

Proceedings of Meetings on Acoustics

Volume 19, 2013

<http://acousticalsociety.org/>



ICA 2013 Montreal

Montreal, Canada

2 - 7 June 2013

Interdisciplinary

Session 2aID: Plenary Lecture: Basics and Applications of Psychoacoustics

2aID1. Basics and applications of psychoacoustics

Hugo Fastl*

***Corresponding author's address: TU München, AG Technische Akustik, MMK, Arcisstr.21, München, 80333, Bavaria, Germany, fastl@mmk.ei.tum.de**

The field of psychoacoustics studies relations between acoustic stimuli, defined in the physical domain, and the hearing sensations elicited by these stimuli. The lecture will address questions of stimulus generation and presentation. Both traditional methods and more recent methods like wave field synthesis will be touched. Concerning hearing sensations, basic magnitudes as for example absolute thresholds or loudness, but also advanced topics like pitch strength will be covered. Applications of psychoacoustics will include rather different fields such as music, noise evaluation or audiology, and address also cognitive effects as well as audio-visual interactions.

Published by the Acoustical Society of America through the American Institute of Physics

CONCEPT OF PSYCHOACOUSTICS

The goal of psychoacoustics is to establish relations between stimuli, defined in the physical domain, and the related hearing sensations. This concept is illustrated by means of figure 1.

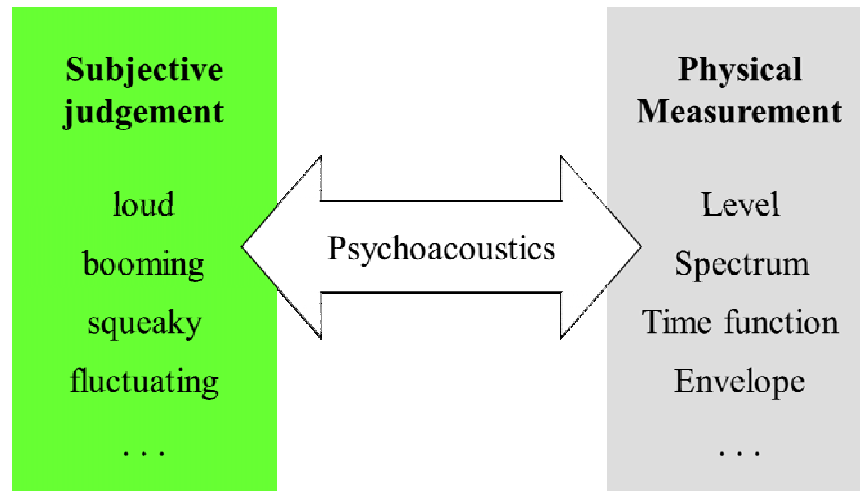


FIGURE 1. Concept of psychoacoustics acting like a bridge between the world of physical measurement on the one side, and the world of subjective judgement on the other side.

Engineers frequently describe acoustic phenomena in physical terms like level, spectrum, time function etc., whereas psychologists may describe sounds in perceptual terms as loud, booming, fluctuating etc. Psychoacoustics acts like a bridge between these two separated worlds. The engineering approach of psychoacoustics eventually aims to develop algorithms which when fed with physical magnitudes of sounds allow for an “average person” a prediction of hearing sensations elicited by these sounds (e.g. Stevens and Davies 1938, Zwicker and Feldtkeller 1967, Blauert 1997, Terhardt 1998, Hartmann 1998, Hellbrück and Ellermeier 2004, Fastl and Zwicker 2007, Moore 2012).

STIMULI AND SENSATIONS

Synthesis, Recording and Presentation of Sounds

Synthesis of Sounds

These days, synthetic stimuli for psychoacoustic experiments frequently are generated using MATLAB. For D/A-conversion at least 44.1 kHz sampling frequency and 16 bit uniform amplitude coding are used which represents the traditional Compact Disc (CD) format. For the succeeding product of the CD, a dichotomy shows up: on the one hand, high sampling frequencies up to 192 kHz and amplitude coding by 24 bit are used, e.g. for DVD-Audio with bit rates around 1.4 Mbit/s. On the other hand, however, using perceptual coding (e.g. Blauert and Tritthart 1975, Krahe 1985, Theile et al. 1987, Stoll 1993, Brandenburg 1988, Painter and Spanias 1997, van de Par et al. 2008) in implementations like MP3 or AAC extremely low bit rates of 32 kbit/s and even lower are used with acceptable results for specific applications. While 16 bit with about 96 dB dynamic range are somewhat limited in psychoacoustics leaving no head- and footroom, 24 bit with about 144 dB dynamic range are not so easily attained

in the analog domain. However, in many cases for psychoacoustic experiments stimuli with “true” 120 dB dynamic range (i.e. some 20 bit resolution) are sufficient.

Recording of Sounds

When natural or technical sounds are recorded for psychoacoustic experiments by a (condenser) measurement microphone, the result is not much influenced by the brand of microphone used. However, for recordings with artificial heads, the brand applied can have a distinct influence on the sound obtained. As an example, in figure 2 free-field responses are given for five artificial heads, which are frequently applied these days in psychoacoustic experiments, both for basic research and practical applications.

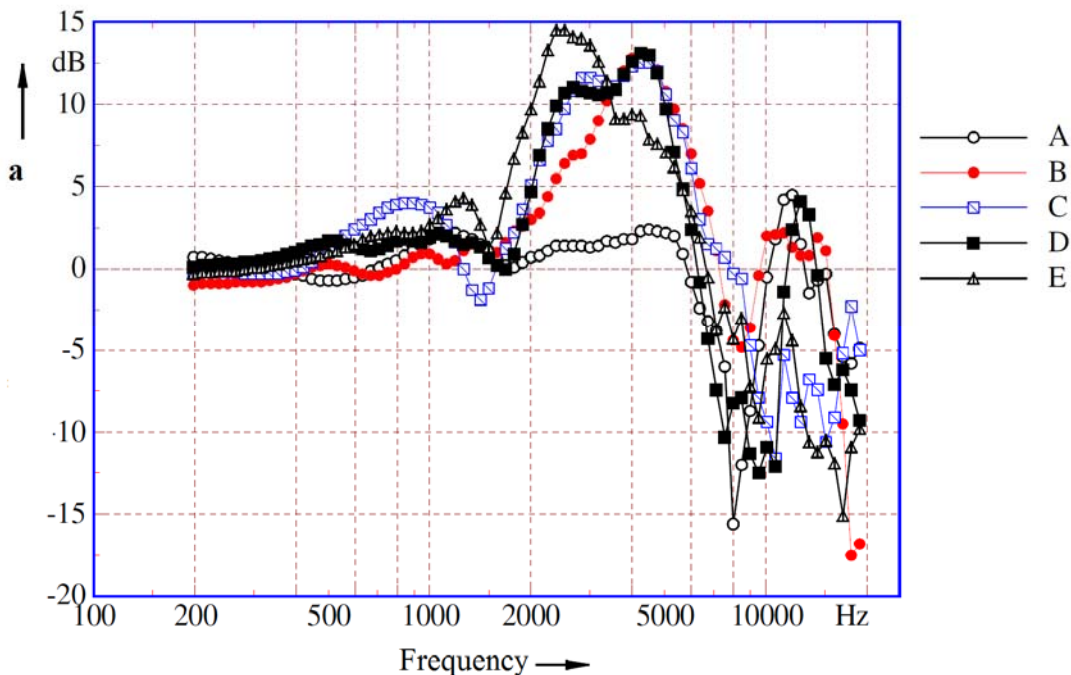


FIGURE 2. Free-field responses of five artificial heads, frequently used in psychoacoustic experiments for both basic research and practical applications (after Daniel et al. 2007).

The results displayed in figure 2 illustrate that one artificial head (circles) has fixed diffuse-field equalization and therefore differs substantially from the responses of the other four products. However, also the four artificial heads with free-field equalization differ in detail in their frequency response, in particular around the resonance of the outer ear and at high frequencies above 7 kHz. These differences, which significantly influence the timbre of sounds, have to be remembered when evaluating sounds recorded by different artificial heads. One step towards a solution of this problem would be, if all manufacturers of artificial heads would use the same shell (cf. Fastl 2004, Genuit and Fiebig 2007).

Presentation of Sounds

In “classic” psychoacoustics, sounds frequently were presented via headphones. In comparison to loudspeaker presentation this has the advantage that room influences are avoided and low distortions of typically only 0.1 % show up (Haar 1951). On the other hand, in the head localization occur with headphones (e.g. Blauert 1997), and “classic” headphones act like bandpass-filters on the sounds. As an example, figure 3 shows the frequency response

of electrodynamic headphones as measured by loudness balances (e.g. Zwicker and Maiwald 1963, Fastl and Fleischer 1978, Fastl and Zwicker 1983).

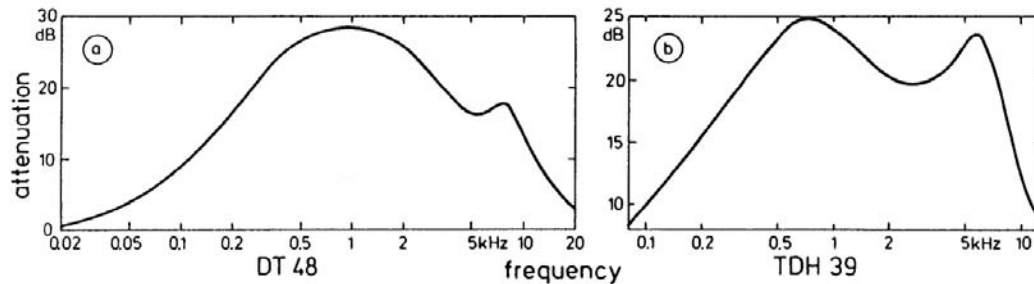


FIGURE 3. Free-field responses of two headphones, used in “classic” psychoacoustic experiments as measured by loudness balances (after Fastl and Zwicker 2007).

Since the frequency responses displayed in figure 3 suggest strong attenuation at both low and high frequencies, usually these headphones are used in combination with an appropriate equalizer (Zwicker and Maiwald 1963, Fastl and Zwicker 1983). When using these headphones without equalizer, rather misleading results may be obtained for broadband sounds (e.g. Hellman and Zwicker 1987).

The problems of in the head localization can be largely overcome when using personalized head related transfer functions (HRTF). A procedure which has proven successful for such an approach is the (physical) measurement of sounds at the entrance of the blocked ear canal by tiny electret microphones (Møller 1992, Hammershøi and Hoffmann 2011, Völk 2011). To arrive at optimum results, not all headphones are equally well suited for this application. As an example, figure 4 shows the frequency response of two high quality headphones in such an application (Völk 2012).

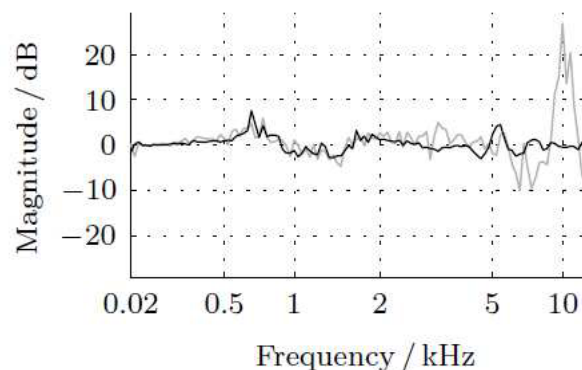


FIGURE 4. Frequency response of two high quality headphones, when measured by tiny electret microphones at the entrance of the blocked ear canal (after Völk 2012).

The results displayed in figure 4 suggest that both headphones perform rather similar for frequencies up to about 2 kHz. However, at high frequencies above 8 kHz large differences in excess of 20 dB can show up. Therefore, it is advisable to check whether even high quality headphones are suitable for optimum sound presentation when for the equalization the procedure with blocked ear canals is used (cf. Völk 2012).

Sometimes wearing headphones can be unpleasant or headphone presentation can be impracticable, e.g. in combination with hearing aids. Therefore, using a kind of wave field synthesis (e.g. Berkhout 1988, Wittek 2004), in cooperation with the German Broadcast Research Lab (IRT), a “virtual headphone” was developed and realized as displayed in figure 5 (Menzel et al. 2005, 2006, Theile 2007).



FIGURE 5. "Virtual headphone": Ring of 22 loudspeakers plus a woofer to project focused sources onto the ears using a kind of wave field synthesis (after Menzel et al. 2005).

The results displayed in the left panel of figure 6 suggest that the "virtual headphone" faithfully reproduces the azimuth in the horizontal plane. However, the right panel illustrates that there is room for improvement concerning the elevation. All virtual sources are perceived higher than intended. This effect is extremely pronounced for sources below ear level like source number 3.

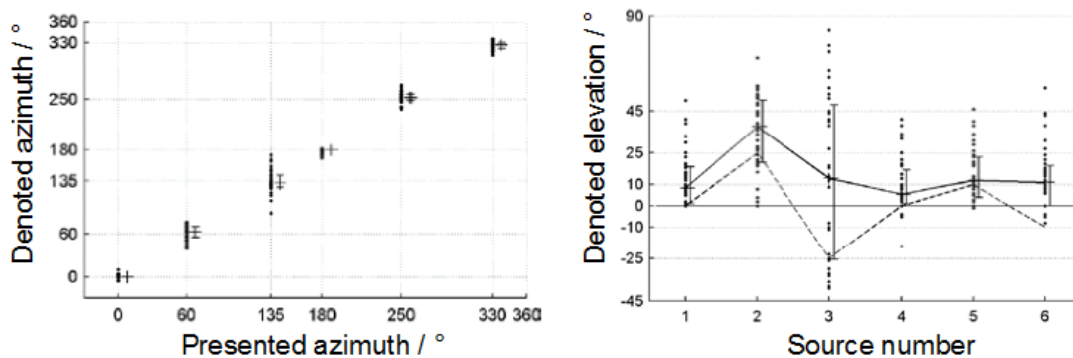


FIGURE 6. Reproduction of intended versus heard azimuth (left) and elevation (right) by the "virtual headphone" illustrated in figure 5 (after Menzel et al. 2006).

Hearing Sensations

In this paragraph, hearing thresholds of applicants for the profession of Tonmeister as well as a "classic" and a "modern" hearing sensation, i.e. loudness and pitch strength, respectively, will be discussed.

Hearing Thresholds

Figure 7 shows the hearing thresholds of applicants for the profession of sound recording engineer (Tonmeister), collected in cooperation with the Bavarian Broadcasting Corporation (BR). The hearing loss is given as a function of frequency (Völk et al. 2008).

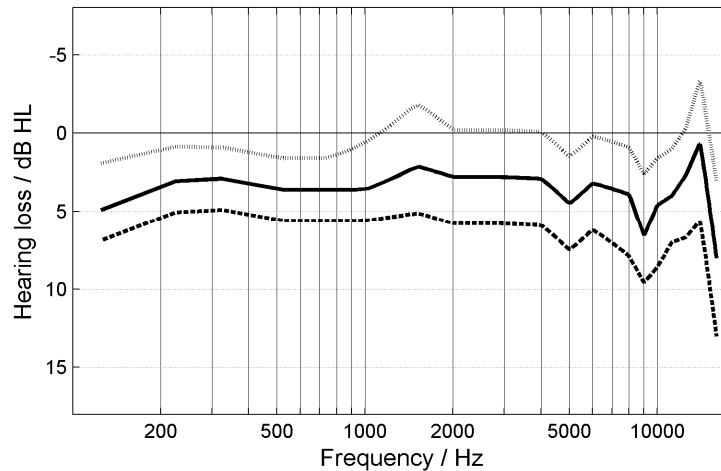


FIGURE 7. Hearing thresholds of applicants for the profession of sound recording engineer (Tonmeister) (after Völk et al. 2008).

In contrast to expectations, the data displayed in figure 7 suggest that applicants for the profession of Tonmeister on the average show a hearing loss of almost 5 dB. Since Tonmeisters usually are considered as having “golden ears”, negative values of hearing loss, i.e. better hearing than average would have been expected. The reason for this discrepancy is not clear: On the one hand, the standard values for absolute thresholds might shift with the years towards higher values, and/or on the other hand, the applicants might have acquired slight hearing loss, possibly by listening for long hours to loud music.

Loudness

Psychoacoustic experiments on loudness perception have a long tradition (e.g. Fletcher and Munson 1933, Stevens and Davies 1938, Zwicker 1958), and algorithms to simulate loudness perception for stationary sounds have been standardized already. Pertinent examples are DIN 45631 from 1991 and ANSI S3.4 from 2007. For the reference sound, i.e. a pure tone at 1 kHz, both standards produce almost identical results, but predictions for pure tones at other frequencies can differ substantially. Perhaps even more important, for broad band sounds, which are more relevant in practical applications, significant differences show up (e.g. Fastl et al. 2009, Sottek 2010).

As an example, figure 8 shows the loudness of pink noise as measured in psychoacoustic experiments, in comparison to predictions by DIN 45631 as well as ANSI S3.4 (Schlittenlacher et al. 2011).

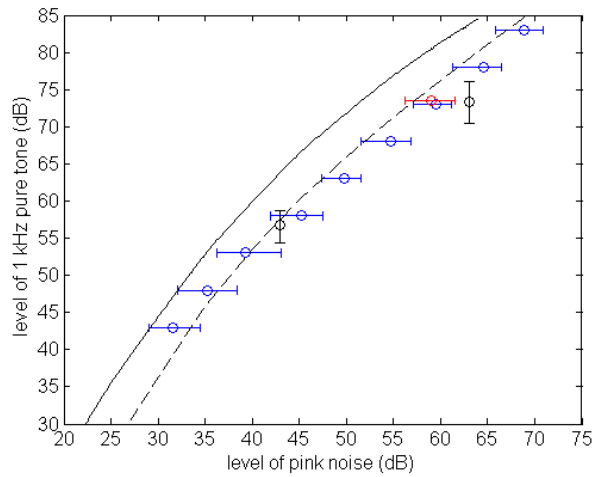


FIGURE 8. Loudness evaluation of pink noise in psychoacoustic experiments (symbols) in comparison to predictions by DIN 45631 (dashed) and ANSI S3.4 (solid) (after Schlittenlacher et al. 2011).

Regarding the results displayed in figure 8 it becomes clear that both standards mimic the curved dependence of loudness on the level of pink noise. The predictions from DIN 45631 (dashed) are somewhat closer to the subjective data than the values from ANSI S3.4 (solid) which usually predicts too high loudness. Interestingly, when the data from ANSI S3.4 are shifted by 5 dB towards higher levels, both standards predict almost the same values for the loudness of pink noise on its level (see Fastl et al. 2009), in line with subjective evaluation.

To predict the loudness of time varying sounds, several algorithms have been proposed (e.g. Zwicker 1977, Moore et al. 1997, Chalupper and Fastl 2002, Rennie et al. 2010) and a corresponding German standard (DIN 45631/A1) was published in 2010. Most algorithms for calculation of loudness proposed so far essentially are based on “classic” work by Fletcher and Munson (1933) or Zwicker (1958, 1977).

As an example, figure 9 shows the flow diagram of a dynamic loudness model (DLM) proposed by Chalupper and Fastl (2002)

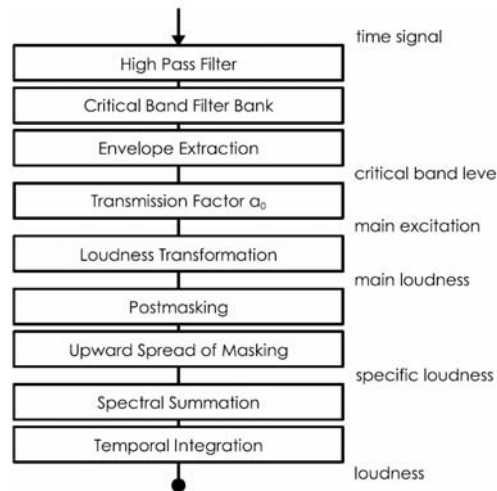


FIGURE 9. Flow diagram of a dynamic loudness model (DLM) proposed by Chalupper and Fastl (2002).

The loudness model illustrated in figure 9 is a typical Zwicker-Type-Model with critical band filters, post masking, upward spread of masking, spectral summation, and temporal integration. An interesting feature of the model is that it can be used to predict loudness perception of persons with normal hearing as well as persons with (slight) hearing deficits. In order to switch from normal hearing to impaired hearing only the block “loudness transformation” has to be changed somewhat to account for growth of loudness in impaired hearing (cf. Buus and Florentine 2001).

The data plotted in figure 10 enable a comparison of loudness measured in psychoacoustic experiments by category units (CU) (cf. Hellbrück and Ellermeier 2004) and predictions by the dynamic loudness model.

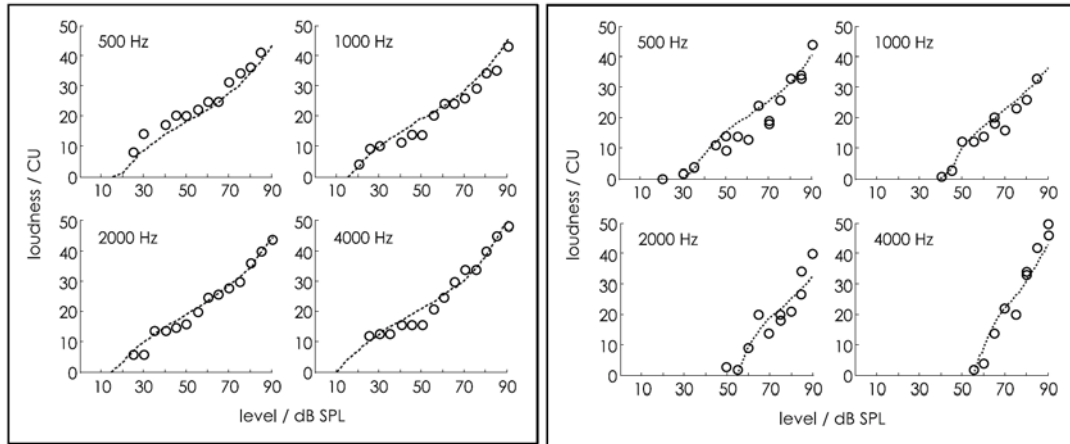


FIGURE 10. Comparison of loudness measured in psychoacoustic experiments by category units (CU, circles), and related predictions by the dynamic loudness model (curves). Left: Normal hearing subject. Right: Subject with hearing deficit at high frequencies (after Fastl and Zwicker 2007).

The results displayed in figure 10 suggest that the dynamic loudness model (DLM) can quantitatively simulate loudness scaling by normal hearing subjects as well as subjects with hearing deficits.

Pitch Strength

Traditionally, pitch is evaluated along a scale extending from low to high. However, at same pitch height, sounds can differ in the strength of their pitch. For example, a low-pass noise with 1 kHz cut-off and steep spectral skirts may elicit a faint pitch sensation (e.g. Fastl 1971) compared to the strong, distinct pitch produced by a pure tone at 1 kHz. While pitch height has been studied since at least 2 500 years (e.g. Pythagoras in ancient Greece), systematic studies on pitch strength started less than half a century ago (e.g. Rakowski 1977, Yost and Hill 1978, Fastl and Stoll 1979). By and large it can be stated that – in comparison to the relative pitch strength of 100% elicited by pure tones – sounds with line spectra can produce about 50 to 90% pitch strength, and stochastic signals up to about 20 % pitch strength. However, very narrow noise bands (e.g. 10 Hz bandwidth) can also lead to rather strong pitches around 80% relative pitch strength (cf. Fastl and Zwicker 2007).

Figure 11 illustrates the model of pitch strength proposed by Fruhmann (2006) for the example of a complex tone with 500 Hz fundamental frequency.

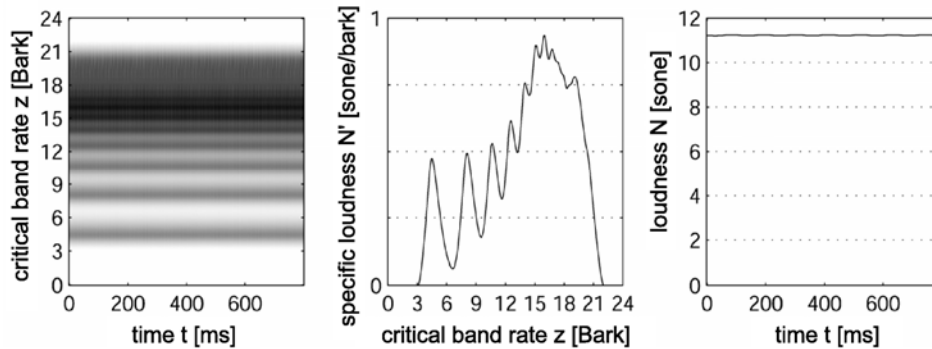


FIGURE 11. Illustration of the model of pitch strength for a complex tone with 500 Hz fundamental frequency (after Fruhmann 2006).

The left panel in figure 11 illustrates the first step in Fruhmann's model of pitch strength, i.e. a spectral analysis using the Fourier Time Transform (FTT, Terhardt 1985). The spectral lines at 5 Bark (500 Hz), 8.5 Bark (1000 Hz) etc are clearly visible. Before the next step – the transformation into a loudness pattern – the sound is divided into a “pitch component” and “noise components”. In the middle panel of figure 11, the pitch component would correspond to the peak at 5 Bark, and the other peaks would be considered as “noise”, reducing the pitch strength of the 500 Hz-pitch-component. The right panel in figure 11 illustrates the third step in the model, i.e. the loudness-time-function of the complex tone, which is steady state, as expected. If, however, fluctuations show up, this means that pitch strength is reduced. For more detail about the model the reader is referred to the original publication by Fruhmann (2006).

Results plotted in figure 12 enable a comparison of data from psychoacoustic experiments on the pitch strength of narrow band noise with different bandwidth and center frequency, and predictions by the model of Fruhmann (2006).

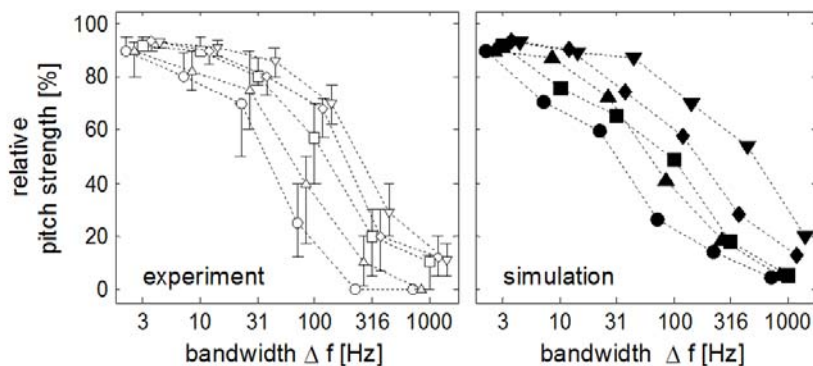


FIGURE 12. Relative pitch strength of narrow band noise centered at 250 Hz (circles), 500 Hz (upward triangles), 1 kHz (squares), 2 kHz (rhombs), and 4 kHz (downward triangles) as a function of bandwidth. Left: Data from psychoacoustic experiments. Right: Corresponding model data. (after Fruhmann 2006).

The data shown in figure 12 reveal that pitch strength decreases with increasing bandwidth of narrow band noise (cf. Wiesmann and Fastl 1991). At the same bandwidth, e.g. 100 Hz, narrow band noise centered at 4 kHz produces a pitch strength which is about a factor of three larger than the pitch strength produced by narrow band noise centered at 250 Hz. The model of Fruhmann generally can predict the dependence of pitch strength produced by narrow band noise on both bandwidth and center frequency. For periodic sounds like pure or complex tones, the predictions by the model usually are within the inter-quartile ranges of the psychoacoustic results (Fruhmann 2006). Since tonal noises can be pretty annoying, Fruhmann's model of pitch strength is of relevance not only for basic science (e.g. Fingerhuth and Parizet 2008), but also for practical applications (e.g. More and Davies 2010).

PRACTICAL APPLICATIONS OF PSYCHOACOUSTICS

In this paragraph examples for practical applications of psychoacoustics will be discussed for music, audiology, noise evaluation, cognitive effects as well as audio-visual interactions.

The optimum Tuning Standard: 432 Hz vs. 440 Hz

As an example for the application of psychoacoustics in music, a quarrel about the “optimum” tuning standard shall be discussed. These days, the official tuning standard for a_4 is fixed at 440 Hz, although many famous orchestras use higher tuning up to some 445 Hz. In contrast, a rather active group of music lovers in Germany (Weng 2001) advocates to use a lower tuning as for example 432 Hz. At first sight, the arguments may be regarded striking, since at the time when most classic music was composed and originally performed, the tuning standard in fact was below 440 Hz.

In order to contribute some facts to the discussion, psychoacoustic experiments were performed in cooperation with the Deutsches Museum, a famous science museum in Munich. Their collection comprises a Welte-Steinway grand piano and a library of rolls of punched paper, which enable to re-invoke performances of famous artists like Claude Debussy, Edwin Fischer, Max Reger to name just a few.

For the experiments, recordings with an artificial head were performed after about two weeks of settling time when the instrument was tuned to 440 Hz or 432 Hz, respectively. Table 1 gives an overview of the artists and the music presented.

TABLE 1. Overview of the artists and music excerpts used in the experiments on the “optimum” tuning standard.

Sound	Title	Composer	Artist
A	Romanze E-moll	F. Chopin	Wilhelm Backhaus
B	Scenes from Tiefland	Eugen d’Albert	Eugen d’Albert
C	Parita B-Dur No. 1	J. S. Bach	Walter Guicking
D	Preludes	Claude Debussy	Claude Debussy
E	Fantasia c-moll	W. A. Mozart	Edwin Fischer
F	Silhouette D-Dur	Max Reger	Max Reger

Subjects had to indicate their ranking using the method of random access (Fastl 2000) when presented the recorded piano music via headphones (Fastl et al. 2003). Results are given in figure 13.

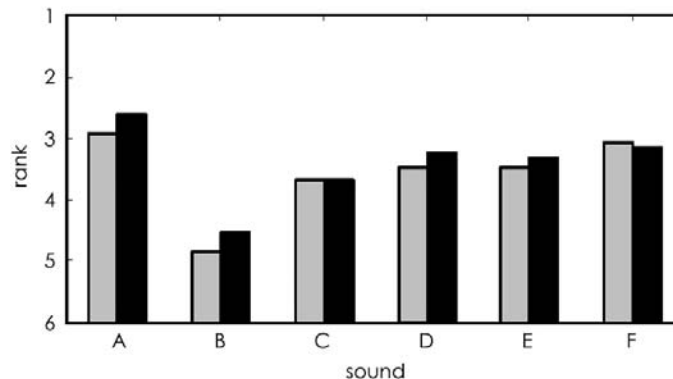


FIGURE 13. Preference for music excerpts as listed in Table 1 when presented in 432 Hz vs. 440 Hz tuning of the Welte-Steinway grand piano. Black columns: 440 Hz tuning, grey columns: 432 Hz tuning (after Fastl and Zwicker 2007).

The results displayed in figure 13 allow no clear cut decision, whether a tuning to 432 Hz or 440 Hz should be preferred. An argument put forward by the advocates of 432 Hz tuning is that our subjects were students who are familiar with the 440 Hz tuning used these days. One reason for utilizing 440 Hz or even higher tuning standards may be that the sound of the instruments gets more “brilliant” which these days is preferred by most Tonmeisters responsible for the timbre of current music recordings.

Cochlear Implants

As an example for the application of psychoacoustics in audiology, the rehabilitation of hearing deficits by cochlear implants shall be discussed. From early attempts to elicit by electric impulses hearing sensations (e.g. Simmons et al. 1965) cochlear implants have developed to the adequate medium for rehabilitation of profound hearing loss (e.g. Clark 2003). Given the extreme reduction of input features compared to normal hearing, it is astonishing how well “star-patients” perform tasks of speech understanding or localization.

Figure 14 shows as an example localization performance of a “star-patient” fitted with two cochlear implants (Seeber et al. 2004).

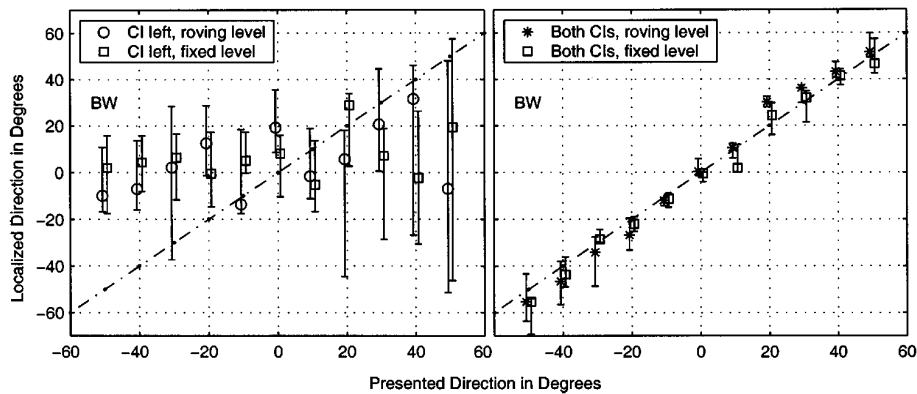


FIGURE 14. Localization ability of a “star-patient” fitted with two cochlear implants. Left: only one implant active; right: both implants active (after Seeber et al. 2004).

As expected, the results displayed in figure 14 reveal that with only one cochlear implant (left panel) localization is practically impossible. However, when both implants are activated (right panel), the localization ability is almost as good as for normal hearing persons (cf. Henning 1975)!

Moreover, results displayed in figure 15 indicate that “star-patients” show almost normal localization ability not only in “real” free-field conditions (asterisks), but also in sound fields realized by virtual acoustics (Seeber and Fastl 2008).

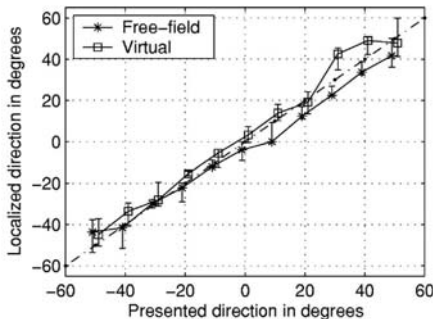


FIGURE 15. Localization ability of a “star-patient” fitted with two cochlear implants in a “real” free-field (asterisks) or a sound field realized by virtual acoustics (squares) (after Seeber and Fastl 2008).

Traditionally, for persons with hearing deficits, speech in noise is tested with stationary noise (e.g. Kollmeier 1996). However, in order to assess also temporal processing, a fluctuating noise has been proposed (Fastl 1987), which simulates on the average spectral and temporal features of fluent speech as illustrated in figure 16.

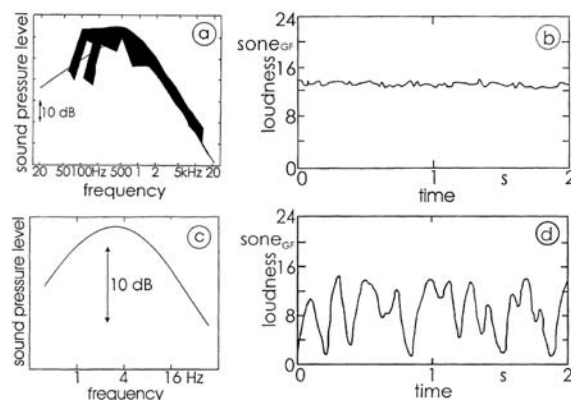


FIGURE 16. Background noises for speech audiometry. (a) Spectrum of carrier signal. (b) Loudness-time-function of stationary noise according to CCITT Rec. G 227, (c) Spectral distribution of modulation signal, (d) Loudness-time-function of fluctuating noise (after Fastl and Zwicker 2007).

The modulation signal displayed in figure 16c with a maximum at 4 Hz mimics the amplitude modulation in fluent speech at “normal” speaking rate with about four syllables per second. It is also in line with the dependence of the hearing sensation fluctuation strength on modulation frequency (cf. Fastl and Zwicker 2007).

Results plotted in figure 17 allow a comparison of speech intelligibility by normal hearing persons and cochlear implant patients when presented stationary versus fluctuating background noise.

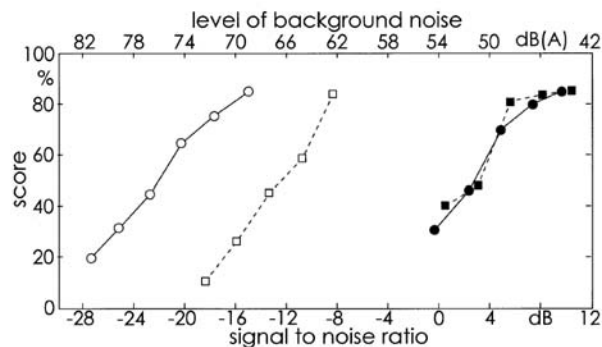


FIGURE 17. Speech intelligibility for sentences in stationary (squares) vs. fluctuating (circles) background noise. Data for normal hearing persons (open symbols) vs. cochlear implant patients (filled symbols) (after Fastl and Zwicker 2007).

The data displayed in figure 17 reveal that the normal hearing persons can profit from the “valleys” in fluctuating noise by “listening into the dips” whereas the cochlear implant patients lack this advantage and need the same speech level for stationary and fluctuating background noise (Fastl et al. 1998, Jespersgaard and Dau 2012).

More recently, cochlear implants called EAS are used which stimulate high frequencies electrically, but low frequencies acoustically. They seem to have advantages in perception of music (Rader et al. 2011) as well as speech in noise. Figure 18 shows some results for speech in stationary noise (OLnoise and CCITT-noise) as well as fluctuating noise (Fastl-noise). Data are given for speech and noise from the same direction (S0N0) as well as speech in front and noise from four surrounding loudspeakers (MSNF) (Rader et al 2009).

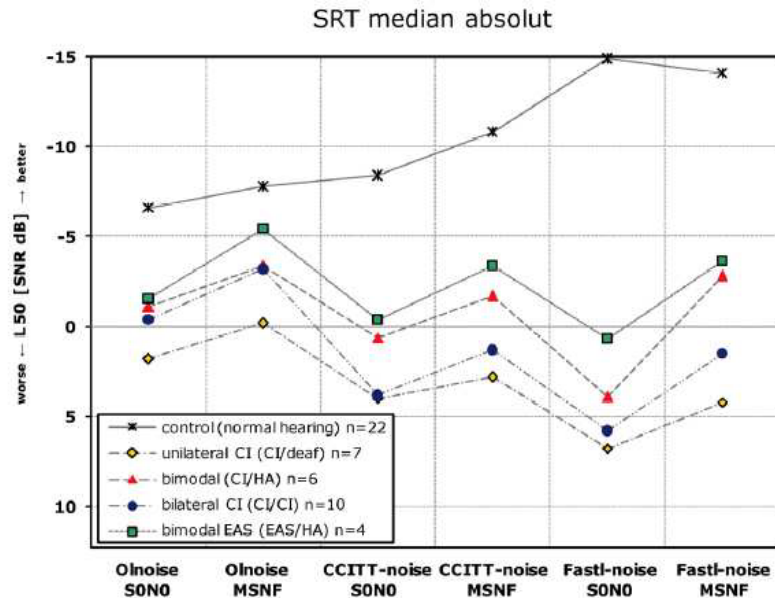


FIGURE 18. Speech in noise for normal hearing persons (stars) as well as patients fitted with different kinds of cochlear implants. Stationary (OLnoise, CCITT-noise) or fluctuating noise (Fastl-noise). Speech and noise from the same direction (S0N0) or speech from front loudspeaker and noise from four surrounding loudspeakers (MSNF) (after Rader et al 2009).

The results displayed in figure 18 reveal again that normal hearing persons – but not cochlear implant patients – can profit from dips in the fluctuating noise. All cochlear implant patients perform worse than the normal hearing persons. However, EAS patients seem to perform better than the other groups of patients (e.g. Rader et al. 2012).

Level-Thermometer versus Loudness-Thermometer

In order to explain to the broad public the meaning of dB-values, a level-thermometer has been publicized (Brüel and Kjaer 1984). As illustrated in the right part of figure 19, sounds are ordered with respect to their A-weighted level, analog to the temperature readings on a thermometer. As shown in the left panel of figure 19, it is proposed to use a similar display for the loudness values in sone, called a loudness-thermometer (Fastl et al. 2006) or sone-meter (Fastl and Florentine 2011).

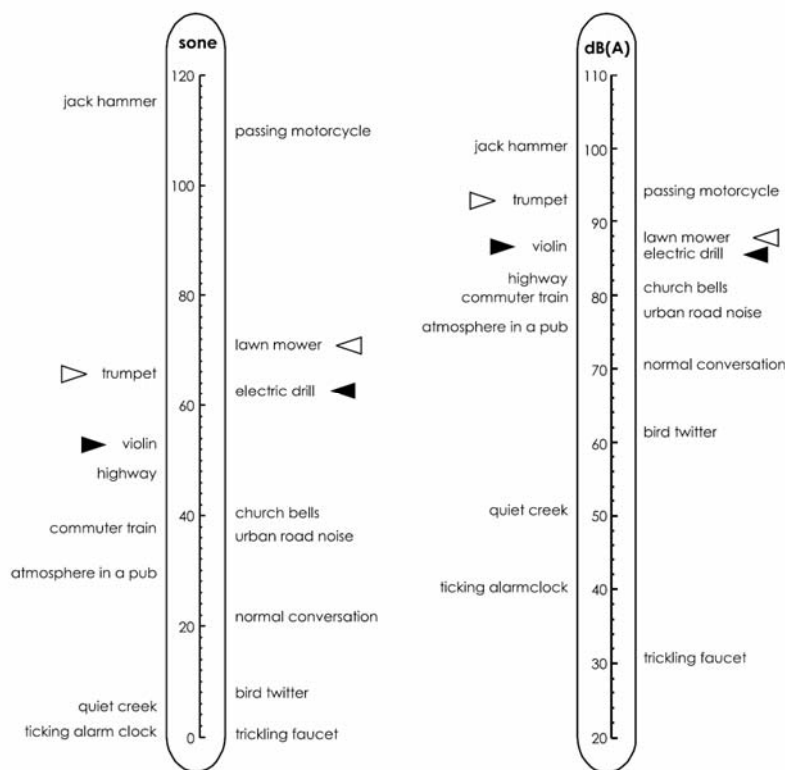


FIGURE 19. Level-Thermometer (left) and Loudness-Thermometer or sone-meter (right) illustrating typical dB(A)-values or sone values for everyday sounds (after Fastl et al. 2006).

Although the level-thermometer gives a first indication about the meaning of dB(A)-values of sounds, because of the logarithmic magnitude, the numbers can be misleading for the general public. Of course, experts know that 60 dB(A) is not twice as much as 30 dB(A) in power, pressure, loudness or what have you. In contrast, the sone-values directly correspond to loudness perceived by an average person. For example, indeed, 40 sone is twice as loud as 20 sone. Another advantage of the sone scale is that it takes into account differences in timbre. For example, as illustrated in figure 19 by the filled arrows, at same dB(A)-value, an electric drill is perceived as being louder than a violin. On the other hand, as illustrated by the unfilled arrows, a lawn mower can be *louder* than a trumpet despite a *lower* dB(A)-value.

Noise Limits for Electric Vehicles (EV) or Hybrid Cars

While for many decades the goal of sound engineering in transportation noise has been to reduce the sounds (e.g. Genuit 2010), with the advent of hybrid and electric vehicles another challenge has to be faced: Electric vehicles or hybrid vehicles in electric mode can be so soft that they can be dangerous for pedestrians. Unfortunately, even fatalities of blind people have to be mourned, who did not hear early enough approaching vehicles in electric mode. Therefore, psychoacoustic studies were performed with the goal to arrive at proposals for sound features necessary for soft vehicles to be heard early enough to prevent collisions between e.g. vehicles and pedestrians (Kerber 2008). To this end, the concept of “critical distance” has been introduced, which is defined as the minimum distance necessary for safe collision avoidance.

The critical distance depends on the speed of the vehicle and the reaction time of both driver and pedestrian. Figure 20 gives the critical distance as a function of driving speed for different reaction times.

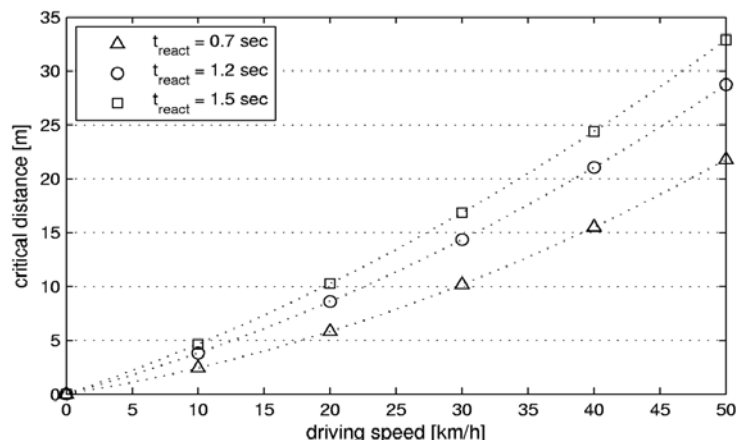


FIGURE 20. Critical distance as a function of driving speed for different reaction times (after Kerber and Fastl 2007).

At a speed of 30 km/h, typical for traffic-calmed areas in Europe, for alert driver and pedestrian and 0.7 s reaction time, the critical distance amounts to 10 m. This means that the pedestrian has to hear the vehicle when it is 10 m apart that a collision can be avoided. If however, the pedestrian is distracted, e.g. by using the mobile phone, the reaction time is increased to typically 1.5 s, and the vehicles has to be heard at a distance of about 17 m for collision avoidance.

Of course, the background noise and the sound produced by the vehicle crucially influence the distance at which a vehicle can be perceived. In essence, in psychoacoustic terms this is a classic masking experiment with the vehicle sound as test signal and the background noise as masker. In addition, the speed of the vehicle is of importance to assess the best case, i.e. alert driver and pedestrian. The results of corresponding psychoacoustic experiments can be predicted by a model illustrated in figure 21.

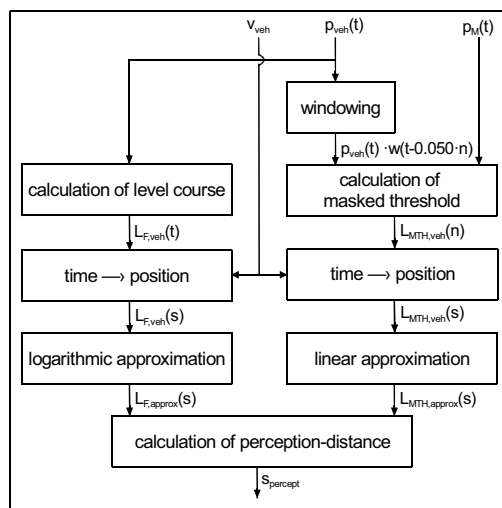


FIGURE 21. Model to predict the distance between vehicle and pedestrian necessary to avoid collision (after Kerber 2008).

The inputs to the model are the speed v_{veh} and sound pressure time-function $p_{veh}(t)$ of the vehicle as well as the sound pressure time-function $p_M(t)$ of the (masking) background noise. Using a 200 ms window, the time course of masked threshold is calculated and with the aid of the speed v_{veh} transformed in positions of critical distance. For masked threshold values a linear approximation, and for the dependence of level on distance a logarithmic approximation are used.

Figure 22 enables the comparison of psychoacoustically measured data and data calculated according to the model illustrated in figure 21.

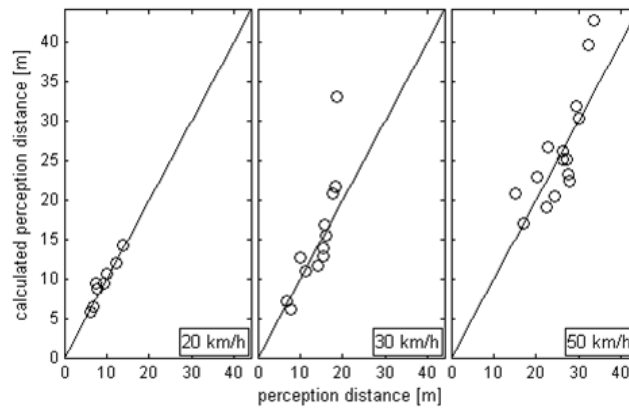


FIGURE 22. Comparison of psychoacoustically measured data for the minimum distance at which a vehicle sound must be perceived to avoid collision, and data calculated according to the model illustrated in figure 21 (after Kerber 2008).

As a rule, the experimentally determined and predicted distances necessary for collision avoidance as displayed in figure 22 are in line. Larger differences between measurement and prediction are on the safe side, i.e. the predicted distance in which the vehicle must be heard is larger than the experimentally determined distance. This means that the model and the can be used to predict vehicle sounds necessary to be heard early enough at vehicle speeds and background noises typical for traffic-calmed areas.

Obscuring the Information about the Sound Source

In practical applications the meaning of a sound can play an important part for its evaluation. This can be of relevance e.g. in cross-cultural studies (e.g. Fastl et al. 1986, Fastl and Yamada 1986, Kuwano et al. 1986, Namba et al. 1987, Kuwano et al. 1997, 2007, Weber and Hansen 2007). An example is given in figure 23.

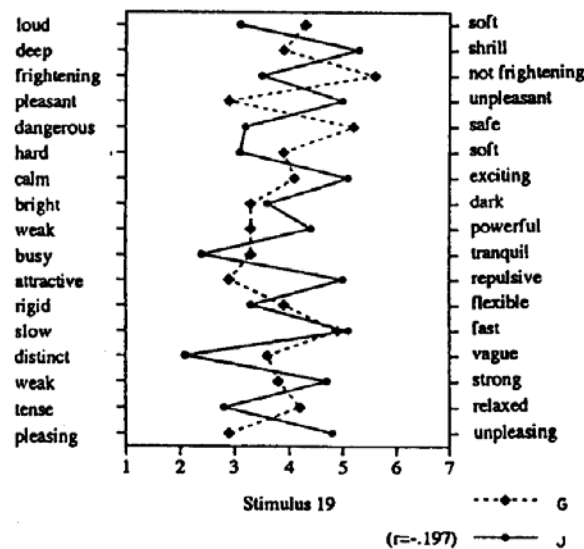


FIGURE 23. Semantic differential for the sound of a bell, evaluated by Japanese or German subjects (after Kuwano et al. 1997).

While for many sounds the evaluation by Japanese or German subjects is rather similar, in this case it is more or less inverse (cf. Kuwano et al. 1997). To the Japanese subjects, the bell sound is frightening, unpleasant, dangerous etc., but to the German subjects it is not frightening, pleasant, safe etc. Since in Japan fire engines produce bell sounds, this specific connotation may have influenced the (negative) evaluation by the Japanese subjects.

In order to minimize such cognitive effects in sound evaluations, a procedure was proposed which largely obscures the information about the sound source (cf. Fastl 2001). A block diagram is shown in figure 24.

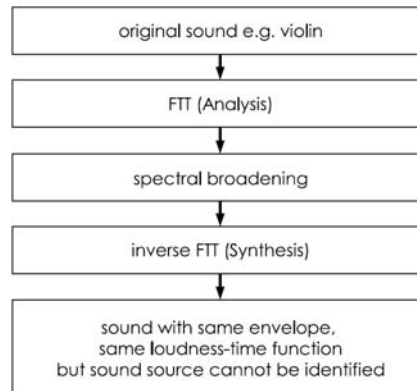


FIGURE 24. Block diagram of the procedure to obscure the information about the sound source (after Fastl 2001).

From an original sound, a Fourier Time Analysis (FTT) is performed in accordance with a proposal of Terhardt (1985). In essence, the signal is analyzed in some 600 channels according to features of the human hearing system. In the next step the spectral resolution is considerably broadened, leading to only 24 channels. By inverse FTT, a sound with the same envelope and almost the same loudness-time-function is obtained, but the sound source can no longer be recognized.

An example is displayed in figure 25. The left panel gives the FTT spectrogram for an instrument, playing a scale. The increase in the pitches as well as the harmonic structure of the sound is clearly seen. The processed sound in the right panel also indicates an increase in pitch and some harmonic structure. However, because of the spectral broadening, the whole image is blurred.

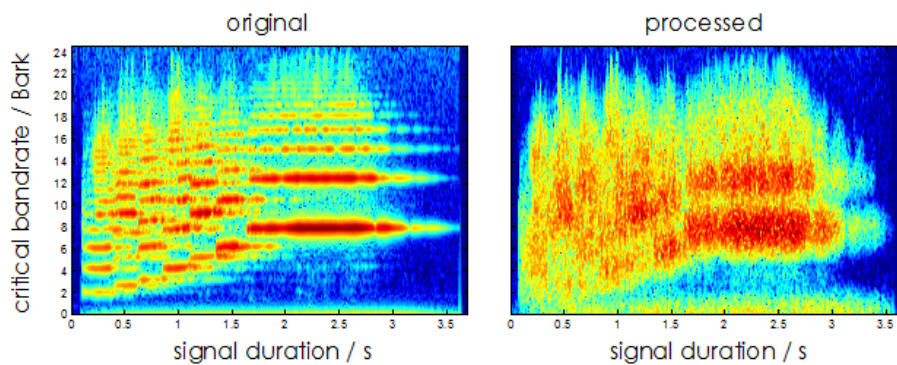


FIGURE 25. Obscuring the information about the sound source. Original FTT spectrogram of a musical instrument playing a scale (left panel) and processed blurred spectrogram (right panel).

Despite the fact that the loudness-time function of both sounds is the same, the processed sound is not attributed to a musical instrument. While it is almost impossible to recognize musical and many technical sounds when processed (Fastl 2001, 2002, Hellbrück et al. 2002, Fastl et al. 2003, Ellermeier et al. 2004a, 2004b, Zeitler et al. 2003, 2004), speech sounds are easily classified correctly.

Audio-visual Interactions

The perception of the loudness of sounds depends not only on acoustic inputs, but also visual inputs can influence auditory perception (e.g. Kohlrausch and van de Par 2005, for an overview and references see e.g. Haverkamp 2012). An example, which in the meantime found its way into popular science media, is the influence of color on the loudness of passing trains (Fastl 2004, Fastl and Patsouras 2004). In the experiment, subjects have to rate the loudness of the sound recording of a passing train, and in addition, images of a train in different color are presented. Figure 26 shows an example for a German high speed train (ICE) in different colors.



FIGURE 26. Images of a German high speed train (ICE) in original color white, and modified colors red, light blue, and light green presented in addition to the pass-by-sound of the train.

The results of the loudness judgments are given in figure 27. The left panel shows data for German subjects, the right panel data for Japanese subjects (cf. Fastl 2004, Rader et al. 2004).

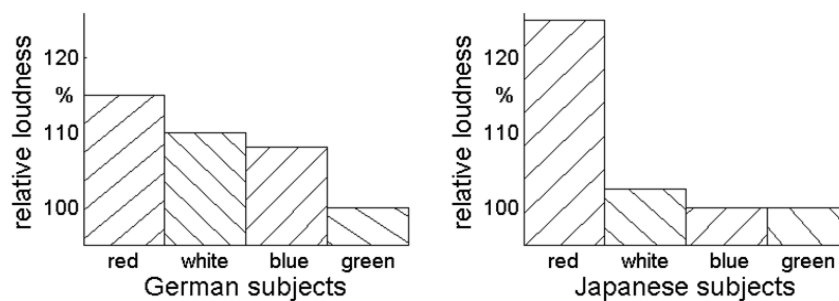


FIGURE 27. Evaluation of the loudness of the sounds from passing trains, when in addition to the sound images of the trains are presented in white, red, blue and green color. Data for German subjects (left) and Japanese subjects (right).

The data displayed in figure 27 indicate that – despite the same sound pressure level – red trains are perceived by 15 to 25 % louder than green trains. This holds for both Japanese and German subjects (Fastl 2004). Similar effects are found for Japanese high speed trains (Shinkansen) or commuter trains (Rader et al. 2004). Also by subjects from the United States, red trains were evaluated louder than green trains (Fastl et al. 2010).

From everyday experience it could be expected that red is a “loud” color, since many powerful sports cars, in particular from Italy, frequently wear red color. However, British sports cars also come in green which seems to be somewhat astonishing, since green was rated as a “soft” color. However, it could be shown (Menzel et al. 2008) that British racing green, which is a dark green (in contrast to the light green used for the trains) is also a “loud” color.

Since it has been shown that the color of products can influence their loudness (e.g. Menzel et al. 2008, 2010a), a hypothesis might be put forward as follows: The sounds from a red iPod would be perceived as louder than sounds from an iPod in “traditional” white. Hence, the level for the red iPod could be reduced to obtain the same perceived loudness. If this were true, because of the level reduction, the hearing systems of the (usually young) listeners would be less endangered.

Figure 28 shows the images of iPods in different color presented together with excerpts of music.



FIGURE 28. Images of iPods in different color presented together with excerpts of music (after Menzel 2011).

Since it was hypothesized (Menzel et al. 2010b) that preferred colors might elicit larger effects of audio-visual interactions, subjects were asked about their preferences. From the colors presented, black was rated as top, and brown as flop. Interestingly, female and male subjects reported different preferences: White got the second rank by the male subjects, but only the sixth rank by the female subjects. On the other hand, pink got rank four by the female subjects, but only rank seven by the male subjects. The color brown got the last rank (ten) from both female and male subjects.

Figure 29 shows the loudness rating of music excerpts when in addition to the sound also images of iPods in different color are presented. The music style included classic (M3), pop (M2), and heavy metal (M6). All loudness data are normalized relative to the loudness perceived for music excerpt M4.

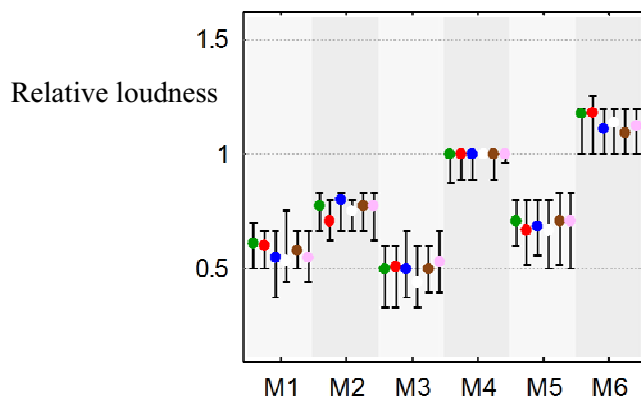


FIGURE 29. Relative loudness of music excerpts when in addition to the sound also images of iPods in different color are presented (after Menzel 2011).

The results displayed in figure 29 suggest that the color of images of iPods presented in addition to the musical sounds has almost no influence on the perceived loudness. All medians are within the inter-quartile ranges. Therefore, unfortunately, “loud” colors do not lead to an increased loudness perception of the music excerpts used. At same level, the music excerpts produce almost the same loudness, irrespective of the color of the image of an iPod presented in addition to the sound.

CONCLUSION

Over the years, for basic hearing sensations, a wealth of data has been assembled. Nevertheless, still challenging questions are open in such “old” topics like pitch or loudness.

As concerns practical applications, fields like audiology (e.g. Kollmeier 1996), automatic speech recognition (e.g. Hemmert et al. 2004) or noise evaluation (e.g. Fastl 2007) have already reached high standards. More recent applications like virtual acoustics generate rewarding questions with respect to both stimuli and perception (e.g. Vorländer 2008a, 2008b, Spors et al. 2008, Blauert et al. 2009, Goossens 2010, Völk 2011a, 2011b, Völk and Fastl 2010, 2011, 2012).

Still the old engineering approach is valid that a meaningful cost/benefit analysis should be based on psychoacoustics: The specifications of e.g. systems for audio communication should be adapted to features of the hearing system: Only improvements which are *audible* are worth the effort!

ACKNOWLEDGMENTS

The author wishes to thank former and current members of his group Technische Akustik for their substantial contributions to the work presented here. Special thanks go to Dipl.-Ing. Florian Völk for stimulating discussions and editorial support. Large parts of the work reported here has been supported over years by the Deutsche Forschungsgemeinschaft (DFG).

REFERENCES

- ANSI S3.4-2007: Procedure for the Computation of Loudness of Steady Sounds. American National Standards Institute (2007)
- Berkhout A. J.: A Holographic Approach to Acoustic Control. *J. Audio Eng. Soc.* 36, 977–995 (1988)
- Blauert J.: *Spatial Hearing – The Psychophysics of Human Sound Localization*. The MIT Press, Cambridge, Massachusetts, London (1997)
- Blauert J., P. Tritthart: Ausnutzung von Verdeckungseffekten bei der Sprachkodierung. In *Fortschritte der Akustik, DAGA'75*, 377–380 (Physik, Weinheim, 1975)
- Blauert J., J. Braasch, J. Buchholz, H. S. Colburn, U. Jekosch, A. Kohlrausch, J. Mourtiopoulos, V. Pulkki, A. Raake: Aural assessment by means of binaural algorithms – The AABBA project. In *International Symposium on Auditory and Audiological Research (ISAAR)* (2009)
- Brandenburg K.: OCF: Coding High Quality Audio with Data Rates of 64 kbit/sec. In *85th AES Convention* (1998) (Preprint 2723)
- Brüel and Kjaer: *Measuring Sound*. (1984)
- Buus S., M. Florentine: Growth of loudness in listeners with cochlear hearing losses: Recruitment reconsidered. *J. Assn. Res. Otolaryngol.* 3, 120–139 (2001)
- Chalupper J., H. Fastl: Dynamic Loudness Model (DLM) for Normal and Hearing-Impaired Listeners. *Acta Acustica united with Acustica* 88, 378-386 (2002)
- Clark G. M.: *Cochlear Implants: Fundamentals and Applications*. Springer, New York (2003)
- Daniel P., H. Fastl, T. Fedtke, K. Genuit, H.-P. Grabsch, T. Niederdränk, A. Schmitz, M. Vorländer, M. Zollner: *Kunstkopftechnik – Eine Bestandsaufnahme*. *Nuntius Acusticus* 6, 1–58 (2007)
- DIN 45631: Berechnung des Lautstärkepegels und der Lautheit aus dem Geräuschspektrum. Verfahren nach E. Zwicker. Deutsche Norm, Beuth, Berlin (1991)
- DIN 45631/A1: Berechnung des Lautstärkepegels und der Lautheit aus dem Geräuschspektrum - Verfahren nach E. Zwicker – Änderung 1: Berechnung der Lautheit zeitvarianter Geräusche; mit CD-ROM. Deutsche Norm, Beuth, Berlin (2010)
- Ellermeier W., A. Zeitler, H. Fastl: Predicting annoyance judgments from psychoacoustic metrics: Identifiable versus neutralized sounds. In *InterNoise 2004* (2004a)
- Ellermeier W., A. Zeitler, H. Fastl: Impact of Source Identifiability on perceived Loudness. In *18th International Congress on Acoustics (ICA)* (2004b)

- Fastl H.: Über Tonhöhenempfindungen bei Rauschen. *Acustica* 25, 350–354 (1971)
- Fastl H.: A Background Noise for Speech Audiometry. *Audiol. Acoustics* 26, 2-13 (1987)
- Fastl H.: Sound Quality of Electric Razors – Effects of Loudness. In *InterNoise 2000* (2000)
- Fastl H.: Neutralizing the meaning of sound for sound quality evaluations. In 17th International Congress on Acoustics (ICA) (2001)
- Fastl H.: Features of neutralized sounds for long term evaluation. In *Forum Acusticum 2002* (2002)
- Fastl H.: Towards a new Dummy Head? In *InterNoise 2004* (2004)
- Fastl H.: Audio-visual interactions in loudness evaluation. In 18th International Congress on Acoustics (ICA) (2004)
- Fastl H.: Psychoacoustics, sound quality and music. In *InterNoise 2007* (2007)
- Fastl H., H. Fleischer: Freifeldübertragungsmaße verschiedener elektrodynamischer und elektrostatischer Kopfhörer. *Acustica* 39, 182–187 (1978)
- Fastl H., G. Stoll: Scaling of Pitch Strength. *Hearing Research* 1, 293–301 (1979)
- Fastl H., E. Zwicker: A free-field equalizer for TDH 39 earphones. *J. Acoust. Soc. Am.* 73, 312–314 (1983)
- Fastl H., Y. Yamada: Cross-cultural study on loudness and annoyance of broadband noise with a tonal component. In *Contributions to Psychological Acoustics*, 341-353 (Kohlrenken, Oldenburg, 1986)
- Fastl H., C. Patsouras: The Sound of Colour. German research, *Magazine of the Deutsche Forschungsgemeinschaft* 2–3, 25–26 (2004)
- Fastl H., E. Zwicker: *Psychoacoustics – Facts and Models*. 3rd edition, Springer, Berlin, Heidelberg (2007)
- Fastl H., M. Florentine: Loudness in daily environments. In *Loudness* 37, 199–222 (Springer, New York, 2011)
- Fastl H., S. Namba, S. Kuwano: Cross-cultural investigations of loudness evaluation for noises. In *Contributions to Psychological Acoustics*, 354-369 (Kohlrenken, Oldenburg, 1986)
- Fastl H., M. Fruhmann, S. Ache: Railway Bonus for Sounds without Meaning? *Acoustics Australia* 31, 99–101 (2003)
- Fastl H., C. Patsouras, T. Rader: Klingt ein Flügel bei 432 Hz-Stimmung besser als bei 440 Hz-Stimmung? In *Fortschritte der Akustik, DAGA '98*, 528–529 (Dt. Gesell. für Akustik e. V., Oldenburg, 2003)
- Fastl H., D. Menzel, W. Maier: Entwicklung und Verifikation eines Lautheits-Thermometers. In *Fortschritte der Akustik, DAGA '06*, 669–670 (Dt. Gesell. für Akustik e. V., Berlin, 2006)
- Fastl H., F. Völk, M. Straubinger: Standards for calculating loudness of stationary or time-varying sounds. In *InterNoise 2009* (2009)
- Fastl H., T. Fleischer, J. Stelkens: Remote psychoacoustic experiments on audio-visual interactions. In 20th International Congress on Acoustics (ICA) (2010)
- Fastl H., W. Oberdanner, W. Schmid, I. Stemplinger, I. Hochmair-Desoyer, E. Hochmair: Zum Sprachverständnis von Cochlea-Implantat-Patienten bei Störgeräuschen. In *Fortschritte der Akustik, DAGA '98*, 358–359 (Dt. Gesell. für Akustik e. V., Oldenburg, 1998)
- Fingerhuth S., E. Parizet: Tonalness perception of harmonic complex tones. *J. Acoust. Soc. Am.* 123, 3685 (2008)
- Fletcher H., W. A. Munson: Loudness, Its Definition, Measurement and Calculation. *J. Acoust. Soc. Am.* 5, 82–108 (1933)
- Fruhmann M.: Ein Modell zur Wahrnehmung der Ausgeprägtheit der Tonhöhe von Schallen. PhD thesis, Technische Universität München, (2006)
- Genuit, K: *Sound-Engineering im Automobilbereich*, Springer, Berlin (2010)
- Genuit K., A. Fiebig: Do we need new artificial heads? In 19th International Congress on Acoustics (ICA) (2007)
- Goossens S.: Simulation und Auralisierung kleiner Räume, DAGA 2010, 19–25 (Dt. Gesell. für Akustik e. V., Berlin, 2010)
- Haar G.: Der Übertragungsfaktor und der Klirrfaktor eines elektrodynamischen Hörers. *Funk und Ton* 5, 17–26 (1951)
- Hammershøi D., P. F. Hoffmann: Control of earphone produced binaural signals. In *Forum Acusticum 2011*, 2235–2239 (2011)
- Haverkamp M.: *Synesthetic Design: Handbook for a Multi-Sensory Approach*. Birkhäuser (2012)
- Hartmann W.: *Signals, Sound and Sensation*. Springer, New York (1998)
- Hellbrück, J., W. Ellermeier: *Hören: Physiologie, Psychologie und Pathologie*. 2nd edition, Hogrefe, Göttingen (2004)
- Hellbrück J., H. Fastl, B. Keller: Effects of meaning of sounds on the loudness judgement. In *Forum Acusticum 2002* (2002)
- Hellman R., E. Zwicker: Why can a decrease in dB(A) produce an increase in loudness? *J. Acoust. Soc. Am.* 82, 1700–1705 (1987)
- Hemmert W., M. Holmberg, D. Gelbart: Auditory-based automatic speech recognition. In *ICSLP – Interspeech '04 – Research Workshop on Statistical and Perceptual Audio Processing*, 1–6 (2004)
- Henning G. B.: Auditory Localization. In *Handbook of Psychobiology* (Academic Press, New York, 1975)
- Jespersgaard C., T. Dau: Relationship between masking release in fluctuating maskers and speech reception thresholds in stationary noise. *J. Acoust. Soc. Am.* 132, 1655–1666 (2012)
- Kerber S., H. Fastl: Perceptibility of approaching vehicles in urban background noise. *InterNoise 2007* (2007)
- Kerber S.: Wahrnehmbarkeit von Fahrzeugaußengeräuschen in Hintergrundgeräuschen. PhD thesis, Technische Universität München (2008)
- Kohlrausch A., S. van de Par: Audio visual interaction in the context of multi-media applications. In *Communication Acoustics*, J. Blauert (ed.), 109–138, Springer, Berlin (2005)
- Kollmeier B.: *Psychoacoustics, Speech and Hearing Aids*, World Scientific Press, Singapore (1996)
- Krahe D.: New Source Coding Method for High Quality Digital Audio Signals. In *NTG Fachtagung Hörrundfunk* (1985)
- Kuwano S., S. Namba, H. Fastl: Loudness evaluation of various sounds by Japanese and German subjects. In *InterNoise '86*, 835–840 (1986)

- Kuwano S., S. Namba, H. Fastl, A. Schick: Evaluation of the Impression of Danger Signals – Comparison between Japanese and German Subjects. In 7. Oldenburger Symposium, 115–118 (BIS, Oldenburg, 1997)
- Kuwano S., S. Namba, A. Schick, H. Höge, H. Fastl, T. Filippou, M. Florentine: Subjective impression of auditory danger signals in different countries. *Acoust. Sci. and Tech.* 28, 360–362 (2007)
- Menzel D.: Zum Einfluss von Farben auf das Lautheitsurteil. Dr. Hut, München (2011)
- Menzel D., E. Faccinelli, H. Fastl: Are absolute thresholds and loudness judgements influenced by different colours? In *Acoustics '08*, 1203–1208 (2008)
- Menzel D., N. Haufe, H. Fastl: Colour-influences on loudness judgements. In 20th International Congress on Acoustics (ICA) (2010)
- Menzel D., T. Schulze, H. Fastl: Zum Einfluss der Farbpräferenz auf die Lautheitsbeurteilung. In *Fortschritte der Akustik, DAGA 2010*, 873–874 (Dt. Gesell. für Akustik e.V., Berlin, 2010)
- Menzel D., H. Wittek, G. Theile, H. Fastl: The Binaural Sky: A Virtual Headphone for Binaural Room Synthesis. In *Intern. Tonmeister Symposium (2005)* http://www.tonmeister.de/symposium/2005/np_pdf/R04.pdf
- Menzel D., H. Wittek, H. Fastl, G. Theile: Binaurale Raumsynthese mittels Wellenfeldsynthese – Realisierung und Evaluierung. *Fortschritte der Akustik, DAGA'06*, 255–256 (Dt. Gesell. für Akustik e.V., Berlin, 2006)
- Menzel D., H. Fastl, R. Graf, J. Hellbrück: Influence of vehicle color on loudness judgments. *J. Acoust. Soc. Am.* 123, 2477–2479 (2008)
- Møller H.: *Fundamentals of Binaural Technology*. *Applied Acoustics* 36, 171–218 (1992)
- Moore B. C. J.: *An Introduction to the Psychology of Hearing*. 6th edition, Emerald Group, Bingley (2012)
- Moore B. C. J., B. R. Glasberg, T. Baer: A Model for the Prediction of Thresholds, Loudness, and Partial Loudness. *J. Audio Eng. Soc.* 45, 224–240 (1997)
- More S., P. Davies: Human responses to the tonalness of aircraft noise. *Noise Control Engineering Journal* 58, 420–440 (2010)
- Namba S., S. Kuwano, H. Fastl: Cross-cultural study on the loudness, noisiness, and annoyance of various sounds. In *InterNoise '87*, 1009–1012 (1987)
- Painter T., A. Spanias: A Review of Algorithms for Perceptual Coding of Digital Audio Signals. In 13th International Conference on Digital Signal Processing, 179–208 (1997)
- Rader T., U. Baumann, H. Fastl: Comparison of Speech Intelligibility by EAS, Bimodal, Uni- and Bilateral Cochlear Implant Patients in a 'Multi-Source Noise Field' (MSNF). In *International Conference on Acoustics (NAG/DAGA)*, 1288–1290 (Dt. Gesell. für Akustik e. V., Berlin, 2009)
- Rader T., H. Fastl, U. Baumann: Speech perception with combined electric acoustic stimulation (EAS) and bilateral cochlear implant in a multi-source noise field. *Ear and Hearing*, 2012
- Rader T., H. Fastl, S. Helbig, U. Baumann: Elektrisch-akustische Stimulation: Funktionelle Ergebnisse im Vergleich zur konventionellen CI Versorgung. *Laryngo-Rhino-Otologie* 90, 709–712 (2011)
- Rader T., M. Morinaga, T. Matsui, H. Fastl, S. Kuwano, S. Namba: Crosscultural Effects in Audio-Visual Interactions. In *Meeting of the Technical Committee on Noise and Vibration of the Acoustical Society of Japan* (2004)
- Rakowski A.: Measurements of pitch. *Catgut. Acoust. Soc. Newsl.* 27 (1977)
- Rennies J., J. L. Verhey, H. Fastl: Comparison of loudness models for time-varying sounds, *Acta Acustica united with Acustica* 96, 383–396 (2011)
- Schlittenlacher J., T. Hashimoto, H. Fastl, S. Namba, S. Kuwano, S. Hatano: Loudness of pink noise and stationary technical sounds. *InterNoise 2011* (2011)
- Seeber B. U., U. Baumann, H. Fastl: Localization ability with bimodal hearing aids and bilateral cochlear implants. *J. Acoust. Soc. Am.* 116, 1698–1709 (2004)
- Seeber B. U., H. Fastl: Localization cues with bilateral cochlear implants. *J. Acoust. Soc. Am.* 123, 1030–1042 (2008)
- Simmons F. B., J. M. Epley, R. C. Lummis, N. Guttman, L. S. Frishkopf, L. S. Harmon, E. Zwicker: Auditory nerve: electrical stimulation in man. *Science* 148, 104–106 (1965)
- Sottek R.: Loudness models applied to technical sounds. *J. Acoust. Soc. Am.* 127, 1880 (2010)
- Spors S., R. Rabenstein, J. Ahrens: The Theory of Wave Field Synthesis Revisited. In 124th AES Convention (2008) (Convention Paper 7358)
- Stevens S. S., H. Davis: *Psychophysiological Acoustics: Pitch and Loudness*. *J. Acoust. Soc. Am.* 8, 1–13 (1936)
- Stevens S. S., H. Davis: *Hearing – Its Psychology and Physiology*. John Wiley & Sons, New York, London, Sydney (1938)
- Stoll G.: A Perceptual Coding Technique Offering the Best Compromise Between Quality, Bit-Rate, and Complexity for DSB. In 94th AES Convention (1993) (Preprint 3458)
- Terhardt E.: Fourier Transformation of Time Signals: Conceptual Revision. *Acustica* 57, 242–256 (1985)
- Terhardt E.: *Akustische Kommunikation*. Springer, Berlin, Heidelberg, New York (1998)
- Theile G.: Neue Anwendung der Wellenfeldsynthese: Binaural Sky, *FKT 10* (2007)
- Theile G., M. Link, G. Stoll: Low-bit rate coding of high quality audio signals. In 82nd AES Convention (1987) (Preprint 2432)
- van de Par S., J. Koppens, A. Kohlrausch, W. Oomen: A new perceptual model for audio coding based on spectro-temporal masking. 124th AES Convention (2008) (Convention Paper 7336)
- Völk F.: System Theory of Binaural Synthesis. 131st AES Convention (2011) (Convention Paper 8568)
- Völk F.: Inter- and Intra-Individual Variability in Blocked Auditory Canal Transfer Functions of Three Circum-Aural Headphones. In 131st AES Convention (2011a) (Convention Paper 8465)

- Völk F.: Headphone Selection for Binaural Synthesis with Blocked Auditory Canal Recording. 132nd AES Convention (2012) (Convention Paper 8677)
- Völk F., H. Fastl: Advantages of binaural room synthesis for research and fitting of hearing aids, cochlear implants, electro-acoustical stimulation, and combined systems. In 20th International Congress on Acoustics (ICA) (2010)
- Völk F., H. Fastl: Locating the Missing 6 dB by Loudness Calibration of Binaural Synthesis. In 131st AES Convention (2011) (Convention Paper 8488)
- Völk F., H. Fastl: Wave Field Synthesis with Primary Source Correction: Theory, Simulation Results, and Comparison to Earlier Approaches. In 133rd AES Convention (2012) (Convention Paper 8717)
- Völk F., H. Fastl, F. Lott, B. Hasford, A. Glink: Mindestanforderungen an das Hörvermögen von angehenden Mediengestaltern Bild und Ton. Fortschritte der Akustik, DAGA 2008, 55–56 (Dt. Gesell. für Akustik e. V., Berlin, 2008)
- Vorländer M.: Virtual Acoustics – Opportunities and limits of spatial sound reproduction. Archives of Acoustics 33, 413–422 (2008a)
- Vorländer M.: Auralization – Fundamentals of Acoustics, Modelling, Simulation, Algorithms and Acoustic Virtual Reality. Springer, Berlin, Heidelberg (2008b)
- Weber R., H. Hansen: Perception of noises with tonal components in Japan, France and Germany. In InterNoise 2007 (2007)
- Weng A.: A weng a' 432 Hertz (2001) <http://www.musik-weng.de>
- Wiesmann N., H. Fastl: Ausgeprägtheit der Tonhöhe und Frequenzunterschiedsschwellen von Bandpaß-Rauschen. Fortschritte der Akustik, DAGA '91, 505–508 (DPG, Bad Honnef, 1991)
- Wittek H.: Spatial perception in Wave Field Synthesis rendered sound fields: Distance of real and virtual nearby sources. In Joint Congress CFA/DAGA, 1039–1040 (Dt. Gesell. für Akustik e. V., Oldenburg, 2004)
- Yost W. A., R. Hill: Strength of the pitches associated with ripple noise. J. Acoust. Soc. Am. 64, 485–492 (1978)
- Zeitler A., H. Fastl, J. Hellbrück: Einfluss der Bedeutung auf die Lautstärkebeurteilung von Umweltgeräuschen. In Fortschritte der Akustik, DAGA '03, 602–603 (Dt. Gesell. für Akustik e. V., Oldenburg, 2003)
- Zeitler A., W. Ellermeier, H. Fastl: Significance of Meaning in Sound Quality Evaluation. In Joint Congress CFA/DAGA, 781–782 (Dt. Gesell. für Akustik e. V., Oldenburg, 2004)
- Zwicker E.: Über psychologische und methodische Grundlagen der Lautheit. Acustica 8, 237–258 (1958)
- Zwicker E.: Procedure for calculating loudness of temporally variable sounds. J. Acoust. Soc. Am. 62, 675–682 (1977)
- Zwicker E., R. Feldtkeller: Das Ohr als Nachrichtenempfänger. Hirzel, Stuttgart (1967)
- Zwicker E., D. Maiwald: Über das Freifeldübertragungsmaß des Kopfhörers. Acustica 13, 181–182 (1963)