Influence of the spectral envelope on the loudness and the preference of multi-tone sounds

Stephan TÖPKEN[1]; Henning SCHEEL[2]; Reinhard WEBER[1]; Steven VAN DE PAR[1]

[1] Oldenburg University – Acoustics Group, Germany

[2] Airbus Operation GmbH, Interior Noise Department, Hamburg, Germany

ABSTRACT

The sound character of multi-tone sounds yields a wide variety of perceptual aspects which are quite different from those obtained from single tonal components. Particularly for a specific class of multi-tone sounds, consisting of two complex tones and additional combination tones, the perceptual effects of the temporal and the spectral sound characteristics are not yet clarified. In this study, the effect of two signal parameters, affecting the spectral envelope of multi-tone sounds, on the loudness and the preference judgments is addressed. Based on a triangular spectral envelope, the peak frequency and the falling slope served as independent variables. The loudness and the preference were determined as levels at the points of subjective equality (PSEs) against one common reference sound, but in separate listening tests. The level difference between the PSEs for preference and the PSE for loudness is used as a quantitative measure, reflecting the additional influence of the sound character on the preference judgment. Clear relationships between the signal parameters and the judgments of loudness and preference were identified, especially for the level difference between the two PSEs. It turned out that the spectral signal variations can be effectively parameterized by the psychoacoustic sharpness and high correlation coefficients between the sharpness values and the judgments are found.

Keywords: Aircraft interior noise, Loudness, Sound Quality, Preference testing
I-INCE Classification of Subjects Number(s): 13.1.6 Aircraft interior noise, 63.1 (Loudness), 63.7 (sound quality), 79.2 (Preference testing)

1. INTRODUCTION

Tonal sounds originating from machinery and engines are a part of transportation noise which is perceived by humans every day. Prominent single frequency tones are often found to be one underlying reason for a higher unpleasantness or annoyance of a sound compared to rather broadband signals with the same loudness, but without prominent tonal components. In the context of aircraft, the cabin noise consists mainly of a broadband background, but also prominent tonal components from the engines and other machinery occur. With regard to new aircraft engine architectures, the number of tonal components and also their prominence in the cabin interior noise is expected to be different from current signatures. For such multi-tone sounds consisting of a large number of partials, the sound character of the tonal part alone is considerably different from a single tone. Especially for sounds resulting from non-linear sound generating mechanisms, consisting of two complex tones and additional interaction tones, many signal parameters may influence the perception of the sound. In addition to the sound character, the variation of the spectral content might also change the loudness of the sounds and both aspects – loudness and sound character – are anticipated to be relevant for the overall appraisal of a resulting sound. However, the particular quantitative contribution of the spectral content to the loudness judgment and the additional influence of the sound character to the appraisal of multi-tone sounds are unclear.

The aim of this study was a determination of loudness judgments and preference evaluations for multi-tone sounds which were varied in terms of their spectral envelope. A generic triangular spectral envelope was used to shape the tonal components. The peak frequency and the slope of the level decrease above the peak frequency were varied as signal parameters.

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

2. METHOD FOR THE DETERMINATION OF PREFERENCE RELEVANT SOUND CHARACTERISTICS

2.1 Measurement approach

The loudness and the preference were measured as points of subjective equality (PSEs) compared to the same fixed reference sound. In separate experiments, the level of a multi-tone test sound was varied adaptively until it was equally loud or equally preferred as the reference sound. The levels at the PSE for loudness L_{loud} and at the PSE for preference L_{pref} were used as quantitative measures of the two judgments, respectively. The level difference between L_{pref} and L_{loud} is interpreted as the additional contribution of the sound character to the preference evaluation, on top of the loudness effect:

$$\Delta L_{soundcharacter} = L_{pref} - L_{loud} \quad (1)$$

Due to the higher unpleasantness of the multi-tone sounds in comparison to the equally loud reference sound, this difference becomes negative, as shown in Fig. 1. This means that equal loudness is not sufficient to make the multi-tone sounds equally preferred to the reference sound. The multi-tone sounds require lower levels to be equally preferred than the levels at the PSE for loudness.

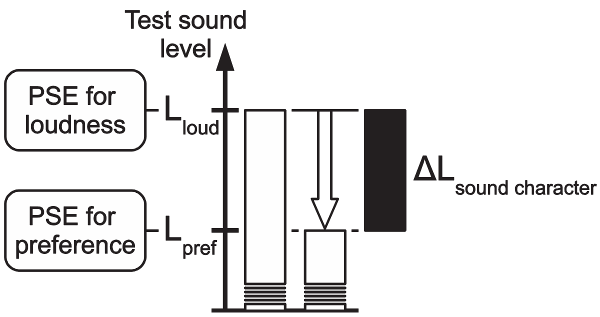

Figure 1 – Schematic results from the two separate listening experiments. With an adaptive procedure the level of the multi-tone test sound was varied until it was equally loud/equally preferred as a reference sound with a fixed level. The subjective judgments are expressed as the level of the test sound at the PSE for loudness (L_{loud}) and the PSE for preference (L_{pref}), each. The level difference $\Delta L_{sound\ character} = L_{pref} - L_{loud}$ is subsuming the additional contribution of the sound character to the preference evaluation on top of the loudness judgment, which has negative values for the investigated sounds, due to their rather unpleasant sound character.

2.2 Listening setup and participants

The listening experiments took place in the anechoic chamber of the University of Oldenburg. The stimuli were presented over one loudspeaker (Mackie HR 824), which was placed directly in front of a single, sitting participant at a distance of 2.25 m. The multi-tone signals and the reference noise were stored on a computer which fed the loudspeaker through an external audio interface (M-Audio, Fast Track Pro). The procedures of the experiments were implemented as a MATLAB-Code. A graphical user interface was presented to the participants on a flatscreen underneath the loudspeaker. The answers of the participants in the adaptive procedures were captured with the aid of a computer keyboard.

The experiments for the two parameter variations were carried out in separate measurement campaigns with N=48 (variation of the slope) and N=34 (variation of the peak frequency) participants, which were aged between 18 and 31 years. Approximately half of the participants were female and the others were male in each campaign.

2.3 Stimuli

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

CX1: $$f_{i,0} = i \cdot f_{10} \qquad i = 1,2,3,...,30 \tag{2}$$

CX2: $$f_{0,j} = j \cdot f_{01} \qquad j = 1,2,3,...,30 \tag{3}$$

CTs: $$f_{i,j} = i \cdot f_{10} + j \cdot f_{01} \qquad i, j = 1,2,3,...,20 \tag{4}$$

The relative levels of the tonal components were derived from a triangular spectral envelope which is defined by a fixed rising slope of 6 dB/oct, a peak frequency f_{peak} and a spectral decline of the components above the peak frequency. In addition to the spectral envelope, the levels of the combination tones were attenuated by 10 dB compared to the partials of the two complex tones. The starting phase information for the partials is taken from one common set of equally distributed random values between 0 and 2π for all stimuli. The peak frequency f_{peak} and the spectral decline in terms of the slope were varied as signal parameters. Two values of the frequency ratio ρ between the fundamentals, which were found to differ considerably in terms of the pleasantness in a previous study (3), were taken as a basis for the current signals. An overview of the parameter configurations for the 28 stimulus signals is given in Tab. 1. The reference sound was a cabin 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 was presented with a constant level of 74 dB(A). The duration of the signals was 5 s in the preference task and 1.5 s in the loudness task.

Table 1: Signal parameters of the multi-tone stimuli

No.	spectral decline *slope* / db/oct	peak frequency f_{peak} / Hz	frequency ratio between the fundamentals $\rho = f_{01} : f_{10}$
1	-6	100	199:150
2	-6	100	4:3
3	-9	100	199:150
4	-9	100	4:3
5	-12	100	199:150
6	-12	100	4:3
7	-15	100	199:150
8	-15	100	4:3
9	-6	100	199:150
10	-6	100	4:3
11	-6	200	199:150
12	-6	200	4:3
13	-6	300	199:150
14	-6	300	4:3
15	-6	400	199:150
16	-6	400	4:3
17	-6	500	199:150
18	-6	500	4:3
19	-15	100	199:150
20	-15	100	4:3
21	-15	200	199:150
22	-15	200	4:3
23	-15	300	199:150
24	-15	300	4:3
25	-15	400	199:150
26	-15	400	4:3
27	-15	500	199:150
28	-15	500	4:3

3. RESULTS

3.1 INFLUENCE OF THE SPECTRAL DECLINE (FALLING SLOPE) OF THE TONAL COMPONENTS ON THE JUDGMENTS

Figure 2 shows the mean results of the judgments as relative levels at the point of subjective equality for loudness (open squares) and at the point of subjective equality for preference (open circles). The two values of the frequency ratio ρ between the fundamentals are indicated by dotted/dashed lines, respectively. For an increasing steepness of the spectral slope and a reduction of high frequency content, the PSEs for loudness and for preference increase, respectively. In general, sounds with a steep slope of the spectral envelope and less energy in the high frequency range can be higher in A-weighted level while being equally loud and also equally preferred than multi-tone sounds with more energy in the high frequency range and a shallow slope. In detail, the effect is more pronounced for the preference PSEs than for the loudness PSEs. The loudness PSEs change only by about 5 dB for a variation of the slope from -6 dB/oct to a slope of -15 dB/oct, while the PSEs for preference change by about 15 dB between these extreme slope values.

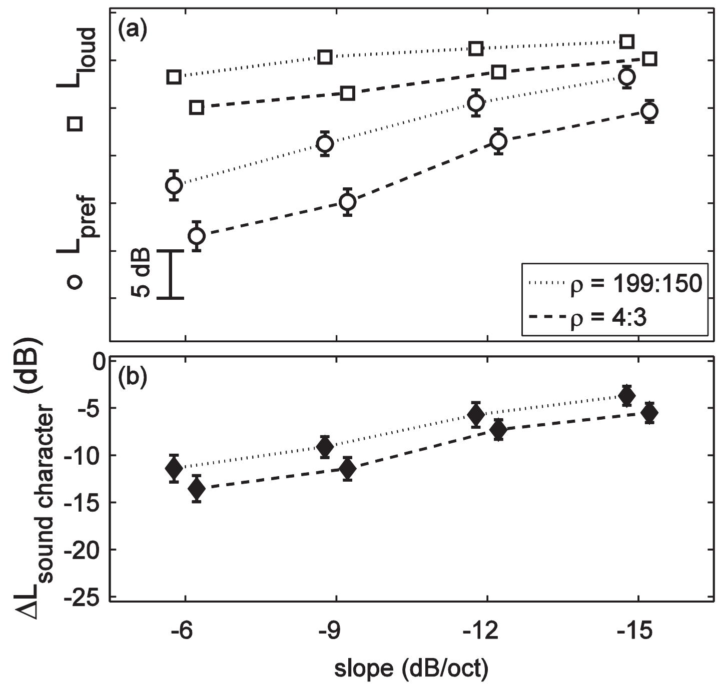

Figure 2: Mean judgments of the loudness (a, open squares) and the preference (a, open circles), each plotted as relative values, and the level difference between the two PSEs, denoted $\Delta L_{sound\ character}$ (b, filled diamonds), plotted over four different slopes towards higher frequencies. The frequency ratio ρ between the fundamentals is indicated by dotted (ρ=199:150) and dashed (ρ=4:3) lines. Error bars show the standard error of the calculated mean (N=48).

The level difference between the PSEs for equal preference and equal loudness, denoted as $\Delta L_{sound\ character}$, is shown in Fig. 2(b). For a slope of -15 dB/oct and a frequency ratio ρ=4:3 the level difference between equal preference and equal loudness is about -6 dB. This means that an additional level reduction of 6 dB is necessary to shift from equal loudness to equal preference. This level difference becomes larger in absolute values, if the spectral envelope becomes shallower and more energy is attributed to the high frequency range. For a slope of -6 dB/oct and a frequency ratio ρ=4:3, a level reduction of -14 dB is necessary to make the multi-tone sound equally preferred compared to the mean PSE for loudness. Similar relationships are also found for a frequency ratio ρ=199:150, but the values are shifted by about 4 dB above the values for frequency ratio of ρ=4:3 for all slope conditions. Thus, for a value of ρ=199:150, the level difference between equal loudness and equal preference seems to be overall smaller than for ρ=4:3.

In general, less energy in the high frequency range is beneficial for the multi-tone signatures, because it allows higher A-weighted levels while being equally loud/preferred than configuration with more high frequency content. A reduction of high frequency content also reduces the difference between the PSEs for equal preference and equal loudness.

3.2 INFLUENCE OF A SPECTRAL PEAK (PEAK FREQUENCY) OF THE TONAL COMPONENTS ON THE JUDGMENTS

The PSEs for loudness and for preference, plotted over the peak frequency of the spectral envelope, are shown in Fig. 3 for a shallow slope of -6 dB/oct. Figure 4 shows the equivalent figure for a steep slope of -15 dB/oct. For a shallow slope of -6 dB/oct (shown in Fig. 3) the PSEs for loudness (open squares) are rather constant over all peak frequencies. Only a slight maximum for a peak frequency of 400 Hz for both values of the frequency ratio ρ is observed. The offset between the two curves for the two values of the frequency ratio ρ is about 3 dB over all peak frequencies. The PSEs for preference (open circles) decrease by about 4 dB for an increase of the peak frequency from 100 Hz to 500 Hz for both values of the frequency ratio ρ in a similar way. Here, the offset between the two curves is about 8 dB. The level difference $\Delta L_{sound\ character}$, shown in Fig. 3(b), becomes more negative for increasing values of the peak frequency. This means that the difference in sound character compared to the reference sound increases with rising peak frequency.

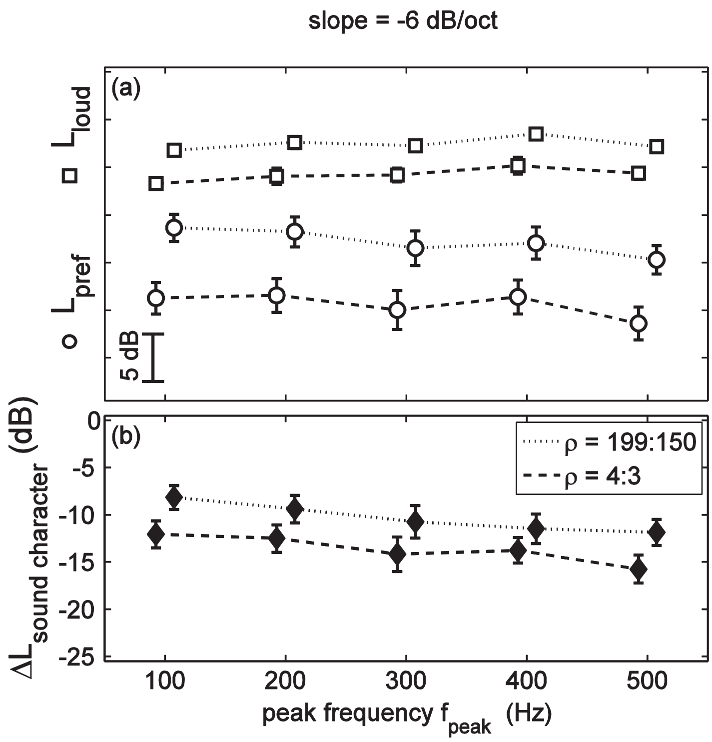

Figure 3: Mean judgments of the loudness (a, open squares) and the preference (a, open circles), each plotted as relative values and the level difference between the two PSEs, denoted $\Delta L_{sound\ character}$ (b, filled diamonds), plotted over the peak frequency. The frequency ratio between the fundamentals is indicated by dotted (ρ=199:150) and dashed (ρ=4:3) lines. Error bars show the standard error of the calculated mean (N=34).

For the steep slope, with a value of -15 dB/oct (Fig. 4), the relationships between the loudness and the preference judgments and the peak frequency are more complicated than for the shallow slope shown Fig. 3. The PSEs for loudness have a local maximum for a peak frequency of 300 Hz (ρ=4:3) and 400 Hz (ρ=199:150), respectively. Similarly, local maxima are also observed for the preference PSEs at a peak frequency of 300 Hz (ρ=4:3) and 200 Hz (ρ=199:150). These maxima indicate optima in terms of the loudness and the preference, each. Signal configurations at these maxima can have a higher A-weighted level while being equally loud or equally preferred than those signals, having higher or lower peak frequencies.

However, the relationships between the peak frequency and the judgments become surprisingly clear, when looking at the level difference $\Delta L_{sound\ character}$ in Fig. 4(b). With increasing peak frequency the level difference $\Delta L_{sound\ character}$, between the PSEs for preference and for loudness, becomes almost linearly more negative for both values of the frequency ratio ρ.

For the two slope conditions shown in Fig. 3 and Fig. 4, a rather constant offset is found between the curves for the two values of ρ. The offset is about 5 dB for the shallow slope of -6 dB/oct (in Fig. 3(b)) and about 3 dB for the steep slope of -15 dB/oct (in Fig. 4(b)).

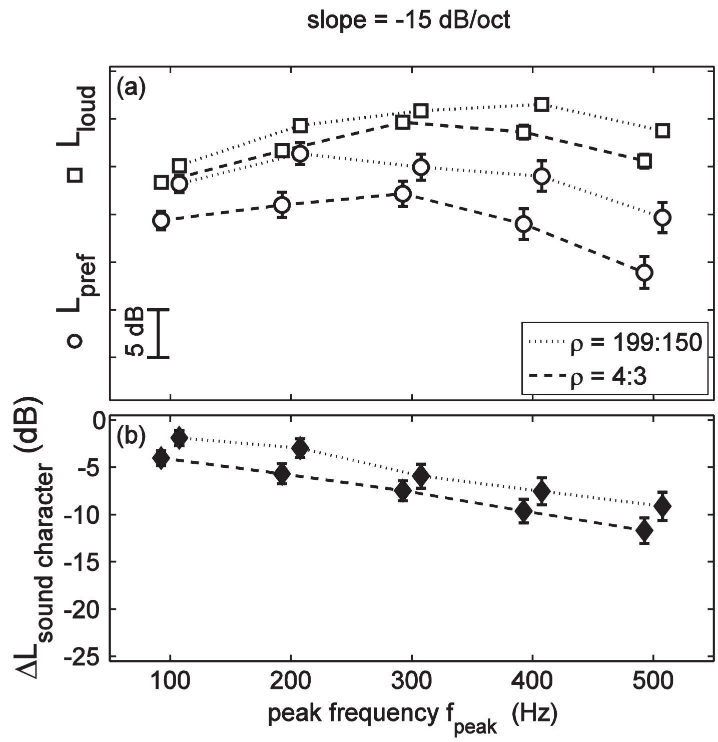

Figure 4: Same as Fig. 3 but for a slope of -15 dB/oct.

Overall, it seems to depend on the steepness of the spectral envelope towards higher frequencies which peak frequency configuration would be least loud or most preferred at an equal dB(A)-level. For the shallow slope of -6 dB/oct, a spectral peak at 100 Hz allows for the highest A-weighted levels while being equally preferred as the reference sound. For the steep slope configuration, the maxima of the PSEs also seem to depend on the frequency ratio between the fundamental frequencies, which are linked to the temporal structure of the signals, in a more complicated way. Rather clear relationships were observed for the level difference $\Delta L_{sound\ character}$ between the PSEs for preference and for loudness for all parameter variations.

4. RELATIONSHIP BETWEEN THE DIN-SHARPNESS AND THE SUBJECTIVE JUDGMENTS

Different psychoacoustic metrics were tested as potential descriptors for the spectral variations of the multi-tone sounds. High correlation coefficients, given in Tab. 2, were found between the sharpness according to the DIN 45692 standard and the PSEs for preference as well as for the level difference $\Delta L_{sound\ character}$. The sharpness according to the DIN standard has the advantage of being loudness and, thus, rather level independent. This avoids spurious correlation between the sharpness as a descriptor variable and the preference evaluations quantified as levels at the PSEs on a dB-scale. Figure 5 shows the relationship between the sharpness and the PSEs for preference. The relationship between the sharpness and the level difference $\Delta L_{sound\ character}$ is given in Fig. 6. The sharpness shares

7%-9% more variance with the level difference $\Delta L_{sound\ character}$ than with the raw values of the PSEs for preference L_{pref}. In addition, also the residuals between the data and linear regressions, indicated by the dashed and dotted lines in Figs. 5 and 6, are smaller for the level difference $\Delta L_{sound\ character}$ (Fig. 6) than for the raw preference evaluations L_{pref} (Fig. 5).

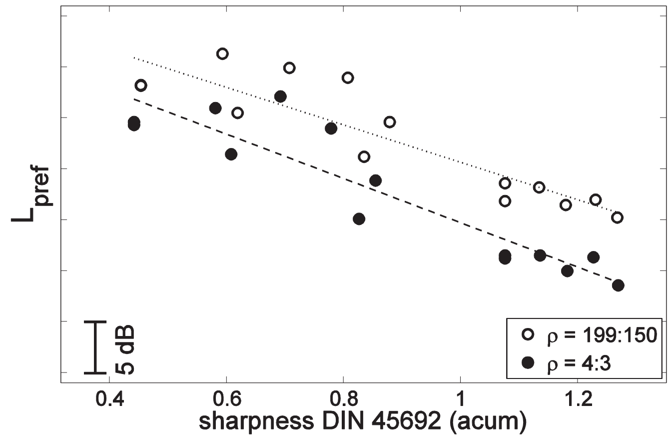

Figure 5: PSEs for preference L_{pref} plotted over the sharpness of the multi-tone signals, calculated according to the DIN 45692 standard. The underlying frequency ratio ρ, affecting mainly the temporal structure of the signals, is indicated by the symbols. Separate linear regressions for the two values of the frequency ratio ρ are shown as dotted (ρ=199:150) and dashed (ρ=4:3) lines.

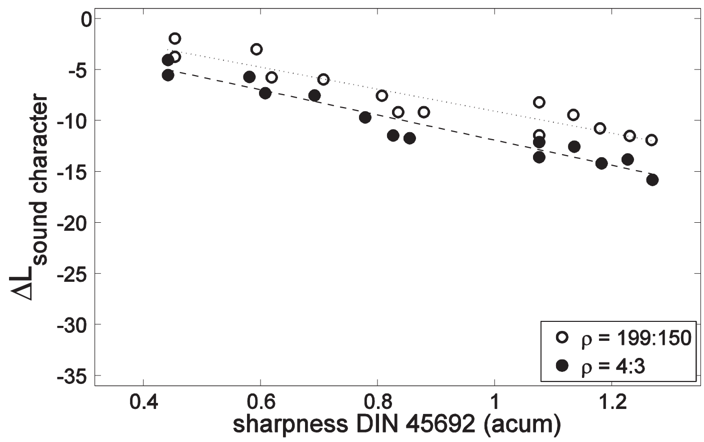

Figure 6: Values of the level difference $\Delta L_{sound\ character}$ between the PSEs for preference and for loudness plotted over the sharpness of the multi-tone signals, calculated according to the DIN 45692 standard. Legend see Fig. 5

Table 2: Pearson's correlation coefficients *r* between the judgments and the sharpness according to the DIN 45692 standard for the two values of the frequency ratio ρ, which are highly significant (p<0.001) in all cases.

Frequency Ratio $\rho = f_{01} : f_{10}$	PSEs for preference L_{pref}	Level difference $\Delta L_{sound\ character}$
4:3	*r* = -0.92**	*r* = -0.97**
199:150	*r* = -0.90**	*r* = -0.94**

4.1 THE COMBINATION OF THE SHARPNESS AND THE REPETITION RATE TO A MODEL FOR $\Delta L_{sound\ character}$

In previous studies it could be shown that the pleasantness of multi-tone sounds is linked to the periodicity of the signals, which resulted from a variation of the frequency ratio ρ (2, 3). In those studies, it turned out that the periodicity of the signals could be effectively characterized by the repetition rate of the signals. High correlation coefficients were found between the repetition rate and the preference evaluations and also for the level difference $\Delta L_{sound\ character}$.

Figure 7 combines the result from that previous study with the current results and shows the level difference $\Delta L_{sound\ character}$ plotted over the sharpness and the repetition rate of the signals. Some additional data, originating from supplementary experiments varying further signal parameters related to the temporal structure and the spectral content of the multi-tone sounds, is also included in Fig. 7 (1). The level difference $\Delta L_{sound\ character}$ becomes more negative with increasing repetition rate and also with increasing sharpness. Combining the repetition rate, as a descriptors of the temporal periodicity of the signals, and the sharpness, as a descriptor for the spectral content of the multi-tone signals, into a purely linear regression model, yields an adjusted r^2=0.89. The regression surface is shown in Fig. 7 as a mesh grid.

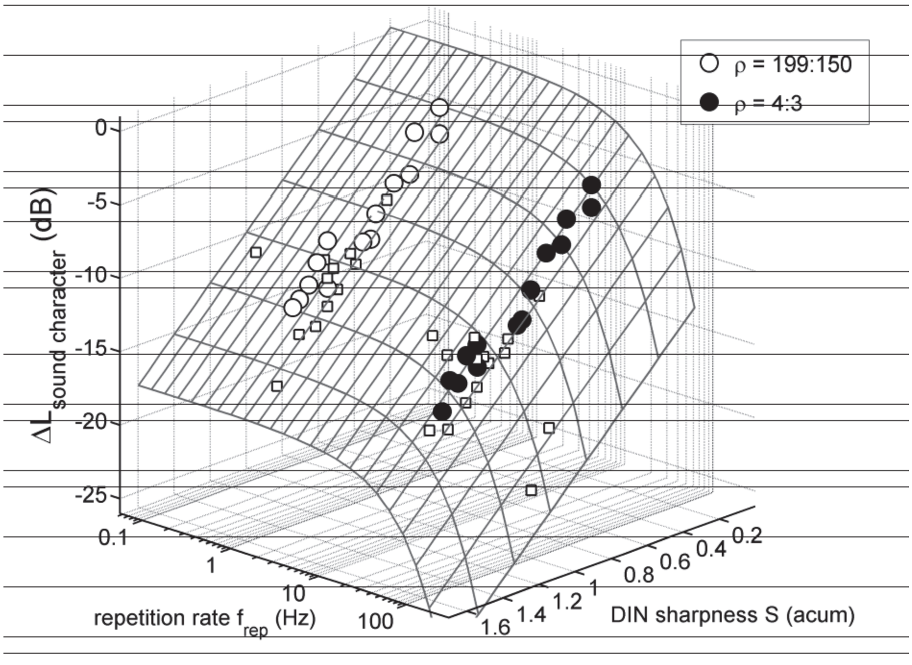

Figure 7: Values of the level difference $\Delta L_{sound\ character}$ between the PSEs for preference and for loudness plotted over the sharpness and the repetition rate of the signals. In general, $\Delta L_{sound\ character}$ becomes more negative with increasing repetition rate and increasing sharpness of the multi-tone sounds. A linear regression based on these two descriptor variables is indicated in the figure by the mesh grid surface.

5. SUMMARY

The influence of the spectral envelope on loudness judgments and preference evaluations was determined for multi-tone signals, consisting of two complex tones and additional combination tones, in the context of aircraft cabin noise. In separate listening experiments, the levels at the points of subjective equality (PSE) for loudness and for preference were determined in comparison to the same fixed reference noise. The level difference between the two PSEs is ascribed to the additional influence of the sound character, relevant in terms of the preference evaluation. The partials of the multi-tone sounds were shaped by a triangular spectral envelope. The peak frequency and the falling slope of the envelope towards higher frequencies were varied as independent variables.

The results indicate that a steep slope towards high frequencies yields lower levels at the PSEs for loudness and for preference than a shallow slope of the spectral envelope. This means that multi-tone sounds with a reduced amount of high frequency content can have higher A-weighted overall levels while being equally loud/preferred than sounds with a more shallow spectral decline. The results further indicate local maxima in the PSEs for loudness and for preference at slightly different values of the peak frequency f_{peak}, which depend on the slope steepness and on the frequency ratio ρ. The particular configurations of the signal parameters at these local maxima allow for higher A-weighted levels than the neighboring ones, while remaining equally loud or equally preferred as the reference, respectively. The relationships between the varied experimental parameters and the level difference $\Delta L_{sound\ character}$, between the PSEs for preference and for loudness, are considerably clearer. The values of $\Delta L_{sound\ character}$ become monotonously more negative with rising peak frequency and also with decreasing steepness of the slope.

It turns out that the psychoacoustic sharpness is a suitable descriptor of the spectral variations in the stimuli and it shares a considerable amount of variance especially with the level difference $\Delta L_{sound\ character}$. In previous studies, the repetition rate of the time signal was already identified as an effective descriptor of the signals' periodicity, which was found to be linked to the preference evaluation and the level difference $\Delta L_{sound\ character}$ (2,3). A purely linear regression model for $\Delta L_{sound\ character}$, which is based on the psychoacoustic sharpness and the repetition rate, accounts for 89 % of the variance in the judgments.

REFERENCES

1. Töpken S, Weber R. Preference and loudness judgments of multi-tone sounds and their relationship to psychoacoustical metrics, Proc. INTER-NOISE 2013, Innsbruck, Austria, 2013. pp. 1-8.
2. Töpken S, Scheel H, Weber R, Signal repetition rates and their relationship to the pleasantness of multi-tone sounds, Proc. INTER-NOISE 2014, Melbourne, Australia, 2014. pp. 1-8.
3. Töpken S, Verhey JL, Weber R. Perceptual space, pleasantness and periodicity of multi-tone sounds. J Acoust Soc Am. 2015;138(1):288-298.