Suitable reference sounds for loudness and preference matching experiments

Stephan Töpken, Steven van de Par, Reinhard Weber
Carl von Ossietzky University Oldenburg, Dept. Med. Physics and Acoustics, Acoustics Group, 26111 Oldenburg, Germany, Email: stephan.toepken@uni-oldenburg.de

Introduction

Matching experiments are a way to indirectly determine auditory sensations (e.g. like loudness) by comparison against a reference sound. Loudness matches, for example, are based on the idea that an increase in level leads to higher perceived loudness. In listening tests, the point of subjective equality (PSE) for loudness between a test and a (common) reference sound can be determined by an adaptive level varying procedure. The level of the reference sound at the PSE for loudness is then an indirect measure of the loudness of the test sound.

However, an increase in level often also leads to a higher unpleasantness, especially for sounds which already yield annoyance. Accordingly, both aspects - loudness and unpleasantness/annoyance - can be measured with a level varying matching procedure [1]. Due to methodological reasons, asking the participants for their preference for one of the two sounds (test or reference) is more suitable than asking for an absolute annoyance judgment, because it can be difficult to create a situation in which a sound becomes annoying, especially in a laboratory environment. In a preference decision between two sounds the relative difference between the two sounds in terms of their unpleasantness is assessed, which is more suitable for a laboratory experiment.

Previous experiments of the authors have shown considerable differences between PSEs for loudness and PSEs for preference for a comparison of multi-tone sounds against a fixed reference [2]. This difference is attributed to preference relevant sound character differences between the test and the reference sound which is denoted $\Delta L_{\text{sound character}}$.

The aim of this study is to determine the ability of a reference sound to differentiate between the loudness and preference judgments and hence its suitability to measure sound character differences. In the current study, two reference sounds (pink noise and white noise) are matched in terms of loudness and preference against five test sounds (pink noise with embedded tonal components) each. The test sounds are constructed in such a way as to verify the current DIN 45681 standard, which prescribes level penalties in the assessment of noise containing tonal components.

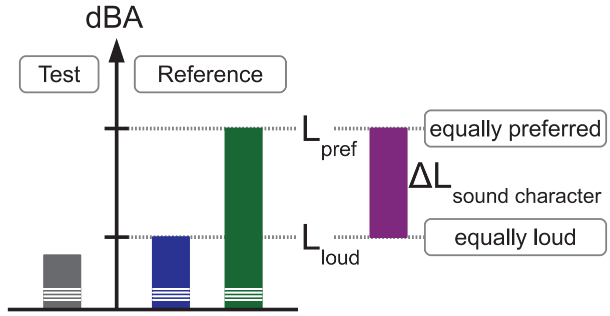

Figure 1: Definition of the levels in the matching experiments. The reference is matched in two separate experiments to equal loudness (L_{loud}) and to be equally preferred (L_{pref}) as a test sound. The difference between these two measures ($\Delta L_{\text{sound character}}$) is attributed to sound character differences between the reference and the test sound which are relvant in terms of preference.

Method

Procedure

In two separate experiments, the loudness and the preference of multi-tone test sounds were evaluated. In paired comparisons of a test and a reference sound, the participants were asked to decide which sound they prefer and which sound is louder. Depending on the answer of the participants, the level of the reference was varied until it becomes equally preferred/equally loud as the test which is kept constant in level. In the loudness task, the level of the reference sound was increased whenever the participant decided that the test sound was louder and decreased if the reference was louder. In the preference task the level of the reference was increased if it was preferred and decreased if the test sound was preferred. In both experiments, the level of the reference was varied with a simple staircase procedure (1-up, 1-down) converging at the 50-percent point of the psychometric function, which is the point of subjective equality. The initial step size of 3 dB was halved after the first upper reversal point of the adaptive procedure. After the second upper reversal point, the mean value over six reversal points was calculated as the PSE. The results of the judgments are thus given as levels of the reference sound at the PSEs for preference (L_{pref}) and loudness (L_{loud}), illustrated in figure 1. The level difference between the PSEs for loudness and preference is attributed to differences in sound character between the test and the reference sound which are relevant in terms of the individual preferences of the subjects. Based on the individual judgments, it is calculated by:

$$\Delta L_{\text{sound character}} = L_{\text{pref}} - L_{\text{loud}} \tag{1}$$

In this sense sound character is defined as everything that differentiates two equally loud sounds with respect to their (un-)pleasantness. If there is no such difference, then equal loudness is equivalent to equal preference and

$$\Delta L_{\text{sound character}} = 0 \text{ dB}.$$

On the other hand, if equal preference is not equivalent with equal loudness due to preference relevant differences in sound character between the two sounds, then

$$\Delta L_{\text{sound character}} \neq 0 \text{ dB}.$$

In this way, $\Delta L_{\text{sound character}}$ is also a measure for the suitability of a reference sound to measure relative sound character differences.

The measurements for the two different reference sounds were carried out in two measurement sessions, separated by at least one day. In each session, the loudness and the preference was judged with the same reference sound each.

Stimuli

Pink noise and white noise were used as reference sounds, which were adjusted in level with the adaptive procedure. The starting level of the reference sounds in the adaptive procedure was 64 dBA (same as the fixed level of the test stimuli) in both cases.

The test stimuli consisted of pink noise with embedded tonal components. Three test sounds contained single components (at 700 Hz, 3500 Hz and 10 KHz) and two test sounds contained two components (at 700 Hz +3500 Hz and 3500 Hz + 10 KHz). The levels of the two partials declined by 3 dB/octave (like the pink noise). The overall tone-to noise ratio was adjusted to $\Delta L = 12$ dB after the DIN 45681 standard, corresponding to a tone penalty of $k_{\text{T}} = 6$ dB for all sounds [3]. All test stimuli had a fixed overall level of 64 dBA.

The test and the reference sounds had a duration of 5 seconds. All stimuli were presented dichotically to avoid strong effects of in-the-head localization. Thus, uncorrelated noise signals were used for left and right ear and different starting phases were used for the left and right signal of the tonal components each.

Listening setup

The listening tests took place in a sound proof booth. The sounds were presented via open headphones (Sennheiser, HD 650), which were driven by an external soundcard (RME Audio, Fireface UC) . The experimental routine (2I-2AFC, 1-up 1-down) and the signal generation were implemented in a Matlab (The Mathworks) routine.

Participants

The listening tests were carried out by 8 participants, with a mean age of 31 years (minimum 25, maximum 38 years). One participant was female.

Results

Figure 2 shows the PSEs for loudness, preference and sound character difference for the pink noise reference.

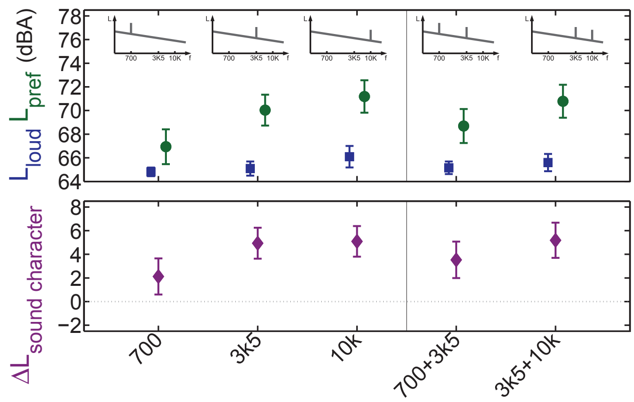

Figure 2: Mean values and standard errors of the PSEs with a pink noise reference for the five different test sounds. Considerable differences between the PSEs for loudness (squares) and the PSEs for preference (circles). Increasing sound character difference (diamonds) with rising frequency of the tonal components.

The level of the pink noise reference has to be increased by 1 dB to 2 dB to make it equally loud as the test sounds. The variation in the PSEs for loudness over the test sounds is only about 1 dB, which is just above the just noticeable differences in level (JNDL). Considerably higher levels are necessary for the pink noise to achieve equal preference compared to the test sounds containing tonal components. The resulting sound character differences are up to 6 dB. In general, the PSEs for preference and the sound character differences increase with a rise in frequency of the single tonal component. For the test sounds with two tonal components, the PSEs for loudness and preference as well as the sound character difference lie in between the PSEs of the respective single components.

Figure 3 shows the PSEs for loudness and preference and sound character difference for the white noise reference. The sound character differences for the white noise reference are considerably smaller than for the pink noise for all five test sounds. The values of $\Delta L_{\text{sound character}}$ are near 0 dB for three sounds and values of about 2 dB are reached for only two sounds. Thus, with the white noise reference it is not possible to measure sound character differences as big as with the pink noise reference for this set of stimuli. For the white noise reference, the PSEs for preference are within the range of the standard error the same as for the pink noise reference. The preference ranking of the test sounds for the pink reference is also reproduced quite well with the white noise reference. However, the PSEs for loudness are about 4 dB higher for the white noise reference than for the pink noise reference for all five test sounds while the differences in between the PSEs for loudness are again about 1 dB over all five test sounds.

Discussion

The DIN 45681 standard prescribes a tone penalty based on the tone to noise ratio of up to 6 dB to cover the an-

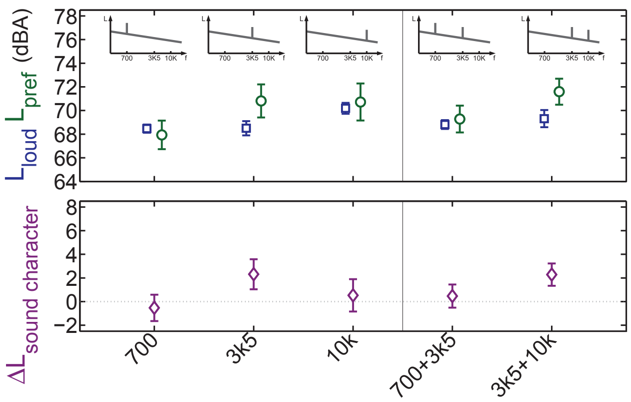

Figure 3: Mean values and standard errors of the PSEs with a white noise reference for the five different test sounds. Smaller differences between the PSEs for loudness (squares) and the PSEs for preference (circles) than for the pink noise reference, resulting also in smaller sound character differences (diamonds).

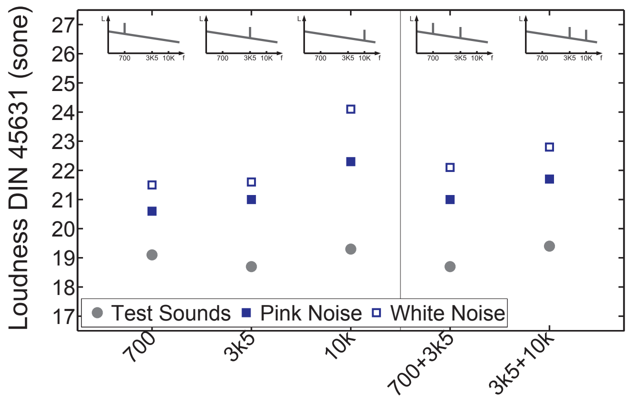

Figure 4: Calculated loudness values for the test sounds (circles) which were fixed at 64 dBA and the reference sounds at equal loudness: pink noise (filled squares) and white noise (open squares).

noyance of tonal sounds [3]. In the current case the tone to noise ratios of the test sounds are adjusted such that penalties of 6 dB after the standard are just reached for all test sounds. Based on the DIN standard the PSEs for preference would be predicted to be at 70 dBA (= 64 dBA sound pressure level of the test sound + 6 dB penalty after the standard). The found PSEs for preference lie between 67 dBA and 71 dBA for pink noise and between 68 and 72 dBA for white noise and are rather well covered by the standard. Only the stimuli containing a 10 kHz component are slightly underestimated by 1...2 dB by the tone penalty after the standard.

The calculated loudness values after the DIN 45631 standard for the test sounds and the two reference sound at the PSEs for loudness are shown in figure 4 [4]. The loudness values of the test sounds (with a fixed level of 64 dBA each) lie between 18.5 sone and 19.5 sone. The loudness values for the reference sounds at their PSEs for loudness are up to 5 sone above the loudness value of the respective test sound, each. The difference in between the the two reference sounds for each test sound is only up to 2 sone. A possible explanation for the differences in the calculated loudness values between the test and the reference sounds is an influence of the timbre on the subjective loudness judgment, which is not reflected in current loudness models [5].

Summary and Conclusion

Sound character differences between pink noise and white noise as reference sound and five test sounds consisting of pink noise with embedded tonal components have been measured. In two separate matching experiments, the level of the reference sound was varied to make it equally loud / preferred as test stimuli with a fixed level of 64 dBA. The level difference between the PSEs for preference and loudness is used as a measure for preference relevant sound character differences between the two sounds.

Over all test sounds, the sound character differences are bigger for the pink noise than for the white noise reference. Thus, pink noise is more suitable for the differentiation between the loudness and preference judgments for this set of test stimuli. This is mainly because the PSEs for loudness are about 4 dB higher for the white noise reference while the PSEs for preference are (within the limits of the standard error) the same for the two reference sounds. The remaining sound character differences for the white noise reference are near 0 dB for three sounds and barely reach 2 dB for only two sounds.

In general, the results from the preference experiment show that all the test sounds containing tonal components are less preferred than either of the two reference sounds. Hence, the reference sounds can be considerably higher in level than the test sounds to be equally preferred / unpleasant. The level difference between the test sounds and the reference sounds at the PSE for preference is quite well estimated by the tone penalty of 6 dB prescribed by the DIN 45681 for the used test sounds.

References

[1] H. Niese: Beitrag zur Relation zwischen Lautstärke und Lästigkeit von Geräuschen (On the relationship between loudness and nuisance of sounds), Acustica **15** (1965) 236–243.

[2] S. Töpken, J. L. Verhey, R. Weber: Preference and loudness of multi-tone sounds, AIA-DAGA 2013, International Conference on Acoustics, Merano, ISBN: 978-3-939296-05-8 (2013) 1269-1272.

[3] DIN 45681:2005-03: Acoustics - Detection of tonal components of noise and determination of tone adjustment for the assessment of noise immisions, Beuth, Berlin, Germany, German industry standard (2005).

[4] DIN 45631/A1:2008-01, Calculation of loudness level and loudness from the sound spectrum - Zwicker method - Amendment 1: Calculation of time-varaint

sound Beuth, Berlin, Germany, German industry standard (2008).

[5] R. D. Malera, L. E. Marks: Interaction among auditory dimensions: Timbre, pitch and loudness, Perception & Psychophysics (1990) **48**, 169-178.