ROUGHNESS OF VIOLIN TONES – THE PERCEPTION OF IRREGULARITIES

Zdenek Otcenasek, Jan Otcenasek

Musical acoustics research centre (MARC), Music and Dance Faculty,
Academy of Performing Arts in Prague, Czech Republic
otcenasek@hamu.cz

ABSTRACT

The article presents the results of a second part of a larger experimental study on the perception of different types of changes in the time courses of rough violin sound signals**.** In the first part of the study, the sounds of a bowed, open violin string (G), played with different bow speeds and force, resulting in sounds differing in roughnesses, were simultaneously documented by a high speed video camera and recorded. The recordings were used as stimuli in a ranking and rating and pair comparison listening test. The roughness dissimilarity ratings and perceived difference verbal attribute descriptors were connected with stimuli positions in a MDS perception space (for the details of the test conduct and evaluation, see [13]). In the presented second part, the revealed multidimensionality of perceived roughness is studied in the context of time course changes in both the audio and the string motion signals. The signal analysis reveals connection of the two of the possible dimensions of roughness (cracked and buzzing percepts) to a signal variability. The cracked percept is connected to irregularities in both signals, and the buzzing percept, to a superposition of regular waveforms of neighboring harmonics (one bark harmonic triads) in two barks of the audio signal.

1. INTRODUCTION

Roughness plays a role in the perception of musical sounds, such as those of a violin, for example. Unpleasantly rough tones of a beginner violinist, or an effort in attaining desirable rough tones of electric guitars might represent instances of a roughness related sound-quality evaluation which can be directly important to musicians.

The roughness psychoacoustic quality is used to describe a complex phenomena associated with subjective perception of different temporal changes in sound. Some roughness concepts regard roughness as a result of envelope fluctuations in time and spectral domain and link auditory roughness to the sound waveform, other have suggested physiological bases of roughness (e.g. interference of multiple harmonics of a complex tone at the basilar membrane [1], higher order auditory grouping [2] or neural bases of roughness [3]). Roughness is also studied in relation to the concept of dissonance, sub-harmonic components of the sound, and with character of aperiodicities in sound signal, e.g. in [4, 5].

Since the causes of roughness can arise from several physical reasons, it is suggested that perceived roughness is multidimensional (e.g. in [6, 7]. Also studies of pathologic voice quality based on listening tests e.g. [7] show results of more than one factor or dimensionality in roughness ratings). The results of listening tests focused on timbre of violin tones have also revealed roughness multidimensionality (buzzing, subharmonicity [8]).

The violin (or the bowed instruments generally) sound due to repeating slip-stick release cycle of the string under the bow. The time courses of the string waveform were explored in detail in multiple studies of bowed violin string motion, e.g. [9]. The string exhibits irregularities or aperiodicities in oscillation caused by irregularities in the timing and magnitude of the string release during play. This results in various time changes in sound signal waveforms which are also joined with distinct perceptual quality. Both periodic and aperiodic modulations and subharmonicity were described in the string movement (various extent of jitter, shimmer and spikes). These time changes are dependent to a large extent on the used bowing technique (on bowing place, speed, pressure force, broadness of the bow hairs in contact, e.g. [10, 11]).

In relation to these results, the goal of the entire experiment was to use different regimes of bowing to obtain a variety of different string motions and roughness forms [9] to explore the multidimensionality of perceived roughness. The goal of the presented research was to investigate the relation between the forms of roughness, the string motion and the radiated sound and relate some forms of roughness (described with suitable verbal descriptors) to the character of the sound signal and string motion near the violin bridge.

2. SUMMARY OF THE EXPERIMENT

2.1 The experiment overview

The experiment was conducted in the first part of the study (for further information on the experiment see [13]). Different types of rough violin sounds were generated by changes in the bowing technique (performed by a musician) in the first part of the experiment. The tones were played on an open G string, on a single violin, and recorded with a Neumann KU100 dummy head. Corresponding string movements were recorded with a Phantom SpeedSense 9060 high speed camera (60kfps). Both recordings were synchronized (see Fig. 1).

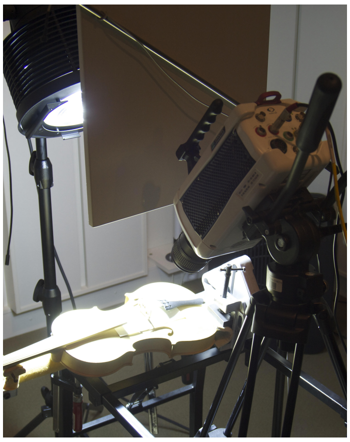

Figure 1: The setup of the sound and string movement recording

An acceptable number of 9 sounds was selected by rejecting similarly sounding recordings in a listening pre-test in order to obtain samples representative of different roughness types. The selected sounds were evaluated by listeners in two listening tests (headphones, mono). Roughness dissimilarity and verbal attribute descriptors were collected in a pair comparison test. Roughness as an entirety was evaluated in a ranking and rating listening test. Tests were done with 20 participating subjects: 10 sound engineers and musicians (the *experts* group), and 10 participants without special musical or listening experience, (the *non-experts* group).

2.2 The experiment results summary

The obtained dissimilarity values were analyzed using a multidimensional scaling method (MDS [12]). Both tests had a statistically significant agreement among all respondents, but the stress values (D) of the *experts* differed from the *non-experts* in solutions with a same dimensionality (see the more distinct elbow for the experts solution line at number 4 in Fig. 2), indicating that experts considered more sound quality attributes: While a 2D solution was sufficient for the non-experts, a solution with less than 4D was not suitable for unbiased preservation of the dissimilarities obtained for the experts (only the more distinct and detailed experts 4D MDS sound configuration is used in the second stage of the experiment in result**,** see chapter 3.)

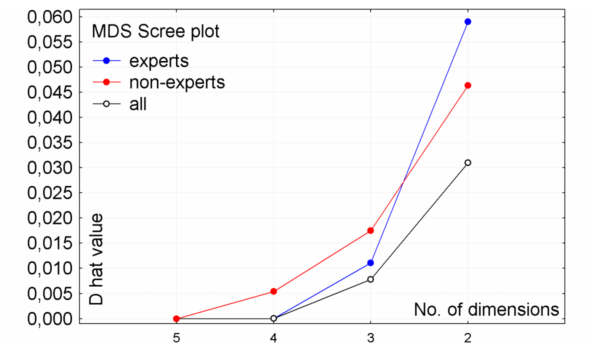

Figure 2: *non-experts*, *experts* and *all* MDS solution scree plots (from 5 to 2 dimensions)

The respondents used slightly different spontaneous verbal descriptors for the same sound qualities in the pair comparison test. Synonymous meanings were merged after post-test consultations with respondents, and the most frequently used descriptor was chosen as a representative label of a certain sound quality attribute. A frequency of occurrence (F_{occ}) of every attribute was computed for each sound stimulus (summarized in Table 1). These values of F_{occ} were embedded to the 4D experts perception space using a linear regression.

The embeddings were used for the interpretation of the the perceptual space configuration and consecutively for a search for the physical aspects of the obtained dimensionality

Fig. 3 displays the spatial directions along which the embedded, separate perceptual attributes increase. The gradients are displayed as colored lines from the positions of sounds with minimal to maximal frequency of occurrence (a small and a large circle) labelled with identical colour in the inscription (the positions are marked with the sound numbers).

Table 1: The overall roughness and F_{occ} of the attributes

Sound. no	Roughness: Rank+rate test			F_{occ} (Pair comparison test)							
	all	experts	non-exp.	Cracked	Strident	Buzzing	Rustle	Blear	Dark-Bright	Narrow	Quinted
1	4.6	4.4	4.8	1	1	1.5	2.5	4	4.5	1	0
2	9.5	9.4	9.5	10	3	0	0.5	1	3	3	0
3	3.3	2.1	4.4	0	3	7.5	0	0	1	0	0
4	1.9	2.4	1.5	0	0	2.5	1.5	4.5	6	6	0
5	6.1	6.4	5.8	4	2.5	0.5	3	4	5	3	0
6	8.3	7.9	8.7	7	8	0	1	0.5	0	1	0
8	5.3	3.9	6.6	0	2.5	7	1	1	1.5	9	0.5
9	2	1.5	2.4	0	0.5	4	0.5	3	3.5	1	0
10	4.4	4.6	4.1	1	1	4.5	1	2	2	3	8

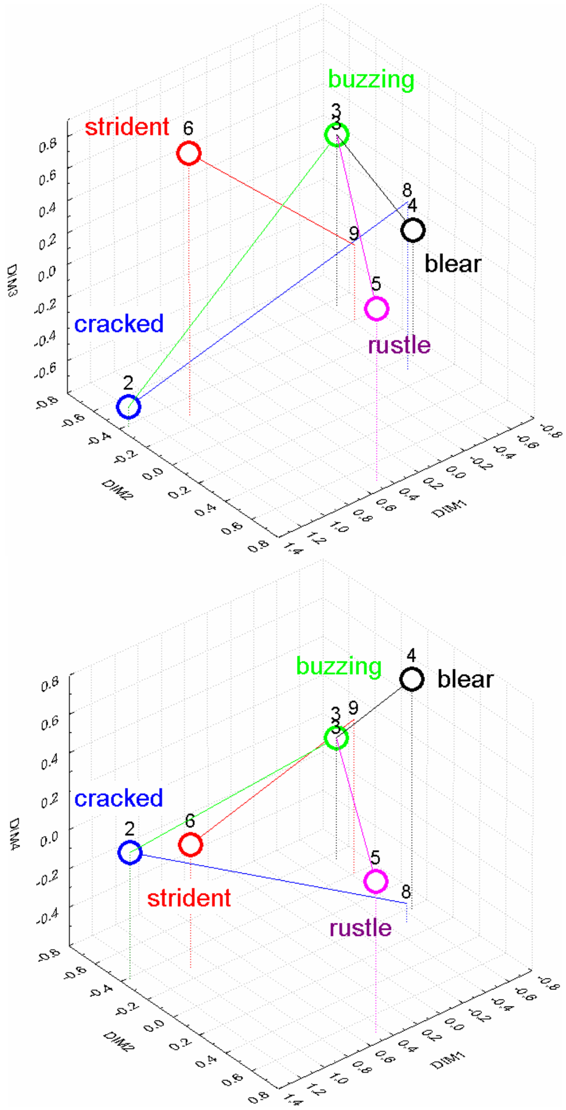

Figure 3: Attribute gradations (marked in different color) from minimum (no circle) to maximum (large circle). Digits mark the positions of sounds in the experts perception space. Top figure: dimensions 1, 2, 3, bottom figure: dimensions 1, 2, 4.

The presented F_{occ} and regression lines were used for the analysis of the roughness causations in the presented second part (see chapter 3).

The obtained generally known psychoacoustic dimensions (the roughness jugged as an entirety in the rating test, the dark-bright and narrow dimensions) were also embedded into this space (see Fig. 4).

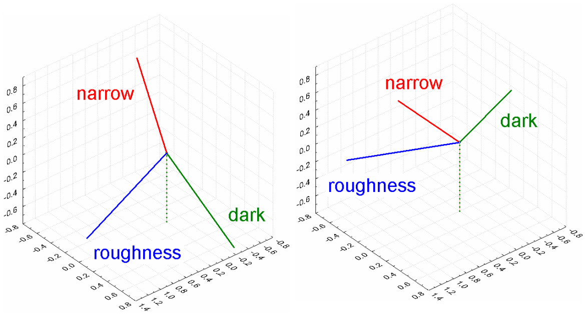

Figure 4: Gradation of regular psychoacoustic dimensions (in different colours), only the positive half-axis from the space midpoint (0, 0, 0, 0) is shown. Left figure shows dimensions 1, 2, 3, right dimensions 1, 2, 4.

Angles between the regression lines were computed to assess the relations between the embedded attributes and to ease the orientation in the 4D perceptual space (for summary of the results, see Table 2). The embedded attribute angles revealed the multidimensionality of roughness (the cracked and buzzing percepts constitute the roughness percept as an entirety) [13].

Table 2: Angles between the attribute regression lines

	strident	buzzing	rustle	blear	dark-bright	narrow	roughness
cracked	65	127	95	100	93	85	32
strident		88	101	129	133	27	61
buzzing			125	130	133	79	116
rustle				39	51	105	82
blear					16	121	104
dark-bright						125	103
narrow							89

The right angles (90±15°, red in Table 2) between buzzing-strident, cracked-rustle, strident-rustle and cracked-bleary attributes indicate these sound attributes are potentially discrete psychoacoustic dimensions or quantities (see Table 3.)

Table 3: Potential discrete psychoacoustic dimensions (left column: sound quality attributes, right: sound quality renditions)

Attributes:	Perceived sound quality description:
bleary	blurred and cloudy sounding
buzzing	buzzing like sound
strident	Piercing, cutting, and gradation of sharpness
cracked	interrupts and time variations in sound

Oblique angles (20~70° and 110~160°, black in Table 2) may indicate non-discrete attributes, but can also result from insufficient number of used stimuli in the 4D perception space (where a certain attribute direction might be inhomogeneously filled-in), and can also arise from a need of higher real space dimensionality than used.

2.3 Sound perception implications summary

The embeddings of all of the attributes into the perception space revealed:

· Roughness (as an entirety) is not collinear with any of the other obtained attributes.

· The cracked and buzzing percept both contribute to the perceived roughness (more in cracked and less in buzzing, 32° and 64° angles respectively).

· Neither rustle or bleary contribute to roughness (90±15° right angles).

· The strident quality increases with an increase in narrowness (27° angle) and brightness (47° angle; brightness is opposite to darkness (133° angle)) and also to roughness (61° angle). This is in accordance with previous studies of sharpness [14, 15], since the strident percept is a gradation of sharpness. In these studies, sharpness was similarly related to: 1) dark, gloomy ↔ clear, bright; 2) narrow ↔ wide, full; 3) rough, harsh ↔ soft, delicate

· Narrowness is independent to buzzing, cracked and rustle (all are angled 90±15° to narrowness) and also is almost opposite to bleary (121°).

· Darkness and bleary are nearly collinear (16° angle). More bleary sounds were also darker.

· Buzzing is partially opposite to bleary (130° angle) and darkness (133° angle). The bleary and dark sounds were less buzzing.

3. THE CAUSATIONS OF ROUGHNESS

One of the results of the described first experiment part is that roughness is a two dimensional phenomena in the used stimuli context, and the perception of roughness judged as an entirety might be predicted from the amount of the cracked and buzzing percepts. The next part of the study focused on the possible causations of these attributes.

The following analysis is based on the time courses of the microphone signal and the string movement near the bridge. The time course of the violin string motion represents a trajectory of a chosen location on the string. The tracking was performed using a software, tracing the best correlating pattern in a defined area between frames (movement of defined pixel patterns on a surface)

3.1 The *cracked* percept and irregularities

The time courses of highly cracked sounds displayed visible periodicity irregularities in both the string motion and in the sound signal.

The values of the autocorrelation function 1st maximum used as a measure of signal invariability (Fig. 5) clearly differentiated the highly cracked sounds. But the lesser cracked and buzzing sounds could not be discriminated using this method (see Fig. 6).

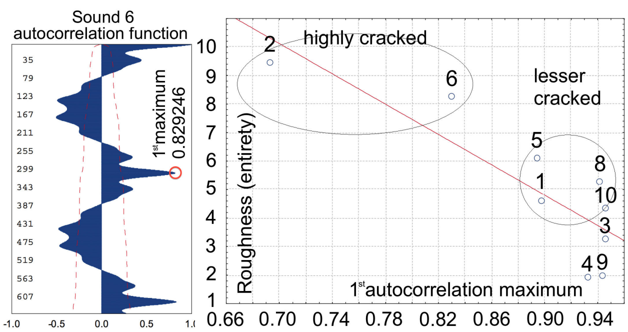

Figure 5: Left: An example of the autocorrelation function and 1st maximum values (the peak in the middle right) Right: The sounds ranked along the values of the 1st maximum of the autocorrelation function and roughness.

The string movement and the sound signals variability was therefore analyzed using a different approach. Sequential correlations of neighbouring time windows with the length of one tone period (SC1P) were used for a between period variability evaluation. The signals were divided into segments with the duration of one period (based on the f_0 of the signal) and the samples of every segment were correlated to the succeeding one using spearman correlation. Time courses with less time changes between the segments had higher correlation coefficients between adjacent periods. The resulting SC1P coefficients are shown in Fig. 6 to 9.
In Fig. 6 and 7 the SC1P values for every sound (colored circles) are ranked on the y-axis according to the overall roughness of the sound, and grouped by the most frequent attribute for every sound (the buzzing sounds are in the left graph, the cracked/other sounds are in right graph). The label displays 2 most frequent attributes for each sound. Sound 4 features the lowest overall roughness rating, and is shown in both graphs.

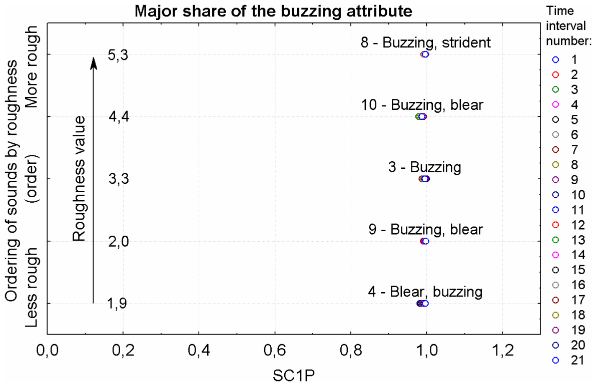

Figure 6: SC1P and sound roughness values for the 21 time segments (each segment represented by a circle) of each sound waveform with a major share of the buzzing percept

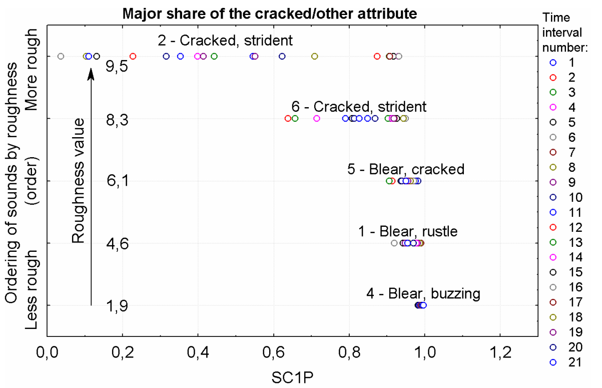

Figure 7: SC1P and sound roughness values for the 21 time segments (each segment represented by a circle) of each sound waveform with a major share of the cracked and other percepts

While the mean SC1P values of sound signals with cracked/other qualities decrease with increasing perceived roughness (and their variance increases), the SC1P values in sound signals with buzzing remain relatively high and with relatively little variance as the values of the roughness rating of the sounds increase.
The SC1P results for the string movements also correspond to these results (see the SC1P values of all sounds in Fig. 8; the cracked sounds 2, 6, 5, 1 have both relatively lower mean and greater variance of SC1P values, and the buzzing sounds 3, 8, 10, 9 show relatively lower variance and higher mean values).
The SC1P correlation values remain relatively high and invariable in the time courses of both signals in sounds with increasing extent of the buzzing attribute, and do not increase with increasing extent of buzzing.

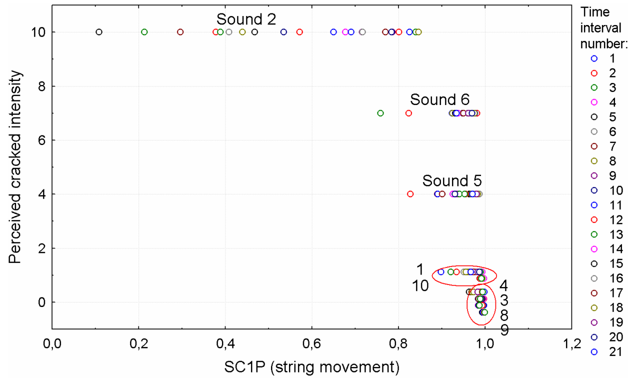

Figure 8: The SC1P values of the string movement for all of the studied sounds (the sounds 4, 3, 8, 9 in red ovals have value 0 of the cracked attribute and the sounds 1, 10 have value 1).

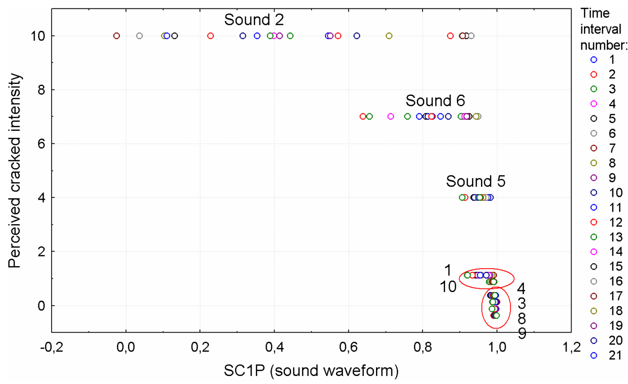

Figure 9: The SC1P values of the sound signals for all of the studied sounds (see legend in Fig.8).

The SC1P values for the string and for the sound signal are similar in values and trends, although there are distribution variations, which have yet to be investigated.

3.2 The *buzzing* percept and harmonics in a bark

The buzzing sounds display very little between period changes in the time courses (see SC1P results in 3.1). Since the spectral waterfalls show more harmonic components with higher amplitudes in the frequency ranges above 3 kHz (when compared with the other sounds) a further approach was used to analyze the signals. The study focused on the interactions of harmonics within a single bark, following to the relationships already reported in previous studies [16, 17]. The sound signals were bandpass filtered using a one bark wide filter. The bark number was chosen with regards to the f_o of the considered sound and the requirement of a three harmonics content (The appropriate bark is centered around the f_o multiples. The 2940 to 3450 Hz band contains the 15^{th}, 16^{th}, 17^{th} harmonics in a 196 Hz open G string tone). Listening to this sound filtered in one bark featuring three harmonics showed that its buzzing considerably differs from an unfiltered one. A broader filtering in the two bark 2950 - 4050 Hz range (TBB) resulted in a more buzzing-like sound, therefore this TBB filtering was used. An example of a TBB signal time course is shown in Fig. 10 Top (shown for sound 5). The lines between the red circles in the graph mark an amplitude of the envelope rise. The length of the line specifies the rise amplitude used as a characteristic of the buzzing extent. The values are gathered in the Table 4 (the first number in the *Rise* row marks the lowest rise found in the TBB bandpass filtered signals / the second marks the highest; *Aver.Rise.* presents the averages of the rise values).

The amplitude variability off TBB signals was also analyzed by a correlation method. A one period autocorrelation window was used likewise to the SC1P. Fig. 10 bottom shows an autocorrelation function example of TBB filtered signal for sound 5 (only the positive autocorrelation values are presented). The displayed red circle (with a value) marks further characteristic employed for buzzing extent estimation (see values in the *Correlat.* row in Table 4).

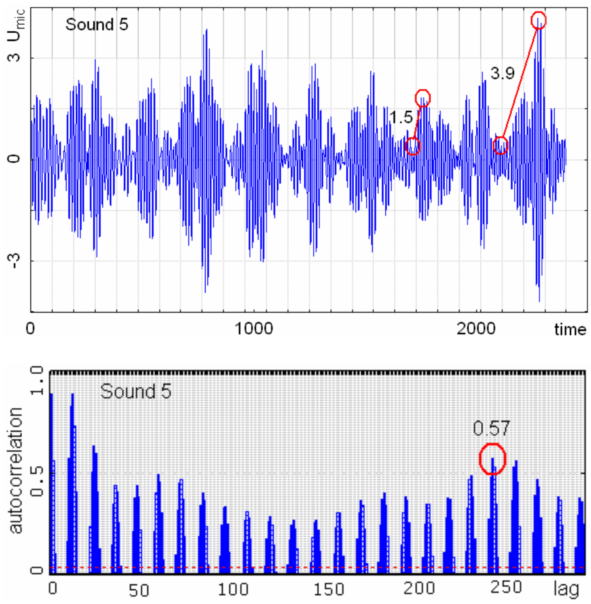

Figure 10 Top: The time course of the sound signal after the two bark bandpass filtering (TBB) in the 2940 to 4040 Hz range. Bottom: Positive values of the autocorrelation function of one 5.1 ms period (196 Hz) of TBB filtered signal.

Table 4: The buzzing value and the characteristic used for the buzzing amount estimation

Sound	**3**	**8**	**10**	**9**	**4**	**1**	**5**	**2**	**6**
Buzzing	**7.5**	**7**	**4.5**	**4**	**2.5**	**1.5**	**0.5**	**0**	**0**
Rise	**4.8 /5.7**	**4.2 /5.6**	**2.3 /3.8**	**2.3 /3.5**	**1.5 /2**	**2.8 /5.2**	**1.5 /3.9**	**3.5 /11.8**	**1.7 /6**
Aver.Rise	**5.3**	**4.9**	**3.2**	**3**	**1.8**	**4**	**2.8**	**4**	**3.5**
Correlat.	**0.86**	**0.87**	**0.86**	**0.89**	**0.8**	**0.66**	**0.57**	**0.37**	**0.26**

The steadiness of the risings (a regularity of the filtered signal) is demonstrated by a low variation in the rise values obtained in different time sections. The values in Table 4 show that the sounds with high correlation value and with smaller average rise value invoke smaller buzzing sensation (compare e.g. the sound 3 and 9). Low correlation values also characterize the cracked percepts. The buzzing sound quality percept is conditioned by regularity of harmonic triads in a bark. A higher unfiltered signal variability could suppress the buzzing causation by disrupting the interactions of the harmonic triads in the two barks by the multiple cracked irregularities (small correlations and large differences between the lowest and the highest rise will occur, see sounds 1, 5, 2, 6 in Table 4). The harmonic triad interactions are also influenced by the harmonic levels (the harmonics in one bark had to have roughly comparable levels (level uniformity)).

4. CONCLUSIONS

The results of the listening tests show that roughness judged as an entirety (in the used violin tone context) might be predicted from the amount of the cracked and buzzing percepts (see [13] and a summary in 2.3). The presented results links the causations of these attributes to a between period variability in both studied signals.
The between period similarity decreases in violin tones with a major share of the cracked percept with increasing perceived roughness in both the string movement course and the sound signal waveform. This decrease in similarity mainly occurs due to irregularities in time courses of the sound signal waveforms. The between-period similarity analyzed by sequential correlations of neighbouring time windows with one tone period length (SC1P) appears to be an overall indicator of the cracked roughness quality (in the cracked dimension).
On the other hand, an increase in the buzzing quality was not directly accompanied by an increase in the SC1P values (sounds with high buzzing quality have high SC1P coefficients already). The causes of the buzzing perception might be explained more reasonably based on the characteristics gathered in Table 4. Three sufficiently stable harmonics in one bark in the time course might form an amplitude modulation by superposition. The results of the two barks filtering analysis (TBB) indicate that the buzzing percept is caused by the modulation interactions in the two neighboring barks. The buzzing sound quality is perceived with more intensity if the pulsations of TBB filtered signals have higher rises and the risings are regular. The average of the rise values obtained in the TBB time sections with duration of one tone period duration of the tone fundamental frequency has shown to be a suitable characteristic of the rising height in this study. The value of the autocorrelation function in the same time section is a suitable characteristic of the risings regularity.

5. ACKNOWLEDGMENTS

Supported by the Ministry of Education, Youth and Sports of the Czech Republic in the Long Term Conceptual Development of Research Institutes grant of the Academy of Performing Arts in Prague: The "Sound quality" project.

6. REFERENCES

[1] H. Fastl, E. Zwicker: Psychoacoustics facts and models, Springer, Berlin (1990)
[2] Ch.G. Tsai: Auditory Grouping in the Perception of Roughness Induced by Subharmonics: Empirical Findings and a Qualitative Model, Proc. Int. Symposium Musical Acoustics. Nara, Japan (2004)
[3] M.J. Tramo, P.A. Cariani, B. Delgutte, L. D. Braida: Neurobiological Foundations for the Theory of Harmony in Western Tonal Music, The Biological Foundations of Music, 92-116 (2001)
[4] J. Kreiman, B.R. Gerratt: Perception of aperiodicity in pathological voice“, J. Acoust. Soc. Am. 117, 2201-2211 (2005)
[5] K. Omori, H. Kojima, R. Kakani, D.H. Slavit, S. Blaugrund: Acoustic characteristics of rough voice: subharmonics, J. of Voice. 11, 40-47 (1997)

[6] C.G. Tsai, L. C. Wang., S.F Wang, Y.W. Shau, T.Y. Hsiao, W. Auhagen: Aggressiveness of the Growl-like Timbre: Acoustic Characteristics, Musical Implications and Biomechanical Mechanisms, Music Perception. 27 (3), 209-221 (2010)

[7] J. Kreiman, B.R. Gerratt, G.S. Berke. The multidimensional nature of pathologic vocal quality, J. Acoust. Soc. Am., 96, 1291-1302 (1994)

[8] J. Stepanek, Z. Otcenasek,: Interpretation of Violin Spectrum Using Psychoacoustic Experiments, Proc. Int. Symposium Musical Acoustics. Nara, Japan (2004)

[9] M.E. McIntire, R. Schumacher, J. Woodhouse: “Aperiodicity in bowed-string motion“, Acta Acustica, 49(1) 13-32 (1981)

[10] J. Woodhouse, P.M. Galluzzo: The bowed string as we know it today, Acta Acoustica, 90, 579-589 (2004)

[11] M.E. McIntire, J. Woodhouse: On the fundamentals of bowed-string dynamics. Acta Acustica, 43, 93-108 (1979)

[12] J. B. Kruskall: Nonmetric multidimensional scaling: a numerical method. Psychometrika 29, 115-129, (1964)

[13] Z. Otčenášek, J. Otčenášek,. Perception of different Types of roughness of violin tones. In: Proceedings of the International Symposium on Musical Acoustics, - Le Mans, France. p. 557-62, (2014) www.conforg.fr/isma2014/cdrom/data/articles/000102.pdf

[14] J.Štěpánek, Interpretation and comparison of perceptual spaces, Proceedings of Inter-Noise 2005, Rio de Janeiro (2005), http://zvuk.hamu.cz/vyzkum/publikacni.php

[15] J. Štěpánek, Z. Otčenášek, Acoustical correlates of the main features of violin timbre perception, Electronic proceedings CIM05 Montréal 2005, Québec, Canada, http://zvuk.hamu.cz/vyzkum/publikacni.php

[16] Z.Otčenášek, J.Štěpánek, "Glossy" and "Buzzing" in timbre of violin sounds, 6 CFA, Lille (2002), str.691-694, http://zvuk.hamu.cz/vyzkum/publikacni.php

[17] J.Štěpánek, Z.Otčenášek, Interpretation of Violin Spectrum Using Psychoacoustic Experiments, CD of Proceedings of the International Symposium on Musical Acoustics (ISMA2004), Nara, Japan, 4-S2-1 (2004), http://zvuk.hamu.cz/vyzkum/publikacni.php