Significance of Meaning in Sound Quality Evaluation

Alfred Zeitler[1], Wolfgang Ellermeier[1], Hugo Fastl[2]

[1] *Sound Quality Research Unit, Dept. of Acoustics, Aalborg University, Denmark*

[2] *AG Technische Akustik, MMK, TU München, Germany*

Introduction

There is agreement that perceptually-adequate sound evaluation has to encompass not only the sensory domain but also cognitive and emotional variables [1]. In particular, the meaning of a sound has been identified as an important factor which may include effects of attitude, prior experience, and user expectations. Further, it is assumed that the effects of meaning are mediated through the prior identification of the sound source in the listening process.

The current investigation employs a new signal-processing method proposed by Fastl [2] which reduces the identifiability of sound sources substantially while preserving the loudness-time functions. In a previous experiment [3] differences in loudness judgments between original and processed versions were found for a number of every-day sounds. The present experiment extends the investigation to annoyance judgements as well, and includes Semantic Differential ratings of all sounds to explain potential effects. Further, a stimulus set is used that covers a larger range of the loudness continuum. Initial results on the loudness scaling portion of the data are reported in [4].

Method

Participants

A total of 100 normal hearing students of Aalborg University aged between 19 and 31 years (median: 23 years) were randomly assigned to one of the four scaling conditions as shown in Table 1.

	Stimuli	
	original	neutralized
Loudness scaling	*N* = 25 (9 female/16 male)	*N* = 25 (12 female/13 male)
Annoyance scaling	*N* = 25 (10 female/15 male)	*N* = 25 (12 female/13 male)

Table 1: Experimental design: tasks and stimuli.

Apparatus and Stimuli

Signals of largely non-stationary character (duration: between 0.7 and 5 s) stemming from a wide range of sources (examples see below, and in [4]) were used.

The original sounds had been recorded in appropriate distances between 0.5 and 7 m from the microphone (type Brüel & Kjær 4165 and 4179) in a sound-insulated listening room complying with IEC 268-13, except for the recordings of car noises. The overall levels ranged from 30 to 80 dB SPL. According to a pilot study, the sounds were highly identifiable (% correct > 80%) in their original condition. The stimulus set also included seven levels of pink noise signals spaced in 10-dB steps between 20 and 80 dB SPL to check for comparability of the subject groups.

A signal processing scheme as proposed by Fastl [2] was applied to reduce source identifiability while preserving the time-loudness functions as well as the spectral envelopes.

Signals were diotically presented to the subjects listening in a double-walled chamber via headphones (Beyer DT 990 Pro).

Procedure

In the first session, participants provided judgements on category-subdivision (CS) scales of either loudness or annoyance. The scales were each comprised of five verbal categories and further subdivided in 10 numerical steps. Danish equivalents of the following labels were used: "very-soft" (1-10), "soft" (11-20), "medium" (21-30), "loud" (31-40), "very loud" (41-50), and "very slightly annoying" (1-10), "slightly annoying" (11-20), "medium" (21-30), "strongly annoying" (31-40), "very strongly annoying" (41-50). The experiment started with practice trials to facilitate the task and to provide orientation of the stimulus series.

After completion of the ratings, the identifiability of the signals was assessed. The participants were asked to to specify what they have heard by writing down a noun and a verb (e.g., "motor - idling").

In the second session, participants completed a concept-specific Semantic Differential (SD) on the same stimulus set. The SD was comprised of 7-point rating scales that were anchored to 12 concept-specific adjective pairs (in Danish) at the extremes (see Table 3).

Results

Identifiability

A percentage-correct scoring scheme was used to determine the degree of identifiability of the sounds. Figure 1 demonstrates that the median score drops from 94% for the original to 14% for the processed sounds. However, it is also shown that the effect of processing varies considerably among sounds.

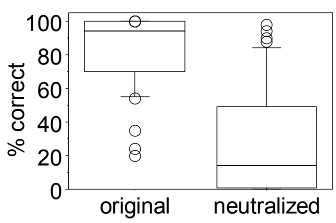

Figure 1: Boxplots of percentage-correct scores. Boxes indicate interquartile range and median, whiskers denote the 10th and 90th percentiles.

Effects on Loudness and Annoyance Scalings

Two-factor repeated-measures analysis of variance were performed to investigate whether the differences between original and neutralized sounds were statistically significant. For the loudness scalings no main effect of processing was found but a significant (stimulus by processing) interaction ($F(11.68,\ 560.39) = 4.80$; $p < .001$). Post-hoc tests were performed, and the alpha-level was adjusted to $p = .00125$ to correct for chance outcomes due to multiple testing (Bonferroni correction). Significant mean differences were thus found for the following sounds: ringing alarm clock ($M_{\text{neutr}} - M_{\text{orig}}$ = -7.76 scale units), buzzing radio alarm (-7.60) and bicycle bell (6.52).

For the annoyance data, a significant main effect of processing was found indicating that the processed sounds were judged more annoying on the average ($F(1,\ 48) = 4.99$; $p < .03$). Further, a significant interaction effect was obtained demonstrating that the effect differs between sounds ($F(16.25,\ 780.01) = 7.92$; $p < .001$). Post-hoct tests yielded significant differences for the following sounds: bicycle bell ($M_{\text{neutr}} - M_{\text{orig}}$ = 12.04 scale units), champagne glass (13.60), coffee-maker (9.36), bouncing coin (11.52), door lock (13.08), toilet flush (13.60), water tap (8.88).

Differences in the Semantic Differentials

Mean ratings on the Semantic Differential scales were considered to explain the aforementioned differences in the loudness and annoyance judgments. Table 2 shows correlations between mean differences in the annoyance or loudness scale means ($M_{\text{neutr}} - M_{\text{orig}}$) and the corresponding differences in the SD ratings.

Adjective scale	CS Annoyance	CS Loudness
calming-agitating	.78**	.76**
dark-bright	.21	.23
dull-sharp	.47**	.49**
flat-rumbling	.45**	.54**
muffled-shrill	.60**	.59**
pure-impure	.13	.10
smooth-rough	-.09	-.11
soft-hard	.47**	.51**
steady-unsteady	-.37*	-.25
ugly-beautiful	-.66**	-.70**
unpleasant-pleasant	-.79**	-.78**
weak-strong	.61**	.65**

* $p < .05$, ** $p < .01$

Table 2: Correlation of mean differences between original and neutralized sounds with differences on the SD scales.

Correlation with Zwicker Loudness

Percentile loudness N_5 values were calculated by means of the software Brüel & Kjær Sound Quality (type 7698, version 3.4.0) and correlated with mean loudness and annoyance judgments (Table 3). The annoyance judgements of the neutralized sounds proved to be more strongly correlated with N_5 loudness values ($p < .01$). For the loudness judgments, the difference between the coefficients failed to reach statistical significance ($p = 0.08$).

	original	neutralized
Mean CS loudness	.91**	.95**
Mean CS annoyance	.81**	.95**

Table 3: Correlations with Zwicker loudness N_5.

Discussion and Conclusions

Significant mean differences between loudness judgments of original and processed sounds were found only for 3 out of 40 stimuli. Thus it may be concluded that source identification was not crucial for the loudness impression. However, a larger variety of source attributions was reflected in the responses to the processed signals, indicating that the stimuli may not necessarily have been meaningless for the subjects.

By contrast, stronger effects were found in the annoyance judgements. Taking into account that psychoacoustic parameters other than loudness may have changed (see [5]), the result is not too surprising. Both psychoacoustic parameters and the meaning are assumed to account for the discrepancies. Further investigation is necessary to uncover the role of psychoacoustic parameters for the differences.

Furthermore, it was shown that the effects – both in loudness and annoyance scaling – can be mapped onto the semantic profiles of the sounds, which are thought to provide a measure of (connotative) meaning.

Moreover, it was shown that the loudness judgements of both original and processed stimuli were very well predicted by "Zwicker loudness". In the case of annoyance, the judgments of the original sounds were less predictable from the loudness metric than those of the processed signals. It is assumed that the effect of meaning was mediated by the degree of source identification, and thus had a more pronounced impact on the judgements of the original sounds.

Acknowledgements

Research supported by the Centercontract on Sound Quality (Aalborg University, Bang & Olufsen, Brüel & Kjær, DELTA Acoustics & Vibration, EFS, STVF).

References

[1] Blauert, J. & Jekosch, U. Sound-quality evaluation - a multi-layered problem. Acta Acustica **83 (1997)** 747-753.

[2] Fastl, H. Neutralizing the meaning of sound for sound quality evaluation. Proceedings 17th ICA **2001**, Rom (CD-ROM).

[3] Zeitler, A., Hellbrück, J. & Fastl, H. Einfluss der Bedeutung auf die Lautstärkebeurteilung von Umweltgeräuschen. DAGA **2003** (CD-ROM).

[4] Ellermeier, W., Zeitler, A., Fastl, H. Impact of source identifiability on perceived loudness. To be presented at ICA **2004**, April 4-9, Kyoto, Japan..

[5] Fastl, H. Features of neutralized sounds for long term evaluation. Proceedings Forum Acusticum, **2002** (CD-ROM).