99NV-54

Subjective Assessment of Roughness as a Basis for Objective Vehicle Interior Noise Quality Evaluation

Martin Pflueger, Robert Hoeldrich
Institute of Applied Systems Technology
Joanneum Research Forschungs-GmbH, Graz, AUSTRIA

Franz K. Brandl, Werner Biermayer
AVL List GmbH
Hans-List-Platz 1, Graz, AUSTRIA

ABSTRACT

This paper focuses on psychoacoustical experiments for the assessment of roughness by using vehicle interior noise. The experimental design has to be carried out carefully to derive reliable data for further analysis with objective parameters. Apart from the acoustical properties of the recording/playback system the different meanings of the word roughness have to be taken into account, because each person has its own interpretation of ‘roughness‘ differing between the phenomenons of roughness, r-roughness, rumble, harshness, fluctuation strength, etc.. An important preparation for psychoacoustical experiments is a clear definition of the sound attribute under investigation by using typical examples. Furthermore, accidental influences of other psychoacoustical parameters like the influence of loudness have to be avoided.

INTRODUCTION

The psychoacoustical sound quality ‘roughness‘ is one of the main quality features of vehicle interior noise. For that reason it is of great importance to take this psychoacoustical parameter into account.

The proposed psychoacoustical experiments for the exploration of roughness were carried out as a part of a research and development programm to establish an onboard analysis tool for vehicle interior noise quality. As stated in a companion paper [1] a new generalized psychoacoustical model of modulation parameters was successfully tested with the obtained roughness rankings [2].

SUBJECTIVE ASSESSMENT OF ROUGHNESS

The sound samples under investigation were recorded with an artificial head recording system [3,4,5,6]. All of the experiments took place in a laboratory which had been adapted to vehicle interior noise [7]. The recordings were played back with headphones and subwoofer simultaneously. By using this playback system it is possible to reproduce frequencies between 10 Hz and 20 kHz. The recording and playback system was tested with vehicle interior noise in previous experiments investigating the psychoacoustical term ‘annoyance’ [2].

The following results are all based on stationary driving conditions 50 kph / 2nd gear, 70 kph / 3rd gear, 70 kph / 4th gear, 100 kph / 4th gear and 130 kph / 5th gear. The experiments were carried out in two steps.

The discussed paired comparison experiments (cf. experiment 1,2 and 3) were originally intended to find an optimal roughness parameter - either available from literature [8,9,10,11] or from instrumentation suppliers - for our noise quality map under development [2].

Unfortunately there was hardly any correlation between the subjective scores ot the test persons and the roughness parameters in test. So we had to develope a new – specific - modulation parameter (roughness) for vehicle interior noise [1]. This decision required further experiments using more test persons and sound samples. These experiments werde carried out with the direct magnitude estimation method (cf. experiment 4 and 5).

PAIRED COMPARISON, EXPERIMENT 1

Experimental Planning

9 vehicle interior noise samples were presented to the test persons in order to evaluate the roughness of the samples. 7 test persons, all of them acousticians and experienced assessors, took part in this experiment. To avoid an irregular impact on the results no instructions were given.

The samples were not chosen accidentally, but represented a range of different variations of roughness within the ensemble.

Results

The results did not show any cyclic triads. This can be interpreted as an indication for the reliability of the assessments. It was obviously easy for the test persons to use the term roughness as an distinguishing feature. The correlation between the test persons is large with one exception (fig. 1).

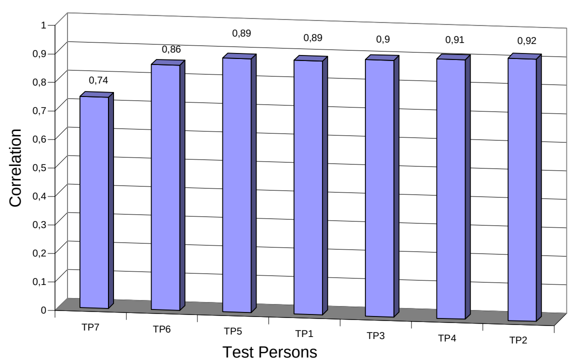


Figure 1: Average correlation between test persons

For 6 out of 7 experienced test persons there is a kind of ‘general agreement’ concerning the roughness ranking of these sounds. Consequently, the assessments of one test person (TP7, fig. 1) were not used for the calculation of the mean ranking.

The most striking result of this experiment was the high correlation between the roughness rankings and loudness parameters like dB(A), loudness (sone), etc..

Furthermore, the experiments have clearly shown that there is no correlation between the subjective roughness rankings and available roughness parameters.

PAIRED COMPARISON, EXPERIMENT 2

Experimental Planning

To avoid the strong impact of loudness on the roughness rankings, the sound samples were amplified to achieve equal loudness. The loudness reference was taken from the sound with the lowest loudness level (12 sone). For experiment 2 all 9 sound samples had 12 sone. The amplification of the sound samples did not change the natural sound quality.

Results

The results contain 2 cyclic triads which points out that the test conditions of the second experiment were harder

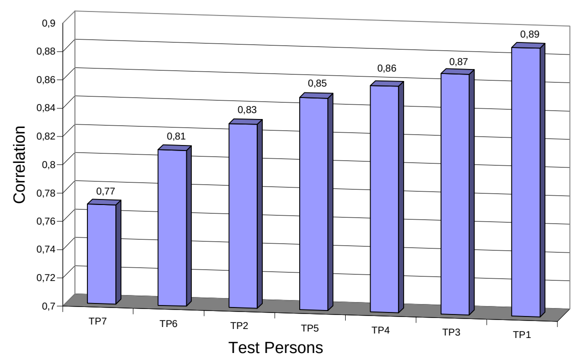


Figure 2: Average Correlation between test persons

t [illegible]) did not correlate with the other test persons. The average correlation between the test persons is smaller than in experiment 1 (fig. 2).

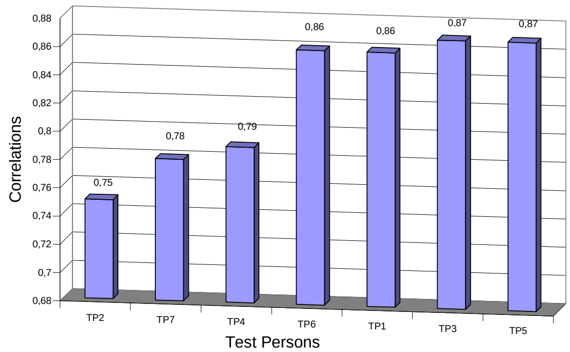
PAIRED COMPARISON, EXPERIMENT 3

Experimental Planning

The experimental planing of experiment 3 corresponded to a large extent with experiment 2. The only difference was that the original sound samples were amplified to achieve the same loudness as the maximum loudness that appeared in the ensemble. As a result all of the sounds had 27 sone.

Results

The results contain 15 cyclic triads which indicate, that several test persons had problems to execute the experiment. TP2 (9 cyclic triads), TP4 (5 cyclic triads) and TP7 (again different ranking than the other test persons but no cyclic triads) were excluded from the calculation of the mean ranking. The correlation between the test persons is not as good as in experiment 1 and 2.



By comparing the rankings of the three paired comparisons, three different types of assessments can be distinguished.
By summing up the ranking differences between experiment 1 (original loudness) and the two equal-loudness experiments (2 and 3) for each test person, it can be seen, that the assessment of some test persons (Tp 2,4,7) is highly influenced by the loudness of the sound samples. On the other hand, there were two test persons (Tp 1,5) whose rankings were hardly influenced by the loudness level (fig 4).

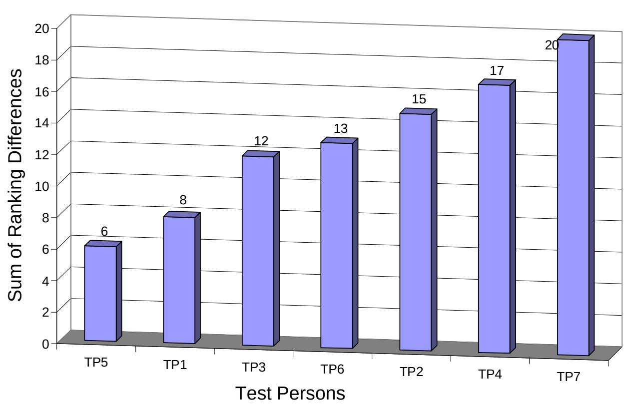
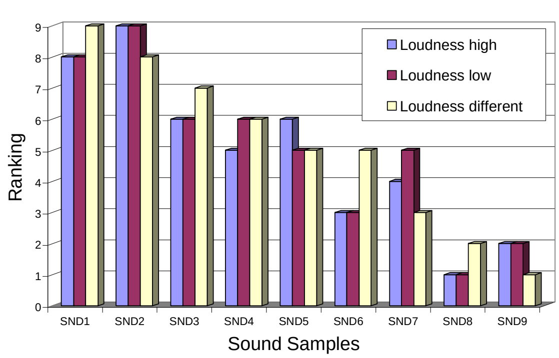


Figure 4: Assessment dependancy on loudness

A comparison of the mean rankings for these 3



experiments shows that the differences between the experiments with equal loudness levels (2 and 3) and the experiment with different loudness levels (1) is larger than the difference between experiment 2 and 3 (fig. 5).

Two types of experiments have to be distinguished. Roughness experiments with equal and different loudness. If the differences in loudness differ too much, the perception of roughness is highly influenced by the loudness:

The correlation of the three average rankings shown in fig. 5 with objective roughness parameters available from instrumentation suppliers did not show any correlation.

Figure 5: Average rankings of the experiments 1,2,3 e e

subjective results of experiment 1. The correlation coefficient R (R = - 0,39) is negative which totally contradicts our psychoacoustical data.

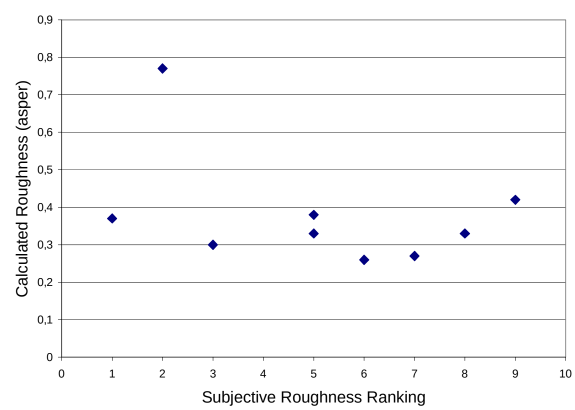


Figure 6: Correlation of subjectively perceived vehicle interior roughness with calculated data.

DIRECT MAGNITUDE ESTIMATION, EXPERIMENT 4

Originally, the experiments using the paired comparison method were carried out to find the best roughness algorithm for the developed our vehicle interior noise quality map. The faultiness of existing roughness algorithms concerning vehicle interior noise forced us to develope a new psychoacoustical model of roughness for objective noise quality evaluation of vehicles [1].

For the verification of the model being at the development stage it was necessary to carry out new experiments using more sound samples and more test persons. These experiments were done by using the direct magnitude estimation method.

Experimental Planning

The selected sound samples had a similar loudness level (low loudness). The samples were not amplified so that they were all original artificial head recordings. 28 test persons took part in the experiments assessing 13 sound samples. In contradiction to the paired comparison experiments discussed above the test persons were not all professional acousticians. Again, the instructions for the test persons did not contain any information about the term ‘roughness’. We decided not to present typical sound examples at the beginning of the test.

Results

The assessments of the 28 test persons are different in many regards. First of all, the assessments of 6 test persons could not be taken into account because these persons did not correlate with the other 22 persons. Some of these persons did clearly not assess roughness but other psychoacoustical quantities like sharpness or loudness.
The assessments of the main group (22 people) shows 2 clusters. Cluster 1 comprises 15 persons and cluster 2 comprises 7 persons. The mean rankings and standard deviations of cluster 1 ( ) and cluster 2 ( ) are shown in fig. 7.

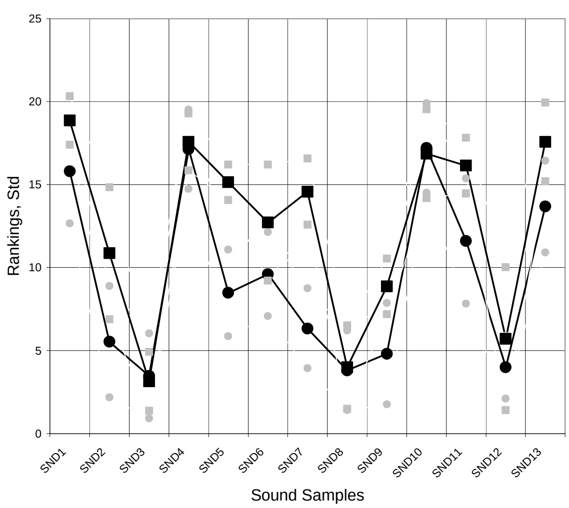


Figure 7: Mean rankings and standard deviations of subjectively perceived roughness.

The data shows that there are certain sounds with large differences in the mean rankings. Ranking differences larger than 5 points (on a scale from 1-20) occur with SND2 (5,3 points), SND5 (6,7 points) and SND7 (8,2 points) (cf. fig. 7). These samples are obviously – to a large extent - responsible for the inhomogeneity of our results.

DIRECT MAGNITUDE ESTIMATION, EXPERIMENT 5

Experimental Planning

The experimental planning of experiment 5 corresponded to a large extent with the concept of experiment 4. The only difference was that the selected sound samples were different from experiment 4 and had a higher loudness level.

Results

Again, the assessments of the 28 test persons are different in many respects. The assessments of 8 test persons could not be taken into account because these persons did not correlate with the other 20 persons. Some of them did not assess roughness but other psychoacoustical quantities. The assessments of the main group (20 people) shows 2 clusters. Cluster 1 comprises 11 persons and cluster 2 comprises 9 persons. The mean rankings and standard deviations of cluster 1 ( ) and cluster 2 ( ) are shown in fig. 8.

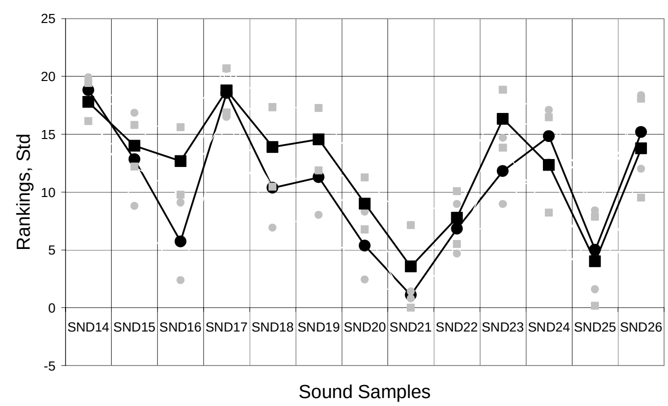


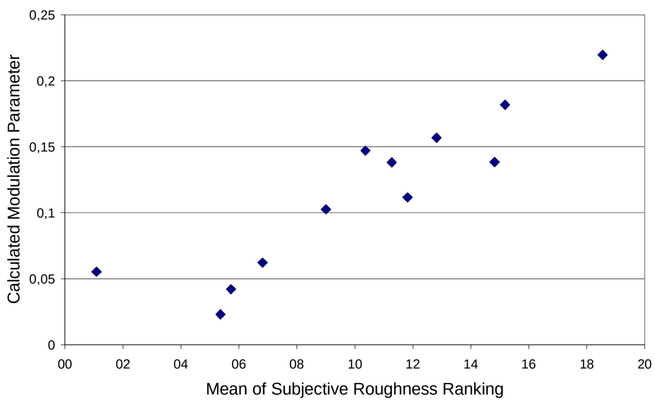
Figure 9: Mean rankings and standard deviations of subjectively perceived roughness.

The biggest differences appearing between the two clusters are with SND16 (6,9 points), SND20 (3,6 points) and SND23 (4,5 points). Again, certain sound samples turned out to be responsible for the inhomogeneity of the subjective rankings.

CORRELATION OF SUBJECTIVE SCORES WITH OBJECTIVE MODULATION PARAMETERS

As reported in a companion paper we developed a generalized psychoacoustical model for modulation parameters which can be adjusted to different types of vehicle sounds. The model offers certain degrees of freedom in the calculation of the excitation-time pattern,

the effective modulation index, the specific roughness and the superposition.
The model was successfully tested with the subjective data of the experiments 4 and 5 (fig. 9).



Compared to available objective roughness parameters there is a significant correlation between the subjective scores and our new modulation parameters. Correlation parameters higher than R = 0,9 can be achieved by dividing the assessments of test persons into clusters.

CONCLUSION

- Sound characteristics caused by modulation play an important role because they contribute significantly to the perceived annoyance.
- Psychoacoustical experiments for the assessment of roughness have to be carried out carefully to derive reliable data for further analysis with objective parameters.
- Two types of roughness experiments have to be distinguished. Experiments with equal and different loudness. If the differences in loudness differ too much, the perception of roughness is highly influenced by the loudness.
- The assessments of about 10 to 20 percent of unexperienced test persons had to be neglected because these persons did not assess roughness but other psychoacoustical quantities.
- It is possible to achieve a high correlation between the mean of subjective scores and objective modulation parameters.

ACKNOWLEDGMENTS

The authors thank sincerely Dr. M. Hussain of Joanneum Research and Dr. A. Ronacher for their contribution to this work.

CONTACT

Dr. Martin Pflueger is running the department of acoustics and psychoacoustics at the JOANNEUM RESEARCH in Austria. One of the main fields of M. Pflueger is the developement of new psychoacoustical measurement systems for vehicle noise.
Email: martin.pflueger@joanneum.ac.at

REFERENCES

[1] R. Hoeldrich, M. Pflueger: A Generalized Psychoacoustical Model of Modulation Parameters (Roughness) for Objective Vehicle Noise Quality Evaluation, SAE 1999, 99NV-55.

[2] F. Brandl, W. Biermayer: A new Tool for the Onboard Objective Assessment of Vehicle Interior Noise quality, SAE 1999.

[3] N.C. Otto: Listening Test Methods for Automotive Sound Quality, AES Preprint, 1997, 4586(L-6)

[4] E. Zwicker, H. Fastl: Psychoacoustics – Facts and Models. – Springer, Berlin, 1990.

[5] M. Zollner: Methodisch bedingte Fehler bei binauralen Hörversuchen, Proc. DAGA 95, pp 779-782.

[6] T. Hashimoto, Takao H.: Subjective Estimation od Running Car Interior Noise, Proc. Inter-Noise 90, pp 1181-1184

[7] M. Hussain, M. Pflüger, F. Brandl: Intercultural Differences in Annoyance Response to vehicle Interior Noise. Euro-Noise 1998, Munich.

[8] W. Aures: Ein Berechnungsverfahren der Rauhigkeit. Acustica Vol. 58, 1985, pp268-281.

[9] P. Daniel, R. Weber: Psychoacoustical Roughness: Implementation of an optimized Model. – Acustica Vol. 83, 1997, pp 113–123.

[10] N. Springer, R. Weber: Existenzbereich der R-Rauhigkeit. – Proc. DAGA 94, pp 1221–1224.

[11] N. Springer, R. Weber: Bewertung von amplitudenmodulierten Schallen im R-Rauhigkeitsbereich. – Proc. DAGA 95, pp 839–842.