Neutralizing the meaning of sound for sound quality evaluations

H. Fastl

Institute for Man-Machine-Communication, Technical University München, Arcisstr. 21, 80333 München, Germany

Sound quality usually can be described by psychoacoustic magnitudes like loudness, sharpness or roughness. However, sometimes the meaning of a sound may strongly influence its subjective evaluation. In order to assess these effects, sounds „without meaning" have to be realized. A frequently applied solution of this problem is to fill the temporal envelope of an original sound with pink noise. However, this procedure has the severe disadvantage that the original sound and the sound with no meaning differ in loudness. Therefore, a procedure is proposed, which can „neutralize" the meaning of sound, and at the same time preserve the original loudness.

INTRODUCTION

For the assessment of sound quality, psychoacoustic magnitudes like loudness, sharpness, fluctuation strength or roughness have proven successful. In many cases, a combination of these basic psychoacoustic magnitudes can predict sound quality ratings by subjects [1]. Despite the fact that the application of these principles usually leads to successful engineering results, sometimes the meaning of a sound may considerably influence its evaluation.

For example it could be verified both in the field (see [2]) and in the laboratory (see [3]) that at same energy equivalent A-weighted level, railway noise is preferred to road traffic noise. From an engineering approach it could be shown [4] that spectral differences between railway noise and road traffic noise are reflected in the specific loudness patterns and can account partly for the advantage of the railway noise. Also the time-structure of noise immissions from railway noise and road traffic noise differ significantly. In addition to these differences, which can be assessed by physical means, there may be also an influence of the different meanings of the sounds [3].

PROCEDURES

In order to neutralize the meaning of sound, in several labs the following procedure is used (cf. Figure 1).

From the original sound, the temporal envelope is extracted, e.g. by rectifying an low-pass filtering. In a modulator, this signal is multiplied with pink noise, leading to a sound which has the same temporal envelope as the original sound but no meaning. However, because loudness depends on the bandwidth of a sound, even if the temporal envelope is faithfully extracted and filled with pink noise, the original sound and the synthesized sound usually differ in loudness.

These differences are of great relevance, since the evaluation of sound quality in many cases crucially depends on the loudness of the sound ([1], [5]).
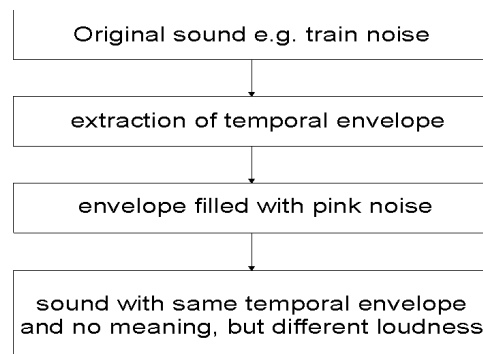

FIGURE 1. Traditional procedure to remove meaning of sound.

The loudness differences to be expected because of different bandwidth are illustrated in figure 2. The left part of figure 2 shows the dependence of the loudness of pink noise (PN) or an 1 kHz-tone on level. At same level, the loudness of the pink noise is much larger than the loudness of the pure tone. For example at 94 dB, as indicated by the vertical double arrow, the loudness of the pink noise is about a factor of 2.5 larger than the loudness of the 1 kHz-tone.

The right part of figure 2 shows original sounds (left column) and sounds with the same temporal envelope, filled with pink noise (right column). A comparison of the left and right column reveals that the temporal structure of original and synthesized sounds is very similar. However, in many cases the loudness of the synthesized sounds is larger. This discrepancy is illustrated in the left part of figure 2 by vertical lines. For example sound (a), when replaced by pink noise with same temporal structure (b), increases from around 40 sone to around 60 sone. Likewise, the

simulations (d) and (f) of the sounds (c) and (e) show a loudness which is approximately 50% larger than the loudness of the original sounds.
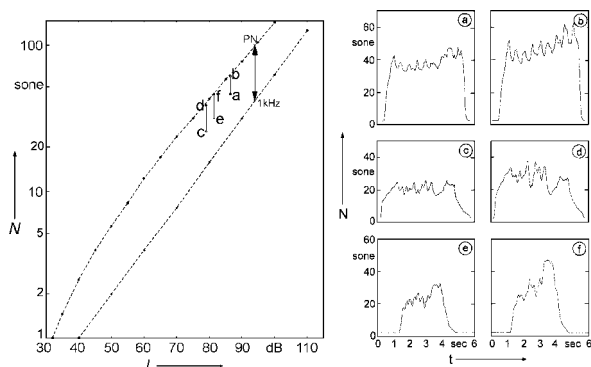

FIGURE 2. Loudness differences of sounds with same level but different bandwidth.

Because loudness is such a dominant cue in sound quality evaluation, we looked for possibilities to preserve the temporal envelope and also the loudness. In this respect, the Fourier-Time-Transform (FTT [6], [7]) proved to be a very helpful tool. The principle of the procedure applied is outlined in figure 3. The original sound first is analysed by the FTT-procedure. After spectral broadening of the elements of the FTT-patterns, sounds are sythesized by an inverse FTT algorithm. As a result we get a synthesized sound, which has the same temporal envelope, no meaning and – in contrast to the procedure illustrated in figure 1 – also the same loudness-time function as the original sound.
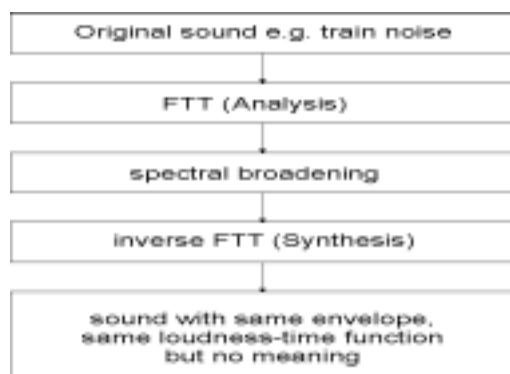

FIGURE 3. New procedure to remove meaning of sound.

Figure 4 shows for comparison original sounds and their counterparts with neutral meaning obtained by the procedure illustrated in figure 3. When comparing loudness-time functions displayed in figure 4a vs. figure 4b, there are almost no differences discernible. This means that the proposed FTT-based procedure not only faithfully reproduces the temporal envelope of the

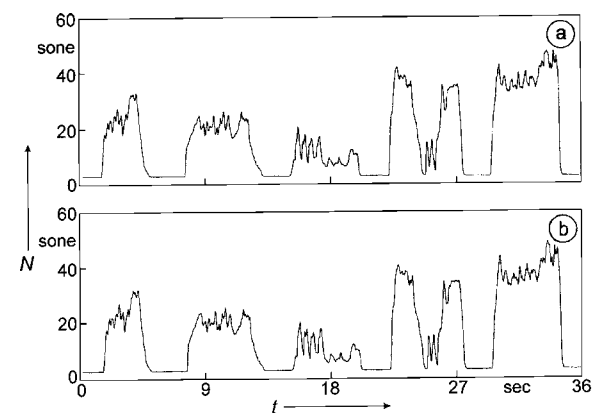original sounds, but also preserves their loudness in great detail.

FIGURE 4. Comparison of original sound (a) and corresponding sound without meaning (b).

OUTLOOK

With the tool illustrated in figure 3, we now have the possibility to study sounds which are as much as possible identical with the crucial difference that one is an original sound with a specific meaning, and the other is the corresponding synthesized sound, from which the meaning was removed.

ACKNOWLEDGEMENTS

REFERENCES

1. Zwicker, E., Fastl, H., *Psychoacoustics. Facts and Models*. 2nd updated ed., Springer-Verlag, Berlin, 1999.
2. Möhler, U., Community response to railway noise: a review of social serveys, J. Sound Vib. 120, 321-331, 1988.
3. Fastl, H., Kuwano, S., Namba, S., Assessing in the railway bonus in laboratory studies. J. Acoust. Soc. Jpn. (E) 17, 139-148, 1996.
4. Fastl, H., Masking effects and loudness evaluation. In: Recent Trends in Hearing Research (H. Fastl et al. Eds.) Bibliotheks- und Informationssystem der Carl von Ossietzky Universität Oldenburg, Oldenburg, 29-50, 1996.
5. Fastl, H., Sound Quality of Electric Razors - Effects of Loudness. In: Proc. inter-noise'2000, CD-ROM, 2000.
6. Terhardt, E., Fourier transformation of time signals: Conceptual revision. Acustica 57, 242-256, 1985.
7. Mummert, M., *Sprachcodierung durch Konturierung eines gehörangepaßten Spektogramms und ihre Anwendung zur Datenreduktion*. VDI Reihe 10, Nr. 522, VDI Verlag, Düsseldorf, 1998.