Signal repetition rates and their relationship to the pleasantness of multi-tone sounds

Stephan TÖPKEN[1]; Henning SCHEEL[2], Reinhard WEBER[1]

[1] Oldenburg University - Acoustics group, Germany

[2]Airbus, Germany

ABSTRACT

The sound character of multi-tone sounds offers a rich bunch of perceptual aspects that can be linked to spectral and temporal properties of the signals. The frequency spacing of the partials is one major underlying determinant. Already subtle changes of the partial spacing can have a big effect on the temporal structure in terms of the signals' repetition rate. For sounds consisting of a superposition of two harmonic complex tones and combination tones this leads to a rich mix of different sensations like beats, fluctuations, roughness, dissonance and tonality which are at least in part know for driving annoyance, unpleasantness or pleasantness. The pleasantness of such sounds has been investigated in listening tests. The ratio between the fundamental frequencies of the two complex tones has been varied systematically and an analysis of the time series of the stimuli shows clear differences in the periodicities and hence repetition rates. In comparison with the subjective judgments high correlation coefficients can be found between the pleasantness and the repetition rate of the signal. A combination of the repetition rate and the psychoacoustic sharpness are able to model the preference relevant sound characteristics of multi-tone sounds.

(192 words)

Keywords: Loudness, Sound Quality, Preference testing
I-INCE Classification of Subjects Number(s): 63.1 (Loudness), 63.7 (sound quality), 79.2 (Preference testing)

1. INTRODUCTION

The unpleasantness of a sound is apart from its loudness also depending on the sound character. In several studies temporal as well as spectral parameters have been systematically varied to measure their influence on the subjective loudness and preference judgments as level dependent points of subjective equality (PSE) against a fixed reference (1). The difference between the preference and the loudness judgments is attributed to the differences in sound character. It turns out that the determination of the sound character as a difference between the individual levels at the PSEs for preference and for loudness yields rather clear relationships between varied stimulus parameters and the sound character measure. A statistically significant correlation coefficient of r = -0.91 has been found between the sound character measure and the sharpness after the DIN standard for variations of the spectral envelope (3, 4). In this paper an efficient way to describe the temporal signature of multi-tone sounds is presented. In combination with the sharpness metric a model for the prediction of the sound character measure is set up.

[1] stephan.toepken@uni-oldenburg.de

2. METHOD FOR THE DETERMINATION OF PREFERENCE RELEVANT SOUND CHARACTERISTICS

To determine the contribution of the sound character to an overall preference judgment towards a sound the PSEs for loudness and preference are measured in separate experiments. In these experiments the level of a test sound is varied until equal loudness / equal preference against a reference sound with a fixed level is achieved (1). The results from these paired comparison experiments are thus subjective judgments expressed as levels at the PSE for loudness or preference. The difference between these two points of subjective equality, which can be individually calculated for each participant and sound, summarizes all aspects relevant for the preference towards a sound which are not already covered by the loudness. In this sense the difference between the level at the PSE for preference L_{pref} and the PSE for loudness L_{loud} is attributed to the sound character (see fig. 1):

$$\Delta L_{soundcharacter} = L_{pref} - L_{loud} \tag{1}$$

The measure $\Delta L_{sound\ character}$ becomes zero, if there are no further differentiating aspects between two sounds other than the loudness. If additional aspects on top of the loudness differences play a role for the preference of a sound then these additional aspects are measured by $\Delta L_{sound\ character}$. In this way the sound character summarizes all attributes other than loudness affecting the preference of a sound (2). Rather high correlation coefficients between $\Delta L_{sound\ character}$ and L_{pref} and rather low correlation coefficients between $\Delta L_{sound\ character}$ and L_{loud} are an evidence for this concept of sound character measurements (4).

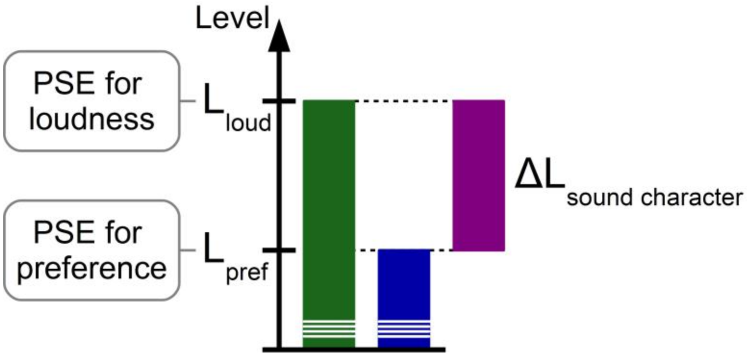

Figure 1 – Results from the two separate listening experiments. With an adaptive procedure the test sound level was varied until it was equally loud/equally preferred as the reference sound with a fixed level. The subjective judgments are expressed as the level of the test sound at the PSE for loudness (L_{loud}) and the PSE for preference (L_{pref}). The level difference L_{pref} - L_{loud} is attributed to the sound character of the test sound.

2.1 Listening setup and participants

The listening experiments take place in the anechoic chamber of the University of Oldenburg. The stimuli are presented with a loudspeaker driven by an external soundcard (M-Audio, Fast Track pro). The loudspeaker is placed in front of the participant. The experiment itself is implemented as a MATLAB routine, which is conducted by the participants with a computer keyboard and a screen underneath the loudspeaker.

The listening tests have been carried out throughout multiple sessions with 34 to 47 participants, aged between 18 and 31 years. Approximately half of the participants are female and the others are male.

2.2 Stimuli

Synthetic multi-tone sounds consisting of up to 460 partials were used as test sounds. The multi tone sounds are built as a superposition of two complex tones (CX1 and CX2) and additional combination tones (CTs). The frequencies components are given by:

$$\text{CX1:} \quad \begin{array}{l} f_{i,0} = i \cdot f_{10} \\ i = 1, 2, 3, \ldots, 30 \end{array} \tag{2}$$

$$\text{CX2:} \quad \begin{array}{l} f_{0,j} = j \cdot f_{01} \\ j = 1, 2, 3, \ldots, 30 \end{array} \tag{3}$$

$$\text{CTs:} \quad \begin{array}{l} f_{i,j} = i \cdot f_{10} + j \cdot f_{01} \\ i, j = 1, 2, 3, \ldots, 20 \end{array} \tag{4}$$

A multi-tone signal consisting of all three components CX1+CX2+CTs will be denoted type CCC in the following. A signal build only from the two complex tones CX1+CX2 is identified as type CC. An overview of all stimuli is shown in table 1.

Table 1 – Overview of the stimuli parameters

signal name	signal type	number of partials	fundamental frequencies		ratio between the fundamental frequencies	spectral parameters
			f_{10} / Hz	f_{01} / Hz	$\rho = f_{01} : f_{10}$	
A_1	CX1	30	100	-	-	fixed
A_2	CX2	30	-	132.66	-	fixed
B_1	CC	60	100	132.66	199:150	fixed
C_1	CTs	400	100	132.66	199:150	fixed
D_1 - D_{22}	CCC	460	100	132.66	199:150	varied
E_1	CCC	460	100	128.57	9:7	fixed
E_2	CCC	460	100	129.30	1293:1000	fixed
E_3	CCC	460	100	130.00	13:10	fixed
E_4	CCC	460	100	132.66	199:150	fixed
E_5	CCC	460	100	133.33	4:3	fixed

The stimuli A_1 and A_2 and consist each of one of the complex tones only. These two stimuli differ in fundamental frequency but they consist both of 30 partials (fundamental frequency and 29 higher harmonics). Stimulus B_1 is a type CC signal based on the two complex tones (A_1 and A_2) and thus contains the 60 partials. The combination tones alone (400 partials) are represented in stimulus C_1. The stimuli D_1 to D_{22} are type CCC signals containing the two complex tones as well as the combination tones. For these stimuli different parameters of the spectral envelope are varied while the ratio between the fundamentals is kept constant. The stimuli B_1, C_1 and D_1-D_{22} are all based on a ratio ρ=199:150. For the stimuli E_1 to E_5 the parameters of the spectral envelope are fixed and the ratio between the fundamentals ρ and thus the temporal signature of the signals is varied:

$$\rho = f_{01} : f_{10} \tag{5}$$

The sounds B_1, C_1, D_1-D_{22} and E_4 have already served as a basis for prior studies on the relationship between subjective judgments and psychoacoustic metrics (3, 4). In all cases the reference sound is a noise signal with a spectral slope of approximately -6 dB per octave up to 1 kHz and -12 dB per octave above 1 kHz. It has a constant level of 74 dB(A).

3. ANALYSIS OF THE TEMPORAL SIGNATURE

One single value describing the temporal structure of the sounds is the repetition rate of the time signal. For the signals A_1 and A_2 each consisting of a harmonic complex tone only the repetition rate is directly determined by fundamental frequency (A_1: f_{10}=100 Hz and A_2: f_{01}=132.66 Hz). All other signals are based on the combination of the two fundamental frequencies. Here the repetition rate can be calculated by the integer fraction of each fundamental:

$$f_{rep} = \frac{\text{fundamental frequency}}{\text{corresponding integer number}} \tag{6}$$

For an exemplary ratio between the fundamental frequencies of ρ = 4:3 the fundamental frequencies f_{10}=100 Hz and f_{01}=132.66 Hz yield a repetition rate

$$f_{rep} = \frac{133.33\ \text{Hz}}{4} = \frac{100\ \text{Hz}}{3} \approx 33.33\ \text{Hz} \tag{7}$$

and a periodic time of

$$t_{rep} = \frac{1}{f_{rep}} = \frac{1}{33.33\text{Hz}} \approx 0.03\ \text{s} \tag{8}$$

Table 2 shows the repetition rate and the periodic time of the signals varied in the ratio between their fundamentals ρ and the two sounds consisting of the complex tones only (A_1 and A_2). The variation of the ratio ρ leads to repetition rates between 0.1 Hz and 33.33 Hz. The lowest repetition rate corresponds to a value of ρ created by a ratio of large integer numbers (ρ = 1293:1000, E_2). Higher repetition rates occur for a values of ρ created by ratios of small integers (e.g. ρ = 4:3, E_5). The highest repetition rates are reached by the two complex tones alone (A_1 and A_2).

Table 2 – Temporal properties of the five stimuli with a variation of the ratio between their fundamentals characterized by the periodic time $\Delta\tau$ and the repetition rate f_{rep}

repetition rate	periodic time	ratio between the fundamentals	fundamental frequencies		signal name
f_{rep} / Hz	t_{rep} / s	$\rho = f_{01} : f_{10}$	f_{10} / Hz	f_{01} / Hz	
0.1	10	1293:1000 ≈13:10	100	129.30	E_2
0.67	1.5	199:150 ≈ 4:3	100	132.66	E_4
10	0.1	13:10	100	130.00	E_3
14.29	0.07	9:7	100	128.57	E_1
33.33	0.03	4:3	100	133.33	E_5
100	0.01	-	100	-	A_1
132.66	0.0075	-	-	132.66	A_2

The calculated repetition rates can also be identified in the autocorrelation function (ACF) of the signals' time series. Figure 2 shows the ACF as a function of time lag τ for the five sounds E_1 to E_5. The lag times for the first unity peak $\Delta\tau$ ranging from 0.03seconds to 10 seconds equal the periodic time t_{rep} calculated as inverse of the repetition rate in table 2.

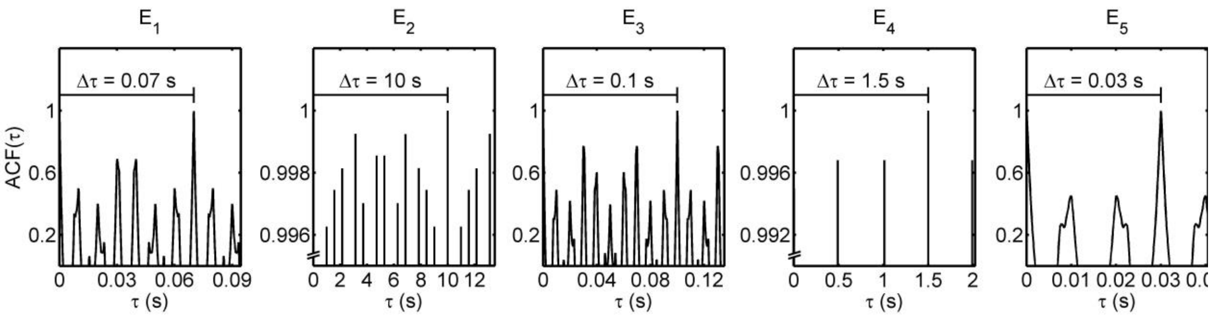

Figure 2 – Autocorrelation function ACF for the five sounds E_1 – E_5 plotted over the lag τ in seconds. The periodic times of the signals, clearly identifiable by the first unity peak of the ACF, range over four orders of magnitudes from 0.03 seconds to 10 seconds.

4. RELATIONSHIP BETWEEN THE PERIODICITY AND THE SUBJECTIVE JUDGMENTS

In figure 3 the mean values of the subjective judgments are plotted over the repetition rate of the signals' time series for the sounds varying in temporal structure (A_1, A_2 and E_1 to E_5). With rising repetition rate the levels of the sounds need to be decreased to make the sounds equally loud and equally preferred. The difference between preference and loudness attributed to the sound character also becomes larger with increasing repetition rate. This difference is biggest for the sounds A1 and A2 each consisting of a harmonic complex tone only. For these sounds the magnitude of the repetition rate is in the region of modulation frequencies where the impression of roughness decreases and an overall tonal character becomes more prominent.

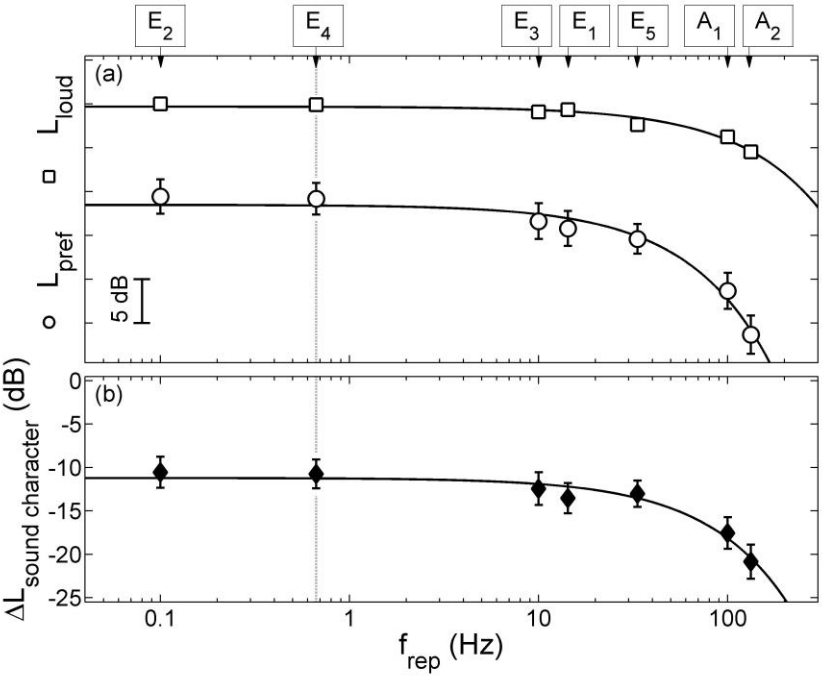

Figure 3 – (a) PSEs for loudness (open square symbols) and preference (open circles) plotted semi-logarithmical over the repetition rate f_{rep} of the stimulus' time signal. Errorbars indicate the standard error of the calculated mean values. Lines indicate linear regressions. (b) same as (a) for $\Delta L_{\text{sound character}}$

5. RELATIONSHIP BETWEEN THE SHARPNESS METRICS AND THE REPETITION RATE FOR THE SOUND CHARACTER MEASURE

Figure 4 shows the mean values of $\Delta L_{sound\ character}$ over the sharpness values (DIN 45692) for all 31 sounds. For the sounds B_1, C_1 and D_1-D_{22} with a fixed ratio ρ=199:150 (indicated by open circles) a rather linear relationship between $\Delta L_{sound\ character}$ and the values of the DIN sharpness is found (3, 4). The sharpness values are quite robust against variations of the parameter ρ which mainly affects the temporal structure of the sounds. All five sounds with a variation of the ratio ρ (E_1-E_5, open square symbols) have nearly the same sharpness of about 1.1 acum. The two sounds consisting of the complex tones only (A_1 and A_2) also have sharpness values of 0.9 acum and 1.1 acum while their differences in sound character compared to the complete multi-tone sounds with similar sharpness values are in the magnitude of 10 dB.

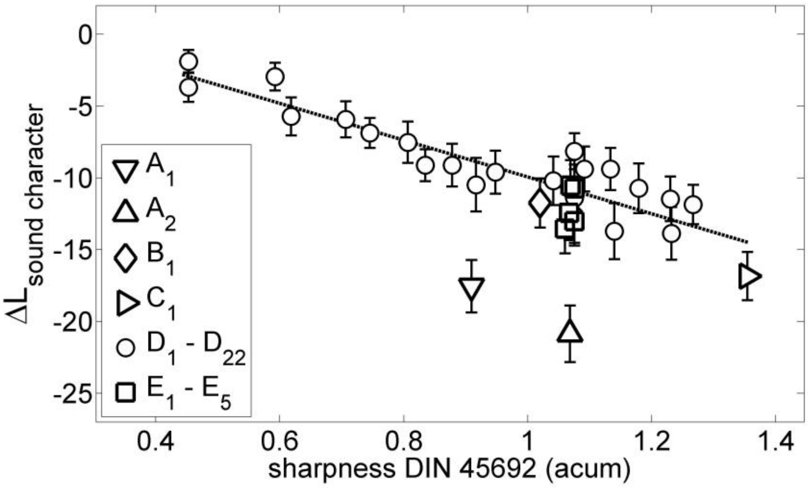

Figure 4 – Relationship between $\Delta L_{sound\ character}$ and the sharpness after the DIN 45692 standard. Errorbars indicate the standard error of the calculated mean values. A statistically significant correlation coefficient of r = -0.91 is found between the sharpness values and $\Delta L_{sound\ character}$ for the sounds B_1, C_1 and D_1-D_{22}. The dotted line represents a linear regression between these two variables. Changes in the temporal structure by a variation of the ratio ρ (E_1-E_5 indicated by open square symbols) do not affect the sharpness values.

6. THE COMBINATION OF SHARPNESS AND REPETITION RATE TO A MODEL FOR ΔL sound character

Figure 5 shows the values of $\Delta L_{sound\ character}$ from the listening experiments plotted over the repetition rate and the sharpness. With increasing values of the sharpness and also with an increase in repetition rate $\Delta L_{sound\ character}$ becomes more negative, which means that the absolute difference between the PSEs for preference and for loudness increases. The two complex tones alone (A_1 and A_2, open triangles) need about 20 dB additional attenuation from being equally loud to become equally preferred.

To describe the combined effects of the temporal and spectral properties on the sound character measure $\Delta L_{sound\ character}$ a model based on the linear superposition of the repetition rate and the sharpness reaches an adjusted r^2=0.89:

$$\Delta L_{sound character} = -12.9\,\tfrac{dB}{acum}\cdot sharpness - 0.08\,\tfrac{dB}{Hz}\cdot f_{rep} + 2.9\ dB \quad (9)$$

The values of this linear superposition model are shown by the grid surface in figure 5a. The residuals of the model (fig. 5b) are below ±2.5 dB.

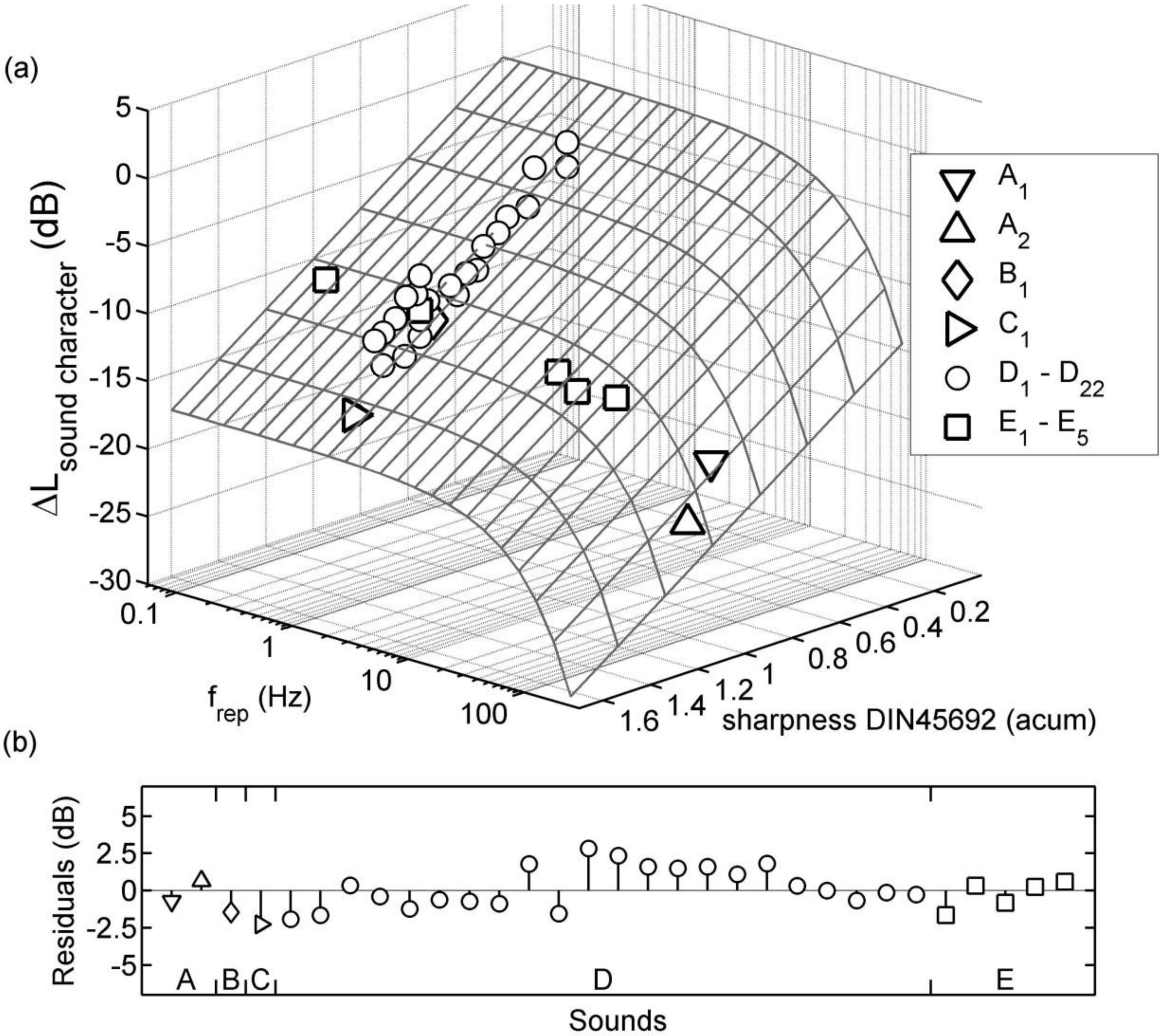

Figure 5 – (a) Values of the sound character measure for a variation of the spectral envelope (open circles) which is best reflected by the sharpness (DIN 45692) and the variation of the temporal structure (open squares and triangles) reflected by the repetition rate. The grid shows a linear superposition model for the sound character measure based on the repetition rate and the sharpness with an adjusted r^2=0.89.

(b) Residuals of the linear superposition model are below ±2.5 dB

Surprisingly this very simple model succeeds in seamlessly including the two sounds A_1 and A_2, each consisting of only one complex tone. These two sounds are rather different to the other sounds due to the much higher repetition rate (about 100 Hz higher) and considerably lower number of tonal components (only fundamental frequency and 29 higher harmonics). Sound C_1 which is missing the two complex tones and thus is lacking some partials in the low frequency range also smoothly fits into this model.

The description of the temporal signature by the repetition rate and the coverage of the spectral properties by the sharpness after the DIN standard seem to build a good basis for a model of the preference relevant sound characteristics in the case of multi-tone sounds.

7. Summary

For multi-tone sounds consisting of two complex tones and combination tones the spacing of the frequency components is one underlying determinant linked to a wide variety of sensations like beats, roughness, dissonance and tonality which are at least in part know for driving the unpleasantness and annoyance or pleasantness. The loudness and the preference have been evaluated in listening experiments by 38 to 47 participants. For a variation of spectral and temporal parameters the levels of the test sound at the point of subjective equality for loudness and for preference have been determined in paired comparisons with a fixed reference sound. The level difference between the preference and the loudness judgments is attributed to the sound character of the test sounds and thus denoted $\Delta L_{sound\ character}$.

An analysis of the signals' time series exhibits clear differences in their repetition rate. A variation of the upper fundamental frequency of only a few Hertz yields differences in the repetition rate ranging over 4 magnitudes.

In comparison with the subjective judgments high correlation coefficients have been found between the repetition rate and the sound character measure for sounds with a variation of temporal characteristics. Previous studies already showed high correlation coefficients between the psychoacoustical sharpness after the DIN standard and the sound character measure for sounds varied in terms of the spectral envelope. A linear superposition model for $\Delta L_{sound\ character}$ based on both parameters - the repetition rate and the sharpness after the DIN 45692 standard - yields an adjusted r^2=0.89 and residuals smaller than 2.5 dB, suggesting that these two parameters are suitable acoustic indicators for the sound character of multi-tone sounds.

REFERENCES

1. Töpken S, Verhey JL, Weber R., Evaluation of multi-tone-sounds, Proc. Internoise 2011, Osaka, Japan, 2011. ISSN: 0105-175x, INCE/J.
2. Pollack, I., Loudness as a Discriminable Aspect of Noise, The American Journal of Psychology, 1949; 62: 285-289.
3. Töpken S, Weber R., Preference and loudness judgments of multi-tone sounds and their relationship to psychoacoustical metrics, Internoise 2013, Innsbruck, Austria, 2013, Austrian Noise Abatement Association.
4. Töpken S, Weber R., Quantitative determination of the sound character and its relationship to psychoacoustical metrics, Fortschritte der Akustik - DAGA 2014, Oldenburg, Germany, 2014, Dega e.V., Berlin.