Akustické listy, **20**(3–4), prosinec 2014, str. 19–26

Roughness Prediction for Complex Acoustic Stimuli

Predikce vjemu drsnosti komplexních akustických signálů

Václav Vencovský[a,b]

[a]Musical Acoustics Research Center, Academy of Performing Arts in Prague, Malostranské nám. 12, 118 00 Praha 1, Czech Republic
[b]Department of Radioelectronics, Czech Technical University in Prague, Technická 2, 160 00 Praha 6, Czech Republic
vaclav.vencovsky@gmail.com

Vencovský [12] introduced a new roughness model and showed its performance in comparison with listening tests for three types of complex acoustics stimuli: amplitude-modulated harmonic complex tones, real samples of pathological voices (sustained vowel /a/), harmonic intervals of the chromatic scale composed of two harmonic complex tones (dyads). This study extends his results. It adds a new type of stimuli (low-frequency harmonic complex tones), slightly changes the method used to estimate roughness, and, for a comparison, depicts results for Daniel and Weber's roughness model [9], and Leman's synchronization index (SI) roughness model [10]. Vencovský's model performed well for all of the stimuli. Daniel and Weber's model performed well for the first type of stimuli, but its results were poor for the rest of the stimuli. The SI model performed well for the first three types of the stimuli but poorly for the low-frequency harmonic complex tones.

1. Introduction

Two pure tones added together create a signal with a fluctuating envelope. The frequency of the fluctuations equals the frequency difference between the tones. We may perceive the fluctuations as periodic changes of the signal loudness when the frequency of the fluctuations is below 30 Hz. When the frequency of the fluctuations slowly increases above 30 Hz, we may perceive a jarring and rough sound sensation described by the term roughness [1]. As the frequency increases, the sensation reaches its maximum and then starts to decrease. We then perceive both tones separately with pitches given by their frequency. This suggests that we perceive the roughness because the ear cannot resolve the spectral components of the sound which are close in frequency [1, 2].

Researchers have employed various psychophysical methods to measure roughness [3–7] and given roughness a unit called asper [6]. Hand in hand with the psychophysical methods, researchers developed various mathematical models predicting roughness (see [2, 6, 8–10]) but none of them have been standardized [11]. One group of the models, known as *curve-mapping* models, detects spectral components of the sound stimuli and maps them into a psychoacoustical curve of roughness [2]. However, the *curve-mapping* models cannot process signals with continuous spectra, for example, noises [10]. The second group of roughness models employs algorithms simulating the function of the peripheral ear [10]. Daniel and Weber's model [9], and Leman's synchronization index (SI) model [10] belong to this group.

The aim of this study is to extend the results of the study of Vencovský [12] which introduced a new roughness model. The model employs an algorithm simulating the function of the peripheral ear and thus belongs to the same group as Daniel and Weber's and the SI roughness models. Vencovský [12] used the model to predict results of listening tests for: amplitude-modulated harmonic complex tones, real pathological voice samples, and harmonic intervals of the chromatic scale. I slightly changed the method used to estimate roughness, verified the roughness model by a new type of stimuli (low-frequency harmonic complex tones), and, for a comparison, depicted predictions given by Daniel and Weber's, and the SI roughness model. Some of the used real pathological voice samples were rough and breathy. Speech synthesizers simulate breathiness by adding noise to the signal [13]. Since the *curve-mapping* roughness models cannot predict roughness for noise stimuli, I did not use any model of this group [10].

2. Roughness models

This section basically describes three different roughness models: Vencovský's model [12], Daniel and Weber's model [9] and synchronization index (SI) model [10].

2.1. Vencovský's model

Vencovský presented the model in the study [12]. The model has a cascade architecture consisting of two stages: a peripheral stage simulating the function of the human auditory system, and a central stage predicting roughness from the simulated neural signal at the output of the peripheral stage.

The peripheral stage – a computational auditory model – transforms an input acoustic stimulus into the simulated neural signal in auditory nerve fibers. It is composed of: an outer and middle ear model, a model of the basilar

Přijato 24. října 2014, akceptováno 11. prosince 2014.

membrane (BM) response and cochlear hydrodynamics developed by Mammano and Nobili [14–16], and algorithms simulating the function of the inner hair cells [17–20]. The signal at the output of the peripheral stage represents the simulated neural signal. Since the peripheral ear conducts spatial-frequency analyses, the output signal has 300 channels: each represents response of the peripheral ear tuned to a certain frequency between 16 Hz and 17 kHz.

The central stage calculates roughness from the simulated neural signal obtained at the output of the peripheral stage. It first extracts the signal envelope and then process it by a 1st-order Butterworth low pass filter with a cut-off frequency of 80 Hz. The filter ensures that the roughness decreases for fluctuations of higher frequencies. The central stage then extracts two features from the parts of the envelope where the signal is rising (rising slopes): duration of the rising slope, and difference between the minimal and maximal value. Roughness is estimated as a product of the frequency calculated from the duration of the rising slopes and the modulation depth calculated from the minimal and maximal value of the rising slopes [12]. This statistics takes into account the shape of the time domain envelope of the simulated neural signal and the roughness model thus predicts roughness for the stimuli not covered by Daniel and Weber's and the SI roughness model [12].

Vencovský [12, 21] designed two versions of the roughness model with a slightly different central stage. The model version described in the study [21] predicts roughness in dependence on parameters of various synthetic stimuli, for example, on the modulation frequency for amplitude-modulated tones. The study [12] presents a model which is more suitable for predicting roughness of various types of complex stimuli. This letter version of the roughness model is used in this study.

2.2. Daniel and Weber's model

Daniel and Weber [9] improved a model designed by Aures [8]. The peripheral part of the model transforms the input stimulus into a spectrum and divides it into bands of a bandwidth equal to the critical bandwidth in Barks [6]. It then turns the spectral representation of the signal in each band to the time domain and estimates the modulation depth of the signal envelope. The model employs a bank of weighting filters and was tuned to account for roughness of amplitude-modulated tones experimentally measured by Aures [22] and presented in the book of Fastl and Zwicker [6]. The model is calibrated to aspers. Daniel and Weber showed that the model accounts for roughness of various types of synthetic stimuli, for example, amplitude and frequency modulated tones, and unmodulated bandpass noises [9]. Wang *et al.* [11] used the model to predict roughness of vehicle noise.

2.3. Synchronization index model

Leman [10] designed the synchronization index (SI) model. The model employs an auditory model which transforms the input acoustic stimulus into bands of the bandwidth equal to the critical bandwidth. The signal in each band represents the simulated neural signal in auditory nerve fibers. The simulated neural signal is then filtered by bandpass filters in order to predict the bandpass characteristics of roughness when it is depicted as a function of the modulation frequency [6]. The model calculates a degree of phase locking of the filtered simulated neural signal to a particular frequency as a ratio between the short-term spectra of the simulated neural signal and the DC component of the simulated neural signal summed across all bands. Leman [10] depicted the model predictions for amplitude-modulated tones and for harmonic intervals of the chromatic scale but did not compare the results with subjective data. Wang [23] used this model to predict roughness of vehicle noise.

3. Experiment 1: Roughness of amplitude modulated harmonic complexes

Experiment 1 estimates the roughness of amplitude-modulated (AM) complex tones. I conducted a rating listening test to measure the roughness of the stimuli and compared the subjective results with predictions of the roughness models.

3.1. Method

Stimuli Harmonic complexes composed of the first three harmonics at frequencies of 300, 600 and 900 Hz were used as stimuli. All three harmonics were amplitude-modulated by the same sinusoidal signal with a frequency of 30, 40, 50, 60, and 70 Hz. Modulation depth – calculated as $20\log_{10} m$, where m is the modulation index ranging from 0 to 1 – was 0, −3, −6, −9 and −12 dB. Amplitude of the first, second and third spectral component was 0, −10 and −20 dB, respectively. Duration of the stimuli was 600 ms and they were ramped on and off with 30 ms raised-cosine ramps. The level of the stimuli was 75 dB SPL. Combination of the modulation frequencies and the modulation depths led to 25 different stimuli.

Listeners: Five experienced listeners - one woman, four men, age ranging between 25 and 44 years, including the author of this study - participated in the experiment. The listeners had normal hearing; pure-tone thresholds below 20 dB HL for frequencies between 250 Hz and 8 kHz.

Procedure and equipment: Roughness of the stimuli was rated on a discrete scale from 1 to 7 in steps of 1, where 1 was for the lowest and 7 for the highest roughness. This procedure was inspired by the study of Patel *et al.* [7] which estimated the roughness of pathological voice

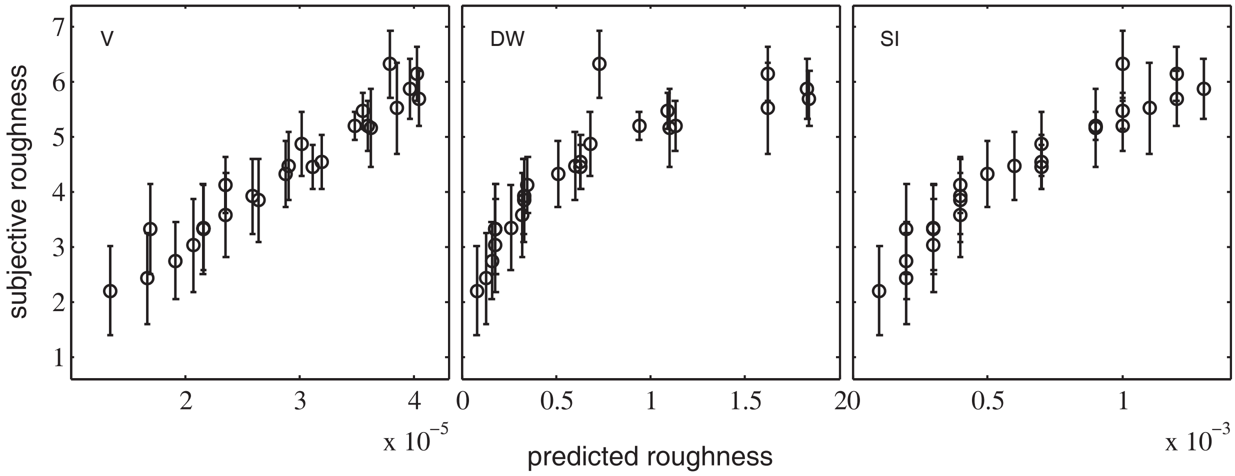

Figure 1: Mean values of the subjective ratings of roughness for amplitude-modulated harmonic complex tones plotted as a function of the roughness predicted by Vencovský's model (left panel), Daniel and Weber's model (center panel), and the SI model (right panel).

stimuli on a 5 point scale. Just noticeable difference of roughness corresponds to about 10% change of the modulation index, m, of a sinusoidally amplitude-modulated tone [6]. The five chosen values of the modulation depths (0, −3, −6, −9 and −12 dB) of the AM complexes should thus cause perceptible changes of roughness. Moreover, the roughness of the stimuli depends as well on the modulation frequency [6]. Thus for the AM complex stimuli, a 5 point scale seemed to be too course. The listeners rated the roughness of 25 different stimuli presented in a random order. Each stimulus was rated separately. The listeners could listen to it as many times as they desired and after assigning the roughness rating, they were presented with the next stimulus. The listening test was composed of 10 sets of randomly ordered 25 stimuli. In other words, each stimulus was rated 10 times giving the overall number of 250 stimuli in the test. The test was conducted on a computer. The stimuli were presented diotically (the same signal to both ears) via Sennheisser HD-600 headphones.

Roughness prediction: The roughness of the AM complexes was predicted by means of three types of roughness models: Vencovský's model [12], Daniel and Weber's model [9] implemented in the sound analyses software PsySound3 [24], and the synchronization index (SI) model [10] implemented in IPEM toolbox [25]. All types of models predict roughness in the time domain and give time-dependent values of roughness. Daniel and Weber's model implemented in PsySound3 and the SI model implemented in IPEM toolbox calculates the resulting value of the predicted roughness as median of the time-dependent values. Since both models calculate medians, this method is in this study also used for Vencovský's roughness model. It differs from the previous study of Vencovský [12] where the maximum of the time-dependent values of roughness was taken. This change did not have any strong effect on the predicted roughness in comparison to the previous study [12].

3.2. Results

Each stimulus was rated ten times, but the first two ratings were not taken into account for the final processing of the results. Cronbach's alpha calculated from the ratings given by each listener was in all cases higher than 0.8 with 5% level of significance, which means that the listeners were reliable. I estimated as well the intersubject reliability between mean ratings for each stimulus. Cronbach's alpha was 0.951 with 5% level of significance.

Fig. 1 shows mean values and standard deviations from the mean calculated across all listeners and ratings plotted as a function of roughness predicted by Vencovský's model [12] (left panel), Daniel and Weber's model [9] (center panel), and the SI model [10] (right panel). Daniel and Weber's model predicts roughness in aspers. The presented roughness model and the SI model predicts roughness in its own model units. Tab. 1 shows Spearman's (s.c.) and Pearson's (p.c.) correlation. All three types of roughness models gave data which correlate with the subjective ratings.

Table 1: Correlation between model predictions and subjective data for amplitude-modulated harmonic complex tones.

		V	DW	SI
Spearman	r:	0.971	0.961	0.977
cor.	p:	$7 \cdot 10^{-16}$	$2.3 \cdot 10^{-14}$	$6.2 \cdot 10^{-18}$
Pearson	r:	0.973	0.863	0.956
cor.	p:	$4.7 \cdot 10^{-16}$	$2.8 \cdot 10^{-14}$	$9.4 \cdot 10^{-14}$

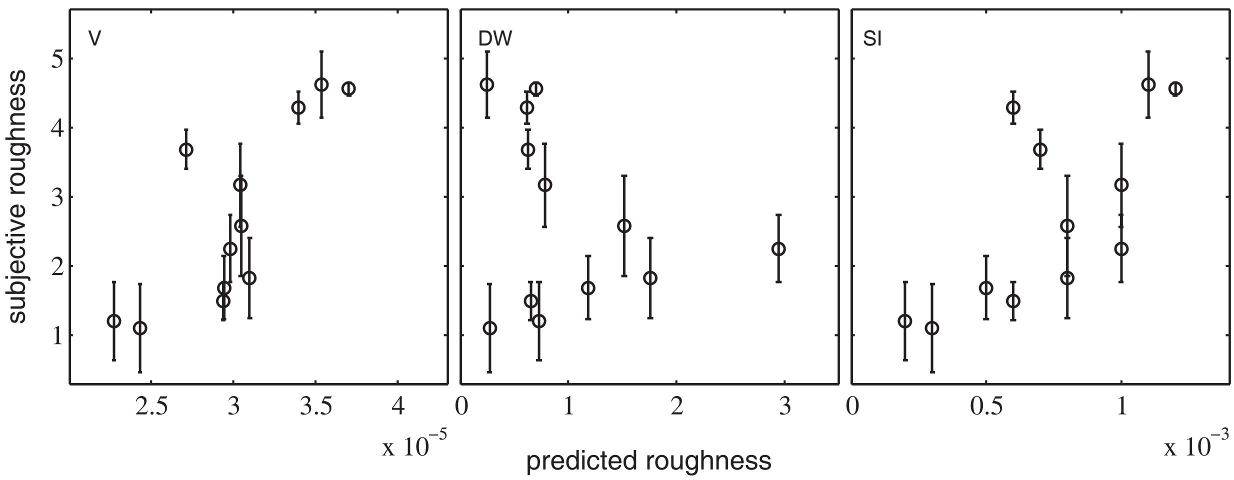

Figure 2: Mean values of the subjective roughness ratings for real voice samples – sustained vowels /a/ – plotted as a function of the roughness predicted by Vencovský's model (left panel), Daniel and Weber's model (center panel), and the SI model (right panel).

4. Experiment 2: Roughness of real voice samples

I have used the same approach as in Experiment 1 to estimate the roughness of real pathological voice samples – sustained vowels /a/.

4.1. Method

Stimuli: The stimuli were 12 real pathological voice samples of a sustained vowel /a/. The samples were extracted from stimuli recorded from 11 different subjects during the scale singing. The subjects had a pathology affecting their larynx. The stimuli differed in the pitch and in the amount of roughness. Duration of the stimuli was 300 ms and they were ramped on and off by 30 ms raised-cosine ramps. The level of the stimuli was 75 dB SPL.

Listeners: Six experienced listeners – men aged between 25 and 36 years, including the author of the study – participated in the experiment. The listeners had normal hearing; pure-tone thresholds below 20 dB hearing level (HL) for frequencies between 250 Hz and 8 kHz.

Procedure and equipment: Roughness was rated on a discrete 5-point scale from 1 to 5 in steps of 1, where 1 was for the lowest and 5 for the highest roughness. The same scale was used in the study of Patel *et al.* [7] estimating roughness of the same type of stimuli – pathological voice samples of a sustained vowel /a/. The procedure and equipment were the same as in Experiment 1. Randomly ordered 12 stimuli were rated 10 times, giving 120 stimuli.

Roughness prediction: Methods used to predict roughness were the same as in Experiment 1.

4.2. Results

As in Experiment 1, each stimulus was rated ten times, but the first two ratings were not taken into account for the final processing of the results. Intrasubject reliability estimated as Cronbach's alpha was for all listeners higher than 0.9 with 5% level of significance. Intersubject reliability estimated as Cronbach's alpha was 0.983 with 5% level of significance. I calculated the final subjective ratings as mean values across all ratings and listeners.

Fig. 2 shows the mean values and standard deviations from the mean of the subjective ratings plotted as a function of roughness predicted by Vencovský's model (left panel), Daniel and Weber's model [9] (center panel), and the SI model [10] (right panel). Daniel and Weber's model predicts roughness in aspers. The presented roughness model and the SI model has its own model units. Tab. 2 shows Spearman's (s.c.) and Pearson's (p.c.) correlation between the subjective and predicted roughness. Daniel and Weber's model predictions do not correlate with the subjective ratings. Vencovský's roughness model predicted roughness data which correlate with subjective results, but the model's performance is poor in the middle of the roughness scale. The results for the SI model are also poor in the middle of the roughness scale.

Table 2: Correlation between the model predictions and subjective data for real pathological voice samples – vowel /a/.

		V	DW	SI
Spearman	r:	0.790	−0.077	0.727
cor.	p:	$3.6 \cdot 10^{-3}$	0.817	$4.9 \cdot 10^{-3}$
Pearson	r:	0.790	0.114	0.680
cor.	p:	$2.3 \cdot 10^{-3}$	0.725	0.011

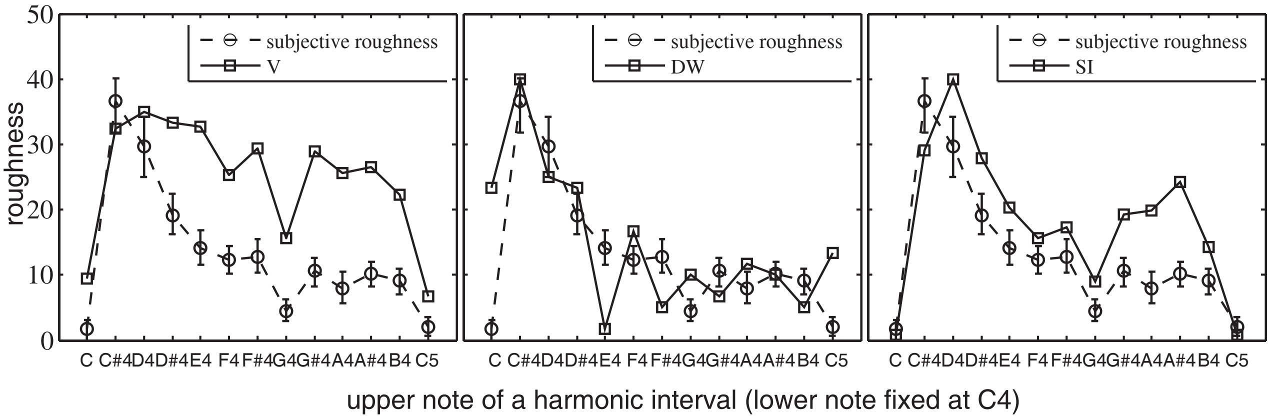

Figure 3: Roughness ratings for the harmonic intervals constructed from synthetic complex tones. Circles connected by dashed lines represent mean values of the subjective data across ten listeners reproduced from the study [2]. Squares are the roughness model ratings for Vencovský's model (left panel), Daniel and Weber's model (center panel), and the SI model (right panel).

5. Experiment 3: Roughness of intervals of the chromatic scale

In Experiment 3, I compared the subjective roughness ratings of harmonic intervals of the chromatic scale with model predictions. I reproduced the subjective ratings of roughness from the study of Vassilakis [2].

5.1. Method

Stimuli: The stimuli were constructed according to the description given in the study of Vassilakis [2]. The harmonic intervals were composed of dyads of harmonic complex tones. The complexes were composed of the first six harmonics with amplitude A_n of the nth harmonic given by the equation $A_n = A_1/n$, where n is the number of the harmonics. The fundamental frequency of the lower harmonic complex tone in the dyad was set to middle C (C4, fundamental frequency 256 Hz, equal temperament). The duration of the stimuli was 1 second and it was ramped on and off by 30 ms raised cosine ramps. The level of the stimuli was 75 dB SPL.

Listeners and procedure: Vassilakis [2] conducted the listening test with ten experienced listeners. They rated roughness on a continuous scale between 0 (not rough) and 42 (rough). The stimuli were presented diotically (the same signal to both ears) via earphones. The listeners' task was to set the position of a scroll according to the perceived roughness.

Roughness prediction: The methods used to predict roughness were the same as in Experiment 1.

5.2. Results

Squares connected by dashed lines in all panels of Fig. 3 show the subjective ratings of roughness of harmonic intervals of the chromatic scale reproduced from the study of Vassilakis [2]. The data are plotted as a function of the frequency of the higher tone in dyads. Circles connected by solid lines represent roughness predicted by Vencovský's model (left panel), Daniel and Weber's model [9] (center panel) and the SI model [10] (right panel). The model predictions were divided by its corresponding maximal values and multiplied by 40 to be in the range between 0 and 40. Tab. 3 shows Spearman's (s.c.) and Pearson's (p.c.) correlation between the subjective and predicted data. The presented roughness model successfully predicts the lowest roughness for the intervals of octave and also reflects the dip for the interval G4. The predicted roughness fits mainly the rank order of the subjective roughness as is reflected by the high Spearman's correlation and lower Pearson's correlation. The data were also well predicted by the SI model. Daniel and Weber's model predictions did not fit the subjective data well.

6. Experiment 4: Low-frequency harmonic complex tones

Miśkiewicz and Majer [26] estimated the perceived roughness of low-frequency harmonic complex tones. The results of their behavioral study are compared with predictions of the roughness models.

Table 3: Correlation between the model predictions and subjective data for harmonic intervals of the chromatic scale.

		V	DW	SI
Spearman	r:	0.918	0.224	0.852
cor.	p:	0	0.46	$3.4 \cdot 10^{-5}$
Pearson	r:	0.764	0.649	0.846
cor.	p:	$2.4 \cdot 10^{-3}$	0.016	$2.7 \cdot 10^{-4}$

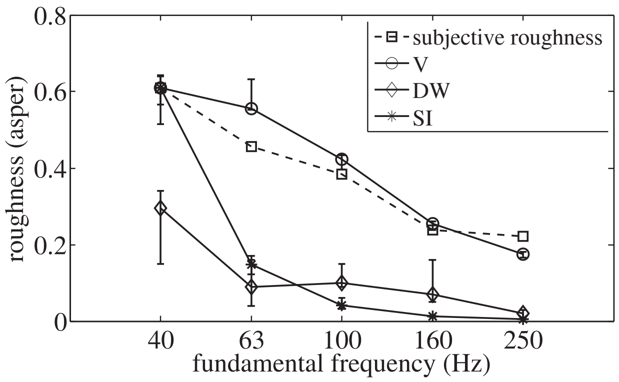

Figure 4: Roughness ratings for low-frequency harmonic complexes. The abscissa denotes a fundamental frequency of the complexes. The ordinate denotes roughness in aspers. Squares connected by dashed lines represent geometric mean values of the subjective data reproduced from the study [26]. Circles, diamonds, and asterisks represent medians and quartiles of the roughness predicted by Vencovský's model, Daniel and Weber's model and the SI model, respectively. Daniel and Weber's model predicts roughness in aspers. The presented roughness model and the SI model data were normalized by its maximal value and multiplied by the maximal subjective roughness.

6.1. Method

Stimuli The stimuli were harmonic complex tones composed of nine harmonics. The fundamental frequency of the harmonics was 40, 63, 100, 160 and 250 Hz. The spectral components were added in random phase. Amplitudes of the harmonics decreased with 6 dB per octave. Loudness of the stimuli was set to 60 phons. It was set by a loudness model designed by Moore *et al.* [27] and implemented in the sound analyses software PsySound3 [24]. Miśkiewicz and Majer [26] used 2.5-second-long stimuli ramped on and off with 2 ms cosine-squared ramps. The duration of the stimuli presented to the roughness models was 1 second. The shorter stimuli did not affected the predicted roughness but decreased computational time.

Listeners and procedure Miśkiewicz and Majer [26] obtained the behavioral data with a group of 10 experienced listeners. They used the method of absolute magnitude estimation, where listeners assign a number according to perceived roughness.

Roughness prediction: The methods used to predict roughness were the same as in Experiment 1.

6.2. Results

Fig. 4 shows the subjective and predicted roughness of low-frequency harmonic complex tones. Squares connected by dashed lines represent geometric means from 50 roughness estimates (10 listeners x 5 estimates) reproduced from the study of Miśkiewicz and Majer [26]. Circles, diamonds, and asterisks represent medians and quartiles of roughness predicted by Vencovský's model, Daniel and Weber's model [9] and the SI model [10], respectively. Since the complex tones were generated with random starting phases of the individual harmonics, the roughness was predicted for ten different harmonic complex tones with the same fundamental frequency. Daniel and Weber's model predicts roughness in aspers, while Vencovsk'y's model and the SI model are not. I normalized the predicted data given by these two roughness models by its corresponding maximal value and multiplied them by the maximal value across the subjective data.

Vencovský's roughness model qualitatively predicted the subjective data for the low-frequency harmonic complex tones. Beside the roughness of the low-frequency harmonic complex tones, Miśkiewicz and Majer [26] showed that low-frequency unmodulated pure tones may also contain roughness. The presented roughness model cannot account for this observation because it predicts roughness from the envelope of the simulated neural signal is without any modulations for unmodulated pure tones. This maybe the reason why the roughness model predicted roughness of the 40 Hz complex tone as almost equal to the roughness of the 63 Hz complex tone. Daniel and Weber's model underestimated roughness of the low-frequency complexes, and the SI model overestimated roughness of the 40 Hz complex tone. The SI model predicts roughness from the energy at low-frequencies which probably caused this overestimation.

I did not calculate Spearman's and Pearson's correlation between the subjective and predicted data because there were only four different stimuli and the agreement between the data is evident from the depicted results in Fig. 4.

7. Conclusion

The aim of this study was to extend the study of Vencovský [12] which introduced a new roughness model and compared its performance with results of listening tests. The study used three types of stimuli: amplitude-modulated harmonic complex tones, real pathological voice samples, and harmonic intervals of the chromatic scale. This study slightly changed the method used to estimate roughness for Vencovský's model. The model processes the input signal in short time frames and gives time-dependent value of the predicted roughness. The study [12] estimated the overall roughness as a maximal value across the time frames, this study calculates median across the time frames. Beside this, I added a new type of stimuli (low-frequency harmonic complex tones), and depicted results for two other roughness models: Daniel and Weber's model [9] and the synchronization index (SI) model [10].

Spearman's and Pearson's correlation described the fit between the predicted and subjective roughness of the used stimuli. Results obtained by Vencovský's roughness model correlated with the subjective data for all of the tested stimuli; Spearman's correlation higher than 0.764, and Pearson's correlation higher than 0.790. The model also predicted the subjective roughness of low-frequency harmonic complex tones (Experiment 4). Daniel and Weber's model performed well only for the first type of stimuli (Experiment 1, amplitude-modulated harmonic complex tones), and gave the worst predictions for the real pathological voice samples (Experiment 2). The SI model performed well for the first three stimuli (Experiment 1 to 3) but could not predict the subjective data for Experiment 4.

Vencovský's model performed best in comparison with Daniel and Weber's, and the SI model. It may indicate that the peripheral stage of Vencovský's model adequately simulates the physiology of the peripheral ear, and that the central stage selects features from the simulated neural signal which cover the perception of roughness. However, the model results for the real pathological voice samples (Experiment 2) were poor especially for the stimuli in the middle of the roughness scale. This should be studied further.

Acknowledgments

I would like to thank to Aleš Vetešník for sending the Matlab scripts of the cochlear model and to Jaroslav Bouše for giving me comments on the text of the study. The research was 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.

References

[1] von Helmholtz, H. L. F.: *On the Sensations of Tone as the Physiological Basis for the Theory of Music*, translated by A. J. Ellis (1885), 4th ed., Dover, New York, 1954.

[2] Vassilakis, P. N.: Auditory roughness as a means of musical expression, *Selected Reports in Ethnomusicology*, 12, Perspectives in Systematic Musicology, p. 119–144, 2005.

[3] Mathes, R. C., Miller, R. L.: Phase Effects in Monaural Perception, *J. Acoust. Soc. Am.* 19(**5**), p. 780–797, 1947.

[4] Terhardt, E.: On the perception of periodic sound fluctuations (Roughness), *Acustica* 30(**4**), p. 201–213, 1974.

[5] Pressnitzer, D., McAdams, S.: Two phase effects in roughness perception, *J. Acoust. Soc. Am.* 105(**5**), p. 2773–2782, 1999.

[6] Fastl, H., Zwicker, E.: *Psychoacoustics: Facts and Models*, Springer, Berlin, Heidelberg, New York, 2007.

[7] Patel, S. A., Shrivastav, R., Eddins D. A.: Identifying a Comparison for Matching Rough Voice Quality, *J. Speech Lang. Hear. Res.*, 55, p. 1407–1422, 2012.

[8] Aures, W.: Ein Berechnungsverfahren der Rauhigkeit, *Acustica*, 58(**5**), p. 268–281, 1985.

[9] Daniel, P., Weber, R.: Psychoacoustical roughness: Implementation of an optimized model, *Acustica* 83(**1**), p. 113–123, 1997.

[10] Leman, M.: Visualization and calculation of the roughness of acoustical musical signals using the synchronization index model (SIM), *Proceedings of the COST G-6 Conference on Digital Audio Effects (D AFX-00)*, Verona, Italy, DAFX 1-DAFX 6, 2000.

[11] Wang, Y. S., Shen, G. Q., Guo, H., Tang, X. L., Hamade, T.: Roughness modelling based on human auditory perception for sound quality evaluation of vehicle interior noise, *J. Sound and Vibration*, 332, p. 3893–3904, 2013.

[12] Vencovský, V..: Modeling roughness perception for complex stimuli using a model of cochlear hydrodynamics, *ISMA 2014*, Le Mans, France, 2014.

[13] Klatt, D.H.: Software for a cascade/parallel formant synthesizer, *J. Acoust. Soc. Am.* 67(**3**), p. 971–995, 1980.

[14] Mammano, F., Nobili, R.: Biophysics of the cochlea: Linear approximation, *J. Acoust. Soc. Am.* 93(**6**), p. 3320–3332, 1993.

[15] Nobili, R., Mammano, F.: Biophysics of the cochlea II: Stationary nonlinear phenomenology, *J. Acoust. Soc. Am.*, 99(**4**), p. 2244–2255, 1996.

[16] Nobili, R., Vetešník, V., Turicchia, L. Mammano, F.: Otoacoustic emissions from residual oscillations of the cochlear basilar membrane in a human ear model, *J. Assoc. Res. Otolaryngol*, 4, p. 478–494, 2003.

[17] Shamma, S. A, Chadwick, R. S., Wilbur, W. J., Morrish, K. A., Rinzel, J.: A biophysical model of cochlear processing: Intensity dependence of pure tone responses, *J. Acoust. Soc. Am.*, 80(**1**), p. 133-145, 1986.

[18] Sumner, C. J., Lopez-Poveda, E. A, O'Mard, P., Meddis, R.: A revised model of the inner-hair cell and auditory-nerve complex, *J. Acoust. Soc. Am.*, 111(**5**), p. 2178–2188, 2002.

[19] Meddis, R.: Auditory-nerve first-spike latency and auditory absolute threshold: A computer model, *J. Acoust. Soc. Am.* 119(**1**), p. 406–417, 2006.

[20] Meddis, R.: Simulation of Mechanical to Neural Transduction in the Auditory Receptor, *J. Acoust. Soc. Am.* 79(**3**), p. 702–711, 1986.

[21] Vencovský, V.: Modeling roughness perception using a model of cochlear hydrodynamics, *Forum Acusticum 2014*, Krakow, Poland, p. 1–6, 2014.

[22] Aures, W.: *Berechnungsverfahren für den Wohlklang beliebiger Schallsignale, ein Beitrag zur gehörbezogenen Schallanalyse*, PhD thesis, TU München, 1984.

[23] Wang, Y. S.: A study on sound roughness evaluation based on an auditory synchronization index model, *Proceedings of the 2nd International Congress on Image and Signal Processing, IEEE*, p. 3612–3616, 2009.

[24] PsySound3: Sound analyses software `http://psysound.wikidot.com/`

[25] IPEM toolbox for perception-based music analysis `http://www.ipem.ugent.be/?q=node/27`

[26] Miśkiewicz, A., Majer, J.: Roughness of low-frequency pure tones and harmonic complex tones, *Forum Acusticum 2014*, Krakow, Poland, 2014.

[27] Moore, B. C. J., Glasberg, B. R., Baer, T.: A Model for the Prediction of Thresholds, Loudness, and Partial Loudness, *Journal of the Audio Engineering Society* 45(**4**), p. 224–240, 1997.