Psychoacoustics and Sound Quality

Hugo Fastl

Technical-Acoustics Group, Department of Human-Machine-Communication,
Technical University of Munich, Munich

Summary. In this chapter psycho-physical methods which are useful for both psycho-acoustics and sound-quality engineering will be discussed, namely, the methods of random access, the semantic differential, category scaling and magnitude estimation. Models of basic psycho-acoustic quantities like loudness, sharpness and roughness as well as composite metrics like psycho-acoustic annoyance will be introduced, and their application to sound-quality design will be explained. For some studies on sound quality the results of auditory evaluations will be compared to predictions from algorithmic models. Further, influences of the image of brand names as well as of the meaning of sound on sound-quality evaluation will be reported. Finally, the effects of visual cues on sound-quality ratings will be mentioned.

1 Introduction

Psychoacoustics as a scientific field has a tradition of more than 2500 years. For example, already around 500 B.C. the Greek philosopher *Pythagoras* – with his monochord – had studied musical consonance and dissonance. These early experiments had all the ingredients of a psycho-acoustical experiment. Psychoacoustics needs a sound stimulus which can be described in the physical domain. For Pythagoras this physical quantity was the length of the string stretched out along his mono-chord and supported by a bridge. By varying the position of the bridge and while plucking both ends of the string he judged with his hearing system whether the resulting musical interval was consonant or dissonant, i. e. he judged on attributes of his auditory percept. In this way, he found out that for simple ratios of the string divisions – like 1:2, i. e. octave, 2:3, fifth, and 3:4, quart – consonant musical intervals were perceived.

In modern psycho-acoustics the procedures applied are very much the same as those that have already been used by Pythagoras. At first, acoustic, i. e. physical, stimuli are produced, nowadays usually with the help of sophisticated digital signal-processing algorithms. After D/A conversion the resulting signals are presented to subjects via headphones or loudspeakers. The subjects, then, are asked to judge upon attributes of what they hear, such as the pitch, the loudness or the tone colour of the perceived sounds.

The same principles are often applied in sound-quality engineering, however, in reversed sequence as follows. During extensive psycho-acoustic stud-

ies an optimum sound for a specific product is "tailored" – a so-called *target sound*. Consequently, it is the task of engineers to modify the physics of sound generation, e. g., in an industrial product, in such a way as to arrive at a sound which comes as close as feasible to the target sound.

In this chapter psycho-physical methods which are useful for both psycho-acoustic research and sound-quality engineering will be discussed. Algorithmic models for the estimation of basic psycho-acoustic quantities like loudness or sharpness as well as compounded metrics will be introduced, and their application to sound-quality design will be elaborated on. Further, some practical examples will be reported, e. g., concerning comparison of results from listening tests with predictions as rendered by algorithmic models, the influence of brand names, and the effect of visual stimuli on sound-quality judgements.

2 Methods

For sound-quality evaluation psycho-physical methods are in use which have already proven successful in psycho-acoustics. From a variety of possible methods, four more important ones have been selected for discussion in this chapter, namely, *ranking methods* – they indicate whether a product sounds better than the product of a competitor, the method of *the semantic differential* – it provides hints on what sounds are suitable to convey an intended message, e. g., as a warning signal – and *category scaling and magnitude estimation* – which can give an indication of how much the sound quality differs among products, which is often of particular relevance for cost/benefit evaluations.

2.1 The Ranking Procedure "Random Access"

A ranking procedure called "random access", which has proven very successful for the investigation of sound-quality [12], is illustrated in Fig. 1. In the example displayed, six sounds, denoted *A* through *F*, have to be ranked with
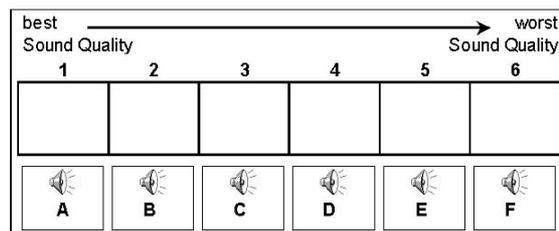

Fig. 1. Example for ranking of the sound quality by the method random access [15]

respect to their sound quality. When clicking on the loudspeaker icon, the respective sound, e. g., an idling motor, is heard. The task of the subject is to shift the icons, *A* through *F*, into one of the empty fields, denoted 1 through 6, in such a way that the sounds are finally ordered with respect to their sound quality. The subjects are free to listen to each individual sound as often as they like and to correct the sequence again and again, until they feel that a final status has been reached. This large freedom of the subjects, who have "random access" to the sounds to be ranked, is one of the reasons for this procedure to be preferred nowadays for ranking of sound quality.

2.2 The Semantic Differential

The method of "the semantic differential" is used to test what sounds are suitable for an intended purpose. In Fig. 2 an example of adjective scales is given, which has been used in an international study on the suitability of signals as warning signals [22]. It goes without saying that warning signals should have high loadings on adjectives like dangerous, frightening and unpleasant.

adjective scales	
loud	soft
deep	shrill
frightening	not frightening
pleasant	unpleasant
dangerous	safe
hard	soft
calm	exciting
bright	dark
weak	powerful
busy	tranquil
conspicuous	inconspicuous
slow	fast
distinct	vague
weak	strong
tense	relaxed
pleasing	unpleasing

Fig. 2. Semantic differential from an international study on warning signals [22]

2.3 Category Scaling

"Category scaling" is a preferred method for the assessment of the loudness of the sounds of products. Five-step scales as well as seven-step scales are usually employed, e. g., [10]). Figure 3 gives examples of five-step scales as well as seven-step scales as used for loudness assessment. In comparison to the five-step scale, the seven-step scale has in addition the steps "slightly
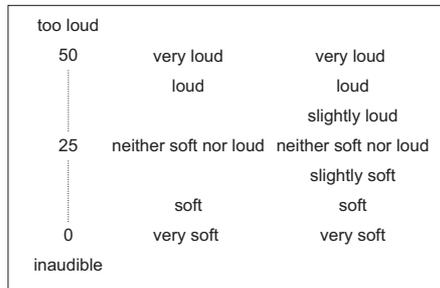
too loud		
50	very loud	very loud
	loud	loud
		slightly loud
25	neither soft nor loud	neither soft nor loud
		slightly soft
	soft	soft
0	very soft	very soft
inaudible		

Fig. 3. Category scaling with seven categories, *right*, five categories, *middle*, and 50 subcategories, *left*

soft" and "slightly loud". For this reason the whole range from "very soft" up to "very loud" shows a finer grading in seven-step scales than it does in five-step ones.

A variant of category scaling which is frequently used in audiology as well as in noise-immission assessment, originates from a five-step scale. However, each step is subdivided into ten subcategories each, such leading to a 50-point scale [17]. The relation between the 50-point scale and the five-category scale is indicated in the left part of Fig. 3. Since the numerical representation of categories may induce a ceiling effect, the categories "inaudible" at the low end and "too loud" at the high end are sometimes added to the 50-point scale. "Inaudible", then, corresponds to zero loudness, while "too loud" may be related to any numbers higher than 50 – see [17]

2.4 Magnitude Estimation

Advantages of the method of "magnitude estimation" are that no ceiling effects shows up and that, theoretically, it has an infinite resolution, e. g., [39]. Magnitude estimation – with anchor sounds – is a frequently used method for sound-quality evaluation. Its procedure is illustrated by means of Fig. 4. Pairs of sounds are presented. The first sound, A, is called "anchor sound" and the second one, B, "test sound". Throughout an experiment the anchor sound is kept constant and the test sound is varied. A numerical value, for instance, 100, is then assigned to a predefined psycho-acoustic quantity of the anchor, A – e. g., to its loudness. The task of the subject, consequently, is to assign a numerical value also to the test sound, B. This value should represent the ratio of the magnitudes of the psycho-physical quantity under observation in the test sound with respect to that in the anchor sound. If, for example, the loudness of a test sound is perceived 20 percent softer than that of the anchor, the subject should give the response 80. Through magnitude estimates a ratio of the magnitudes of psycho-physical quantities is obtained directly, which is often of advantage for cost/benefit analyses.xxx Intra-individual as well as inter-individual differences of magnitude estimates usually come out within
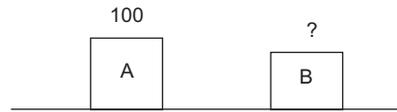
Fig. 4. Illustration of sounds presented for magnitude estimation

10 percent variation. However, sometimes the choice of the anchor sound may influence the results of magnitude estimation significantly. Therefore, it is recommended to use at least two anchor sounds, one with a large magnitude of the psycho-physical quantity in question and the other one with a small magnitude.

Certainly, all psycho-physical methods as mentioned so far have their specific advantages and disadvantages. Random access and the semantic differential are used when a more "qualitative" description is aimed at. If a more "quantitative" assessment of sound quality is the goal, methods like category scaling and magnitude estimation are recommended. They provide data on the level of interval and ratio scales which can easily be processed further with parametric statistics. While traditional category scaling is confined to five or seven response categories, magnitude estimation – in principle – has an infinite resolution and can also provide absolute zeros. However, in magnitude scaling, effects of the frame of reference as well as influences of the choice of the anchor sound(s) have to be taken into account.

3 Modelling of Psychoacoustic Quantities

In sound-quality engineering basic psycho-acoustic quantities like loudness, sharpness, roughness, and fluctuation strength play an important role. Since the evaluation of those quantities in psycho-acoustic experiments can be quite time consuming, models have been proposed which simulate the formation of psycho-acoustic quantities. These models can be used to provide estimates to predict the magnitudes of these quantities from given input data on the physical, i. e. acoustic, level.

3.1 Loudness

As a rule, the "loudness" of a product sound strongly affects the sound quality of the product. Therefore a model of loudness which had been proposed by *Zwicker* already back in 1960 [35] has been improved [37, 38], and been extended in recent years – thereby including its applicability to persons with hearing deficits [3].

The basic features of *Zwicker*'s loudness model, see [39], are illustrated in Fig. 5. Essentially, there are three steps that form the kernel of the *Zwicker* model. In a first step, the physical frequency scale is transformed into the
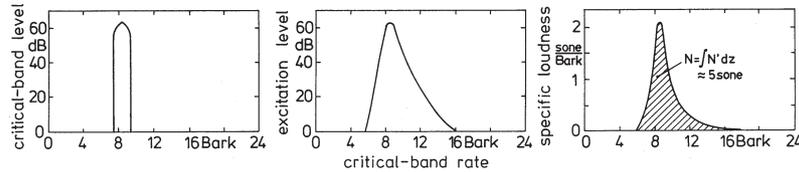
Fig. 5. Illustration of spectral effects as simulated in *Zwicker*'s loudness model [39]

psycho-acoustic *Bark* scale. The name *Bark* was chosen for the unit of this scale to honour the late Professor *Barkhausen* of Dresden for his merits with respect to the basics of loudness measurements. Figure 5 shows, in the left panel, a 1/3-octave-band noise centered at 1 kHz, displayed along the *Bark* scale. In the middle panel of Fig. 5 masking effects are accounted. These masking effects reflect a spectral broadening of the excitation within the cochlea, mainly due to inner ear mechanics. In particular, higher frequencies are masked by lower frequencies, an effect which is nowadays exploited in many practical applications, such as the GSM coding in mobile telephones or *mp*3 coding in consumer electronics. The right panel in Fig. 5 shows a specific loudness/critical-band-rate pattern which is commonly denoted a loudness pattern or "*Zwicker* diagram". Simply speaking, the transition from the masking pattern in the middle of Fig. 5 to the loudness pattern at the right is obtained by taking the square root of sound pressure or the fourth root of sound intensity, respectively. Most important for practical applications is the fact that the area as covered by the loudness pattern, hatched in the figure, is directly proportional to the perceived loudness. This means that with this area being reduced by, say, 30%, it can be predicted that the associated loudness will also be reduced by 30%. This direct proportionality to perceived loudness is unique to this loudness-estimation procedure and cannot be obtained by alternative spectral-analysis systems, such as *Fourier* transforms, 1/3-octave-band analysis, wavelets, gamma-tone filters, etc.

Zwicker's loudness model has been standardized both in international [18] as well as in national [6] standards. In comparison to ISO 532 B of 1975, the latest revision of DIN 45 631, as of 1991, includes improvements with respect to the evaluation of sounds with strong low-frequency components.

Figure 5 illustrates the spectral processing of loudness, while temporal effects of loudness, see [39], are depicted in Fig. 6. The top panel shows the temporal envelope of tone impulses with a 100-ms duration, solid in the figure, or a 10-ms duration, dashed. The middle panel shows the temporal processing of loudness in each of the 24 channels of a loudness analyzer. One can clearly see that the decay of specific loudness is steeper after a short sound is switched off, in comparison to the decay after a longer sound. The lower panel in Fig. 6 shows the time dependency of total loudness, being summed up across all 24 channels. When being presented at the same sound-pressure level, sounds of 100-ms duration give rise to twice the loudness, i. e. 32 sone,
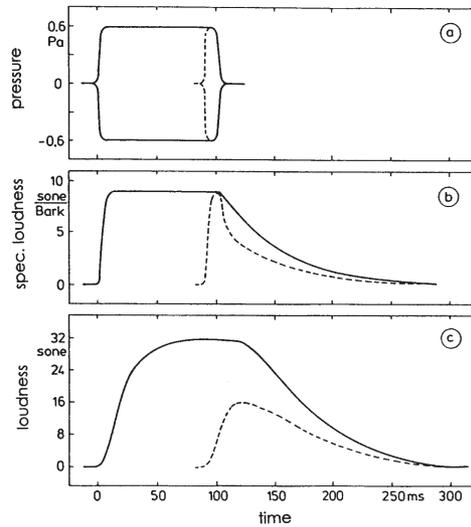
Fig. 6. Illustration of temporal effects in loudness processing [39]

as compared to sounds of 10-ms duration, namely, 16 sone. In contrast to spectral processing of loudness, see Fig. 5, for temporal processing, Fig. 6, it is not the total area under the loudness function, but the peak value of the loudness function which is of relevance.

Figure 7 illustrates an actual implementation [3] of a *Zwicker*-type loudness model. Essentials of spectral processing, as illustrated in Fig. 5, can be found in the critical-band-filter bank, the upward spread of masking and in
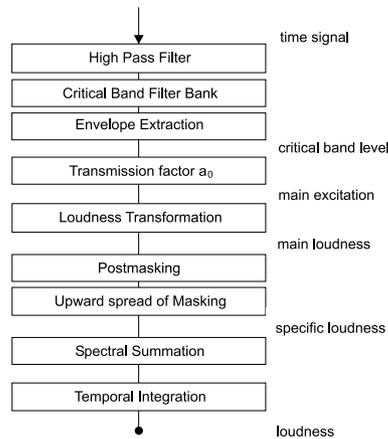

Fig. 7. Block diagram of the dynamic loudness model, DLM, as proposed in [3]

the spectral summation applied. The temporal processing which has been il-
lustrated in Fig. 6 is also represented in Fig. 7, namely, by the blocks marked
envelope extraction, post-masking and temporal integration. Most important
is the block denoted *loudness transformation*. As was discussed already in
connection with Fig. 5 in simplified form, this block represents the fact that
loudness is proportional to the square-root of sound pressure or the fourth
root of sound intensity. A distinguishing feature of the new implementation
– called dynamic loudness model, DLM – is that, by modification of the
loudness-transformation block, loudness perception of both normal-hearing
and hearing-impaired persons can be simulated [3]. This novel feature is of
particular relevance for many practical aspects of sound-quality engineer-
ing, as in the ageing populations of industrialized countries a large part of
prospective customers of a product will show mild to moderate hearing losses.
Further, even a growing percentage of the younger generation has developed
evident hearing deficits these days, frequently due to extremely loud leisure
activities.

Figure 8 provides more details of loudness transformation in normal-
hearing as well as hearing-impaired persons. The dashed curve shows the
relation between level and loudness for normal-hearing persons. The dash-
dotted curve would give the same relation for a person with a 50-dB hearing
loss, provided that processing of loudness in the hearing system were lin-
ear. However, as illustrated by the solid curve in Fig. 8, according to a phe-
nomenon which is known as "recruitment", the following can be observed:
Loudness perception of hearing-impaired people "catches up" at high levels.
This means that for impaired persons the gradient of loudness is very steep
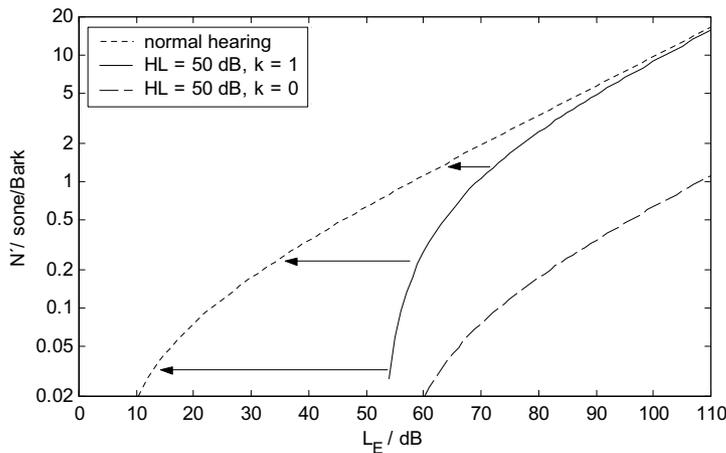

Fig. 8. Relation between level and loudness for normal hearing persons, *dashed*,
and hearing impaired people, *solid*. The *dash-dotted* curve ignores the recruitment
phenomenon [3]

just above the threshold of audibility. The consequence is that a small increase in level of a few Decibels can lead to a drastic increase of loudness, thus causing serious annoyance to the hearing-impaired person. Modern hearing aids try to compensate this effect by digital signal processing. Yet, even the most-advanced hearing instruments can (still) not restore normal hearing completely.

3.2 Sharpness

Besides loudness a further psycho-acoustic quantity, called "sharpness", plays a prominent role in sound quality. Sharpness, among other things, can be regarded as a measure of tone colour [1]. If the right amount of sharpness is added to a sound, e. g., the sound of an industrial product, this will give it a character of powerfulness. However, too much sharpness will render a sound aggressive. If the loudness pattern of a sound is available, its sharpness can be relatively easily estimated by calculation. The corresponding procedure is illustrated in Fig. 9. The left panel depicts the spectral distribution of a narrow-band noise, a broad-band noise and a high-pass noise. The right panel in Fig. 9 shows the loudness pattern as already known from Fig. 5. However, to account for the increased sharpness of high-frequency sounds, a weighting function, g, is to be applied. In order to derive sharpness from the resulting patterns, the first momentum is calculated. The respective values are indicated in the right panel of Fig. 9 by vertical arrows. It becomes clear from Fig. 9 that, when adding low frequencies to a high-pass noise, the center of gravity shifts downwards, thus leading to a smaller value of sharpness – compare the dotted and dashed arrows. This means for practical purposes in sound engineering that the sharpness and, hence, the aggressiveness of product sounds can be reduced by adding low-frequency components.

It should, however, be kept in mind that such an addition of low-frequency components also increases total loudness. Nevertheless, if the loudness of the original sound is not too high, the reduction in sharpness and, hence, aggressiveness can overcompensate the loudness increase in its effect on overall sound quality.
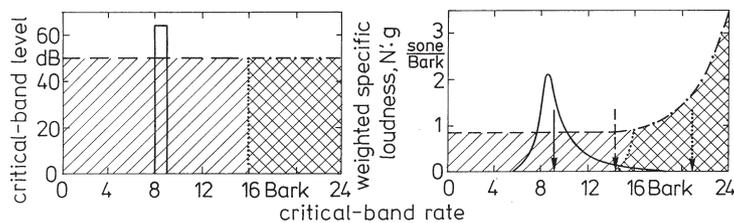

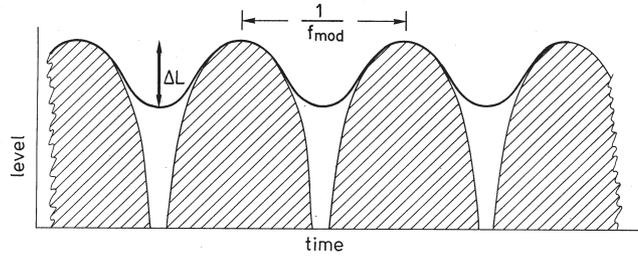
Fig. 9. Illustration of the model of sharpness [39].

Fig. 10. Illustration of the input to the model of roughness [8].

3.3 Roughness

"Roughness", a further psycho-acoustic quantity, is used in sound-quality engineering, e. g., to stress the feature of "sportiness" in a car-engine sound. Roughness is governed by temporal variations of a sound and reaches a maximum for modulation frequencies around 70 Hz [28]. In essence, roughness can be described by the temporal-masking pattern of sounds [8]. This reasoning is illustrated in Fig. 10. The hatched areas show the temporal variation of a sound, modulated in amplitude by a degree of modulation of almost 100%, with the level being displayed as a function of time. Theoretically the troughs between the peaks reach a minimum near minus infinity on the Decibel scale. In practical applications, however, the minimum level is controlled by the dynamics of the hearing system, i. e. the modulation depth of the temporal-masking pattern, ΔL, reaches much smaller values due to the effects of post-masking. Post-masking is represented by the the decay of psycho-acoustic excitation in the hearing system.

   This limited resolution of level is illustrated in Fig. 10 by the solid curve. The temporal distance of the peaks is inversely related to the modulation frequency. In principle, the roughness, R, of a sound can be described by the product of the modulation depth, ΔL, of the temporal masking pattern and the modulation frequency, f_{mod}.

$$R \approx \Delta L \cdot f_{\mathrm{mod}} \tag{1}$$

Since this product carries the unit $[R] = \mathrm{dB/s}$, the hearing sensation roughness is proportional to the speed of the variation of the temporal masking pattern.

3.4 Fluctuation Strength

The psycho-acoustic quantity "fluctuation strength" is similar to roughness. However, fluctuation strength reaches a maximum at modulation frequencies of about 4 Hz. The input to the model of fluctuation strength is the same

as the input to the model of roughness, as displayed in Fig. 10. In addition to the modulation depth, ΔL, of the temporal-masking pattern, the relation of the modulation frequency to a modulation frequency, f_{mod}, of 4 Hz is of relevance. Therefore fluctuation strength, F, can basically be calculated as follows:

$$F \approx \frac{\Delta L}{4\mathrm{Hz}/f_{\mathrm{mod}} + f_{\mathrm{mod}}/4\mathrm{Hz}} \tag{2}$$

Fluctuation strength plays a crucial role in the assessment of human speech for the following reason. The envelope fluctuation of fluent speech also shows a maximum around a modulation frequency of 4 Hz. This roughly corresponds to the number of syllables pronounced per second. As one would have expected from nature, the human speech organ indeed produces speech sounds with dominant envelope fluctuations at a rate that the human hearing system is most sensitive to.

3.5 Composed Metrics

A combination of psycho-acoustic quantities has proven successful for the prediction of the perceived annoyance of sounds from noise emissions as well as immissions [32]. The corresponding formula for this annoyance estimator, PA, reads as follows.

$$PA \approx N_5 \cdot (1 + \sqrt{w_{\mathrm{S}}^2 + w_{\mathrm{FR}}^2}) \tag{3}$$

with
$N_5 \ldots$ percentile loudness in sone

$w_{\mathrm{S}} = (\frac{S}{\mathrm{acum}} - 1.75) \cdot 0.25 \lg(\frac{N_5}{\mathrm{sone}} + 10)$ for $S > 1.75$ acum

describing the effects of sharpness, S, and

$w_{\mathrm{FR}} = \frac{2.18}{(N_5/\mathrm{sone})^{0.4}} (0.4 \cdot \frac{F}{\mathrm{vacil}} + 0.6 \cdot \frac{R}{\mathrm{asper}})$

describing the effects of fluctuation strength, F, and roughness, R.

The units acum, vacil and asper are related to sharpness, S, fluctuation strength, F, and roughness, R, respectively. For details of their definition the reader is referred to [39]. From the formula it becomes clear that loudness is a dominant feature of annoyance. The percentile value, N_5, indicates that a value near the maximum loudness is of importance for sound quality ratings. However, sharpness as well as roughness and fluctuation strength may play an important role as well. When thinking, for example, of a dentist's drill, not only the loudness but, even more so, the sharpness is responsible for the

annoyance. Likewise, the tick-tack of a clock, in particular when heard during the night, is usually not annoying because of its loudness but because of the regularity and, hence, the fluctuation strength of the sound produced.

Although the model of psycho-acoustic annoyance proposed in [33] can account for various practical situations, it is certainly not designed to solve all questions of sound quality. Nevertheless, this model contains some relevant ingredients for sound-quality evaluation, namely, loudness, sharpness, roughness and fluctuation strength. However, the appropriate "recipe" for a mixture of psycho-acoustic quantities may vary for different families of product sounds and different applicational context.

Another composite model based on several psycho-acoustic quantities, as being put forward in [29], has proven successful to rate "sensory pleasantness" of sounds – in particular, the pleasantness of speech and music [30]. However, in this model, clearly audible tonal components receive a bonus, while, when dealing with noise-immission problems, tonal components are rather undesired and thus usually are assigned a penalty instead. Consequently, this model of sensory pleasantness is not recommended to estimate pleasantness related to noise immissions.

4 Sound Quality

Since this book contains a further chapter dealing with sound quality [19], we restrict ourselves here to a small selection of practical sound-quality-evaluation examples from our own laboratory.

Due to the electronics being implemented on modern road vehicles, it is now comparatively easy to control and adjust the engine's state of operation [25]. In our first example, displayed in Fig. 11, a *Diesel* motor was driven in either a "hard" or "normal" manner. Normal, here, means the adjustment as used in the current product. The advantage of the hard motor adjustment is that the engine is more fuel efficient. Obviously, the disadvantage is that it produces more noise. In psycho-acoustic experiments it was assessed whether the acoustic disadvantage of the fuel efficient hard motor adjustment can be reduced by absorptive measures. In one part of the study, frequencies from 1 kHz to 5 kHz where attenuated by 3 to 15 dB in 3 dB steps. In the other set of experiments, the whole spectrum was attenuated by 3 to 15 dB, again in 3 dB steps. The data displayed in Fig. 11 show the ranking of sound quality in medians, circles, as well as the inter-quartiles, bars. The crosses denote loudness predictions from acoustic measurements.

The results, as displayed in Fig. 11, show clearly that the motor with a "hard" motor adjustment obtains the poorest sound quality ranking, i. e. rank 12. However, the motor with the "normal" motor adjustment – as used in series vehicles – attains rank 4 in sound quality. Even better ranks, namely 1 to 3, result when the whole spectrum of the hard motor sound is attenuated by 9 to 15 dB.
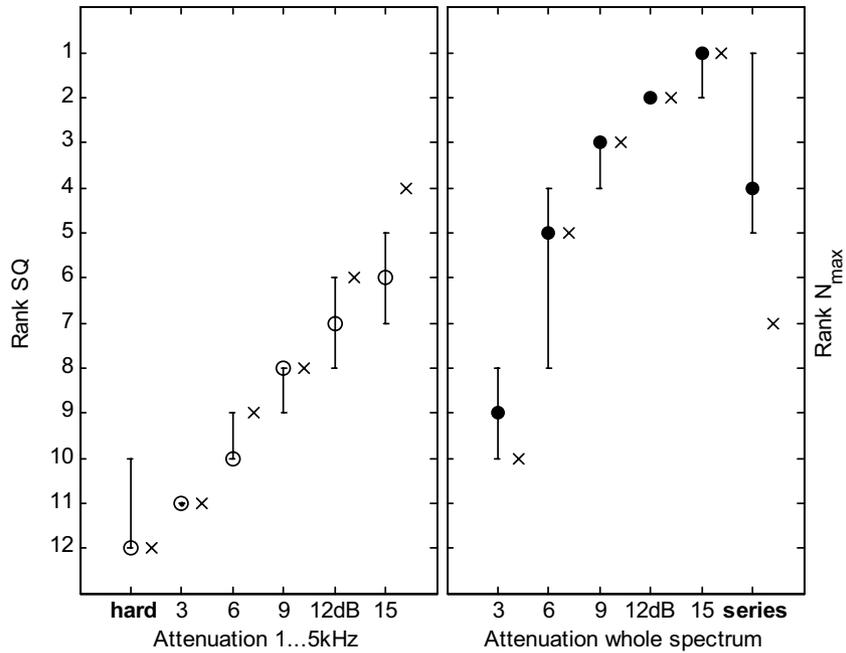
Fig. 11. Improvement of the sound quality of a *Diesel* motor with "hard" motor adjustment and simulated absorptive measures which cause different amounts of attenuation. *Circles* denote subjective sound quality estimates. *Crosses* mark loudness predicted from acoustic measurement [25]

The results depicted in Fig. 11 suggest that even when the sound of a hard motor adjustment is attenuated in the frequency range of 1 to 5 kHz by as much as 15 dB, the sound quality of a "normal" motor adjustment is still not achieved. Rather the complete spectrum of the motor sound for hard motor adjustment would have to be attenuated by as much as about 7.5 dB to just attain the sound quality of a present-day-series vehicle. Although it will not at all be easy, it seems quite worthwhile for engineers to take on this challenge of reducing the sound level, because of the higher fuel efficiency of the motor with hard motor adjustment.

The crosses displayed in Fig. 11 indicate the ranking of the physically-measured maximum loudness, $N_{\max}$, produced by each sound. As a rule there is good agreement of the sound-quality ranking and the ranking of the maximum loudness. However, for an attenuation of 15 dB between 1 and 5 kHz sound quality attains only a rank of 6 while the ranking of maximum loudness attains rank 4. On the contrary, for the series motor, the ranking in loudness attains only rank 7, whereas the sound quality attains rank 4. This means that loudness alone cannot always predict sound-quality ratings. In spite of its larger loudness, rank 7, the sound quality of the motor of the series vehicle
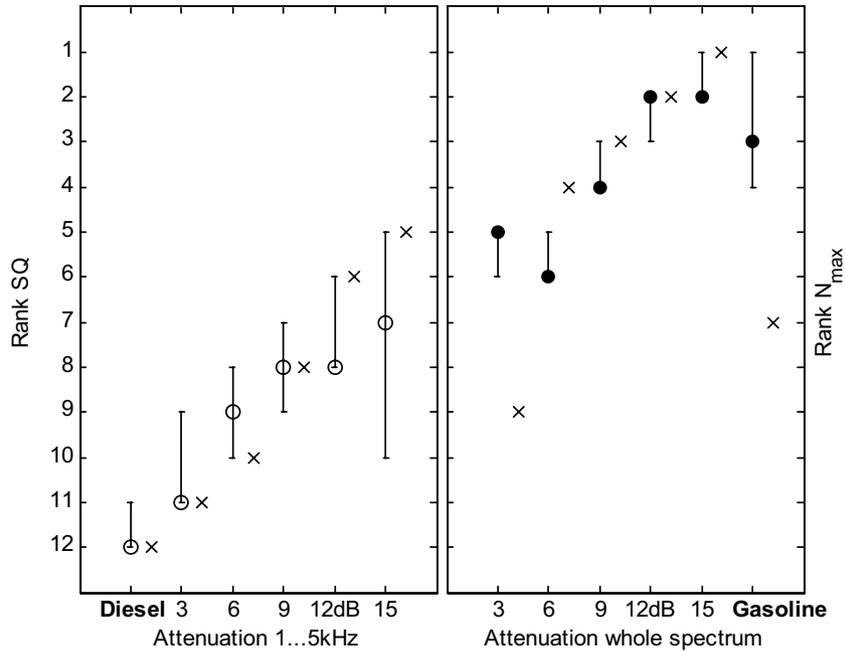
Fig. 12. Comparison in sound quality between a gasoline engine and a "normal" *Diesel* engine with additional absorptive measures applied [25]

is ranked higher in sound quality, rank 4. This is presumably due to its more familiar, "natural" sound. A similar argument holds true for the hard motor when attenuated by 15 dB in the frequency range from 1 to 5 kHz. Although the loudness, then, is significantly reduced, rank 4, the sound quality is still ranked lower, rank 6 – presumably because the resulting stimulus sounds quite "unnatural".

An even more striking example is given in Fig. 12. In this case, a series *Diesel* engine with the same attenuations as described above has been compared in sound quality to a gasoline engine [25]. Again circles indicate subjective rankings, crosses rankings estimated from acoustic measurement.

Again, the physical ratings, crosses in the figure, and the subjective evaluations are frequently in line - at least within the inter-quartile ranges. However, clear differences are obtained for the gasoline engine. While its loudness attains only rank 7, the sound quality attains rank 3. Despite its higher loudness, the sound of the gasoline engine is preferred. Presumably this is due to a particular sound attribute of *Diesel* engines which can be characterized by the term "*Diesel* nailing". The large discrepancy between sound-quality ranking and ranking of loudness for the sound with a 3-dB attenuation of the whole spectrum is not clear. Obviously it is not easy for the subjects to evaluate those sounds since - contrary to expectation - the sound which is only

3 dB attenuated gets a better ranking than the sound with 6 dB attenuation. Starting from a series *Diesel* engine, an attenuation of the whole spectrum of about 10 dB would be necessary to arrive at a sound quality similar to that of a gasoline engine.

5 The Meaning of Sounds

When evaluating sound quality, the meaning as assigned to sounds when listening to them may have an effect on judgements. In a global market it may thus be of some relevance to take possible cultural differences into account. Cross-cultural studies with listeners in Japan and Germany [21] showed that sometimes one and the same sound can be rated differently by subjects from different cultural backgrounds. For example, by German subjects, the sound of a bell was interpreted as the sound of a church-bell, leading to connotations such as "pleasant" or "safe". On the contrary, Japanese subjects were reminded by the bell sounds to sounds of a fire engine or a railroad crossing, leading to feelings as denoted by the terms "dangerous" or "unpleasant". In Fig. 13 [21] the corresponding data from a study with the method of semantic differential are displayed. Data for Japanese subjects are connected by solid lines, data of German ones by dotted lines.

The data for the German subjects suggest their feelings that the bell-sound is not frightening, but pleasant, safe, attractive, relaxed, and pleasing. Japanese subjects, however, feel that the bell sound is shrill, frightening, unpleasant, dangerous, exciting, busy, repulsive, distinct, strong, tense, and unpleasing.
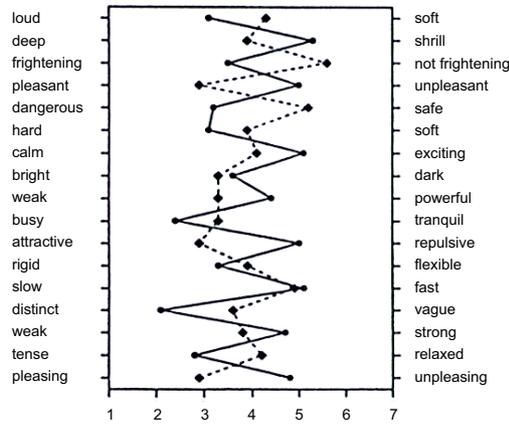

Fig. 13. Semantic-differential data for a bell sound. Data for Japanese subjects are connected by *solid lines*, for German ones by *dotted lines* – adopted [21]
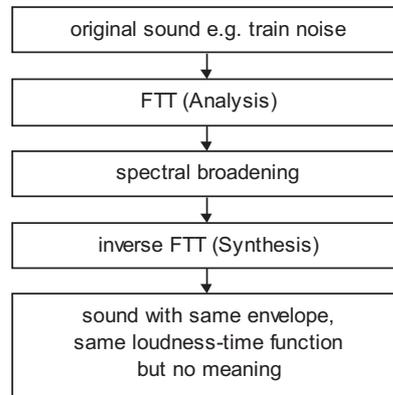
Fig. 14. Block diagram illustrating the procedure to remove information about the sound source from a stimulus [13]

In order to overcome undesired influences of meaning in psycho-acoustic experiments a procedure has been proposed [13] which largely removes the information about the sound source from a stimulus. The block-diagram displayed in Fig. 14 illustrates the correlated procedure. From an original noise, e. g., train noise, a spectral analysis is being performed by means of a *Fourier*-time transform, FTT. The FTT algorithm [31] is a spectral-analysis technique which, in contrast to, e. g., *Fourier* transforms, uses a sliding temporal window corresponding to a frequency-dependent bandwidth which mimics the the frequency resolution of the human hearing system. In the next steps, after spectral broadening and, hence, obscuring the spectral details, the sound is re-synthesized by means of an inverse FTT. In this way a sound with the same spectral and temporal envelope and, such, the same loudness/time function is created from which, however, information about the sound source has been removed.

The data displayed in Fig. 15 enable a comparison of the loudness-time functions of (a) original sounds, (b) the same sounds, but with the information about the sound source being removed.

The results as displayed in Fig. 15 clearly show that the goal of ending at identical loudness-time functions of original sounds and sounds without information about the sound source can well be achieved. With the procedure as outlined in Fig. 14 the information about the sound source can be removed from many signals which are important in our daily life [34]. However, some signals like for example FTT-processed speech sounds still have great similarity to the original sound. It is worth mentioning at this point, that algorithms to remove meaning from sound have also been proposed in speech technology to study prosody – "re-iterant speech", see, e. g., [24] for details.
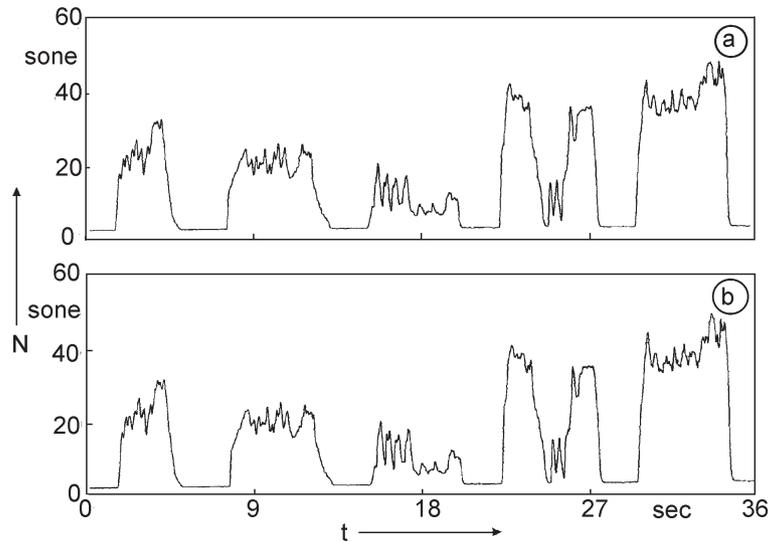
Fig. 15. Loudness-time function of (**a**) original sounds, and (**b**) the same sounds, but with the information about the sound source being removed by the procedure illustrated in Fig. 14 [13]

6 Image of Brand Names

When evaluating the sound quality of passenger cars, the image of the brand name of the cars has been shown to be of relevance. A well-known typical example along these lines is that the quality of a car is judged on the basis of the sound produced when closing a door. If the door sound is "tinny", this suggests that the whole vehicle is cheap and not at all solid. On the contrary, a full and saturated door-closing sound has a connotation of luxury. In a cooperation with colleagues from Osaka University Japan door-closing sounds have been studied with the method of the semantic differential. In addition to evaluating sound quality our subjects have been asked to guess the type of car and to guess the corresponding brand name of the car. Figure 16 shows data which are an excerpt of a larger study [23]. Data are given for a door sound which was allocated to a luxurious sedan, as well as for a door sound judged to stem from an economy sedan.

The data displayed in Fig. 16 suggest that the door sounds of luxury sedans are best described by adjectives such as deep, pleasant, heavy, dull, dark, powerful, calm, smooth, pleasing. On the contrary, adjectives related to the door sounds of economy sedans are best characterized as metallic, unpleasant, noisy, bright, shrill, rough, unpleasing.

Figure 17 gives an example of a histogram of the number of guesses associated to different brand names for luxurious sedans – based on door-closing
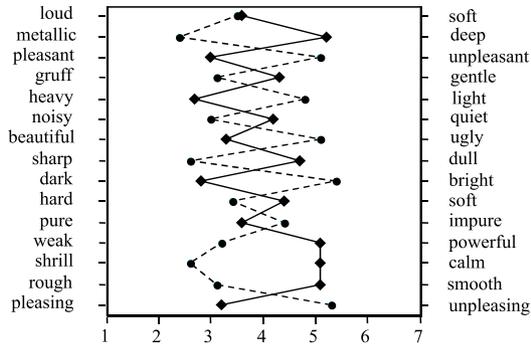
Fig. 16. Semantic differential for car door sounds judged to stem from a luxurious sedan, *solid*, or an economy sedan, *dashed* [23]
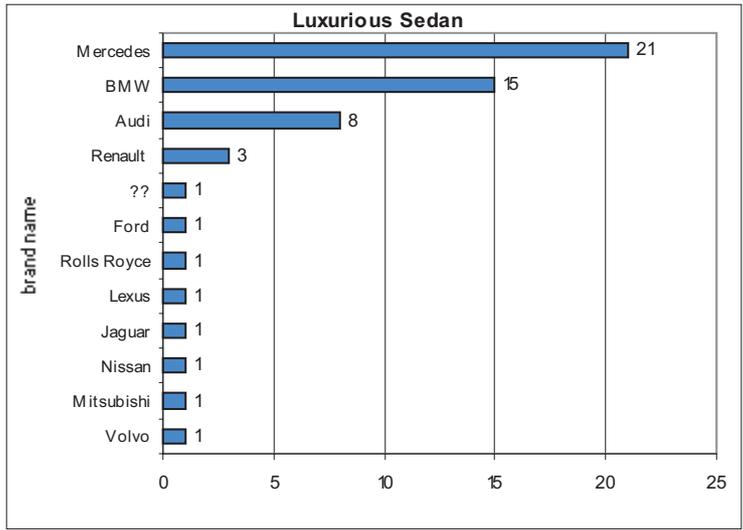

Fig. 17. Distribution of brand names associated with car door sounds of luxurious sedans [7]

sounds. [7]. The twenty German subjects clearly related luxurious sedans to brand names like Mercedes, BMW or Audi.

The German Automobile Association, ADAC, regularly publishes a ranking of car manufactures. In our context it very interesting to compare this ranking of manufactures with the rating of car-door sounds as rendered from subjects in psycho-acoustic experiments [7]. The results are given in Table 1. When regarding these results, it is very interesting to note the strong relation of the ranking of car manufactures by the ADAC to the rating of the sounds

Table 1. Ranking of car manufactures by ADAC and rating of brand names by subjects for the different categories on the basis of sounds of closing car doors [7]

Ranking of car manufacturers	Rating 1-4 in each class by subjects
1. Mercedes	Luxurious 1
2. BMW	Sporty 1, Luxurious 2, Others 3
3. Audi	Luxurious 3
4. Volkswagen	Economy 1, Pick up 3, Others 4
5. Porsche	Sporty 2
6. Toyota	Economy 4
7. Peugeot	-
8. Smart	-
9. Renault	Luxurious 4
10. Ford	Economy 3, Pick up 4
11. Opel	Economy 2
12. Skoda	Others 2

of closing doors of cars. For example, the brand name Mercedes, being ranked first by the ADAC, gets the highest rating in the category of luxurious cars. BMW, which is ranked second by the ADAC, gets the best rating for sporty cars in the psycho-acoustic experiment, the second best for luxurious cars and the third best for others. Audi, number three in the ADAC ranking, gets the third rating of luxurious cars and so forth. Obviously, the brand name of a car strongly triggers the expectations about the sounds produced by a closing door.

7 Audio–Visual Interactions

Sound-quality ratings may depend not only on auditory stimuli but on input from other senses as well, for instance, from the visual system. In the following, two examples to support this view will be given.

The first deals with the influence of a visual image on the sound-quality rating of speech. In a concert hall, speech was radiated from the stage and recorded at different positions in the hall. In a first experiment subjects just listened to the recorded speech and rated its speech quality. In a further experiment, in addition to the acoustic presentation of the speech sounds, subjects where presented photos taken at the respective recording position, depicting the distance between the source and the receiving point. Fig. 18 shows a schematic plan of the ground floor, denoting the sound-source, S, and three positions, 1 through 3, of the receiver. The sounds as sent out by the speaker, S, where recorded on DAT tape at the positions 1, 2, or 3. In addition, photos where taken at these positions, showing the speaker on the stage and enabling the listeners to evaluate their distance from the receiving point.
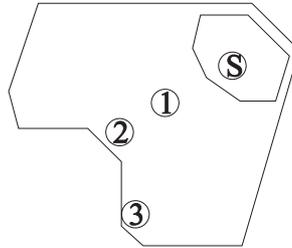
Fig. 18. Schematic plan of the ground floor of a concert hall with indications of the sound source, *S*, and three positions, 1 through 3, of the receiver [15]

Figure 19 gives the ratings of the speech quality for acoustic presentation alone, unfilled symbols in the figure, or with additional visual presentation, filled symbols. The data show that a visual image can influence the rated sound quality of speech. At position 1, which is relatively close to the source, the addition of the visual image causes the rating to degrade from fair to poor, medians taken. This may be due to the effect that the visually perceived distance to the speaker, which is relatively small, calls for a better speech quality which is not degraded by reverberation as in a concert hall. Since the concert hall was, of course, designed for classical music and, hence, has a reverberation time at mid-frequencies around 2 seconds, this reverberation is much too large for speech. For best intelligibility of speech a reverberation time below 1 second would be optimal [5]. At position 2, there is no influence of the visual image on the rating of speech quality. Obviously the subjects think that for such a larger distance from the speaker the quality is fair. Most interesting is the rating at position 3. Without visual information the speech quality is rated fair. However, with additional visual information the speech quality is even rated good. Obviously, given the large distance between the speaker, *S*, and the receiving point, 3, the subjects feel that for such an adverse situation the speech quality can be rated as being relatively good.
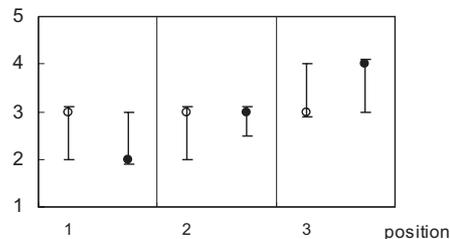

Fig. 19. Rating of speech quality in a concert hall at positions 1, 2, and 3 for acoustic presentation alone, *unfilled symbols*, and acoustic plus visual presentation, *filled symbols* [15]
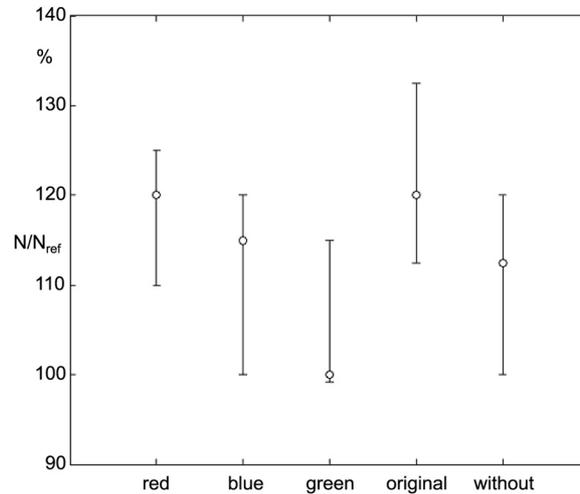
Fig. 20. Loudness rating of the sound from a passing train when photos of the train in different color are presented together with the sound [26]

Our second and last example deals with the influence of colour on the rating of the loudness of sound sources [26]. Sounds of a passing train have been presented either without visual input or with pictures of the same train, but painted in different colours. The related data are displayed in Fig. 20. Despite identical acoustic stimuli, the train sound is perceived as being softest when the train is painted in a light green. The loudness rating of this setting is taken as a reference in the further course of the experiment. According to the results displayed in Fig. 20, a train painted in red is perceived as being 20 percent louder than the same train in green.

The authentic painting of the train – a German high-speed, ICE, train – is white with a red stripe. In this case the relative loudness reached is also 120 percent of the reference. If the train sound is presented without visual input, it is perceived as somewhat softer than the presentation of sound plus original image – very similar to a train painted in light blue. In summary then, the colour of a product can indeed influence the loudness and, hence, the quality of its sound to some extent. Comparable cross-modal influences have been shown before for other kinds of visual cues [16, 20], and other auditory quality features – see, e. g., [27].

8 Outlook

The application of psycho-acoustic principles to sound engineering and sound-quality design has only recently become accepted as a useful analysis and

design method. Although a solid psycho-acoustic basis had been available for quite a while, e. g., [36], applications of psycho-acoustics to noise and product-sound evaluation used to be rather sparse in the past (as noted, e. g., in [2]). Yet, since roughly a little more than a decade, application of knowledge from psycho-acoustics, e. g., [11], or even from musical acoustics, e. g., [15], has increased substantially in this context. Among other reasons, this situation is driven by economic necessities. In a global market with many competing products having almost the same functionality, the sound attached to a product can well become a decisive quality feature. Moreover, from the quality of the sound that a product produces, the user may extrapolate to the quality of the whole product – more or less consciously. It is thus to be expected that the application of psycho-acoustics in sound-quality evaluation and design will further increase.

Acknowledgment

References

1. von Bismarck G (1974) Sharpness as an attribute of the timbre of steady sounds. Acustica 30:159–172
2. Blauert J (1986) Cognitive and aesthetic aspects of noise engineering. Proc Internoise'86 vol I:5–14
3. Chalupper J, Fastl H (2002) Dynamic loudness model. ACUSTICA/acta acustica 88:378–386
4. Cederlöf R, Jonsson R E, Kajland A (1981) Annoyance reactions to noise and from motor vehicles – an experimental study. Acustica 13:270–279
5. Cremer L, Müller H A (1978) Die wissenschaftlichen Grundlagen der Raumakustik. Hirzel, Stuttgart
6. DIN 45631 (1991) Berechnung des Lautstärkepegels und der Lautheit aus dem Geräuschspektrum. Verfahren nach Zwicker E
7. Filippou T G, Fastl H, Kuwano S, Namba S, Nakamura S, Uchida H (2003) Door sound and image of cars. Fortschr. Akust. DAGA'03, Dt. Ges. f. Akustik, Oldenburg
8. Fastl H (1982) Beschreibung dynamischer Hörempfindungen anhand von Mithörschwellen-Mustern. Hochschul-Verlag, Freiburg
9. Fastl H (1987) A background noise for speech audiometry. J Audiol Acoust. 26:2–13
10. Fastl H, Kuwano S, Namba S (1996) Assessing the railway-bonus in laboratory studies. J Acoust Soc Japan (E) 17:139–148
11. Fastl H (1998) Psychoacoustics and sound quality metrics. In: Davies P, Ebbitt G (eds) Proc Sound-Quality Symp 3-10, Ypsilanti MI

12. Fastl H (2000) Sound quality of electric razors - Effects of Loudness. Proc Internoise 2000, Nizza, CD-ROM
13. Fastl H. (2001) Neutralizing the meaning of sound for sound-quality evaluations. In: Proc 17th Int Congr Acoust, ICA, Rome, CD ROM
14. Fastl H (2002a) Psychoacustics and sound quality. Fortschr Akust, DAGA'02, 765–766. Dt Ges f Akustik, Oldenburg
15. Fastl H (2002b) Sound design of machines from a musical perspective. In: Davies P, Ebbitt G (eds) Proc Sound-Quality Symp, Dearborn MI, CD-ROM
16. Fastl H (2004) Audio–visual interactions in loudnesss evaluation. 18th Int Congr Acoust, ICA, Kyoto, 1161-1166
17. Hellbrück J (1993) Hören. Hogrefe, Göttingen
18. ISO 532B (1975) Method for calculating loudness level
19. Jekosch U (2004) Assigning meaning to sounds: semotics in the context of product-sound design. Chap 8 this vol
20. Kaska J , Noack R, Mau U, Maas P (1985) Untersuchung der Wirkung von visuellen Umweltmerkmalen des Strassenbildes auf die Lästigkeit von Verkehrslärm an schweizer und deutschen Anwohnern. In: A Schick (ed) Rep 4th Oldenburg Symp Psychoacoust, Univ Oldenburg, Oldenburg
21. Kuwano S, Namba S, Fastl H, Schick A (1997) Evaluation of the impression of danger signals - comparison between Japanese and German subjects. In: Schick A, Klatte M (eds) Contributions to Psychological Acoustics. 115–128. BIS, Oldenburg
22. Kuwano S, Namba S, Schick A, Hoege H, Fastl H, Filippou T, Florentine M, Muesch H (2000) The timbre and annoyance of auditory warning signals in different countries. In: Proc. Internoise 2000, Nizza, CD ROM
23. Kuwano S, Fastl H, Namba S, Nakamura S, Uchida H (2002) Subjective evaluation of car door sound. In: Proc. Sound-Quality Sym., Dearborn MI
24. Mersdorf J (2000) Sprecherspezifische Parametrisierung von Sprachgrundfrequenzverläufen. Doct diss, Ruhr-Univ Bochum, Shaker, Aachen
25. Patsouras C, Fastl H, Patsouras D, Pfaffelhuber K (2002a) How far is the sound quality of a Diesel powered car away from that of a gasoline powered one? Proc Forum Acusticum Sevilla 2002, NOI-07-018-IP, CD-ROM
26. Patsouras Ch, Filippou T G, Fastl H (2002b) Influences of colour on the loudness judgement. Proc Forum Acusticum Sevilla 2002, PSY-05-002-IP, CD-ROM
27. Steven H (1981) Subjektive Beurteilung von Geräuschemissionen von Lastkraftwagen. In: Forsch Inst Geräusche Erschütterungen Aachen (ed) Report 105-05-104/02, Umweltbundesamt, Berlin
28. Terhardt E (1968) Über akustische Rauhigkeit und Schwankungsstärke. Acustica 20:215–224
29. Terhardt E (1984) Wohlklang und Lärm aus psychophysikalischer Sicht. In: Schick A, Walcher KP (eds) Ergebnisse des 3. Oldenburger Symposions zur Psychologischen Akustik. 403–409. P Lang AG, Bern,
30. Terhardt E (1985) Fourier transformation of time signals: Conceptual revision. Acustica 57:242–256
31. Terhardt E (1987) Audiokommunikation, Springer, Berlin-Heidelberg
32. Widmann U (1995) Subjektive Beurteilung der Lautheit und der Psychoakustischen Lästigkeit von PKW-Geräuschen. Fortschr Akust, DAGA'95, 875–878, Dt Ges Akust, Oldenburg,
33. Widmann U (1998) Aurally adequate avaluation of Sounds. In: Fastl H, Scheuren J (eds) Proc. Euro-Noise'98. 29–46, Dt Ges Akust, Oldenburg

34. Zeitler A, Fastl H, Ellermeier W (2004) Impact of source identifiability on perceived loudness. 18th Int Congr Acoust, ICA, Kyoto, 1491–1494
35. Zwicker E (1960) Ein Verfahren zur Berechnung der Lautstärke. Acustica 10:304–308
36. Zwicker E (1982) Psychoakustik. Hochschultext, Springer, Berlin
37. Zwicker E, Fastl H (1983) A portable loudness meter based on ISO 532 B. In: Proc. 11th Int Congr. Acoustics, ICA, vol 8:135–137, Paris
38. Zwicker E, Deuter K, Peisl W. (1985) Loudness meters based on ISO 532 B, with large dynamic range. Proc Internoise'85 vol II:1119–1122
39. Zwicker E, Fastl H (1999) Psychoacoustics. Facts and models. 2nd ed, Springer, Berlin–Heidelberg–New York