EFFECTS OF MEANING OF SOUND ON LOUDNESS JUDGEMENTS

PACS: 43.66.+y

Hellbrück, Jürgen[1]; Fastl, Hugo[2]; Keller, Barbara[1]
[1]Environmental and Health Psychology, Catholic University of Eichstätt-Ingolstadt
Ostenstrasse 26
D-85071 Eichstätt, Germany
juergen.hellbrueck@ku-eichstaett.de
[2]AG Technische Akustik, MMK, TU München, Arcisstrasse 21
D-80333 München, Germany
fastl@mmk.e-technik.tu-muenchen.de

ABSTRACT

Loudness sensation depends mainly on the intensity, bandwidth, and duration of sounds. However, it is often said that context effects such as the meaning of sound can influence loudness judgments. In our study, the loudness of a noise scene of a duration of 20 min. was continuously judged by use of a category scale. In addition, the global loudness of the whole noise was evaluated. Twelve subjects participated in this experiment. In the same way, another sample of eleven subjects judged the loudness of a noise that mirrored in detail the loudness fluctuations of the former noise; however this signal did not contain any meaning. This meaningless noise was gained by applying an algorithm introduced by Fastl [1], which does not affect loudness, but which largely eliminates meaning. An additional meaningless sound was created by filling the temporal envelope of the original sound with pink noise. This sound was judged by a third sample of 10 subjects. One of the results of the study was that differences of the continuous loudness judgments were shown, though only at rather low sound pressure levels. In this case loudness of meaningless sound was found to be systematically lower. Loudness judgments of meaningless sound gained by Fastl's method correlate better with original sound than sound synthesized by pink noise. The global loudness, however, did not reveal any significant differences.

INTRODUCTION

Loudness sensation is a psychoacoustic magnitude that corresponds very closely to sound intensity, but depends also on bandwidth and duration of sound. Loudness can be measured by various procedures of magnitude estimation, i.e. direct psychophysical scaling based on sensation ratios. Loudness is a crucial psychoacoustic factor in subjective sound evaluation, and has a substantial effect on annoyance. In some cases, however, the meaning of sound is believed to influence the subjective evaluation. For example, it has been often found that railway noise is preferred to road traffic noise, though both sounds have the same energy equivalent A-weighted sound pressure level. This could be demonstrated in field studies [2] as well as in laboratory studies [3], and the related phenomenon is designated as railway bonus. This preference for railway noise is sometimes explained by the different meaning of sounds. For example, one might suggest that subjects prefer railway noise since trains are more accepted than road traffic, i.e. subjects evaluate the sound source, not the sound. On the other hand, it could be shown that railway noise and road traffic noise dffer remarkably in frequency spectrum and therefore have different specific loudness patterns. Moreover, time-structure is also different between these two noises.

The question whether the meaning of sound has an influence on subjective sound evaluation arises not only in the context of debates on environmental noise assessment but also in debates on sound quality evaluation. For example, the image of a product may be assumed to have an effect on the sound evaluation. In general, from the psychoacoustic viewpoint the answer to the question whether judgments of psychoacoustic magnitudes are biased by meaning, is crucial.

In order to study the effect of meaning of sound on loudness, it is necessary to compare subjective evaluations of meaningful sounds with those of sounds which have the same loudness but no meaning. Removing the meaning of sound is usually realized by extracting the temporal envelope, and multiplying this signal with pink noise. In this way a sound is generated which has the same temporal structure as the original one but which is devoid of meaning. Loudness, however, depends not only on sound intensity but also on bandwidth. Therefore, the loudness of such a synthesized sound usually differs from the original sound. Such differences may reach an amount of up to 20 sone [1].

In order to avoid such problems Fastl has proposed a new procedure. The original sound is first analysed by Fourier-Time-Transform (FTT). Elements of the FTT-patterns are then spectrally broadened, and sounds are synthesized by an inverse FTT-algorithm. Consequently, using this procedure a sound is synthesized with the same temporal structure and the same loudness-time function [1].

In our experiment, we compared loudness judgments based on the time sequence of a long-term traffic noise with two synthesized versions of this noise, one realized with help of the FTT-based procedure, the other with the conventional procedure using a pink noise. Additionally, the overall loudness of the traffic noise scene as well of the synthesized noises was judged.

If loudness judgments depend only on the psychoacoustic magnitude loudness, then loudness judgments of the original sound and those of the synthetic sound neutralized by Fastl's procedure should not significantly differ. Between the pink noise and the original sound, however, a difference is to be expected. If meaning does have an effect, however, then loudness judgments of original noise and the FTT-processed noise should reveal significant differences. Because of the lack of literature reporting respective findings we have no hypotheses of the way in which such differences caused by meaning would occur. Does meaning affect judgments in a general way, or are judgments affected in a specific way by certain meanings? Moreover, meaning not only might by dependent on objects but also upon the individual subject and the context. Because no well-founded specific hypotheses can be made, our experiment is part of a pilot study.

METHOD

Experimental Design

The experimental design should give an answer to the question whether loudness judgments of neutralized sound differ from judgments of the original meaningful sound. To this end, we compared the original sound with two synthesized meaningless sounds. The original sound was neutralized to gain the meaningless sound. Two versions of neutralizations were realized, one by applying the FTT-based method as outlined above, and the other one by the conventional method using pink noise. Under each condition different subjects performed loudness judgments. They were randomly assigned to the experimental conditions. In the Table 1 the experimental conditions are illustrated:

Subjects

In total, 33 subjects (average age: 41) participated in the experiment. All of them were employees of the Bavarian Environmental Protection Agency. None of them was experienced in loudness scaling. All subjects reported normal hearing abilities.

Table 1: Experimental Conditions

Original Sound	Neutralized Sound by FTT Method	Neutralized Sound by Pink-Noise Method
12 Subjects average age: 39	11 Subjects average age: 44	10 Subjects average age: 39

Stimuli

The original sound we used was a traffic noise scene (20 min. in duration), recorded near a village on a main road at a gated level crossing. The railway crossing involved was very busy. The trains that passed were goods trains, regional and Intercity trains. Cars and lorries regularly queued up in front of the closed gate at the crossing and most of the drivers switched off their engines in accordance with a sign near the crossing. After a train had passed and the gate had been opened, the noise-situation was dominated by engines starting up and vehicles pulling away. Once the gate was open vehicles drove past one by one at a speed suited to the railway crossing. The background noise consisted of moderate wind noise, rustling leaves, occasional birdsong, and sounds associated with the rural environment and proximity of the village.

The noise scene was recorded with artificial-head technique (HeadAcoustic HRS II.2) and played back with 2 loudspeakers (Canton Ergo 92 DC) in the unechoic chamber of the Bavarian Environmental Protection Agency in Augsburg (Germany). Sound pressure level was calibrated by measuring a calibration broadband noise at the position of subject's head. The mean energy level of the 20-minute noise was L_{Aeq} = 76.3 dB [4,5].

Scaling method

We used the Category Subdivision Scale (CS Scale) which has proven successful for continuous loudness scaling of traffic noise [4]. The CS Scale is a combination of category scale and number allocation. It comprises five verbally distinguished categories. Each category contains ten steps allowing the observer to make fine gradings: 1-10 ("very quiet"), 11-20 ("quiet"), 21-30 ("medium"), 31-40 ("loud"), and 41-50 ("very loud"). The scale allows the possibility of going beyond 50 to express noise perceived as painfully loud.

CS-scaling is a direct scaling procedure based on categorical judgements which has been used thus far predominently for loudness scaling in audiometry and hearing-aid fitting, but which is not limited in its application to these procedures. It could also prove advantageous in noise assessment. Unlike other procedures of direct scaling which are based solely on number allocation, verbal categorisation provides additional information, such as which noise level is neither soft nor loud, and thus "medium loud". This type of information is very useful if one wants to know how people feel and speak about the noise [4].

Procedure

Each subject was first exposed to the whole noise, and instructed just to listen to the noise. After the sound had been fully presented, the subjects were requested to judge the overall loudness (OL) of the whole sound by Category Subdivision Scale (CS Scale).

In the second trial the noise was interrupted by 5s-pauses every 15 s. In each pause the subjects were required to judge the loudness of the respective 15s-noise interval. In total, eighty 15s-noise intervals were consecutively judged [5].

RESULTS

In Table 2 the means and standard deviations of the overall loudness judgments are shown.

Table 2: Means and standard deviations of overall loudness judgments

Original Sound	Neutralized Sound by FTT Method	Neutralized Sound by Pink Noise
M = 33.08 SD = 4.58 N = 12	M = 33.64 SD = 5.94 N = 11	M = 35.60 SD = 6.15 N = 10

The loudness judgments of each of the eighty short-term intervals were averaged. The means and corresponding standard deviations are shown in Figure 1. At first glance, one can see that FTT-processed sounds are judged to be softer than the original sound, but only at low intensity levels. At high levels there is a ceiling effect. A direct comparison is shown in Figure 2 and Figure 3, where CS-loudness judgments of the meaningful original sounds are compared with FTT-processed sounds and Pink Noise sounds respectively. The coefficient of the product-moment correlation between loudness judgments of the original sounds and FTT-processed sounds ($r = .963$) is significantly higher than the coefficient of the correlation between loudness judgments of original sounds and Pink Noise sounds ($r = .877$) ($\alpha = .01$; after Fisher's Z-Transformation).

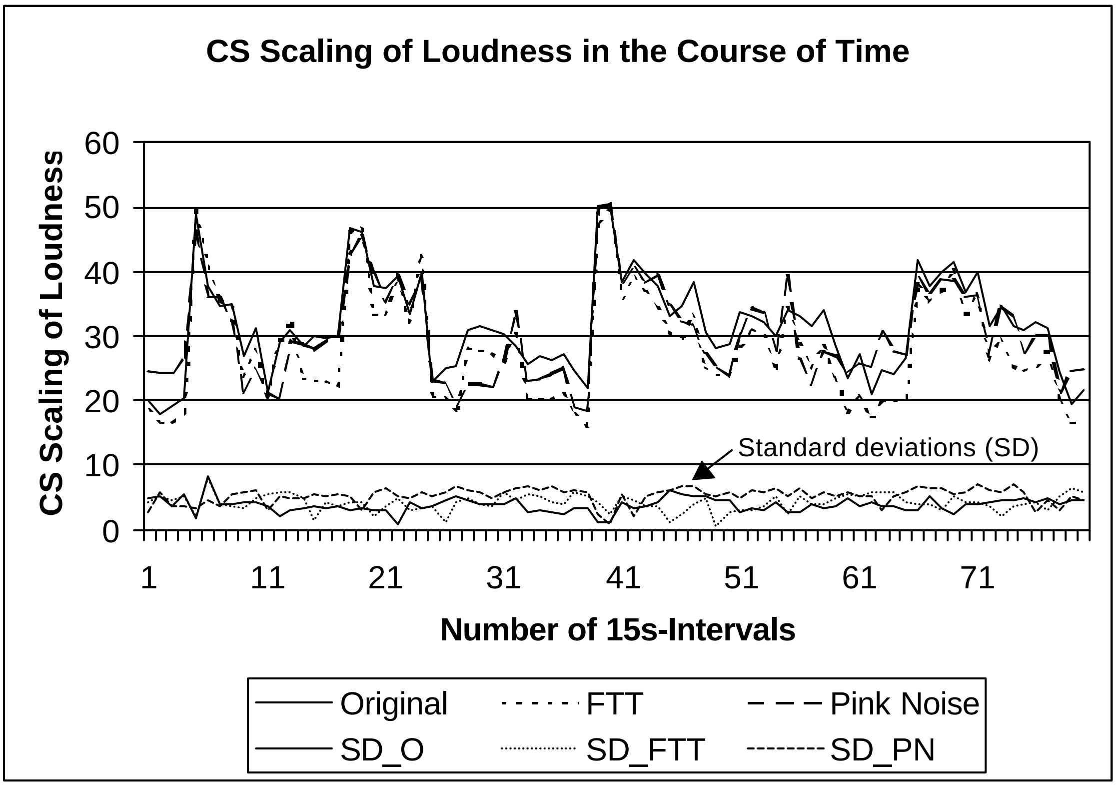

Figure 1. Means and standard deviations (SD) of CS-loudness scaling.

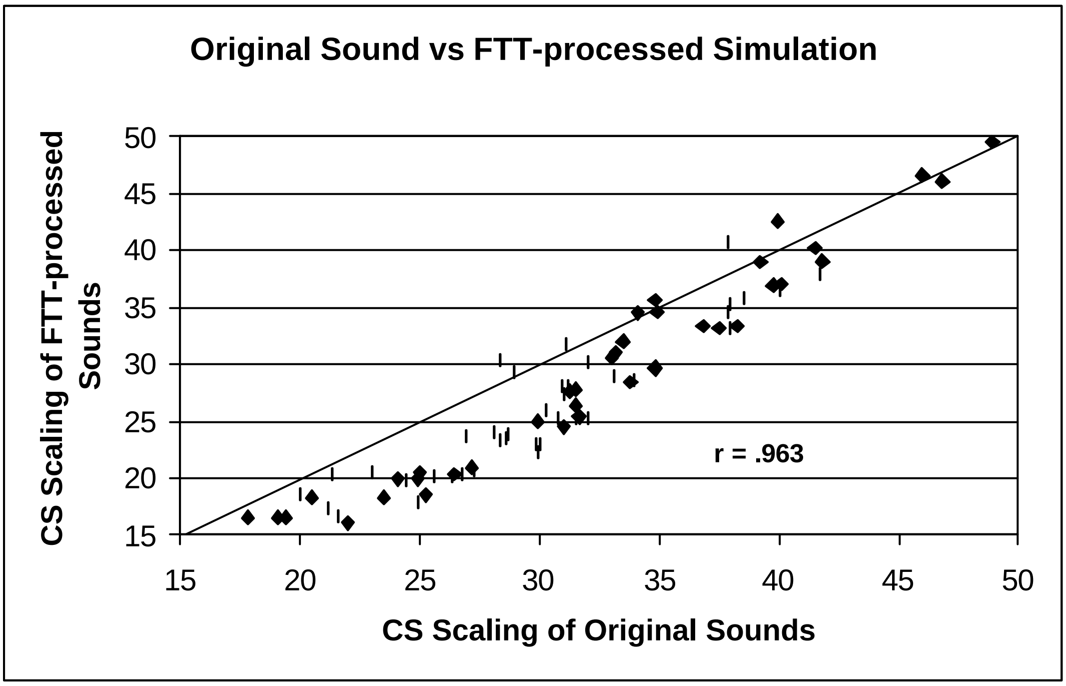

Figure 2. Comparison between loudness judgments of original sounds and FTT-processed sounds for each of the respective eighty 15s-intervals. Product-moment correlation: r = .963; 1%-confidence interval: r = .934 – .979.

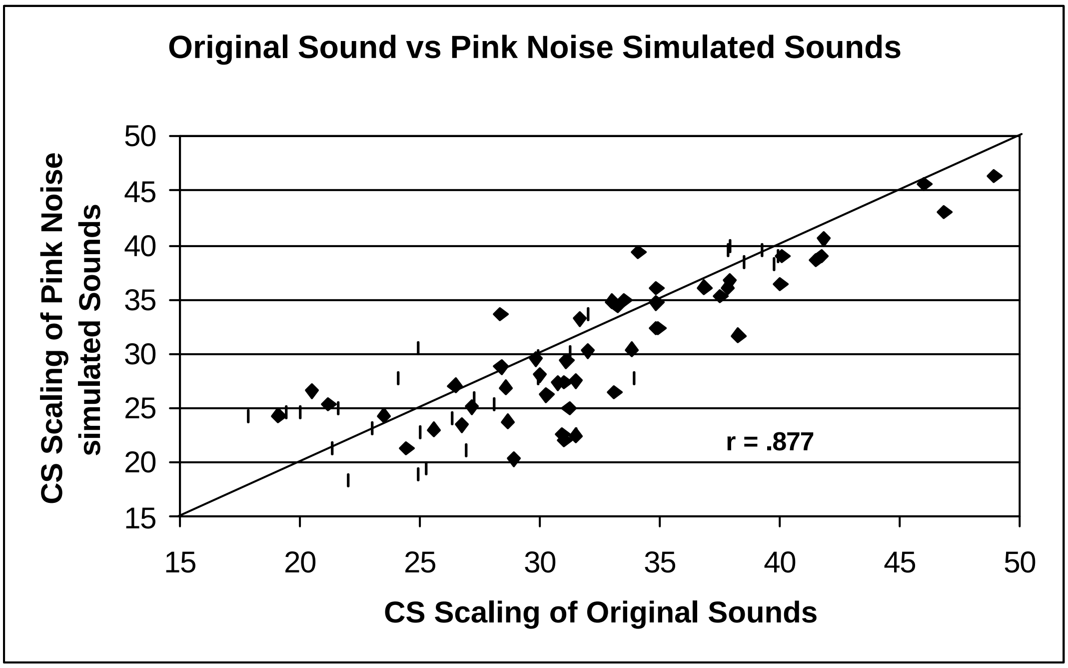

Figure 3. Comparison between loudness judgments of original sounds and pink-noise sounds for each of the respective eighty 15s-intervals. Product-moment correlation: r = .877; 1%-confidence interval: r = .789 – .930.

DISCUSSION

The results are clear-cut. First, there is practically no significant difference between the mean overall loudness judgments. Second, the correlation between the consecutive loudness judgments of 15s-intervals of the original sound and those of the FTT-processed sound is considerably higher than the correlation between respective loudness judgments of the original sound and the pink noise sound. Furthermore, one can clearly see that loudness judgments between original sound and FTT-processed sound differ systematically. At low intensity levels FTT-processed sounds are judged to be softer than the respective original sounds. The pink noise shows the same tendency, but, due the higher scattering of the data not as clearly.

The high correlation between FTT-processed sounds and original sounds indicates that the FTT method reflects the psychoacustic parameters of the original sound very faithfully, at least with regard to loudness judgments. By contrast, the pink-noise procedure is apparently less reliable.

The systematic deviation between FTT-processed sound and original sound can be explained by the influence of meaning. But why are the loudness judgments of original sound systematically higher, in general, than those of synthesized sounds? It might be suggested that the recognition of sound sources sharpens attention by isolating such sounds from the background, and that therefore, such sounds appear to be louder. This might be the case especially at low intensity levels, while at high levels loudness dominates all other factors. Thus, our results might reflect effects of selective attention in hearing. Auditory attention plays an important role in detection and discrimination of stimulus attributes, in the frequency domain and in the intensity domain as well. This has been shown in several experiments carried out by Scharf and co-workers as well as other researchers (for an overview cf. [6]).

ACKNOWLEDGEMENTS

The authors wish to thank Bernhard Seeber for support in generating the sounds, and Wolfgang Vierling and Rüdiger Borgmann of the Bavarian Environmental Protection Agency in Augsburg, Germany, for support in realizing this study.

REFERENCES

[1] Fastl, H., Neutralizing the meaning of sound for sound quality evaluations, In: Proc. 17th Intern. Congr. Acoustics 2001, CD-ROM (2001).

[2] Möhler, U., Community response to railway noise: a review of social surveys, J. Sound Vib. 120, 321-331 (1988).

[3] Fastl, H., Kuwano, S., Namba, S., Assessing the railway bonus in laboratory studies. J. Acoust. Soc. Jpn. (E) 17, 139-148 (1996)

[4] Hellbrück, J., Category subdivision scaling - A powerful tool in audiometry and noise assessment. In H. Fastl et al. (Eds.), Recent trends in hearing research. Festschrift for Seiichiro Namba (pp. 317-336). Oldenburg: BIS (1996).

[5] Hellbrück, J. Memory effects in loudness scaling of traffic noise – how overall loudness of short-term and long-term sounds depend on memory. J. Acoust. Soc. Jap. (E) (21), 329-332 (2000).

[6] Scharf, B., Auditory attention. In H. Pashler (Ed.), Attention (pp. 75-117). Hove: Psychology Press (1998).