Preference and loudness of multi-tone sounds

Stephan Töpken[1], Jesko L. Verhey[2], Reinhard Weber[1]
[1] *Carl von Ossietzky University Oldenburg - Acoustics Group, 26111 Oldenburg*
[2] *Otto von Guericke University Magdeburg - Dept. of Exp. Audiology, Leipziger Str. 44, 39120 Magdeburg*

Introduction

Noise generated by the rotary parts of a machine often consists of prominent complex tones and broadband noise. In special cases where nonlinear sound generating mechanisms lead to additional combination tones the resulting sounds can evoke a rather unpleasant sensation. In a previous study, the preference and loudness of this type of signals were evaluated using synthetic sounds composed of two complex tones and additional combination tones [1]. The points of subjective equality (PSE) for loudness and preference were measured compared to a reference sound using an adaptive paired comparison paradigm varying the level of the multi-tone sound.

The aim of the current study is to identify the contributions of the constituent basic sound elements, here the two complex tones and the combination tones, to the preference and loudness judgments of the entire sound.

Procedure

The points of subjective equality (PSE) for preference and loudness are measured in two separate experiments with an adaptive level varying paradigm. In a paired comparison the multi-tone sound is compared to a constant reference sound. The level of the multi-tone sound is changed according to the listener's preference or loudness judgment. The results are then given as level difference ΔL at the PSE which are a measure reflecting the relative preference and the relative loudness of the multi-tone sound compared to the reference sound.

The measurement of the preference is based on two underlying assumptions (see Fig. 1):

1. If the reference and the multi-tone sounds have the same A-weighted SPL, then the multi-tone sounds will not be preferred to the reference sound due to its unpleasant multi-tone sound character (Fig. 1, top).
2. If the level of the multi-tone sound is considerably attenuated compared to the level of the reference sound, then it will be preferred, obviously because the attenuated level increases the pleasantness (Fig. 1, bottom).

Somewhere between these two extremes lies the PSE for preference (Fig. 1, middle).

In a paired comparison of reference and multi-tone sound the participants are asked which sound they prefer. Depending on the answer of the participants the level of the multi-tone sound is varied adaptively by the simple up-down or staircase method [2]: It is reduced, if the signal is not preferred and raised if the signal is preferred.

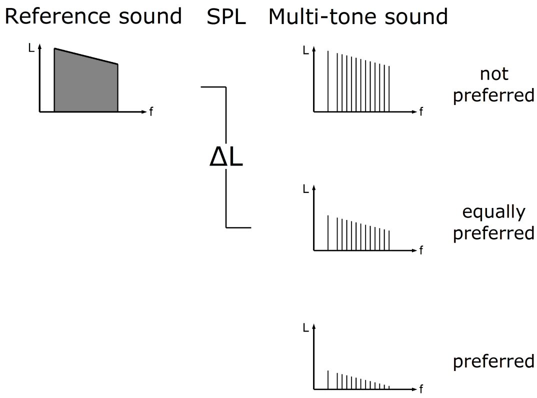


Figure 1: General assumption for the determination of preference judgments by a level variation of the multi-tone (test-) sound. If the level of the multi-tone and the reference sound are the same, then the multi-tone sound is not preferred due to its unpleasant sound character. If the level of the multi-tone sound is reduced considerably, then the multi-tone sound is preferred simply because it is less loud than the reference sound. Between these two levels lies the point of subjective equality for preference, where the multi-tone sound is equally preferred as the reference sound.

Figure 2 shows an example of a sequence of test sound levels during an adaptive track. At the beginning of each adaptive track the level of the multi-tone test sound is the same as the reference sound level, here L_A=74 dB(A).

The measurement of the loudness is carried out using the same paradigm, but now asking "Which sound is louder?" Previous studies have shown (e.g., [1]) that, in a condition, where both sounds have the same A-weighted overall level, the multi-tone sound is generally perceived as the louder one. If the level of the multi-tone sound is considerably attenuated, the reference sound is louder. Hence, the PSE for loudness will be somewhere in

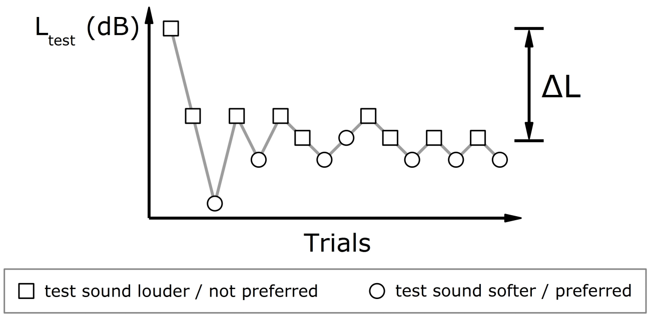


Figure 2: Exemplary sequence of the test sound level over trials. The symbols indicate the answer of the participant leading to the successive level change of the test sound. The level difference ΔL after the convergence of the adaptive paradigm is the point of subjective equality (PSE).

between these two level conditions. The level difference between the reference sound and the test sound is the PSE for loudness.

For the preference task, the level of the test stimulus is initially varied with a step size of 6 dB that is halved after each upper reversal of the level curve until the minimum step size of 1.5 dB is reached, where the measurement phase starts. The two sounds have a duration of 5 seconds each and they are separated by a pause of 1 second. A pause of 2 seconds is introduced between the answer and the presentation of the next stimulus pair.

For the loudness task, the step size is 3 dB at the beginning and then halved after the second upper reversal of the level curve again down to the step size of 1.5 dB. Between the presentations of the two sounds of a pair a pause of half a second and between the answer and the presentation of the next stimulus pair a pause of 1 second is inserted. The stimuli now have a duration of 1.5 seconds each. In both experiments (preference and loudness) the PSE of each participant is calculated as mean of the levels at the six reversal-points of the measurement phase (with a step size of 1.5 dB). The presentation order of reference and test sound within each trial and the order of the multi-tone test sounds over trials is randomized. The experiments for the five stimuli were measured interleaved with other multi-tone test sounds.

Stimuli

In this study superpositions of multi-tone sounds and their constituents are investigated as test sounds. The entire multi-tone sound is a superposition of three main elements - two complex tones (CX1 and CX2, with frequency components f_{i0} and f_{0j}) and additional combination tones (CTs with frequency components f_{ij}). The frequency components are calculated as shown in equations (1)-(3) with fundamental frequencies given in equation (4) and (5):

$$\text{CX1:} \quad f_{i0} = i \cdot f_{10} \qquad i = 1...30 \qquad (1)$$

$$\text{CX2:} \quad f_{0j} = j \cdot f_{01} \qquad j = 1...30 \qquad (2)$$

$$\text{CTs:} \quad f_{ij} = i \cdot f_{10} + j \cdot f_{01} \qquad i, j = 1...20 \qquad (3)$$

$$f_{10} = 100 \text{ Hz} \qquad (4)$$

$$f_{01} = 132.66 \text{ Hz} \qquad (5)$$

The entire multi-tone sound (CX1+CX2+CTs) therefore consists of 460 partial tones. The starting phases of the partials are taken from a fixed set of equally distributed random numbers $(0 \ldots 2\pi)$. They are the same for all test sounds. All partials decrease in level with a slope of 6 dB per octave. In the complete sound consisting of all elements, the partials of the combination tones are attenuated by 10 dB relative to those of the complex tones. To investigate the contribution of the constituting elements on the overall judgment, five stimuli out of the basic elements are prepared: (i) The entire multi-tone sound consisting of the two complex tones and the combination tones (CX1+CX2+CTs), (ii) a superposition of the two complex tones only, without the combination tones (CX1+CX2), (iii) the combination tones only without the two complex tones (CTs), (iv) the complex tone with the lower fundamental only (CX1) and (v) the complex tone with the upper fundamental only (CX2). An overview of the five test stimuli is provided in table 1. 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.

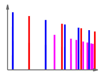
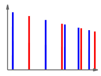
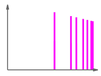
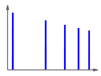
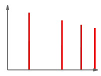
No.	Description	Elements	
i	The entire multi-tone sound consisting of booth complex tones and the combination tones	CX1 +CX2 +CTs	
ii	The two complex tones only, no combination tones	CX1 +CX2	
iii	The combination tones only, no complex tones	CTs	
iv	The lower complex tone only ($f_{10} = 100$ Hz)	CX1	
v	The higher complex tone only ($f_{01} = 132.66$ Hz)	CX2	

Table 1: Description of the constituent sound elements (see eq. 1 - 3) of the five stimuli used in the listening tests.

Listening Setup

The experiments were conducted in the anechoic chamber of the University of Oldenburg (lower limiting frequency: 50 Hz). The task itself was implemented as a Matlab (The Mathworks) routine on a computer. An external soundcard (M-Audio, Fast Track Pro) supplied the audio signals to an active loudspeaker (Mackie, HR 824) positioned in front of the participant who was seated inside the anechoic chamber. The experimental routine is operated by the participant via a computer keyboard and the TFT-screen, placed underneath the loudspeaker. The setup was calibrated with a handheld sound level meter (B&K 2226) at the listening position above the empty chair. The calibration was checked before the first and after the last experiment each day during the measurement phase. The setup is schematically shown in figure 3.

Participants

47 paid volunteers (23 females, 24 males), mainly from the University of Oldenburg, participated in the experiments. The mean age of the participants was 23 years (min=19 years, max=31 years). All participants reported no hearing difficulties.

Results

Figure 4 shows the mean level difference ΔL between the multi-tone test sound and the reference sound at

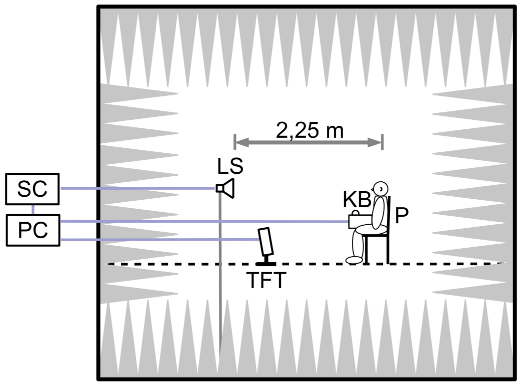


Figure 3: Schematic view of the setup in the anechoic chamber used for the listening tests: SC soundcard, PC computer, LS loudspeaker, TFT flatscreen, KB keyboard, P participant.

the level of the PSE for loudness (top panel) and the level of the PSE for preference (bottom panel) for the five test sounds. The error bars indicate the interindividual standard deviation of the judgments of the 47 participants.

All five multi-tone sounds require a considerable level attenuation to elicit the same loudness or preference sensation as the reference sound. Thus the multi-tone sounds are considerably louder and also less preferred than the reference sound when presented at the same A-weighted sound pressure level as the reference. The attenuation of the test sound at equal preference as the reference sound is larger than that to become elicit the same loudness sensation. The level differences ΔL between the multi-tone sounds and the equally-loud reference sound range from -5 dB to -11 dB. An attenuation of 11 dB is required for stimulus CX2, i.e., the second complex tone. The level differences to reach the PSE for preference are in the range from -18 dB to -31 dB. Hence, an attenuation of the multi-tone sounds

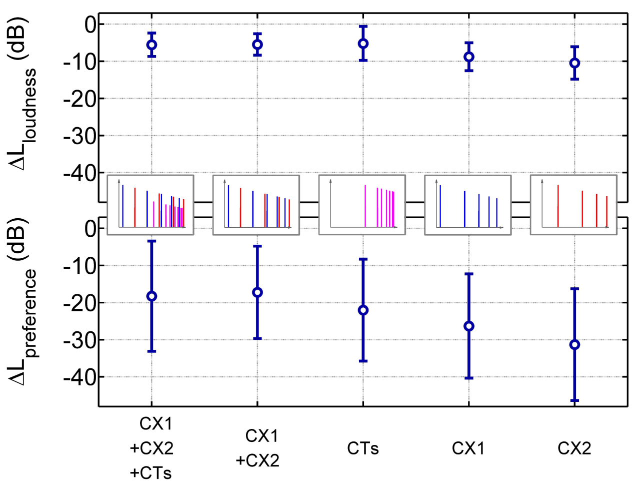


Figure 4: Mean results over 47 participants for the five stimulus configurations. Error bars indicate the interindividual standard deviation. The superposition of the two complex tones (CX1+CX2) is most preferred. The complex tones alone (CX1, CX2) are least preferred and loudest. The loudness is judged more uniformly by the participants than the preference.

of up to 31 dB for the least preferred sound (CX2) is necessary to produce the same preference as the reference sound. The interindividual standard deviations for the loudness task are about 4 dB and for the preference task about 14 dB. This indicates a higher inter-subject variability for the preference judgments than for the loudness judgments.

The results for the different sound elements, shown in Fig. 4, exhibit the following properties:

- The measurement of the loudness and the preference as a level difference ΔL compared to a constant reference sound provides a clear quantitative differentiation between the loudness and the preference judgment.
- The superposition of both complex tones (CX1+CX2) is slightly more preferred than the complete multi-tone sound (CX1+CX2+CTs) and more than the combination tones alone (CTs).
- A complex tone alone (CX1 or CX2) is considerably less preferred than the superposition of them (CX1+CX2). The level difference is about 9dB for CX1 and 14dB for CX2 at equal preference.
- The complex tone with the higher fundamental (CX2, $f_{10} = 132.66$ Hz) is also less preferred than the complex tone with the lower fundamental (CX1, $f_{01} = 100$ Hz).

Summary

The preference and the loudness of synthesized multi-tone sounds are determined as the point of subjective equality (PSE) compared to a reference sound as a function of the constituent sound elements. Five stimuli consisting of up to two complex tones with different fundamental frequencies (CX1, CX2) and additional combination tones (CTs) are evaluated by 47 participants. During the listening tests the level of the multi-tone sounds is varied according to the answer of the participants with an adaptive paired comparison paradigm. It is lowered, if the multi-tone sound is louder/not preferred and increased if the multi-tone sound is softer/preferred compared to the reference. The level difference at the PSE is a measure for the preference/loudness of the multi-tone sound. This PSE-method provides a clear quantitative distinction between the preference and the loudness judgments. The results show that all five multi-tone stimuli are generally louder and less preferred than the reference sound at the same dB(A)-level. In summary, the single complex tones (CX1, CX2) are less preferred and louder than the combination tones (CTs) or superpositions of these elements (CX1+CX2 or CX1+CX2+CTs). Nevertheless a superposition of the two complex tones (CX1+CX2) leads to the most preferred sound out of the five stimuli, which can be more than 8 dB higher in A-weighted SPL at the PSE for preference compared to the single complex tones alone (CX1, CX2).

References

[1] Töpken, S., Verhey, Jesko L., Weber, R.: Präferenz und Lautheit bei Multitonsignalen.
Fortschritte der Akustik - DAGA 2012, Darmstadt, Germany (2012)

[2] Levitt, H.: Transformed Up-Down Methods in Psychoacoustics.
J. Acoust. Soc. Am., JASA **49** (1971), 467-477