# Binaural Loudness for Artificial-Head Measurements in Directional Sound Fields*

**VILLE PEKKA SIVONEN****
(ville.sivonen@oulu.fi)

*Department of Otorhinolaryngology, University Hospital of Oulu, Oulu 90029 OYS, Finland*

**AND**

**WOLFGANG ELLERMEIER****
(ellermeier@psychologie.tu-darmstadt.de)

*Institute of Psychology, Technische Universität Darmstadt, 64283 Darmstadt, Germany*

The effect of the sound incidence angle on loudness was investigated for fifteen listeners who matched the loudness of sounds coming from five different incidence angles in the horizontal plane to that of the same sound with frontal incidence. The stimuli were presented via binaural synthesis by using head-related transfer functions measured for an artificial head. The results, which exhibited marked individual differences, show that loudness depends on the direction from which a sound reaches the listener. The average results suggest a relatively simple rule for combining the two signals at the ears of an artificial head for binaural loudness predictions.

## 0 INTRODUCTION

Acoustical measurements of sound fields increasingly rely on the use of artificial heads. It is unclear, however, how such measurements should be used to predict loudness. The standardized loudness model ISO 532 [1] for the (frontal) free field, or for the diffuse field, utilizes the sound signal measured with a monophonic microphone in the absence of a listener, whereas with an artificial head the sound signals at each ear are measured, as they are obstructed by the physical dimensions of the dummy's head and torso.

The problem of reconciling artificial-head measurements with unobstructed monophonic signals can be solved by employing head-related transfer functions (HRTFs; [2]–[5]). These functions characterize the transfer of sound from an unobstructed free field to the listener's two ears for a given sound incidence angle. To utilize artificial-head measurements for loudness predictions in ISO 532 [1], the inverse of an HRTF can be used to convert an at-ear signal to the corresponding signal in the center of the head with the head absent. The model of Moore et al. [6], as modified by Glasberg and Moore [7] (now standardized in [8]), explicitly enables the use of at-ear measurements for predicting loudness. This is achieved by effectively separating the HRTF filtering stage, which is an integral part of ISO 532 [1], from the loudness computations of the model.

The conversion of measurements on real ears or using an artificial head to the corresponding free- or diffuse-field exposure has recently been standardized in [9], [10]. Besides loudness predictions, this allows for assessing at-ear measurements, for example, in terms of noise exposure limits, against well-established criteria based on measurements in the absence of a listener [11].

In addition to compensating for the measurement point, the binaural nature of sound exposure has to be taken into account when making artificial-head measurements. In a sound field both ears of a listener receive a signal, and the listener's brain combines the neural representations of

*Manuscript received 2007 September 5; revised 2008 April 21 and May 6. Portions of this work were presented at Euronoise 2006, the 6th European Conference on Noise Control, Tampere, Finland.

**The work was carried out while the authors were with the Sound Quality Research Unit (SQRU), Aalborg University, Aalborg, Denmark.

these signals to a single binaural percept. In a diotic situation the signals at the two ears are the same, and therefore either of the measured at-ear signals can be converted to the corresponding free- or diffuse-field signal[1] in the absence of the head, and subsequently be utilized for loudness computations in [1], or in [6]. Free- and diffuse-field HRTFs are similar for the two ears due to symmetry, and assuming that both ears are equally sensitive as transducers to loudness, the conversion using either at-ear signal results essentially in the same outcome in the diotic case.

Both models [1] and [6] are applicable to steady-state sounds and utilize average long-term spectra measured, for example, on a one-third-octave resolution for loudness computations. ISO 532 [1], however, is not defined for situations where the spectra at the two ears are different. Different at-ear spectra tend to be the rule, rather than the exception, when making acoustical measurements in sound fields. This is due to sound sources being located nonsymmetrically with respect to the ears of a head or reverberation having nonsymmetrical effects on the signals arriving at the ears. It is thus important to investigate how, for a sound source located in space, the two at-ear signals should be combined to yield a single prediction of binaural loudness.

In the model proposed by Moore et al. [6], binaural loudness for different at-ear spectra is handled by assuming that the loudness values evoked by the signal at each ear are summed to give the overall loudness. Hence the loudness of a monaural stimulus of given level is half the loudness of the same stimulus presented diotically at the same level, which is supported by earlier headphone studies on binaural loudness summation (see, for example, [12], [13]). Binaural loudness is thus computed as a simple sum of two monaural values, $N_{\mathrm{binaural}} = N_{\mathrm{left}} + N_{\mathrm{right}}$. To better account for more recent data, the binaural loudness computation has been modified in [14] using parameters that predict the binaural-to-monaural ratio to be 1.5. Note, however, that the model and its modification are mainly based on headphone studies, which neglect the effect of HRTF filtering, and thus produce auditory events localized inside the listener's head. In addition to the inside-the-head localization, binaural signal combinations that would never reach the listener's ears in a real sound field may have been generated.

By obtaining loudness matches of anechoic sounds coming from different directions in a listening test, and by measuring the listeners' HRTFs for the same incidence angles, it can be investigated how the natural changes in the at-ear signals due to the listener's head, torso, and pinnae may influence perceived loudness. Since HRTFs describe the transfer from a free field to the listeners' ears, they can be used for converting the matched sound pressure levels (SPLs) of the sound sources measured in the absence of the head to equal-loudness at-ear SPLs. These at-ear SPLs can then be used to model binaural loudness summation of directional sounds arriving at the listeners' ears from various incidence angles, while having the auditory events localized outside the listener's head.

Two investigations used this strategy [15], [16], acquiring both directional loudness matches between sounds from various incidence angles and sounds from a reference direction, and then relating the matches to the changes in the at-ear stimulation as a function of incidence angle. Both described the directional loudness matches as agreeing with a relatively simple combination of the left-ear ($L_{\mathrm{left}}$) and right-ear ($L_{\mathrm{right}}$) SPLs [in decibels (dB)], expressed by equivalent monaural levels,

$$L_{\mathrm{mon}} = g * \log_2(2^{L_{\mathrm{left}}/g} + 2^{L_{\mathrm{right}}/g}). \tag{1}$$

However, the binaural gain constants $g$ were different, the former study [15] reporting a 6-dB binaural-summation rule ($g = 6$) to fit the data best, whereas in the latter study [16] a 3-dB gain ($g = 3$) was found to describe the mean data much better. The binaural gain constant of Eq. (1) describes how many dB a monaural sound would have to be higher in level to be perceived as equally loud as the same sound presented binaurally.[2] It is worth noting that the gains of 6 and 3 dB of [15] and [16] correspond to binaural-to-monaural loudness ratios of 1.5 and 1.2, respectively.

In the study of Sivonen and Ellermeier [16] the modeling was also performed on an individual basis. Even though the effects of individual HRTFs were taken into account, large individual differences remained when fitting the parameter $g$. However, since the sound signals arriving at the listeners' ears owing to physical differences in the shapes of pinnae and heads were highly idiosyncratic as well, fairly large interindividual differences in stimulation resulted, undermining the validity of comparisons across individuals.

When stimuli are generated using nonindividual human, or generic, artificial-head HRTFs, the physical differences between listeners cannot play a role. Directional sounds generated with such HRTFs produce the same directional effects in at-ear exposures for each listener. Thus the comparison of the directional loudness matches across individuals is more straightforward, since the effect of individual HRTF filtering does not have to be accounted for. Nonindividual or generic HRTFs are played back using binaural synthesis over headphones. The use of binaural synthesis is justified in the directional loudness paradigm. Individual binaural synthesis has been shown to result in essentially the same directional loudness matches as listening to real sound sources [17]. However, it is worth noting that when listeners are presented with sounds filtered through HRTFs other than their own, the fidelity of

[1]The inverse of the HRTF used for the conversion and the sound field in the loudness computations must be of the same type—either free field or diffuse field.

[2]Instead of a logarithmic base of 2, Eq. (1) can also be formulated with a base of 10; $L_{\mathrm{mon}} = a * \log_{10}(10^{L_{\mathrm{left}}/a} + 10^{L_{\mathrm{right}}/a})$, where $a = g/\log_{10}(2)$; or alternatively, it can be represented with linear pressure values $p_{\mathrm{mon}} = p^b_{\mathrm{left}} + p^b_{\mathrm{right}}$, where $b = 20 * \log_{10}(2)/g$. The constants $a$ and $b$, however, appear less informative than the binaural gain constant $g$ in Eq. (1).

the playback may be degraded, as has been shown, for example, for the localization performance of binaural recordings [18].

The goals of the present investigation are threefold.

1) To use generic, rather than individual HRTFs to investigate whether holding this source of interindividual variation constant will reduce the variability seen in directional loudness matches.

2) To run both a naive sample of listeners and an expert group that had previously gone through numerous conditions of real-source and binaurally synthesized directional loudness experiments in order to explore the generality of the findings.

3) To sketch a model for predicting binaural loudness using artificial-head measurements based on the best account of the average data.

## 1 METHOD

### 1.1 Subjects

Five listeners (four male, one female; ages between 26 and 48 years; median 28 years) took part, and they all had participated in earlier experiments in a real sound field [16] or using binaural synthesis with individual HRTFs [17]. These "expert" listeners were familiar with the task of comparing the loudness of sounds coming from different directions. An additional group of 10 listeners (seven male, three female; ages between 22 and 31 years; median 26 years), considered "naive" to the task, were recruited from a student population.

All listeners were screened for normal hearing, with the requirement that none of the hearing thresholds exceed 20 dB hearing level re ISO 389-1 [19].

### 1.2 Apparatus

An artificial head (Brüel & Kjær head and torso simulator type 4100), which is commercially available and commonly used in the field, was used in acquiring the generic HRTFs. The HRTFs were measured using the same procedure as reported in [16], with the exception that the sound pressures both in the absence of the artificial head ($p_1$ pressure, see [4]) and at the entrance to the ear canal ($p_2$ pressure, see [4]) were measured using the same type of microphones (built-in microphones of the artificial head; Brüel & Kjær type 4190). Despite using the same microphone type, a correction for the differences in the frequency responses of the dedicated left- and right-ear microphones was determined by measuring the same sound field simultaneously with the two microphones placed approximately 3 mm apart. The subsequent signal processing for obtaining the HRTFs was as described in [16].

Generic headphone transfer functions (PTFs) were measured for a pair of headphones (Beyerdynamic DT-990) using the artificial head in order to equalize for their response in the binaural playback. The playback setup and the rest of the apparatus were as reported in [17].

The listeners were seated inside an anechoic chamber in a chair with a headrest to restrict head movements. The listeners' responses were collected with a two-button response box, and a model of the box having larger indicator lights was placed straight ahead of the listeners, to avoid that the listeners would tilt their heads downward to the response box in their hands. Loudspeakers mounted inside the anechoic chamber at the intended incidence angles were visible to the listeners to improve the plausibility of the binaural synthesis.

### 1.3 Stimuli

One-third-octave noise bands centered at 1 and 5 kHz were used as stimuli. The length of each sound was 1 s, including 20-ms raised-cosine rise and decay ramps. The same noise bands as in the earlier study [16] in a real sound field were used to allow comparison between investigations. Furthermore, narrow-band signals, as opposed to more real-life-like wide-band signals, were used to maximize the directional effects on loudness [20].

The two noise bands were convolved with the generic HRTFs for six incidence angles in the left hemisphere of the horizontal plane. The incidence angles were 0°, 30°, 60°, 90°, 135°, and 180° of azimuth at an elevation of 0°, that is, from ahead to behind the artificial head, at the height of the ears of the head.

The left- and right-ear magnitude spectra of the generic HRTFs are plotted in Fig. 1 to illustrate the effective stimulation for each incidence angle. The differences in the responses of the left- and the right-ear microphones are compensated for. As seen in Fig. 1, interaural level differences (ILDs) for the incidence angles ahead (0°) and behind (180°) the artificial head are small due to symmetry. By contrast, ILDs of over 30 dB can be observed for some other incidence angles. Note that the same holds for the interaural time difference (ITD), where the ITD for the direction on the side (at ±90° of azimuth) reaches its maximum, whereas the ITDs for sound sources ahead and behind the head are close to zero. The ILD, however, is more relevant than the ITD for directional loudness processing [15], [16], [20], and loudness models [1], [6], [14] only take the magnitude spectrum as an input to computations.

In addition to convolving the two noise bands with the HRTFs, minimum-phase equalization filters for the headphones, as specified in [17], were applied separately for the left- and right-ear signals in the binaural synthesis. Generic filters based on the PTF measurements using the artificial head were applied. Thus the equalization was done for PTFs measured exactly at the same position in the ear canal as the HRTFs, but individual coupling of the headphones to each listener's ears was not accounted for. The coupling, however, has no effect on the directional changes that the HRTFs impose on the at-ear SPLs. Furthermore, when inspecting the magnitude spectra of measured PTFs for each listener and the artificial head, the differences between individual and generic transfer functions turned out to be small (less than 4 dB, affecting both ears symmetrically) in the frequency range used in the present study.

In the listening test proper, the stimuli were played back at an overall level corresponding to 65 dB SPL in the free field. To derive the effective at-ear SPLs for the various incidence angles, the gains in the one-third-octave bands caused by the HRTFs are listed in Table 1. The 65-dB overall level combined with the left and right gains yields the SPLs at the ears of the artificial head. Furthermore the ILD column confirms that the SPLs at the two ears are similar for the sound sources ahead and behind the head, whereas a wide range of ILDs is observed for the other incidences.

### 1.4 Procedure

An adaptive, two-alternative, two-interval, forced-choice procedure employing a one-up, one-down rule [21], [22] was used to obtain directional loudness matches between a sound of a given center frequency with synthesized frontal (reference) incidence and the same sound with one of the other five (comparison) incidence angles.

Loudness matches were obtained for 10 conditions (five comparison incidence angles × two center frequencies), and eight replications were collected in each condition. To investigate the baseline variability of the loudness matches, the naive group also had to match the loudness of the frontal reference to itself, just as the experienced listeners had done in [16], [17].

The listener's task in each trial was to judge which of the sounds in the pair was louder. The level of the frontal reference sound was fixed (corresponding to 65 dB SPL in the free field) and the level of the comparison sound was varied by the adaptive procedure. Half the replications started +10 dB and the other half −10 dB from the level of the frontal reference. The initial step size of an adaptive track was 4 dB, and after two reversals it was decreased to 1 dB. A track was terminated after eight reversals, and the levels of the last six reversals were averaged to give an estimate of the loudness match.

The adaptive tracks were interleaved in blocks of eight, and the order of the adaptive tracks was randomized within one set of replications. Each 1-hour listening session included either three or four blocks, each lasting approximately 10 min, with rest breaks between the blocks. In total the listening test lasted approximately three hours per listener, the expert listeners participating in 10 and the naive listeners in 11 blocks.

## 2 RESULTS AND DISCUSSION

### 2.1 Directional Loudness Sensitivities

The listeners matched the loudness of a synthesized sound from each of the comparison incidence angles to the

Table 1. One-third-octave-band levels (in dB Pa/Pa) for generic HRTFs and ILD at two stimulus center frequencies.

| | 1 kHz | | | 5 kHz | | |
|---|---|---|---|---|---|---|
| Azimuth | Left | Right | ILD | Left | Right | ILD |
| 0° | −0.1 | −0.2 | 0.1 | 13.1 | 12.2 | 0.9 |
| 30° | 2.7 | −4.8 | 7.5 | 17.1 | 4.2 | 12.9 |
| 60° | 4.3 | −3.8 | 8.1 | 17.5 | −1.5 | 19.0 |
| 90° | 4.9 | −1.5 | 6.4 | 17.2 | −9.8 | 27.0 |
| 135° | 4.1 | −5.6 | 9.7 | 8.5 | 0.7 | 7.8 |
| 180° | 1.1 | 1.2 | 0.1 | 7.1 | 5.6 | 1.5 |

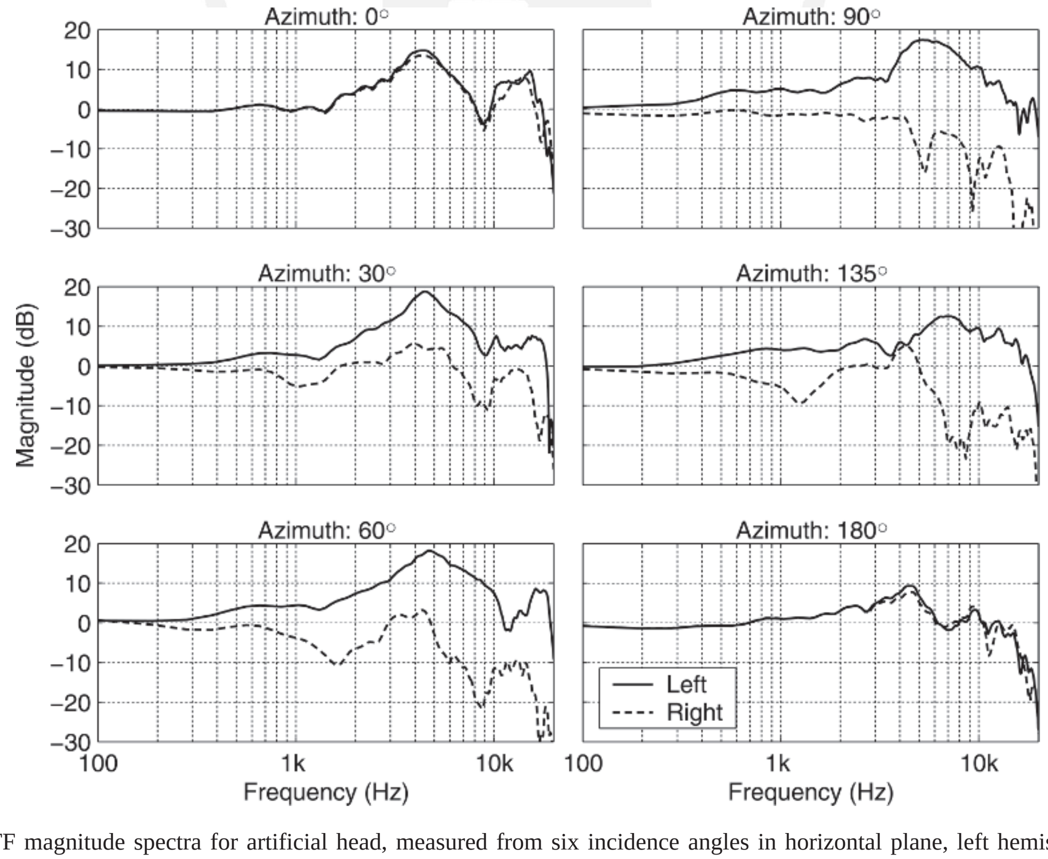


Fig. 1. HRTF magnitude spectra for artificial head, measured from six incidence angles in horizontal plane, left hemisphere (0° of azimuth—ahead; 180°—behind).

same sound with frontal incidence. These matches yielded sound pressure level adjustments for each incidence angle, meaning that if a sound emanating with a given angle had to be attenuated relative to the frontal reference, this incidence angle was perceived as louder. Therefore the relative level adjustments were inverted to yield directional loudness sensitivities, in agreement with the earlier studies [16], [17].

#### 2.1.1 Expert Listeners

The directional loudness sensitivities are plotted in Fig. 2 for the five expert listeners. For a given individual, the loudness-sensitivity curves vary as a function of the incidence angle over ranges of up to 4 and 10 dB at 1 and 5 kHz, respectively. At 1 kHz the data points are generally above the 0-dB line, indicating that loudness is increased relative to the frontal reference. At 5 kHz the data are both above and below the (0-dB) level of the frontal reference. A two-factor (incidence angle × center frequency) repeated-measures analysis of variance (ANOVA) indicated that both the main effect of the incidence angle [$F(4, 16) = 50.10$; $p < 0.001$] and its interaction with the center frequency [$F(4, 16) = 39.48$; $p < 0.001$] were highly significant. This implies that loudness is not constant as a function of incidence angle, and that the directional loudness-sensitivity curves are different for the two center frequencies (left and right panels of Fig. 2).

Along with the subjective data, the changes in the at-ear SPLs are plotted in Fig. 2, normalizing the one-third-octave-band levels of Table 1 by subtracting the levels of the frontal reference from them. Thus the objective changes in the at-ear stimulation can be compared with the directional loudness sensitivities. It is evident in Fig. 2 that for both center frequencies the directional loudness sensitivities largely follow the shape of the curve tracing level at the (left) ear receiving the higher SPL. However, individual differences can be observed, even though the changes in at-ear SPLs as a function of incidence angle were the same for all listeners. These individual differences reach up to 6 dB; see, for example, the 5-kHz data at an azimuth of 90° in Fig. 2. The statistical significance of the differences between listeners is confirmed by the three-way interaction between incidence angle, center frequency, and listener: $F(16, 350) = 8.28$; $p < 0.001$. As seen in Fig. 2, the differences between the participants' curves are much larger than the confidence intervals characterizing the replicability of a given experimental condition for a given listener.

#### 2.1.2 Naive Listeners

The directional loudness sensitivities for the 10 naive listeners are plotted in Fig. 3. As was the case for the expert listeners in the earlier studies [16], all naive listeners were able to match the synthesized frontal reference to itself; see the data at 1 kHz at 0° azimuth, the leftmost data points in Fig. 3. For the remaining conditions the range of variations in the individual directional loudness sensitivities at both center frequencies is similar to what was observed for the expert listeners. More importantly, the differences between the 10 individual listeners seem to be at

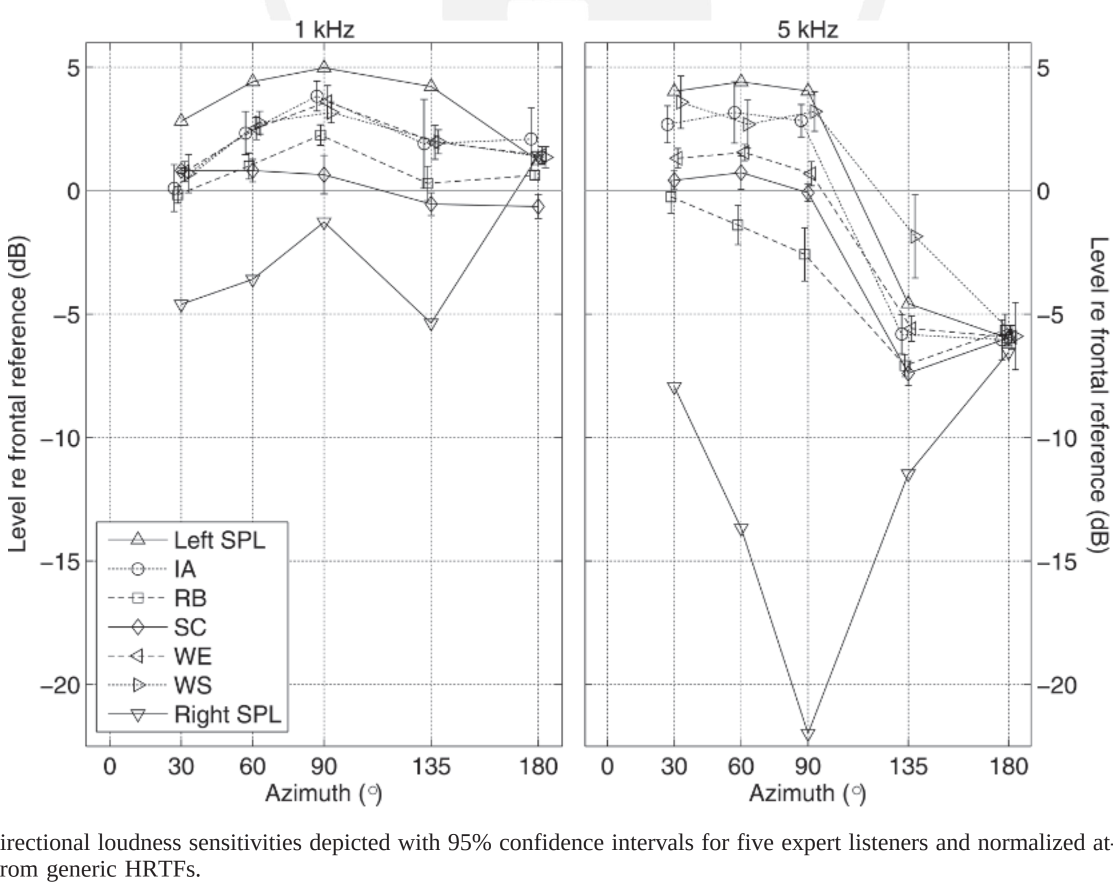


Fig. 2. Directional loudness sensitivities depicted with 95% confidence intervals for five expert listeners and normalized at-ear SPLs derived from generic HRTFs.

least on the same order of magnitude as obtained for the group of experienced listeners, implying that the idiosyncracies in the data of the expert listeners were replicated for another independent sample of listeners.

A statistical analysis of the naive listeners' data yielded the same significant effects (of incidence angle and its interaction with center frequency) as for the experts, including the significance of differences between listeners: $F(16,700) = 2.59$; $p < 0.001$.

### 2.1.3 Mean Data

To further explore the group differences between expert and naive listeners, a three-factor (incidence angle × center frequency × group) mixed ANOVA was performed. The ANOVA confirmed that neither the group factor nor any of its interactions with the other factors were significant (all $p > 0.25$).

## 2.2 Binaural Loudness Summation

For both groups of listeners individual differences in the listeners' directional loudness sensitivities exist, even though the directional changes in the at-ear SPLs were the same for all listeners since generic HRTFs were used (see Figs. 1 and 2). In purely sensory terms, this appears to imply that the listeners sum the signals at their ears in an individual manner to yield a binaural loudness.

In [16] binaural loudness summation was modeled based on individual directional loudness sensitivities and at-ear SPLs derived from individual HRTFs, utilizing Eq. (1) and estimating the binaural gain factor. Considerable differences between listeners were obtained, the amount of binaural summation ranging from close to 0 dB to far above 10 dB. This kind of modeling was also performed with the present data based on generic HRTFs. The one-third-octave-band SPLs listed in Table 1 were entered as left- and right-ear SPLs. Minimizing the sum of squares of the errors between the prediction $L_{mon}$ and the directional loudness sensitivity obtained, the amount of binaural summation needed was estimated to best predict each expert listener's data (see Fig. 2). For details of the modeling, see [16].

The listener-specific binaural loudness-summation values [$g$ in Eq. (1)] are listed in Table 2. The individual values are taken from [16], where the modeling was based

Table 2. Amount of binaural loudness summation (binaural gain in dB) estimated from data with individual (Ind.; [16]) and generic (Gen.; present data) HRTFs at an overall level of 65 dB SPL.

| | $f_c = 1$ kHz | | $f_c = 5$ kHz | | Best Fit Across $f_c$ | |
|---|---|---|---|---|---|---|
| Subject | Ind. | Gen. | Ind. | Gen. | Ind.* | Gen. |
| IA | 0.1 | 2.7 | 0.7 | 1.7 | 0.1 | 2.0 |
| RB | 4.0 | 8.7 | 10.0 | 8.3 | 9.1 | 8.4 |
| SC | 13.1 | 20.4 | 3.1 | 4.8 | 3.8 | 5.2 |
| WE | 0.1 | 2.3 | 4.9 | 3.5 | 2.8 | 3.2 |
| WS | 1.7 | 2.4 | 0.1 | 0.3 | 0.7 | 1.3 |
| Average data—expert | | | | | 2.6 | 3.5 |
| Averaged data—naive | | | | | | 5.2 |

* The best fit included a third $f_c$ at 0.4 kHz.

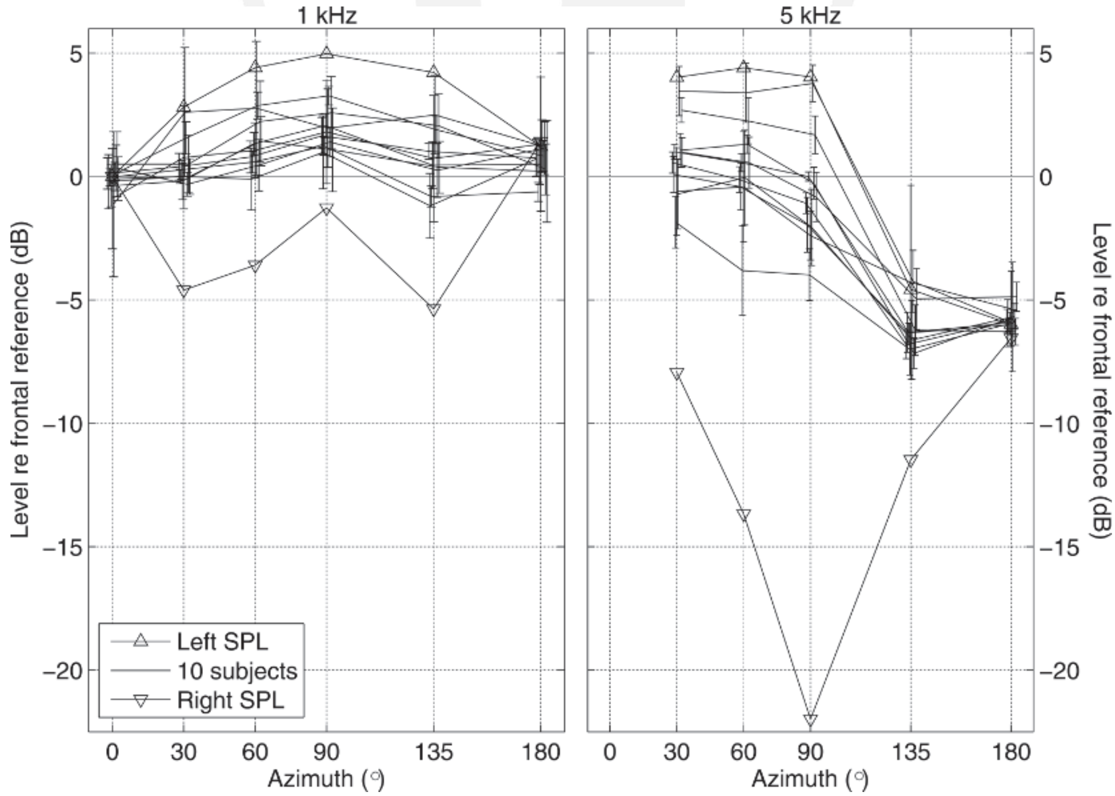


Fig. 3. Directional loudness sensitivities depicted with 95% confidence intervals for ten naive listeners and normalized at-ear SPLs derived from generic HRTFs.

on individual HRTFs, and the generic values are based on the modeling of the present data. The modeling was also carried out by aggregating the data across center frequencies, and finally, by averaging the directional loudness sensitivities across listeners for both groups and aggregating over center frequencies (the rightmost column and the two bottom rows in Table 2, respectively).

As seen in Table 2, there is considerable variation in the amount of binaural loudness summation across listeners. This is true both for the data obtained using individual HRTFs and for the present data utilizing generic HRTFs in binaural synthesis, even though in the latter case the same directional effects were played back to all listeners. The variation could be due to individual weighing of the left- and right-ear inputs in binaural judgments of loudness, or to cognitive effects, such as some listeners basing their judgments to a certain extent also on the loudness at the source, in addition to the loudness of the signals at their ears. (For a detailed discussion, see [16].)

Moreover, when comparing the summation values between the two studies in Table 2 for each listener and at each center frequency, or for the best fit across center frequencies, marked similarities are observed. Even though the actual values may be numerically different, the rank order among listeners is largely preserved; see Table 2. It thus seems that with a few exceptions, the individual binaural summation process of the at-ear SPLs is retained when generic HRTFs are used.

The two bottom rows in Table 2 show the amount of summation for the averaged directional loudness-sensitivity data, aggregated over the two center frequencies. The estimate for the expert listeners comes fairly close to a 3-dB summation for both studies. For the naive listeners, for whom only the generic HRTFs were investigated, the estimate is slightly larger.[3] According to the statistical analysis, the small group difference in the binaural gain (1.7 dB) in Table 2 is most likely due to chance.

## 2.3 Binaural Loudness for Artificial-Head Measurements

Despite the apparently robust individual differences, the mean data may be utilized for developing a binaural loudness model. Note that mean data have typically been used for modeling loudness perception, even though the signals at the ears of a listener may be very individual due to the listener's pinnae, head, and torso interacting with a real sound field or his/her ears coupling to the headphones.

The model proposed in the present study is based on the following assumptions.

1) The loudness of sound fields is "energetic," that is, loudness is based on signals reaching the listener's two ears [15], [16], [23]. Thus "informational" cognitive effects on loudness, such as reported in [24], are disregarded.

2) The "transducer sensitivities" for loudness are the same for the two ears, that is, the same signal is perceived as equally loud at each ear. This has been explored in [25], and for mean data, verified to be the case.

3) The binaural gain is constant over the SPL, and therefore the summation of binaural loudness is independent of level. In the directional loudness paradigm, this seems to hold over a moderate range [16], although occasionally a level dependency has been reported in headphone investigations without spatial synthesis [26].

4) Loudness is based on the magnitude spectra at the two ears, thus neglecting potential effects of other interaural parameters on binaural loudness, such as correlation [27] or the ITD. Note that this assumption also holds for the binaural computation models for loudness of Moore and coworkers [6], [14]. It is, however, worth noting that the modified model of Moore and Glasberg [14] attempts to compute binaural loudness by modeling each of the stages of processing that actually take place in the auditory pathway, whereas the model proposed here treats those as a "black box," and merely aspires to predict the net effect of how the spatial location of a source affects binaural loudness.

Despite its limitations, the binaural loudness model proposed in the present study is applicable to measurements carried out with an artificial head in any type of sound field, resulting in diotic or dichotic at-ear signals. Since the model is based on at-ear signals, no assumptions about the sound field need to be made, in contrast to measurements carried out with a monophonic microphone.

A block diagram of the binaural loudness model for artificial-head measurements is depicted in Fig. 4. Sound pressure signals are first measured at the left and right ears of the head [$p_{2_L}(t)$ and $p_{2_R}(t)$, respectively]. The magnitude spectra of these signals [$|P_{2_L}(f)|$ and $|P_{2_R}(f)|$, respectively] are then analyzed using a given frequency resolution, such as, in one-third-octave bands.

The 3-dB summation of the SPLs corresponds to a power summation of the sound at each ear. Thus a power sum of the left- and right-ear magnitude spectra is subsequently computed. The loudness of this binaural power sum corresponds to the loudness of the same signal presented monaurally [$|P_{2_M}(f)|$], that is, to one ear only.

The standardized loudness model [1] utilizes measurements carried out with a monophonic omnidirectional microphone in the absence of a listener in either the free or the diffuse field [$|P_1(f)|$ in Fig. 4]. The listening to the two types of sound fields, however, is always binaural (the long-term spectra at the two ears being roughly the same due to symmetry). Therefore the monaural at-ear signal resulting from the power sum must be converted to the corresponding diotic signal. In case of the binaural power summation, the conversion for the linear magnitude spectra is simply a division by $\sqrt{2}$, corresponding to subtracting a binaural gain of 3 dB from the monaural at-ear SPL.

The diotic at-ear signal [$|P_{2_D}(f)|$] is then converted to the corresponding signal in the absence of the head [$|P_1(f)|$] using the inverse of an HRTF. Here the same indicence (free or diffuse) must be used for both the inverse HRTF and the sound-field type in the loudness model.

[3]Note that the larger the amount of summation, the larger the effect of the ear with the lower SPL on binaural loudness. Thus for the naive sample of listeners the (right) ear with the lower SPL input "pulls down" the binaural, directional loudness curve slightly more than for the expert listeners.

In [16] the performance of the present binaural model was contrasted with the loudness model of Moore et al. [6] for anechoic narrow-band stimuli. For externalized sound sources located in space, the present model performed considerably better than the model in [6], which overestimated the effect of the ear receiving the lower sound pressure level (see [16, fig. 8]). Since perfect summation of loudness in sones (as assumed in [6]) for the narrow-band stimuli of the present study corresponds fairly closely to a 10-dB binaural gain, the analyses rendered in Table 2, which generally show much smaller gains, also argue against this conceptualization. The same conclusion was drawn in [20] for reverberant, narrow-band, as well as for anechoic wide-band stimuli.

In Fig. 5 the predictions made by the present model are compared to those made with the model of Moore et al. [6] and also with its recent modification [14]. For all three models the inputs were the binaural at-ear spectra on one-third-octave-band resolution. For the model of the present study, a diotic signal was first determined as a power sum of the binaural inputs and then utilized in computing loudness using [6]. For the two models based on loudness summation, the binaural inputs were used as such in the predictions, the main difference being that in [6] a perfect summation of loudness is assumed, whereas in [14] the summation is less than perfect because of interaural inhibition (loudness sum and inhibited loudness sum in Fig. 5, respectively). A detailed description of utilizing at-ear spectra in predicting directional loudness data can be found in [16].

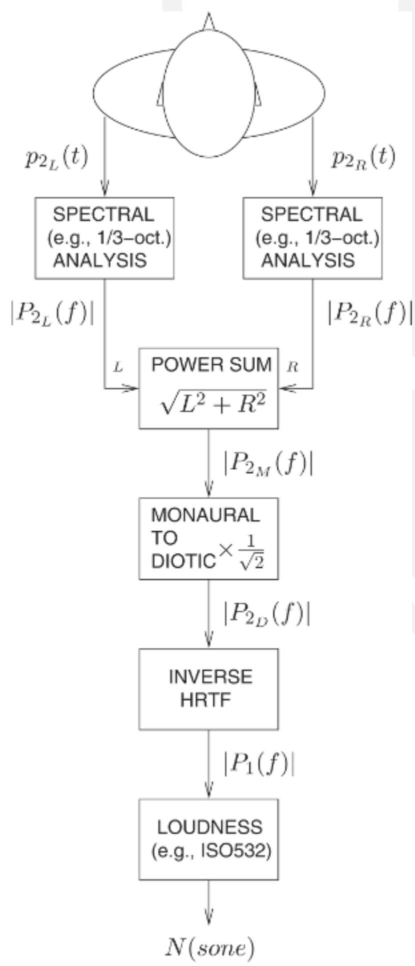


Fig. 4. Binaural loudness model for artificial-head measurements. Sound pressure signals are measured at both ears. A power sum of at-ear magnitude spectra is computed and converted to the corresponding diotic signal. An inverse HRTF is utilized to determine the sound signal in the absence of the head, which is subsequently used, for example, as an input to the standardized loudness model [1].

As seen in Fig. 5, for both center frequencies simply summing the at-ear loudnesses is clearly at odds with the measured data, which are the mean over all 15 listeners participating in the present study. At 5 kHz both the inhibited loudness sum and the power sum fare approximately equally well. At 1 kHz, where the ILD is up to 10 dB (see Table 1), the power sum, however, gives the best prediction.

It thus seems that when investigating the binaural loudness summation of directional sounds, the assumption of perfect summation in sones does not hold. The inhibited loudness sum is generally closer, whereas the power summation still makes a slightly better prediction. Based on this finding, the binaural gain for directional sounds is approximately 3 dB, corresponding to a binaural-to-monaural loudness ratio of 1.2.

## 3 CONCLUSION

The effect of the sound incidence angle on loudness was investigated in the horizontal plane using generic HRTFs in binaural synthesis.

1) Loudness matches to a frontal reference were by and large indistinguishable from earlier data obtained via individual binaural synthesis, as well as in a real sound field. They showed the same directional dependencies and interindividual variations.

2) Naive and expert listeners produced results that were statistically indistinguishable with respect to both mean values and the presence of reliable interindividual differences. This argues for the generality of the findings, and against experts developing peculiar, idiosyncratic listening strategies.

3) The fact that derived binaural summation parameters largely preserved the same rank order across individuals, no matter whether the synthesis was individual (as in earlier studies) or generic (as in the present experiment), suggests that a large part of the interindividual differences is due to the way binaural inputs are weighted and combined, and not to peculiarities in HRTF filtering.

4) The crucial feature of a model describing the mean results is a power summation of the sound pressures at the two ears. A proposal for how to implement this model when making artificial-head measurements to compute binaural loudness is outlined.

## 4 ACKNOWLEDGMENT

This research was carried out as part of the Centercontract on Sound Quality, which establishes participation in and funding of the Sound Quality Research Unit (SQRU) at Aalborg University. The participating companies are Bang & Olufsen, Brüel & Kjær, and Delta Acoustics & Vibration. Further financial support comes from the Ministry for Science, Technology, and Development (VTU), and from the Danish Research Council for Technology and Production (FTP). The authors would like to thank Pauli Minnaar for his comments on the experimental details and on an earlier version of this manuscript. Brian C. J. Moore and an anonymous reviewer are thanked for their valuable suggestions during the review of this manuscript. Brian C. J. Moore is further thanked for giving us access to a modified version of his loudness model. Finally, the participants in the listening test are thanked for their effort.

## 5 REFERENCES

[1] ISO 532, "Acoustics—Method for Calculating Loudness Level," International Organization for Standardization, Geneva, Switzerland (1975).

[2] E. A. G. Shaw, "Transformation of Sound Pressure Level from the Free Field to the Eardrum in the Horizontal Plane," *J. Acoust. Soc. Am.*, vol. 56, pp. 1848–1861 (1974).

[3] F. L. Wightman and D. J. Kistler, "Headphone Simulation of Free Field Listening. I: Stimulus Synthesis," *J. Acoust. Soc. Am.*, vol. 85, pp. 858–867 (1989).

[4] H. Møller, M. F. Sørensen, D. Hammershøi, and C. B. Jensen, "Head-Related Transfer Functions of Human Subjects," *J. Audio Eng. Soc.*, vol. 43, pp. 300–321 (1995 May).

[5] J. Blauert, *Spatial Hearing: The Psychophysics of Human Sound Localization* (MIT Press, Cambridge, MA, 1997).

[6] B. C. J. Moore, B. R. Glasberg, and T. Baer, "A Model for the Prediction of Thresholds, Loudness, and Partial Loudness," *J. Audio Eng. Soc.*, vol. 45, pp. 224–240 (1997 Apr.)

[7] B. R. Glasberg and B. C. J. Moore, "Prediction of Absolute Thresholds and Equal-Loudness Contours Using a Modified Loudness Model," *J. Acoust. Soc. Am.*, vol. 120, pp. 585–588 (2006).

[8] ANSI S3.4-2007, "Procedure for the Computation of Loudness of Steady Sounds," American National Standards Institute, New York (2007).

[9] ISO 11904-1, "Acoustics—Determination of Sound Immission from Sound Sources Placed Close to the Ear—Part 1: Technique Using a Microphone in a Real Ear (MIRE Technique)," International Organization for Standardization, Geneva, Switzerland (2002).

[10] ISO 11904-2, "Acoustics—Determination of Sound Immission from Sound Sources Placed Close to the Ear—Part 2: Technique Using a Manikin (Manikin Technique)," International Organization for Standardization, Geneva, Switzerland (2004).

[11] ISO 1999, "Acoustics—Determination of Occupational Noise Exposure and Estimation of Noise-Induced Hearing Impairment," International Organization for Standardization, Geneva, Switzerland (1990).

[12] R. P. Hellman and J. Zwislocki, "Monaural Loudness Function at 1000 cps and Interaural Summation," *J. Acoust. Soc. Am.*, vol. 35, pp. 856–865 (1963).

[13] L. E. Marks, "Binaural Summation of the Loudness of Pure Tones," *J. Acoust. Soc. Am.*, vol. 64, pp. 107–113 (1978).

[14] B. C. J. Moore and B. R. Glasberg, "Modeling Binaural Loudness," *J. Acoust. Soc. Am.*, vol. 121, pp. 1604–1612 (2007).

[15] D. W. Robinson and L. S. Whittle, "The Loudness of Directional Sound Fields," *Acustica*, vol. 10, pp. 74–80 (1960).

[16] V. P. Sivonen and W. Ellermeier, "Directional Loudness in an Anechoic Sound Field, Head-Related Transfer Functions, and Binaural Summation," *J. Acoust. Soc. Am.*, vol. 119, pp. 2965–2980 (2006).

[17] V. P. Sivonen, P. Minnaar, and W. Ellermeier, "Ef-

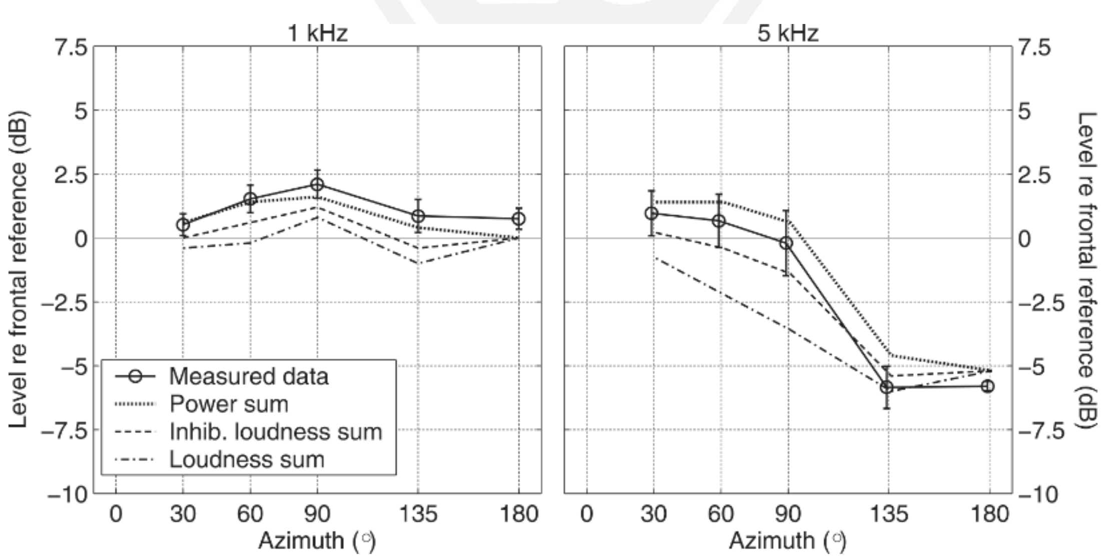


Fig. 5. Mean directional loudness sensitivities depicted with 95% confidence intervals for all 15 listeners and predictions based on power summation, inhibited loudness summation, and perfect loudness summation.

fect of Direction on Loudness in Individual Binaural Synthesis," presented at the 118th Convention of the Audio Engineering Society, *J. Audio Eng. Soc.* (*Abstracts*), vol. 53, p. 709 (2005 July/Aug.), convention paper 6512.

[18] P. Minnaar, S. K. Olesen, F. Christensen, and H. Møller, "Localization with Binaural Recordings from Artificial and Human Heads," *J. Audio Eng. Soc.*, vol. 49, pp. 323–336 (2001 May).

[19] ISO 389-1, "Acoustics—Reference Zero for the Calibration of Audiometric Equipment—Part 1: Reference Equivalent Threshold Sound Pressure Levels for Pure Tones and Supraaural Earphones," International Organization for Standardization, Geneva, Switzerland (1998).

[20] V. P. Sivonen, "Directional Loudness and Binaural Summation for Wideband and Reverberant Sounds," *J. Acoust. Soc. Am.*, vol. 121, pp. 2852–2861 (2007).

[21] H. Levitt, "Transformed Up–Down Methods in Psychoacoustics," *J. Acoust. Soc. Am.*, vol. 47, pp. 467–477 (1971).

[22] W. Jesteadt, "An Adaptive Procedure for Subjective Judgments," *Percept. Psychophys.*, vol. 28, pp. 85–88 (1980).

[23] D. W. Robinson, L. S. Whittle, and J. M. Bowsher, "The Loudness of Diffuse Sound Fields," *Acustica*, vol. 11, pp. 397–404 (1961).

[24] P. Zahorik and F. L. Wightman, "Loudness Constancy with Varying Sound Source Distance," *Nat. Neurosci.*, vol. 4, pp. 78–83 (2001).

[25] V. P. Sivonen, "Directional Loudness Perception—The Effect of Sound Incidence Angle on Loudness and the Underlying Binaural Summation," Ph.D. dissertation, Aalborg University, Aalborg, Denmark (2006).

[26] G. S. Reynolds and S. S. Stevens, "Binaural Summation of Loudness," *J. Acoust. Soc. Am.*, vol. 32, pp. 1337–1344 (1960).

[27] B. A. Edmonds and J. F. Culling, "Interaural Correlation and Loudness," *J. Acoust. Soc. Am.*, vol. 119, p. 3235 (2006).

## THE AUTHORS

V. Sivonen

W. Ellermeier

Ville Sivonen received an M.Sc. degree in electrical engineering with a major in signal processing from the University of Oulu, Oulu, Finland, in 2001. As a part of his studies he spent a semester as an exchange student at the Department of Applied Acoustics, Chalmers University of Technology, Gothenburg, Sweden. He received a Ph.D. degree from Aalborg University, Aalborg, Denmark, in 2006 with a dissertation on "Directional Loudness Perception—The Effect of Sound Incidence Angle on Loudness and the Underlying Binaural Summation."

Starting in 2001 he worked as an acoustical design engineer at Nokia Mobile Phones in Oulu and in 2002 he joined Brüel & Kjær in the Sound Quality Research Unit at Aalborg University, Denmark. Since 2006 he has been employed as an audiological engineer in the Department of Otorhinolaryngology, Oulu University Hospital, Finland. His research interests are in binaural and spatial hearing, technical audiology, and cochlear implants.

●

Wolfgang Ellermeier majored in psychology at the University of Würzburg, Würzburg, Germany. Subsequently he started his doctoral studies in the psychoacoustics laboratory of David S. Emmerich at the State University of New York at Stony Brook, NY, and in 1988 he received a Ph.D. degree in experimental psychology. He spent his postdoctoral years at the University of Regensburg, Regensburg, Germany, where he completed his Habilitation (a postdoctoral lecture qualification) on auditory profile analysis in 1995.

In 2001 he was appointed to head an interdisciplinary research program investigating basic issues of sound quality assessment and sound design at Aalborg University, Aalborg, Denmark. This Sound Quality Research Unit was cosponsored by leading Danish acoustics companies/agencies such as Bang & Olufsen, Brüel & Kjær, and Delta. Between 2001 and 2006 it produced a sizable number of publications, patents, and, most importantly, young researchers trained in both the engineering and the psychoacoustics aspects of hearing science. In 2007 he became a professor of applied cognitive psychology at the Technische Universität Darmstadt, Darmstadt, Germany. His principal research interests are psychoacoustics, psychophysical methodology, and mathematical psychology.