popsány i úpravy záznamů před konečným vyhodnocením. Námí stanovený postup by tak mohl být užitečný pro budoucí standardizaci měření dynamického rozsahu lidského hlasu.

2. Metoda

2.1. Subjekty

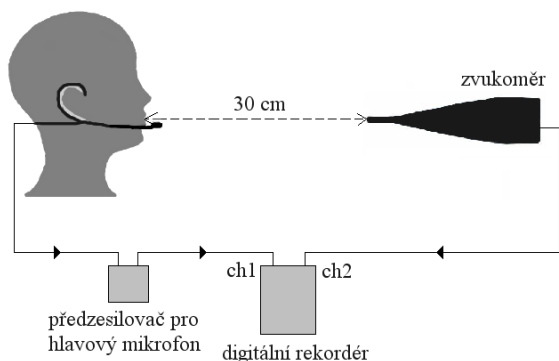
Všichni vyšetřovaní (38 osob) vyplnili nejprve dotazník Voice Handicap Index (VHI) [20] pro zjištění subjektivních hlasových potíží. Pro vlastní experiment byly vybrány pouze osoby, které uvedly stupeň hlasových potíží „žádné“ či „mírné“. Osoby uvádějící stupeň potíží „střední“ či „velké“ byly z měření vyřazeny. Normalita hlasu byla ověřena také pomocí celkové hodnoty VHI. Povoleny byly hodnoty VHI maximálně 38, což je v rozmezí normativních hodnot (hodnoty pod 38,7 jsou v rámci dvou směrodatných odchylek normální populace [20]). Z pů-

vodního souboru nesplnila kritéria normality jedna osoba. Bylo změřeno 37 osob: 23 žen ve věku průměrně 23,9 let (od 15 do 49 let) a 14 mužů ve věku průměrně 29,9 let (od 15 do 44 let). Z toho 20 osob (13 žen a 7 mužů) byli členové pěveckého sboru, dále se jednalo o 10 studentů (7 žen a 3 muži) a 7 pedagogů (3 ženy a 4 muži) na Přírodovědecké fakultě Univerzity Palackého v Olomouci. Byly vybrány subjekty se zdravým normálním hlasem, ať již trénované, či nikoliv. Pro účely této studie nebyl brán zřetel na rozlišení hlasu podle trénovanosti, vzhledem k tomu, že cílem práce bylo určit dynamický rozsah pokrývající všechny možné kategorie hlasu. Vzhledem k fyziologickým rozdílům byly ženské a mužské hlasy vyhodnoceny odděleně a porovnány.

2.2. Vybavení

Část měření se členy sboru proběhla v jejich zkušební místnosti (velikost cca 145 m³, hladina akustického tlaku pozadí A 23 dB, doba dozvuku 0,35–1,30 s v pásmu 0,2–10 kHz, odpovídající dozvuková vzdálenost 0,6–1,15 m). Druhá část se studenty a pedagogy proběhla ve speciální akusticky upravené místnosti Českého rozhlasu Olomouc (velikost 48,75 m³, hladina akustického tlaku pozadí A 18 dB, doba dozvuku 0,06–0,14 s v pásmu 0,2–10 kHz, dozvuková vzdálenost 1–1,6 m).

Pro měření dynamického rozsahu byly využity dva mikrofony. Pro nejnižší fonace to byl hlavový mikrofon DPA (typ 4066) umístěný ve vzdálenosti přibližně 5 cm stranou od úst ve spojení s předzesilovačem DPA (typ MMA 6000, zesílení 15 dB). Pro záznam nejhlasitějších fonací byl použit 1/2" kondenzátorový mikrofon Brüel & Kjaer (typ 4188) ve vzdálenosti 30 cm od úst ve spojení s předzesilovačem zvukoměru Brüel & Kjaer (typ 2238 Mediator, nefiltrovaný AC výstup). Pro kalibraci byl použit kalibrátor Brüel & Kjaer (typ 4231). Měření bylo zaznamenáno pomocí digitálního rekordéru M-Audio (typ Microtrack II) s nastavenou vzorkovací frekvencí 48 kHz a rozlišením 24 bitů. Přístrojová sestava je na obrázku 1. Záznamy z měření byly zpracovány v počítači pomocí programu Multi-Speech (model 3700, verze 3.2.0., KayPENTAX,



Obrázek 1: Přístrojová sestava při měření

USA) a pro filtraci a matematické výpočty bylo využito programu MatLab (7.0.1., MathWorks).

2.3. Postup vyšetření

Vyšetření hladin akustického tlaku probíhalo podle následujícího postupu:

1. vyslovení jména a příjmení vyšetřovaného a datum měření
2. dvoustupňová kalibrace obou mikrofونů [19]:
 - (a) záznam kalibračního tónu (94 dB) z kalibrátoru nasazeného na mikrofon zvukoměru
 - (b) fonace samohlásky [a:] na pohodlné výšce tónu po několik sekund
3. co nejnižší fonace samohlásky [a:] na pohodlné výšce tónu
4. co nejhlasitější fonace – zvolání „Hej!, Hou!, Haló!“

Záznam měření pak byl zpracován pomocí počítačových programů Multi-Speech a MatLab.

2.4. Zpracování signálů záznamu

Dvoustupňová kalibrace Smyslem dvoustupňové kalibrace je umožnit určení hladin akustického tlaku hlasu snímaných současně zvukoměrem a hlavovým mikrofونem v různých vzdálenostech.

V první fázi byl zkalibrovan signál zvukoměru pomocí známého signálu vydávaného akustickým kalibrátorem (na hladině 94 dB). Ve druhé fázi byla nastavena hladina akustického tlaku signálu z hlavového mikrofону tak, aby byla rovna hladině signálu zvukoměru (ve vzdálenosti 30 cm od úst) při vyslovení samohlásky [a:]. Úroveň takto zkalibrovaného signálu z hlavového mikrofону pak odpovídá hladině akustického tlaku ve vzdálenosti 30 cm. Vzhledem k tomu, že fyzická vzdálenost hlavového mikrofону je menší než 30 cm (kolem 5 cm), zlepšuje se tím odstup signálu od šumu místnosti, což je výhodné zejména pro měření nejnižších fonací. Vzdálenost hlavového mikrofону nemusí být známa přesně, ale po dobu měření daného subjektu musí být konstantní. Možný efekt blízkého pole je v tomto případě zanedbáván [18].

Ze záznamů dvoustupňové kalibrace byly zjištěny kalibrační faktory c_1 pro signál z hlavového mikrofону a c_2 pro signál z mikrofону zvukoměru:

$$c_2 = p_0 \cdot 10^{L_r/20} \cdot rms^{-1}, \quad (1)$$

kde p_0 je referenční akustický tlak (20 μ Pa), L_r je rovno kalibrační hodnotě hladiny akustického tlaku – 94 dB, produkované kalibrátorem, rms (z angl. *root mean square*) je vypočtená časově průměrovaná hodnota ze záznamu kalibračního signálu ze zvukoměru:

$$rms = \sqrt{\frac{\sum_{k=1}^K y^2(k)}{K}}, \quad (2)$$

kde $y(k)$ je původní signál zaznamenaný mikrofonom a K je délka záznamu odpovídající celkovému počtu vzorků. Kalibrační faktor c_1 pak je roven:

$$c_1 = \frac{c_2 \cdot rms_2}{rms_1}, \quad (3)$$

kde rms_1 je hodnota ze záznamu samohlásky [a:] ze signálu z hlavového mikrofону a hodnota rms_2 z téhož záznamu, ale z mikrofónu zvukoměru.

Faktor c_1 upravuje úroveň signálu z hlavového mikrofónu tak, aby měl během fonace stejné hodnoty akustického tlaku p jako zkalibrovaný signál mikrofónu zvukoměru. Hladiny akustického tlaku hlavového mikrofónu tak odpovídají hladinám akustického tlaku, které jsou ve vzdálenosti zvukoměru (tj. 30 cm) [19].

Kalibrační faktory c_1 a c_2 umožňují převést hodnoty původního signálu z hlavového mikrofónu y_1 a zvukoměru y_2 na akustický tlak v pascálech ve vzdálenosti 30 cm od úst pomocí jednoduchých vztahů:

$$p_1 = c_1 \cdot y_1, \quad (4)$$

$$p_2 = c_2 \cdot y_2. \quad (5)$$

Nejhlasitější fonace V programu Multi-Speech byla vystržena část záznamu s nejhlasitější fonací. Pomocí programu MatLab byla detekována špičková hodnota hladiny akustického tlaku ze signálu mikrofónu zvukoměru L_{2peak} (dB) podle rovnice:

$$L_{2peak} = 20 \cdot \log \left(\frac{p_{2peak}}{p_0} \right), \quad (6)$$

kde p_0 je referenční akustický tlak a p_{2peak} je změřený špičkový akustický tlak (Pa), který je roven:

$$p_{2peak} = c_2 \cdot y_{2peak}, \quad (7)$$

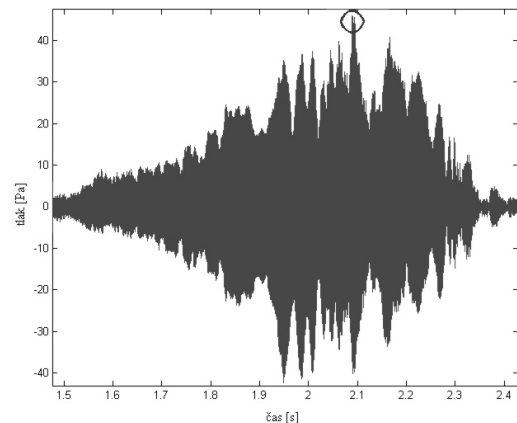
kde c_2 je vypočtený kalibrační faktor pro signál ze zvukoměru a y_{2peak} je:

$$y_{2peak} = \max |y_2|, \quad (8)$$

kde y_2 je původní signál ze záznamu ze zvukoměru.

Záznam nejhlasitější fonace s vyznačenou polohou špičkového akustického tlaku je na obrázku 2.

Šum Při určování nejtišší fonace bylo nutné zjistit, zda nedošlo k ovlivnění měření šumem. K tomu může dojít, je-li nedostatečný odstup signálu od šumu (doporučený odstup je více než 15 dB [18]). Hladinu šumu lze snížit vhodnou filtrací, je-li energie šumu koncentrována na jiných frekvencích než energie vlastního signálu. Pomocí programů Multi-Speech a MatLab byly zjištěny hladiny šumů a na záznamy s nejtiššími fonacemi pak aplikovány standardní váhové filtry A, C, Z a speciální hornopropustný filtr hp, viz dále. Váhový filtr A zeslabuje frekvence mimo oblast 1–6 kHz [8], váhový filtr C zeslabuje



Obrázek 2: Záznam nejhlasiťjší fonace s vyznačením polohy špičkového akustického tlaku, ze kterého byla poté vypočtena špičková hladina akustického tlaku

frekvence mimo oblast 50 Hz až 4 kHz, váhový filtr Z původní signál nijak neovlivní [8]. V programu MatLab bylo pro tento účel využito definic váhových filtrů typu A a C podle IEC/CD 1672 převzaté z [6].

Pro specifické filtrování šumu pozadí byl dále použit hornopropustný filtr typu „butterworth“ 5. řádu (dále označován symbolem hp), jehož parametry byly nastaveny tak, aby nedošlo k ovlivnění hodnoty tlaku nejnižší fonované frekvence o více než 1 dB a byl pouze potlačen nízkofrekvenční šum. Prvním krokem bylo zjištění základní frekvence nejtišší fonace F_0 pomocí spektrální analýzy metodou Fourierovy transformace v programu MatLab. Nejnižší zjištěná F_0 u 37 změřených osob byla frekvence 97 Hz, pod níž byl u všech záznamů již pouze šum. Zlomová frekvence filtru byla stanovena na 70 Hz, při níž filtr zeslabuje tóny se základní frekvencí 97 Hz o pouhých 0,2 dB.

Pro zjištění hladiny šumu pozadí byla v programu Multi-Speech vystržena část se záznamem obsahujícím pouze šum místnosti. V programu MatLab byly vypočteny časově průměrované hladiny akustického tlaku L_{XYT} (tj. L_{AT} , L_{CT} , L_{hpT} , L_{hpAT} , L_{hpCT}) před filtrací a po ní. Hladiny L_{XYT1} pro signál z hlavového mikrofónu a L_{XYT2} pro signál ze zvukoměru, byly vypočteny dle rovnic:

$$L_{XYT1} = 20 \cdot \log \left(\frac{c_1 \cdot rms_1}{p_0} \right), \quad (9)$$

$$L_{XYT2} = 20 \cdot \log \left(\frac{c_2 \cdot rms_2}{p_0} \right), \quad (10)$$

kde T značí časově průměrovanou hodnotu, p_0 je referenční akustický tlak, c_1 a c_2 jsou vypočtené kalibrační faktory podle rovnic (1) a (3) a rms_1 a rms_2 jsou hodnoty vypočtené ze signálů šumu pomocí rovnice (2) v programu MatLab. Indexy X a Y zde označují druh první a druhé filtrace aplikované na signál. V případě, že byla provedena pouze jedna filtrace signálu, je index Y vypuštěn. Výsledné

hodnoty šumu L_{XYT1} a L_{XYT2} byly použity pro zjištění odstupů signálu nejnižších fonací od šumu.

Nejnižší fonace V programu Multi-Speech byla vystřižena část záznamu s nejnižší fonací. Pro výpočet nejnižší fonace bylo nutné mít na záznamu pouze části se znělým hlasem. Pro zviditelnění přechodu fonace ze znělé do neznělé části byl nejprve odfiltrován šum pomocí váhových filtrů A, C, nebo hp. Neznělé části, včetně přechodových stavů fonace-šum byly manuálně vystřiženy a takto upravený záznam byl použit pro zjištění nejnižší fonace. V programu MatLab byla vypočtena časově vážená hladina akustického tlaku L_{XYF} (tj. L_{AF} , L_{CF} , L_{hpF} , L_{hpAF} nebo L_{hpCF}) podle definice specifikované normou IEC/61672-1 [8]:

$$L_{XYF}(t) = 20 \cdot \log \frac{\sqrt{\frac{1}{\tau} \int_{-\infty}^t p_{XY}^2(\xi) e^{-(t-\xi)/\tau} d\xi}}{p_0}, \quad (11)$$

kde písmeno F značí použití standardního rychlého časového vážení (viz dále, [7]), p_0 je referenční akustický tlak, $p_{XY}(\xi)$ je okamžitý (Pa) akustický tlak zaznamenaný hlavovým mikrofonem, zkalibrovaný podle rovnice (4) a vážený frekvenčními filtry X a Y, ξ je pomocná proměnná časové integrace a τ je časová konstanta, která je pro rychlé časové vážení (F) rovna 0,125 s. Časové vážení bylo podle doporučení IEC/61672-1 realizováno pomocí filtru, který odpovídá exponenciálnímu poklesu

$$U = e^{-t/\tau} \quad (12)$$

a zlomové frekvenci

$$f_0 = \frac{1}{2\pi\tau}. \quad (13)$$

Z takto vypočtených hladin akustického tlaku pak byla pro každý záznam nalezena hladina minimální. Záznam nejnižší fonace s vyznačenou minimální hladinou akustického tlaku je na obrázku 3.

3. Výsledky

3.1. Nejhlasitější fonace

U žen byla naměřena nejvyšší špičková hladina akustického tlaku 125 dB a u mužů 127 dB ve vzdálenosti 30 cm od úst. Průměrná hodnota byla u žen 118 ± 5 dB a u mužů 121 ± 4 dB. Všechny naměřené hodnoty jsou uvedeny v tabulce 1. Pomocí statistického t -testu bylo zjištěno, že rozdíl mezi výslednými hodnotami u žen a u mužů je statisticky významný ($p < 0,01$).

3.2. Šum

Byly vypočteny časově průměrované hladiny akustického tlaku šumu L_{XYT1} a L_{XYT2} pro signály z hlavového mikrofonu a ze zvukoměru před filtrací a po ní. Průměrné zjištěné hodnoty šumu a jejich směrodatné odchylky pro

všechny použité druhy filtrací jsou uvedeny v tabulce 2. V tabulce 3 jsou uvedeny všechny odstupy fonací od šumu a také jejich střední hodnoty pro daný druh filtrace a jejich směrodatné odchylky. Při použití filtrů hp byly průměrné hodnoty odstupů nejnižší fonace od šumu a jejich směrodatné odchylky (SD) 22 ± 4 dB u žen, resp. 23 ± 3 dB u mužů, 21 ± 4 dB, resp. 15 ± 4 dB při filtraci A, 11 ± 5 dB, resp. 12 ± 4 dB při filtraci C, 22 ± 4 dB, resp. 15 ± 4 dB při filtraci hpA a 23 ± 3 dB, resp. 23 ± 4 dB při filtraci hpC.

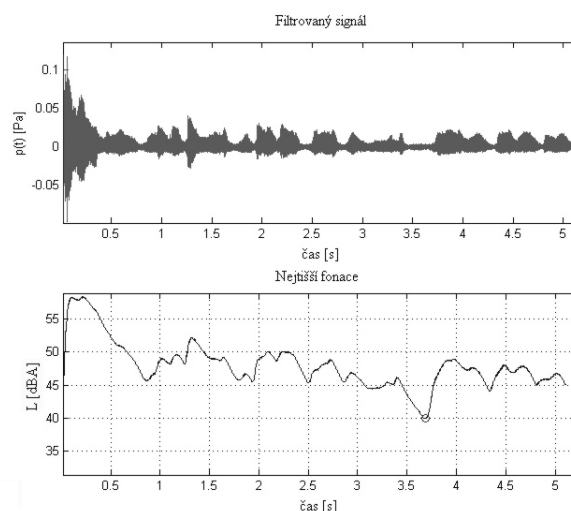
3.3. Nejnižší fonace

U žen byla naměřena nejnižší hladina akustického tlaku 39 dB (filtrace hpA) a u mužů 31 dB (filtrace hpA) pro vzdálenost 30 cm od úst. Průměrná hodnota byla u žen 45 ± 4 dB (filtrace hpA) a u mužů 37 ± 5 dB (filtrace hpA). Pomocí statistického t -testu bylo zjištěno, že rozdíl mezi výslednými hodnotami u žen a u mužů je statisticky významný ($p < 0,01$). Všechny naměřené hladiny akustických tlaků fonací při použití různých filtrů jsou uvedeny v tabulce 3.

4. Diskuse

4.1. Nejhlasitější fonace

Nejvyšší hladina akustického tlaku 127 dB ve vzdálenosti 30 cm od úst (bez filtrace) byla naměřena u 35letého člena pěveckého sboru. U žen byla nejvyšší naměřená hladina o 2 dB nižší, tedy 125 dB (bez filtrace), a to u 22leté členky pěveckého sboru. Zjištěná průměrná hladina u mužů byla 121 ± 4 dB (SD), u žen pak 118 ± 5 dB. Z těchto výsledků vyplývá, že muži byli schopni dosáhnout vyšších hladin



Obrázek 3: Záznam nejnižší fonace filtrované pomocí standardního váhového filtru A, nahoře jako závislost akustického tlaku na čase a dole jako závislost hladin akustického tlaku na čase (při použití standardního rychlého časového vážení) s vyznačenou nejnižší hladinou

Č. s.	Ženy	Muži
	L_{peak} (dB)	L_{peak} (dB)
1	120	120
2	115	126
3	118	126
4	119	124
5	119	127
6	110	124
7	123	118
8	125	118
9	117	116
10	116	120
11	118	117
12	113	116
13	111	123
14	115	125
15	113	–
16	103	–
17	120	–
18	116	–
19	121	–
20	117	–
21	113	–
22	118	–
23	117	–
	118 ± 5	121 ± 4

Tabulka 1: Naměřené špičkové hodnoty nehlasiťších fonací L_{peak} v jednotkách dB ve vzdálenosti 30 cm od úst pro 23 žen a 14 mužů. Zvýrazněná hodnota je nejvyšší zjištěná hodnota, zvýrazněné hodnoty v posledním řádku jsou průměrné hodnoty \pm směrodatná odchylka. Č. s. je označení čísla měřeného subjektu. Uváděné hodnoty byly obdrženy bez použití váhových filtrů a bez časového vážení, a to pomocí rovnice (6)

akustického tlaku než ženy. Tento rozdíl se ukázal jako statisticky významný (t -test, $p < 0,01$).

Hladiny nehlasiťší fonace uváděné v již publikované literatuře se různí: např. Sulter v [16] uvádí průměrné hladiny vážené filtrem A nižší u mužů $100,7 \pm 8,54$ dB než u žen $104,1 \pm 7,75$ dB. Coleman [15] podle citace [3] naměřil extrémní hodnoty hladin akustických tlaků, po přepočtu z 15,26 cm na standardní vzdálenost 30 cm od úst, u mužů 120 dB, u žen pouze 116 dB. Přepočet byl nutný také pro výsledky z práce Leina [12], kde jsou uvedeny hodnoty pro vzdálenost 40 cm, hodnoty u mužů odpovídaly 122 dB a u žen 118 dB pro 30 cm. Heylen v práci [10] uvádí pouze grafické znázornění hlasového pole, přesto lze průměrné naměřené hodnoty odhadovat jako 99 dB u mužů a 98 dB u žen (vážené filtrem A).

Jedním z faktorů, který ovlivňuje výsledky, je způsob fonace: pro náš experiment byla od vyšetřovaných osob vyžadována nehlasiťší fonace jako zvolání, kdežto ve většině jiných publikací je způsob nehlasiťší fonace uvá-

děn jako zpěv nebo co nehlasiťší fonace samohlásky [a:]. Často jde o zjišťování hladin akustického tlaku pouze ze čtení určitého úryvku textu nebo fonaci samohlásky [a:] v přirozené poloze. V práci Awana [1] lze porovnat výsledky fonaci samohlásky [a:] v přirozené poloze, při minimální a maximálně hlasité fonaci (ovšem bez křiku nebo zvolání). Při přirozené fonaci uvádí nejvyšší naměřenou hodnotu jako 110,33 dB a při maximální fonaci jako 112,5 dB. V práci [2] zjišťuje autor hladiny akustického tlaku ze čtení úryvku textu a zde uvádí jako nejvyšší naměřenou hodnotu 78,5 dB. Hacki například v práci [9] porovnává hladiny akustického tlaku získané zpěvem, křikem (shouting) nebo ze čtení textu. Výsledky v jeho práci jsou ovšem také publikovány pouze jako grafické znázornění hlasového pole a konkrétní hodnoty je tedy nutné odečíst z grafů. Přesto lze využít alespoň přibližných hodnot, při čtení okolo 97 dB, zvolání okolo 110 dB a zpěv okolo 112 dB (vážené filtrem A). Lamarche se v práci [11] zabývá rozdílem v hladinách akustických tlaků při zpěvu, a to podle rejstříků. Jako průměrnou nejvyšší hodnotu uvádí $115,5 \pm 2,12$ dB u mezzosopranistek. Leino se pak v práci [12] zaměřil na porovnání výsledků při čtení, hlasitém čtení a při zvolání (shouting). U mužů byly extrémní hodnoty při zvolání až 124 dB, u žen pak až 127 dB (což je jediná publikovaná hodnota srovnatelná s hodnotami zjištěnými v této práci). Schneiderová v práci [13] také porovnává pohodlnou fonaci, zpěv a zvolání. Při pohodlné fonaci je extrémní naměřená hodnota 68 dB, při zpěvu je to 112 dB a při zvolání pak 110 dB (vážené filtrem A). V práci [14] uvádí průměrnou hodnotu při zpěvu $95,2 \pm 4,0$ dB, při zvolání pak $90,7 \pm 7,7$ dB.

V porovnání s literaturou se námi aktuálně naměřené hodnoty jeví jako nejvyšší zjištěné. To může být způsobeno několika faktory. Svou roli mohl hrát například výběr subjektů či typ fonace, tedy zvolání. Také způsob analýzy – výběr špičkové hladiny akustického tlaku spíše než jeho průměrované či časově vážené hodnoty – mohl být příčinou vyšších zjištěných hodnot. Jak již bylo řečeno, většina autorů využívala buďto automatických programů, nebo odečítala hodnoty ze zvukoměru, navíc často s použitím standardního váhového filtru A, což mohlo ovlivnit naměřené hodnoty. Námi použitý způsob určení nehlasiťší fonace byl zaměřen speciálně na určení nejvyšší okamžité (nepřůměrované) hladiny akustického tlaku, a to bez použití váhového filtru A. Tato hodnota je užitečná pro stanovení standardních požadavků na mikrofony, které lze použít pro záznam hlasu ve všech možných situacích.

4.2. Šum

Bylo zjištěno, že hladina šumu na záznamech byla průměrně 50 ± 4 dB na kanálu 1 z hlavového mikrofону a 63 ± 3 dB na kanálu 2 z mikrofону zvukoměru. Tyto hladiny bylo nutné pomocí filtrací snížit. Pro nalezení ideální filtrace byly využity standardní váhové filtry A, C a Z a speciálně navržený hornopropustný filtr hp. Průměrně byly hladiny šumu sníženy pomocí filtru hp o 21 dB na

	Z		hp		A		C		hpA		hpC	
	L_1	L_2	L_1	L_2	L_1	L_2	L_1	L_2	L_1	L_2	L_1	L_2
Průměr ± SD	50 ± 4	63 ± 3	29 ± 3	51 ± 1	23 ± 3	42 ± 1	41 ± 6	58 ± 2	23 ± 3	42 ± 1	28 ± 4	50 ± 1

Tabulka 2: Průměrné hodnoty hladin akustického tlaku šumu ± směrodatné odchylky v jednotkách dB ve vzdálenosti 30 cm od úst filtrované přes kombinace filtrů hp, A, C a Z, L_1 a L_2 odpovídají hodnotám pro hlavový mikrofon, respektive mikrofon zvukoměru

kanálu 1. Průměrný odstup hladiny fonace od šumu byl 22 a 23 dB u žen a u mužů. Pomocí filtrace A se šum snížil o 27 dB. Průměrný odstup byl 21 a 15 dB u žen a u mužů (při vážení filtrem A). Pomocí filtrace C se šum snížil o 9 dB. Průměrný odstup byl 11 a 12 dB u žen a u mužů (při vážení filtrem C). Pomocí kombinace filtrů hp a A se šum snížil průměrně o 27 dB. Průměrný odstup byl 22 a 15 dB u žen a u mužů (při vážení filtrem A). Pomocí kombinace filtrů hp a C se šum snížil o 22 dB. Průměrný odstup byl 23 a 23 dB u žen a u mužů (při vážení filtrem C). Z výsledků odstupů hladin fonací od šumu na kanálu z hlavového mikrofonu (viz tabulka 3) je zřejmé, že například filtrace C způsobuje průměrně jen velmi malý odstup, a fonace jsou tedy většinou šumem ovlivněny. Ostatní filtrace již způsobují průměrné odstupy fonací větší než 15 dB, a jsou proto vhodnější. Pro snížení hladiny šumu se nejúčinněji jevila kombinace filtrů hp a A, s jejíž pomocí byla hladina šumu snížena o 18–33 dB a odstup fonace od šumu byl také uspokojivý. Tato kombinace filtrů také vykázala nejextrémnější, t.j. nejnižší změřené hladiny u nejnižších fonací.

Hladiny akustického tlaku šumu v místnostech, kde probíhalo vyšetřování, byly změřeny (za úplného ticha) také pomocí odečtu hodnoty ze zvukoměru s nastaveným standardním A váhovým filtrem. Hodnoty 18 dB v místnosti Českého rozhlasu a 22 dB v místnosti pro zkoušky sboru jsou výrazně nižší než námi zjištěné hodnoty šumu ze záznamů z mikrofonu zvukoměru – průměrně 42 dB (váženo filtrem A). Tento zdánlivý nesoulad souvisí s limitací dynamického rozsahu použitých záznamových zařízení a šumovou hladinou předzesilovačů, které poněkud zvýšily celkový šum záznamu (aby bylo možno sejmout nejhlasitější fonace, byl dynamický rozsah zvukoměru při záznamu nastaven na oblast 60–140 dB, zatímco při měření šumu místnosti byl tento nastaven na 0–80 dB).

4.3. Nejnižší fonace

Nejnižší naměřená hladina akustického tlaku A byla 31 dB, a to u 42letého učitele. Odstup signálu od šumu byl v tomto případě cca 8 dB. Tento odstup byl nižší než doporučovaných 15 dB [18] a lze předpokládat, že šum navýšil změřenou hladinu akustického tlaku o 0,5 až 1 dB [4]. U žen byla nejnižší naměřená hladina A o 8 dB vyšší než u mužů, 39 dB, a to u 22leté studentky. Průměrná hladina A nejnižších fonací u mužů byla 37 ± 5 dB a u žen 45 ± 4 dB. Z těchto výsledků je zřejmé, že muži byli schopni dosáhnout nižších hladin akustického tlaku

než ženy. Tento rozdíl se ukázal jako statisticky významný (t -test, $p < 0,01$).

Tento závěr je opačný než výsledky Sultera v [16], který uvádí průměrnou hladinu u mužů $45,7 \pm 4,43$ dB (SD) a u žen hladinu nižší, $44,2 \pm 4,35$ dB (při A vážení). Stejně tak Coleman v [5] podle citace v [3] naměřil extrémní hodnoty u mužů vyšší – 45 dB a u žen 42 dB (po přepočtu na standardní vzdálenost). Také Leino v [12] uvádí vyšší extrémní hodnoty u mužů 47,5 dB – u žen pak 42,5 dB (po přepočtu na standardní vzdálenost 30 cm). Z práce Heylena [10] lze hodnoty u mužů odhadovat z grafů naopak jako nižší – 46 dB, u žen pak 47 dB (při vážení filtrem A).

Odlíšnost námi změřených výsledků od údajů v literatuře lze vysvětlit jinou metodikou analýzy a také výraznějším potlačením šumu, které umožnilo detekování nižších hladin akustického tlaku. Významným faktorem, který ovlivňuje hladinu akustického tlaku, je také volba frekvenčního vážení. Standardní filtrace typu A tlumí více hluboké tóny, a tedy hluboké mužské hlasy jsou po filtraci slabší než hlasy ženské, které jsou položeny přibližně o oktávu výše. Použití filtrace typu A pro měření hlasu je doporučeno Uníí evropských foniatrů [15], nicméně toto doporučení bývá některými autory považováno za kontroverzní a není univerzálně respektováno, a proto výsledky různých autorů bývají těžko porovnatelné. Dalším faktorem je volba typu časového vážení. Konstanta časového vážení ovlivňuje stabilitu výsledné hladiny akustického tlaku při primárně nestabilním signálu. Časové vážení typu S (slow) vykazuje menší fluktuace hladiny akustického tlaku než časové vážení typu F (fast). Vzhledem k tomu, že nejnižší fonace vykazují, i při odstranění neznělých pasáží, nestability, rychlé časové vážení umožňuje detekci nižších hladin akustického tlaku než vážení pomalé. V našem případě bylo zvoleno rychlé časové vážení, neboť nás zajímaly hodnoty nejnižších fonací, které ve většině případů trvaly kratší dobu než 1 sekundu. Tato volba je také doporučována Uníí evropských foniatrů [15]. Nicméně problém časového vážení, případně časového průměrování v případě nejnižších fonací si zaslouhuje hlubší pozornost v budoucnu.

Dalším z možných faktorů ovlivňujících výsledky měření nejnižší fonace je také volba detekce prahu znělého hlasu, která ovlivňuje výsledné naměřené hodnoty, zejména v případech, kdy fonace nejsou dostatečně stabilní. V literatuře doposud neexistuje jasné doporučení, jakým způsobem detekci prahu znělého hlasu provádět. Někteří autoři využívají manuálního odečtu ze zvukoměru, jiní programů pro automatickou akustickou analýzu, u kterých však často

Č. s.	Ženy					Muži				
	hp	A	C	hpA	hpC	hp	A	C	hpA	hpC
1	50 [18]	43 [19]	52 [7]	41 [17]	51 [20]	46 [20]	39 [19]	47 [7]	34 [15]	46 [20]
2	58 [24]	49 [21]	58 [11]	49 [21]	58 [25]	59 [26]	48 [20]	59 [11]	46 [18]	59 [26]
3	55 [26]	48 [26]	55 [12]	48 [27]	55 [27]	52 [22]	37 [15]	52 [7]	36 [14]	52 [22]
4	51 [18]	45 [21]	52 [5]	45 [21]	51 [19]	56 [26]	42 [16]	56 [15]	39 [13]	56 [26]
5	53 [21]	47 [24]	54 [8]	47 [25]	53 [22]	54 [19]	38 [13]	55 [10]	36 [12]	54 [20]
6	56 [22]	50 [26]	55 [8]	51 [28]	55 [22]	52 [21]	37 [15]	53 [7]	35 [13]	51 [22]
7	50 [17]	46 [23]	51 [5]	42 [20]	50 [18]	45 [20]	33 [16]	47 [8]	32 [17]	46 [22]
8	55 [22]	49 [22]	58 [10]	47 [20]	57 [24]	52 [20]	46 [19]	54 [8]	46 [19]	53 [21]
9	49 [16]	43 [18]	50 [4]	43 [19]	49 [17]	52 [27]	40 [19]	52 [19]	40 [20]	52 [28]
10	62 [30]	57 [32]	62 [15]	57 [33]	62 [30]	49 [23]	32 [8]	49 [15]	31 [8]	49 [23]
11	56 [22]	46 [18]	56 [8]	46 [18]	56 [22]	49 [25]	35 [17]	49 [16]	35 [17]	49 [26]
12	52 [23]	44 [22]	53 [10]	44 [22]	52 [24]	48 [24]	33 [13]	48 [15]	33 [13]	48 [25]
13	53 [22]	45 [23]	53 [8]	45 [24]	53 [23]	46 [19]	32 [6]	40 [7]	32 [6]	40 [13]
14	55 [22]	47 [21]	56 [9]	47 [23]	55 [23]	51 [24]	42 [20]	51 [16]	42 [21]	51 [25]
15	49 [24]	45 [25]	50 [18]	45 [25]	49 [25]	–	–	–	–	–
16	49 [19]	41 [17]	49 [12]	41 [17]	49 [20]	–	–	–	–	–
17	47 [18]	40 [14]	47 [11]	40 [14]	47 [20]	–	–	–	–	–
18	53 [27]	45 [23]	53 [22]	45 [23]	53 [28]	–	–	–	–	–
19	49 [22]	39 [15]	49 [15]	39 [15]	49 [22]	–	–	–	–	–
20	49 [22]	40 [18]	49 [14]	40 [19]	49 [23]	–	–	–	–	–
21	53 [28]	45 [25]	53 [19]	45 [26]	53 [29]	–	–	–	–	–
22	51 [26]	42 [22]	51 [17]	42 [22]	51 [28]	–	–	–	–	–
23	48 [21]	39 [17]	48 [13]	39 [19]	48 [21]	–	–	–	–	–
	52 ± 4 [22 ± 4]	45 ± 4 [21 ± 4]	53 ± 4 [11 ± 5]	45 ± 4 [22 ± 4]	52 ± 4 [23 ± 3]	51 ± 4 [23 ± 3]	38 ± 5 [15 ± 4]	51 ± 5 [12 ± 4]	37 ± 5 [15 ± 4]	51 ± 4 [23 ± 4]

Tabulka 3: Hodnoty hladin akustického tlaku nejtišších fonací v jednotkách dB ve vzdálenosti 30 cm od úst filtrované přes kombinace filtrů hp, Z, A a C pro 23 žen a 14 mužů. Hodnoty v hranatých závorkách jsou hodnoty odstupu hladin akustického tlaku fonace od hladin šumu. Zvýrazněná hodnota je nejnížší naměřená hodnota pro muže a ženy, zvýrazněné hodnoty v posledním řádku jsou hodnoty průměrné (pro hladiny fonací a v hranatých závorkách pro odstupy od šumu) ± směrodatná odchylka

chybí informace o použité metodě a kritériích detekce. Zde bylo využito standardního časového vážení typu F a manuálního ořezání signálu pouze na epizody znělého hlasu. Ruční editaci záznamů lze považovat za jistou limitaci naší studie, neboť zahrnuje nebezpečí nepřesností při odstraňování částí záznamů bez fonace a přechodových dějů. Pro účely této pilotní studie však byla manuální detekce zvolena jako nejpraktičtější varianta, neboť automatická detekce prahu znělého hlasu je značně komplexním problémem, který zasluhuje více času a pozornosti v budoucnu.

5. Závěr

Nejvyšší naměřená hladina akustického tlaku byla 127 dB ve vzdálenosti 30 cm od úst. Nejnižší naměřená hladina akustického tlaku A byla 31 dB ve vzdálenosti 30 cm od úst. Výsledky tohoto experimentu jsou poměrně dosti odlišné od hodnot mnoha různých autorů publikovaných již dříve a ukazují na větší dynamický rozsah lidského hlasu, než bylo doposud uvažováno. To může souviset s odlišnou metodou analýzy – detekcí špičkových hladin akustického tlaku a záměrným nepoužitím vážení signálu u nejhlasiv-

tějších fonací, či způsobem potlačení šumu u nejtišších fonací. Cílem práce ovšem bylo především stanovit postup pro získání extrémních dynamických limitů lidského hlasu a tomu odpovídá i způsob analýzy dat. Zjištění těchto limitů je důležité především pro efektivní stanovení parametrů mikrofónů schopných zaznamenat lidský hlas v celém jeho dynamickém rozsahu.

Poděkování

Práce byla podpořena do roku 2010 projektem GAČR 101/08/1155, v roce 2010 studentským projektem Univerzity Palackého PrF_2010_050 a v roce 2011 studentským projektem Univerzity Palackého PrF_2011_024 a projektem Evropského sociálního fondu CZ.1.07/2.4.00/17.0009 „Partnerství a věda“.

Reference

- [1] Awan S. N.: „Phonetographic Profiles and Fo-SPL Characteristics of Untrained Versus Trained Vocal Groups.“ *Journal of Voice* 5: p. 41–50, 1991

- [2] Awan S. N.: „Superimposition of Speaking Voice Characteristics and Phonetograms in Untrained and Trained Vocal Groups.“ *Journal of Voice* 7: p. 30–37, 1993
- [3] Baken R. J., Orlikoff R. F.: „Clinical measurement of speech and voice.“ *Singular Publishing Group, USA*, 2000
- [4] Brüel&Kjaer: „Measuring sound.“ *K. Larsen & Son A/S, Naerum, Denmark*, 1984
- [5] Coleman R. F., Mabis J. H., Hinson J. K.: „Fundamental Frequency – Sound Pressure Level Profiles of Adult Male and Female Voices.“ *Journal of Speech and Hearing Research* 20: p. 197–204, 1977
- [6] Couvreur Ch.: Octave (Matlab scripts). Matlab Central, <http://www.mathworks.com/matlabcentral/fileexchange/69>, Belgium, 1997
- [7] ČSN 01 1600: „Akustika – Terminologie.“ Český normalizační institut, 2003
- [8] ČSN EN 61672-1: „Elektroakustika – Zvukoměry – část 1: Technické požadavky.“ Český normalizační institut, 2003
- [9] Hacki T.: „Comparative speaking, shouting and singing voice range profile measurement: physiological and pathological aspects.“ *Logopedics Phoniatrics Vocology* 21: p. 123–129, 1996
- [10] Heylen L., Wuyts F. L., Mertens F., De B. M., Van de Heyning P. H.: „Normative voice range profiles of male and female professional voice users.“ *Journal of Voice* 16: p. 1–7, 2002
- [11] Lamarche A., Ternstrom S., Pabon P.: „The Singer’s Voice Range Profile: Female Professional Opera Soloists.“ *Journal of Voice* 24: p. 410–426, 2010
- [12] Leino T., Laukkanen A. M., Ilomaki I., Maki E.: „Assessment of Vocal Capacity of Finnish University Students.“ *Folia Phoniatrica et Logopaedica* 60: p. 199–209, 2008
- [13] Schneider B., Bigenzahn W.: „Influence of glottal closure configuration on vocal efficacy in young normal-speaking women.“ *Journal of Voice* 17: p. 468–480, 2003
- [14] Schneider B., Bigenzahn W.: „Vocal risk factors for occupational voice disorders in female teaching students.“ *European Archives of Oto-Rhino-Laryngology* 262: p. 272–276, 2005
- [15] Schutte H. K., Seidner W.: „Recommendation by the Union of European Phoniatrists (UEP) – Standardizing Voice Area Measurement Phonetography.“ *Folia Phoniatrica* 35: p. 286–288, 1983
- [16] Sulter A. M., Schutte H. K., Miller D. G.: „Differences in phonetogram features between male and female subjects with and without vocal training.“ *Journal of Voice* 9: p. 363–377, 1995
- [17] Šrámková H.: „Výšetření dynamického rozsahu akustického tlaku lidského hlasu.“ Univerzita Palackého v Olomouci (diplomová práce), 2010
- [18] Švec J. G., Granqvist S.: „Guidelines for selecting microphones for human voice production research.“ *American Journal of Speech – Language Pathology* 19: p. 356–368, 2010
- [19] Švec J. G., Popolo P. S., Titze I. R.: „Measurement of vocal doses in speech: experimental procedure and signal processing.“ *Logopedics Phoniatrics Vocology* 28: p. 181–192, 2003
- [20] Švec J. G. a kol.: „Česká verze dotazníku Voice Handicap Index pro kvantitativní hodnocení hlasových potíží vnímaných pacientem.“ *Otorinolaryngologie a Foniatrie* 3: p. 132–139, 2009

Akustické listy: ročník 17, číslo 3 říjen 2011
Vydavatel: Česká akustická společnost, Technická 2, 166 27 Praha 6
Počet stran: 20 Počet výtisků: 200
Redakční rada: M. Brothánek, O. Jiříček, J. Kozák, R. Čmejla, J. Volín
Jazyková úprava: R. Svobodová
Uzávěrka příštího čísla Akustických listů je 30. listopadu 2011.

ISSN: 1212-4702
Vytisklo: Nakladatelství ČVUT, výroba

© ČsAS
NEPRODEJNÉ!