Using Virtual Acoustic Space to Investigate Sound Localisation

Laura Hausmann and Hermann Wagner
RWTH Aachen, Institute of Biology II
Germany

1. Introduction

It is an important task for the future to further close the gap between basic and applied science, in other words to make our understanding of the basic principles of auditory processing available for applications in medicine or information technology. Current examples are hearing aids (Dietz et al., 2009) or sound-localising robots (Calmes et al., 2007). This effort will be helped by better quantitative data resulting from more and more sophisticated experimental approaches.

Despite new methodologies and techniques, the complex human auditory system is only accessible in a restricted way to many experimental approaches. This gap is closed by animal model systems that allow a more focused analysis of single aspects of auditory processing than human studies. The most commonly used animals in auditory research are birds (barn owls, chicken) and mammals (monkeys, cats, bats, ferrets, guinea pigs, rats and gerbils).

When these animals are tested with various auditory stimuli in behavioural experiments, the accuracy (distance of a measured value to the true value) and precision (repeatability of a given measured value) of the animal's behavioural response allows to draw conclusions on the difficulty with which the animal can use the stimulus to locate sound sources. An example is the measurement of minimum audible angles (MAA) to reveal the resolution threshold of the auditory system for the horizontal displacement of a sound source (Bala et al., 2007). Similarly, one can exploit the head-turn amplitude of humans or animals in response to narrowband or broadband sounds as a measure for the relevance of specific frequency bands, as well as binaural and monaural cues or perception thresholds (e.g. May & Huang, 1995; Poganiatz et al., 2001; Populin, 2006).

The barn owl (*Tyto alba*) is an auditory specialist, depending to a large extent on listening while localising potential prey. In the course of evolution, the barn owl has developed several morphological and neuronal adaptations, which may be regarded as more optimal solutions to problems than the structures and circuits found in generalists.

The owl has a characteristic facial ruff, which amplifies sound and is directionally sensitive for frequencies above 4 kHz (Coles & Guppy, 1988). Additionally, the left and right ear openings and flaps are asymmetrically with the left ear lying slightly higher than the right one. This asymmetry creates a steep gradient of interaural level differences (ILDs) in the owl's frontal field (Campenhausen & Wagner, 2006). These adaptations to sound localisation are one of the reasons why barn owl hearing was established as an important model system during the last decades.

This chapter will focus on the application of a powerful technique for the investigation of sound processing, the virtual auditory space technique. Its basics, relevance and applications will be discussed for human listeners as well as in barn owls, supplemented by a comparison with other species.

Sound localisation is based on extraction of physical cues of the sound reaching the eardrums. Such physical cues are the monaural spectral properties of the sound as well as differences between the sounds reaching the left and rights ears, leading to binaural cues. These cues vary systematically with sound source position relative to an animal's head.

A sound originates from a source and travels through the air until it reaches the eardrums of a listener. Several distortions (reflection, attenuation) are imposed on the sound along its path. Sound parameters may be measured at or close to the eardrum. The comparison of the measured sound at the eardrum with the sound emitted by the source allows for a determination of the distortion and is unique for each individual. The resulting transfer functions are called the head-related transfer functions (HRTFs) referring to the major influence of head shape in the process of distortion. HRTFs carry information about the location of a sound source. Note that the term HRTF refers to the frequency domain, whereas one speaks of the head-related impulse response (HRIR) when the signal is represented in the time domain. Both signals may be transformed from one domain to the other by means of a Fourier transformation (Blauert, 1997).

In monaural spectra the large decreases in amplitude, termed notches, carry information about sound source direction due to their systematic directional variations. Animals and humans use this information during sound localisation, in particular when resolving front-back confusions (Gardner & Gardner, 1973, Hebrank & Wright, 1974). The comparison of the HRTFs measured at the two ears yields two major binaural parameters: Interaural time difference (ITD) and interaural level difference (ILD). The ITD depends on the angle of incidence as well as on the distance between the two ears. This cue may be further divided into envelope and carrier ITDs.

Envelope ITDs occur specifically at the onset and end of a sound and are then called onset ITDs, whereas ITDs derived from the carrier occur in the ongoing sound and are, therefore, called ongoing ITDs. ITD is constant along a circle describing a surface of a cone, termed "cone of confusion" because for sound sources along this cone surface, the identical ITDs do not allow unambiguous localisation of narrowband stimuli (cf. Blauert, 1997). This leads to ambiguities with respect to front and back, and, therefore, this cone is also known as cone of confusion. ILDs arise from the frequency and position depended attenuation of sound by the pinna, the head and the body that typically differs between the two ears.

2. Investigation of sound localisation - current approaches and problems

The simplest approach to find out more about the relevance of the sound parameters is to replay natural sounds from a loudspeaker and measure the subject's reaction to the sounds. These experiments are typically carried out in rooms having walls that strongly suppress sound reflections. If the distance between source and listener is large enough, we have a free-field situation, and the approach is called free-field stimulation.

Free-field sounds have a major disadvantage: the physical cues to sound location cannot be varied independently, because a specific ITD resulting from a given spatial displacement of the sound source also involves a change in the ILD and the monaural spectra. This renders it difficult or even impossible to derive the contribution of single cues to sound localisation.

On the other hand, free-field sounds contain all relevant cues a subject may use in behaviour. Although free-field stimulation allows for an investigation of how relevant specific sound characteristics are, such as the frequency spectrum, the limits of this technique are obvious. Since this chapter focusses on the virtual space technique, we will not review the results from the numerous studies dealing with free-field stimulation.

One way to overcome the problems inherent in free-field stimulation is the dichotic stimulation via headphones, allowing the independent manipulation of ITDs or ILDs in the stimulus. Dichotic stimulation was used to prove that humans use ITDs for azimuthal sound localisation for frequencies up to 1.5 kHz and ILDs for frequencies above 5 kHz (reviewed in Blauert, 1997). The upper frequency limit for ITD extraction seems to be determined by the ability of neurons to encode the phase of the signal's carrier frequency, which in turn is necessary to compare phase differences between both ears.

The lower border for ILD extraction, likewise, seems to be related to the observation that the head of an animal only creates sufficiently large ILDs above a certain frequency. These conclusions are supported by data from animals such as the cat, the ferret, monkey and the barn owl (Koeppl, 1997; Koka et al., 2008; Moiseff and Konishi 1981; Parsons et al., 2009; Spezio et al., 2000; Tollin & Koka, 2009). The use of both ITDs and ILDs in azimuthal sound localisation is known as duplex theory (Blauert, 1997; Macpherson & Middlebrooks, 2002; Rayleigh, 1907).

In the barn owl, the filtering properties of the facial ruff together with the asymmetrical arrangement of the ear openings and the preaural flaps in the vertical plane cause ILDs to vary along an axis inclined to the horizontal plane. This allows the barn owl to use ILDs for elevational sound localisation (Moiseff 1989, Campenhausen & Wagner 2006; Keller et al. 1998). In contrast, mammals use ILDs for high-frequency horizontal localisation (reviewed in Blauert, 1997).

The ability of the owl's auditory neurons to lock to the signal's phase within almost the whole hearing range (Köppl, 1997) – again in contrast to most mammals – together with the use of ILDs for elevational localisation is one of the reasons that make the barn owl interesting for auditory research, despite the mentioned differences to mammals.

With earphone stimulation, binaural cues can be manipulated independently. For example, the systematic variation of either ITDs or ILDs while keeping the other cue constant is nowadays a commonly used technique to characterise neuronal tuning to sound or to investigate the impact of the cues on sound localisation ability (reviewed in Butts and Goldman, 2006). Another example is the specific variation of ongoing ITDs, but not onset ITDs (Moiseff & Konishi, 1981; von Kriegstein et al., 2008) or a systematic variation of the degree of interaural correlation in binaurally presented noises (Egnor, 2001).

Although dichotic stimulation helped to make progress in our understanding of sound localisation, one disadvantage of this method is that human listeners perceive sources as lying inside the head (Hartmann & Wittenberg, 1996; Wightman & Kistler, 1989b) rather than in outside space. Consequently, when only ITDs or ILDs are introduced, but no spectral cues, the sound is "lateralised" towards a direction corresponding to the amplitude of ITD or ILD, respectively. For human listeners, this may yield a horizontal displacement of the sound image sufficient for many applications. However, both vertical localisation and distance estimation are severely hampered, if possible at all. In contrast, a free-field sound source or an appropriately simulated sound is really "localised".

This means that dichotic stimuli do not contain all physical cues of free-field sounds. A method to overcome the problems of dichotic stimulation as described so far, while

preserving its advantages, is the creation of a virtual auditory space (VAS), the method and implementation of which is the topic of this chapter. The work of Wightman and Kistler (1989a,b) and others showed that free-field sources could be simulated adequately by filtering a sound with the personal head-related transfer functions (HRTFs). Bronkhorst (1995) reported that performance degraded when subjects were stimulated with very high frequency virtual sounds. This observation reflects the large difficulties in generating veridical virtual stimuli at high-frequencies.

3. The virtual space technique

While dichotic stimulation does not lead to externalisation of sound sources, the use of HRTF-filtered stimuli in a virtual auditory space does (Hartmann & Wittenberg, 1996; Plenge, 1974; Wightman & Kistler, 1989b). For that reason, numerous attempts have been made to develop virtual auditory worlds for humans or animals. The main goal there is that virtual sound sources in VAS should unambiguously reflect all free-field sound characteristics.

A second goal, especially in human research, is to create virtual auditory worlds that are universally applicable across all listeners. This requires a trade-off between the realistic simulation of free-field characteristics and computational power, that is, one wants to discard nonessential cues while preserving all relevant cues. For that purpose, knowledge is required on which cues are the relevant cues for sound localisation and which are not.

The method involved in creating VAS originated in the 1950ies when systematic experiments using artificial head manikins were undertaken (reviewed by Paul, 2009). However, it is only the computational power developed within the last two decades that allows for elaborate calculations and manipulations of virtual auditory stimuli.

Measuring the HRTFs is usually done by inserting small microphones into the ear canals of the subject, as sketched in Figure 1. The sound impinging on the eardrum is measured. Sound is replayed from a free-field loudspeaker (see Fig. 1a). The loudspeaker signal should contain all relevant frequencies within the hearing range of the subject. It has been shown that measurement at or close to the eardrum is adequate, because the measured signal contains the important information (Wightman & Kistler, 1989a,b).

When the signal arrives at the eardrum, it has been filtered by the outer ear, the head and the body of the subject. The amplitude and phase spectra at the eardrum represent the HRTF for the given ear and the respective position. The monaural spectrum of a specific sound at a given position may be obtained by filtering the sound with the respective HRTF. ITDs and ILDs occurring at a given position are derived by comparing the respective measured HRTFs at the two ears.

The procedure of replaying a free-field sound and recording the resulting impulse response at the subject's eardrum is usually carried out for representative spatial locations, i.e., the free-field speaker is positioned at a constant distance at different locations, for example by moving it along a circular hoop (Fig. 1a). In this way, the desired spatial positions in both azimuthal and elevational planes may be sampled (Fig. 1b+c).

One may use click stimuli (Dirac pulses, see Blauert, 1997) as free-field sounds. In this way, a broad range of frequencies can be presented in a very short time. However, such stimuli do not contain much energy and as a consequence have to be repeated many times (typically 1000) in order to increase the energy provided to the listener (reviewed in Blauert, 1997; see also Poganiatz & Wagner, 2001; Poganiatz et al., 2001).
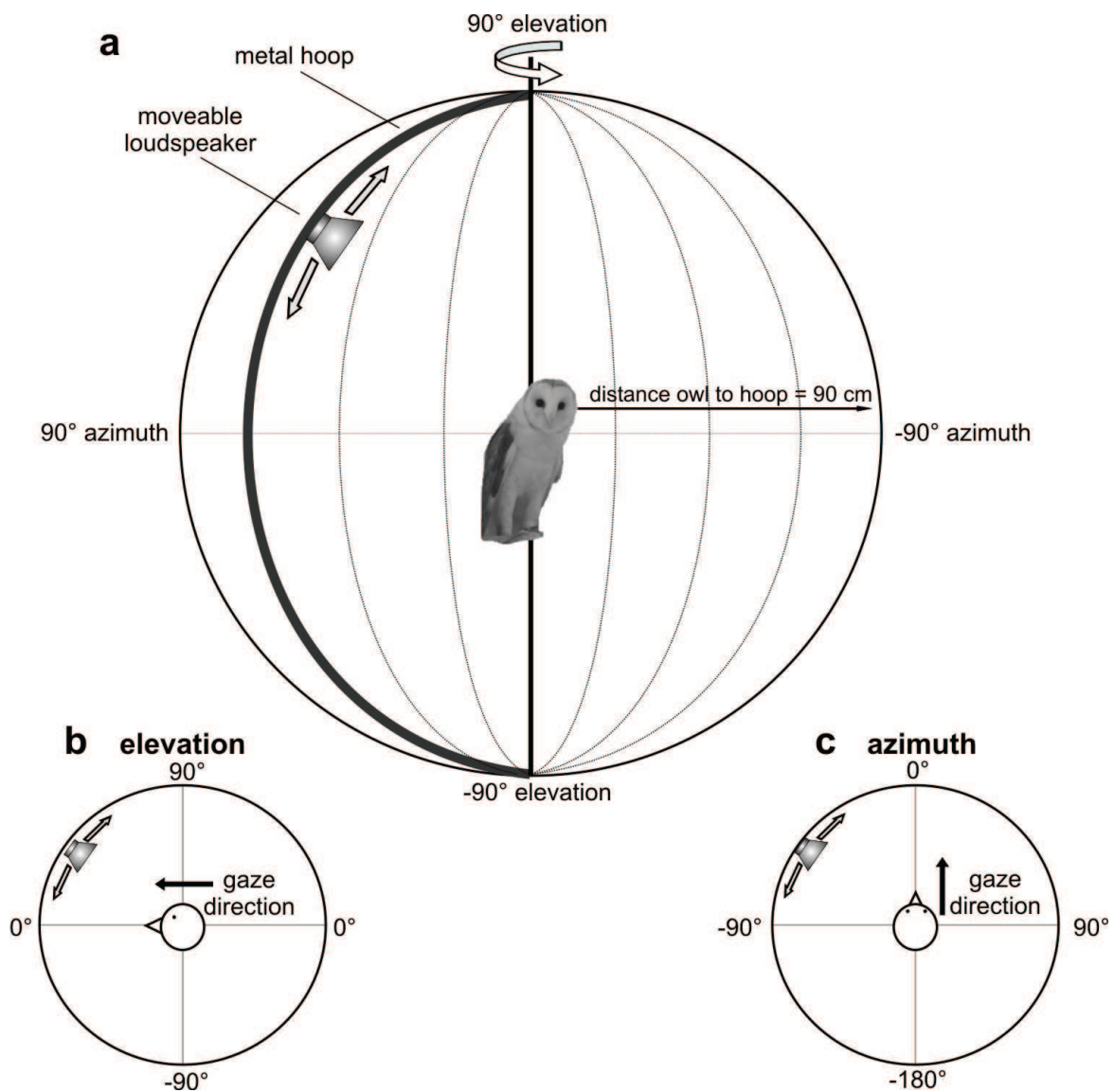
After Hausmann et al. (2010).

Fig. 1. **Schematic of a setup for HRTF measurements.** A) During HRTF measurements, the anesthetised owl is fixated with the help of a cloth jacket in the center of a metal hoop. A loudspeaker can be moved upwards or downwards along the hoop, allowing variation of the vertical stimulus angle as shown in panel B). The hoop can be rotated about its vertical axis, which allowed positioning of the hoop at various azimuthal values, with 0° being directly in front of the owl as shown in panel C).

Other stimuli are so-called sweep signals, which run from low to high frequencies or vice versa in a given time interval. For example, logarithmically rising sweeps have successfully been used for HRTF recordings in the owl (Campenhausen & Wagner, 2006; Hausmann et al., 2010). Such sweep signals have the advantage that a small number of repetitions of sound emissions suffices to yield reproducible measurements, while containing energy in all desired frequencies within the subject's hearing range.

The quality of HRTF recordings measured with both types of stimuli is comparable, as demonstrated by similar shape of HRTFs and localisation performance in the owl for HRTF-filtered stimuli recorded either during application of click noise (Poganiatz & Wagner, 2001; Poganiatz et al., 2001) or of sweeps (Campenhausen & Wagner, 2006; Hausmann et al., 2009; Hausmann et al., 2010). Short click stimuli have also commonly been used for HRTF measurements in other animals and humans, leading to localisation performance comparable to free-field stimulation (Delgutte et al., 1999; Musicant et al., 1990; Tollin & Yin, 2002; Wightman & Kistler, 1989b).

The impulse responses recorded at the subject's eardrum are influenced by the individual transfer functions not only of the subject itself, but also of the equipment used for the recordings such as the microphones, loudspeaker and hardware components. In order to provide an accurate picture of the transfer characteristics, all impulse responses recorded with the subject (specific for each azimuthal (α) and elevational (ε) position) have to be corrected for the transfer characteristics of the system components (T_{sys}).

The correction can be easily done by transforming each impulse response into the frequency domain via Fast Fourier transformation (FFT), and then divide each subject-specific FFT ($H_{\alpha\varepsilon}$) by the reference measurement recorded for the system components including the microphone, but without the subject (T_{sys}) following equation 1.

$$H_{\alpha\varepsilon} = \frac{H_{\alpha\varepsilon} \cdot T_{sys}}{T_{sys}} \tag{1}$$

In both behavioural and electrophysiological experiments, HRTF-filtered stimuli open a wide range of possible manipulations to analyse single characteristics of sound processing and allow for prediction of localisation behaviour based on HRTF characteristics (humans: Getzmann & Lewald, 2010; Hebrank & Wright, 1974; cat: Brugge et al., 1996; May & Huang, 1996; guinea pig: Sterbing et al., 2003; owl: Hausmann et al., 2009; Poganiatz et al., 2001; Witten et al., 2010).

Virtual auditory worlds can be created for all animals whose HRTFs are measured. An advantage of using the barn owl rather than many mammalian species is that the owl performs saccadic head-turns towards a sound source when sitting on a perch (Knudsen et al., 1979), while the eyes or pinnae can barely be moved (Steinbach, 1972). In contrast, many mammals may move their eyes and pinnae. This allows for example cats or monkeys to locate sound sources even with restrained head to a certain extent (Dent et al., 2009; Populin, 2006; Populin & Yin, 1998). The owl's saccadic head-turn response allows to use the owls' head-turn angle as a measure for the perceived sound source location (Knudsen & Konishi, 1978).

The next section will review how HRTF-filtered stimuli have been implemented in the barn owl as a model system to tackle specific issues of sound localisation which are also relevant for human sound localisation.

4. Virtual auditory space and its applications in an auditory specialist

One of the first applications of VAS for the barn owl was the work of Poganiatz and coworkers (2001). The authors conducted a behavioural study in which individualised HRTFs of barn owls were manipulated in that the broadband ITD was artificially set to a specific value, irrespective of their natural ITD. This artificial ITD was either -100 µs, corresponding to a position of approximately -40° of azimuth (by definition left of the

animal) based on a change of 2.5 μs per degree (Campenhausen & Wagner, 2006), or to +100 μs, corresponding to +40° of azimuth (by definition right of the animal).

All other cues such as the ILD and monaural spectra were preserved. That is, the stimuli were ambiguous in that the ITD might point towards a different hemisphere than did all the remaining cues. The authors of the study predicted that the owl should turn its head towards the position encoded by the ITD if the ITD was the relevant cue for azimuthal sound localisation. Similarly, the owl should turn towards the position encoded by ILD and monaural spectra if these cues were relevant for azimuthal localisation.

When these manipulated stimuli were replayed via headphones to the owls, the animals always turned their heads towards the position that was encoded by the ITD and not by the remaining cues. From these findings, Poganiatz et al. (2001) concluded that the owls used exclusively the ITD to determine stimulus azimuth. As we will show below, this may hold for a large range of auditory space. However, the resolution of spatially ambiguous ITDs in the frontal and rear hemispheres requires further cues.

The same approach of manipulating virtual stimuli was used for investigating the role of ILD for elevational sound localisation by setting the broadband ILD in HRTF-filtered stimuli to a fixed value (Poganiatz & Wagner, 2001). Such experiments showed that barn owls' elevational head-turn angles depend partly, but not exclusively on ILDs. The role of ILDs and other cues for elevational localisation will be tackled in more detail below.

Thus, these earlier studies did not resolve the cues needed to resolve front-back confusion or localisation of phantom sources that occur at positions that can be predicted from a narrowband sound's period duration and ITD. Both phenomena are commonly known problems in humans especially for localisation in the median plane (Gardner & Gardner 1973; Hill et al. 2000; Wenzel et al. 1993; Zahorik et al. 2006). Furthermore, it is still unclear which cues, apart from broadband ILD, contribute to elevational sound localisation. The owl's ability to locate sound source elevation is essentially based on its ear asymmetry and facial ruff. Going a step further and utilise the morphological specialisations of the barn owl, a possible application for humans might thus be to mimic an owl's facial ruff to achieve better localisation performance in humans.

We extended the use of HRTF-filtered stimuli to answer some of the above raised questions. The method introduced by Campenhausen & Wagner (2006) allowed us to measure the influence of the barn owl's facial ruff for a closer analysis of the role of external filtering as well as of the interplay of the owl's asymmetrically placed ears with the characteristically heart-shaped ruff.

Using VAS enabled us to analyse the contribution of the facial ruff and the asymmetrically placed ear openings independently from each other, an important aspect if one wants to implement the owls' specialisations for engineering of sound localisation devices.

Virtual ruff removal (Hausmann et al., 2009) was realised by recording HRTFs for anesthetized barn owls a) with intact ruff of the animal that was tested in behavioural experiments later on (individualised HRTFs), b) for a reference animal with intact ruff (reference owl, normal non-individualised HRTFs) and c) for the same reference animal after successive removal of all feathers of the facial disk, leaving only the rear body feathers intact (see also Campenhausen & Wagner, 2006), named "ruffcut" condition.

The advantage of simulating ruff removal rather than actually removing the ruff of the behaving owls consisted in a better reproducibility of stimulus conditions over the course of the experiments, as the feathers regrow after removal and thus stimulus conditions change. Furthermore, responses to the stimuli were comparable between subjects since the stimulus

conditions were equal for all three owls included in the study. And third, virtual ruff removal is a more animal-friendly approach than real removal of the feathers, since the behaving owls are not hampered in their usual localisation behaviour, as would be the case if one actually removed their facial ruff.

The measurements yielded three sets of HRTFs. In behavioural experiments, broadband noise (1-12 kHz) was filtered with these HRTFs to simulate ruff removal for three owls. The former two stimulus conditions with intact ruff were required for comparison of the normal localisation performance with that in response to simulated ruff removal. In parallel, the changes of binaural cues were analysed. Virtual ruff removal resulted not only in a reduction of the ITD range in the periphery (Fig. 2B), but also in a corresponding reduction of azimuthal head-turn angles for ruffcut versus normal stimuli (Fig. 2A). That is, the virtual ruff removal induced a change in the localisation behaviour that could be correlated with the accompanying changes of the ITD as the relevant cue for azimuthal localisation (see also Poganiatz et al., 2001).

Virtual ruff removal influenced behaviour in two major ways. First, it caused the ILDs in the frontal field to become smaller, and the ILDs did no longer vary with elevation, as is the case in HRTFs recorded with intact ruff (see also Keller et al., 1998). Correspondingly, the owls lost their ability to determine stimulus elevation.

Second, while owls having a normal ruff could discriminate stimuli coming from the rear from those coming from the front even if the stimuli had the same ITD (Hausmann et al., 2009), this ability to distinguish between front and back in HRTFs having the same ITD was lost after virtual ruff removal. This finding implies that the ITD is indeed the only relevant cue for azimuthal localisation in the frontal field, as suggested by Poganiatz et al. (2001), but stimulus positions with equal ITD in the front and in the rear, respectively, may not be discriminated based on the ITD alone.

Hence, the ruff provides cues other than ITD to resolve position along the cone of confusion (see Blauert, 1997). Potential candidates for the cues provided by the ruff for front-back disambiguation are ILDs and monaural spectral cues, both of which are altered after ruff removal (Campenhausen & Wagner, 2006; Hausmann et al., 2009).

The role of ILDs and spectral cues can be investigated by keeping the ILD in virtual auditory stimuli constant, while the ITD and spectral cues vary with location according to their natural amplitude. Such an approach was pursued in an earlier study by Poganiatz & Wagner (2001), where ILDs in virtual acoustic stimuli were set to a fixed value of either -6 dB (left ear louder) or +6 dB in the frequency range from 4 to 10 kHz. In response to those manipulated stimuli, the owls responded with a positive elevational head-turn to the +6 dB stimuli and with a negative head-turn to the -6 dB stimuli. When the stimulus ILD was set to +6 dB, the owls' head-turn was directed to a relatively constant elevational position. In response to stimuli whose ILD was set to -6 dB, however, the elevational head-turn amplitude was constant at positive stimulus azimuth but increased with incrementally negative stimulus angle, or vice versa. The localisation behaviour depended on the stimulus position, meaning that the elevational localisation was not exclusively defined by the mean broadband ILD.

The study of Poganiatz & Wagner (2001) argued against monaural spectral cues as this additional cue, since the spectra had been preserved according to the natural shape. However, the owls' elevational head-turn angles did not follow a simple and clear relationship, which renders conclusions on the contribution of single binaural and monaural cues difficult. The question remained of whether owls needed broadband ILDs to determine

the elevation of virtual sound sources, or whether ILDs in single frequency bands could be used as well.
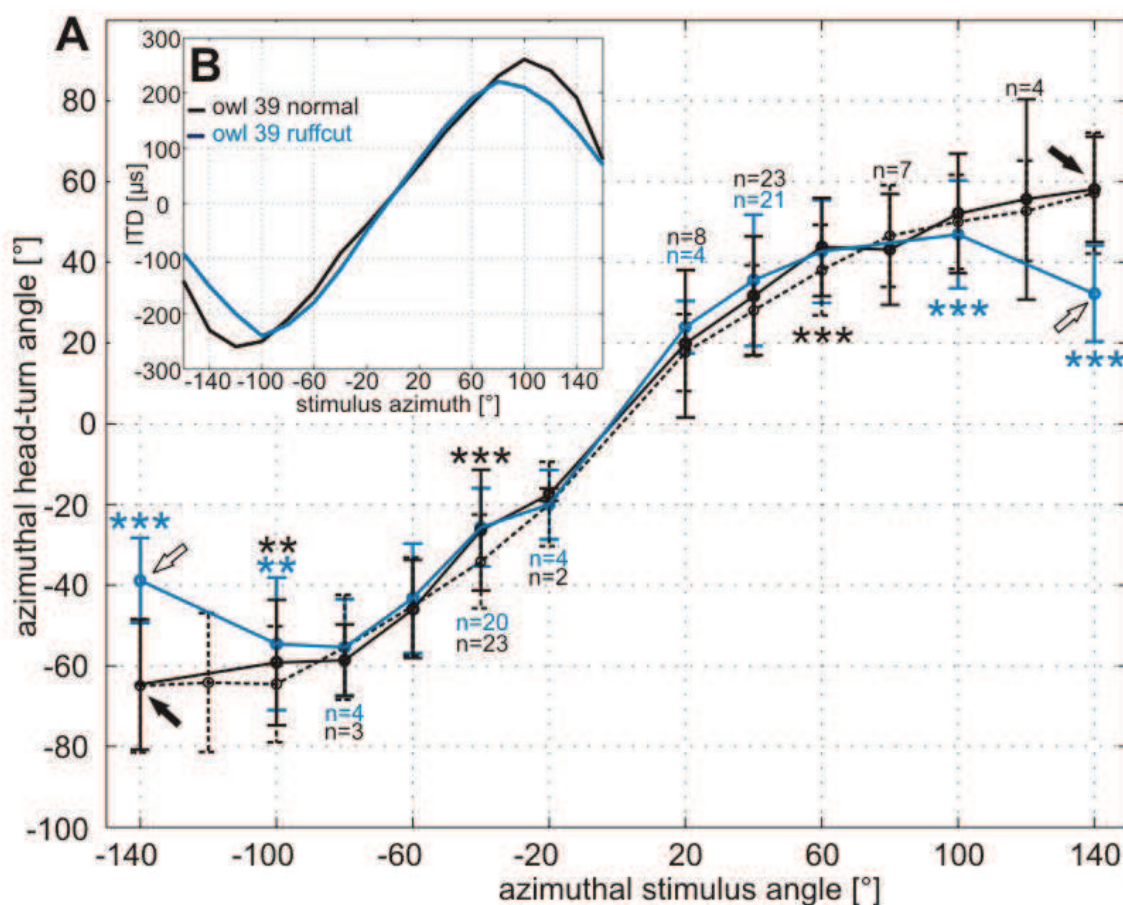
After Hausmann et al. (2009)

Fig. 2. **ITDs and azimuthal head-turn angle under normal and ruffcut conditions.** A) The azimuthal head-turn angles of owls in response to azimuthal stimulation (x-axis) with individualised HRTFs (dotted, data of two owls), non-individualised HRTFs of a reference animal (normal, black, three owls) and to the stimuli from the reference owl after ruff removal (ruffcut, blue, three owls). Arrows mark ±140° stimulus position in the periphery, where azimuthal head-turn angle decreased for stimulation with simulated ruff removal, in contrast to stimulation with intact ruff (individualised and reference owl normal) where they approach a plateau at about ±60°. Significant differences between stimulus conditions are marked with asterisks depending on the significance level (**p<0.01, ***p<0.001) in black (individualised versus reference owl normal) respectively in blue (reference owl normal versus ruffcut). Each data point includes at least 96 trials, unless indicated otherwise by the number of trials (n). B) The ITD in µs contained in the HRTFs at 0° elevation is plotted against stimulus azimuth in degree for the reference owl normal (black) and ruffcut (blue). Note the sinusoidal course of the ITD and the smaller ITD range after ruff removal. ITDs decrease at peripheral azimuths for both intact and removed ruff.

Due to the complex variations of ILDs with both elevation and azimuth in the barn owl, the influence of specific cues on elevational localisation is difficult to investigate. Furthermore, as we have just seen, elevational localisation is influenced by cues other than the ILD, which stands in contrast to the exclusive dependence of azimuthal head-turn angle on ITDs at least in the frontal field (but see Hausmann et al. 2009 for azimuthal localisation in the rear).

Since ILDs are strongly frequency-dependent, the next step we took was the stimulation of barn owls with narrowband stimuli to investigate elevational localisation, so to narrow down the range of relevant frequencies used for elevational localisation. Again, the virtual space technique allowed for a manipulation of stimuli in which ILD cues are preserved for each narrow frequency band, while spectral cues are sparse.

This stimulus configuration may answer the question of whether owls can make use of narrowband spectral cues. If they do, their localisation behaviour should resemble that for non-manipulated stimuli of the same frequency. On the other hand, if monaural narrowband spectra cannot be used, the owls' localisation behaviour for stimuli with virtually removed ILD should differ from that to stimuli containing the naturally occurring ILD. We tested barn owls in the proposed stimulus setup.

We first created narrowband noises. The ILD in such stimuli was then set to a fixed value of zero dB ILD, similar to the approach of Poganiatz & Wagner (2001), without changing the remaining localisation cues. In response to those stimuli, barn owls exhibited elevational head-turn angles that varied with stimulus elevation, indicating that narrowband ILD was sufficient to discriminate sound source elevation.

In addition, the owls were able to resolve azimuthal coding ambiguities, so-called phantom sources, when the virtual stimuli contained ILDs, but not when the ILD was set to zero. This finding implied that owls may use narrowband ILDs to determine the hemisphere a sound originates from, or in other word, to resolve coding ambiguities. The formation of phantom sources will be reviewed in more detail in the following.

5. Coding ambiguities

Coding ambiguities arise if one parameter occurs more than once in auditory space. Coding ambiguities lead to the formation of phantom sources. Many animals perceive phantom sound sources (Lee et al. 2009; Mazer, 1998; Saberi et al., 1998, 1999; Tollin et al. 2003). The main parameter for azimuthal localisation in the frontal hemisphere is the ITD. In the use of ITD, ambiguities occur for narrowband and tonal stimuli when the period duration of the center frequency or tone is shorter than the time that the sound needs to travel around the head of the listener.

For narrowband and tonal stimuli, ITD is equivalent to the interaural phase difference. The sound's phase at one ear can either be matched with the preceding (leading) phase or with the lagging phase at the other ear. Both comparisons may yield valid azimuthal sound source positions if the ITD corresponding to the interaural phase difference of the stimulus falls within the ITD range the animal can experience. For example, a 5 kHz tone has a period duration of 200 μs. In the owl, stimulation from -40° azimuth (i.e., 40° displaced to the left side of the owl's midsagittal plane) corresponds to about -100 μs ITD, based on a change of about 2.5 μs per degree (Campenhausen & Wagner, 2006). In this case, the 5 kHz tone is leading at the owl's left ear by 100 μs, which would result in calculation of the correct sound source azimuth.

However, it is also possible to match the lagging phase at the left ear with the next leading phase at the right ear, resulting in a phantom source at +40° azimuth in the right

hemisphere. A study by Saberi et al. (1998) showed that in case of ambiguous sound images, the owls either turned their heads towards the more frontal sound source, be it a real or a phantom source, or else they turned towards the more peripheral sound source.

With increasing stimulus bandwidth, the neuronal tuning curves for the single frequencies are still cyclic and, therefore, ambiguous as we have just seen. However, there is always one peak at the real ITD, while the position of the phase multiples (side peaks) is shifted according to the period duration, which varies with frequency (Wagner et al., 1987).

Integration, or summation, across a wider band of frequencies thus yields a large peak at the true ITD and smaller side peaks. Hence, for wideband sounds, integration across frequencies reduces ITD coding ambiguities via side-peak suppression in broadband neurons (Mazer, 1998; Saberi et al., 1999; Takahashi & Konishi 1986; Wagner et al., 1987). Sidepeak suppression reduces the neuronal responses to the phantom sources (corresponding to the phase equivalents of the real ITD) compared to the response to the real ITD. Mazer (1998) and Saberi et al. (1999) showed in electrophysiological and behavioural experiments that a bandwidth of 3 kHz was sufficient to reduce phase ambiguities and to unambiguously determine the real ITD.

Thus, in many cases, a single cue does not allow to determine the veridical spatial position unambiguously. This was also shown by electrophysiological recordings of the spatial receptive fields for variations in ILD, but constant ITD (Euston & Takahashi, 2002). In this stimulus configuration, ILDs exhibited broad regions where the ILD amplitude was equal, thus ambiguous.

Aross-frequency integration also reduces such ILD ambiguities, which are based on the response properties of single cells for example in the external nucleus of the inferior colliculus (ICX). Such neurons respond to a narrowband stimulus having a given ITD but varying ILDs with an increased firing rate at wide spatial regions. That is, this neuron's response does not code for a single spatial position, but for a variety of positions which cannot be distinguished based on the neuronal firing rate alone. Only the combination of a specific ITD with a specific ILD results in unambiguous coding of spatial positions and results in the usual narrowly restricted spatial receptive fields (Euston & Takahashi, 2002; Knudsen & Konishi, 1978; Mazer, 1998). In the case of the owl, the natural combinations of ITD and ILD that lead to sharply tuned spatial receptive fields are created by the characteristic filtering properties of the ruff (Knudsen & Konishi, 1978).

To summarise the preceding sections, the ruff plays a major role for the resolution of coding ambiguities. However, it is only the interaction of the ruff with the asymmetrically placed ear openings and flaps that creates the unique directional sensitivity of the owl's auditory system (Campenhausen & Wagner, 2006; Hausmann et al., 2009). This finding should be taken into account if one wants to mimic the owl's facial ruff in engineering science

It is interesting that humans can learn to listen and localise sound sources quite accurately when provided with artificial owl ears (Van Wanrooij et al., 2010). The human subjects in that study wore ear moulds that were scaled to the size of the listener, during an uninterrupted period of several weeks. The ear moulds were formed to introduce asymmetries just as observed in the barn owl. The ability of the subjects to localise sound sources in both azimuth and elevation was tested repeatedly to measure the learning plasticity in response to the unusual hearing experience. At the beginning of the experiments, localisation accuracy in both planes was severely hampered. After few weeks, not only azimuthal localisation performance was close to normal again, but also elevational localisation of broadband sounds, and only these. That is, the hearing performance

apparently underlies a certain plasticity, meaning that a listener can learn to locate sounds accurately even with unfamiliar cues, which opens interesting fields of application.

Similar plasticity was observed in ferrets whose ears were plugged, who learned to localize azimuthal sound sources accurately again after several weeks of training (Mrsic-Flogel et al. 2001).

These experiments underline that auditory representations in the brain are not restricted to individual species, but rather that humans or animals can learn new relationships between a specific combination of localisation cues and a specific spatial position. Despite this plasticity, in everyday applications, it may not seem feasible when listeners need a long period of time to learn a new relationship. However, when familiarity to sound spectra is established via training, localisation performance is improved, a fact that is amongst others exploited for cochlear implant users (Loebach & Pisoni 2009).

Now what are the implications of the above revised findings for the creation of auditory worlds for humans?

First, it is crucial to preserve low-frequency ITDs in virtual stimuli, since these are not only required, but also seem to be dominant for azimuthal localisation (reviewed in Blauert, 1997 for humans; owl: Witten et al., 2010).

Second, ILD cues are necessary in the high-frequency range for accurate elevational localisation in many animal species including humans (e.g. Blauert, 1997; Gardner & Gardner, 1973; Huang & May, 1996; Tollin et al., 2002; Wightman & Kistler, 1989b). In the low-frequency range, the small attenuation by the head results in only small ILDs that hardly vary with elevation (human: Gardner & Gardner 1973; Shaw 1997; cat: May & Huang 1996; monkey: Spezio et al., 2000; owl: Campenhausen & Wagner, 2006; Keller et al., 1998; Hausmann et al., 2010), which makes ILDs a less useful cue for low-frequency sound localisation. However, a study by Algazi et al. (2000) claims that human listeners could determine stimulus elevation surprisingly accurate even when the stimulus contained only frequencies below 3 kHz, although the listeners' performance was degraded compared to a baseline condition with wideband noise. These two cues allow for relatively accurate determination of sound source position in the horizontal plane in humans (see Blauert 1997). However, ITD and ILD variations alone may as well be introduced to dichotic stimuli presented via headphones, without the requirement of measuring the complex individual transfer functions. That is, as long as pure lateralisation (Plenge 1974; Wightman & Kistler 1989a,b) outside the median plane suffices to fulfil a given task, it should be easier to introduce according ITDs and ILDs to the stimuli. However, for a sophisticated simulation of free-field environments, as well as for unambiguous allocation of spatial positions to the frontal and rear hemispheres, one should use HRTF-filtered stimuli. This holds the more as ILD cues seem to be required for natural sounding of virtual stimuli in human listeners (Usher & Martens, 2007).

Since an inherent feature of HRTFs is the fact that they are individually different, the question arises of whether HRTF-filtered stimuli are feasible for general application, that is, if they can in some way be generalised across listeners to prevent the necessity of measuring HRTFs for each potential listener individually. The latter would be critical anyway because for numerous applications, the future user of the virtual auditory space is unknown in advance. The issue of the extent to which HRTFs can be used for stimulation of different subjects without loosing informational content will be tackled in the following section.

6. Localisation with non-individualized HRTFs – does everybody hear differently?

Meanwhile, there are many studies that attempt to generate sets of "universal" HRTFs, which create the impression of free-field sound sources across all (human) listeners. Such HRTFs eliminate the inter-individually different characteristics which are not crucial for accurate localisation while preserving all relevant characteristics. Even though the listener's performance should not be impaired by the presence of naturally occurring, but unnecessary cues in virtual stimuli, discarding those cues may be advantageous. The preservation of the cues that are indispensable for sound localisation, while eliminating the cues which are not crucial, minimises the effort and time required for computing stimuli.

Across-listener generalised HRTFs intend to prevent the need for measuring the HRTFs of each individual separately, and thereby simplify the creation of VAS for numerous fields of application. At the same time, it is important to prevent artifacts such as front-back confusions, one of the reasons which justify the extended research in the field of HRTFs and virtual auditory spaces.

Whenever HRTF-filtered stimuli are employed, the problem arises of how inter-individually different refractional properties of the head or pinna or differences in head diameter affect localisation performance in response to virtual stimulation. It would be of no use to possess sophisticated virtual auditory worlds, if these were not reliably perceived as being externalised, or else if the virtual space did not unambiguously simulate the intended free-field sound source. A global application of, for example, virtual auditory displays can only be achieved when VASs are really listener-independent to a sufficient extent.

Hence, great efforts have been made to develop universally applicable sets of HRTFs across all listeners, but discarding cues that are not required. An even more important aspect, of course, is to resolve any ambiguities that occur with virtual stimuli but not with natural stimuli. HRTF-filtered stimuli have been used to investigate whether the use of individualised versus non-individualised HRTFs influenced localisation behaviour in various species (e.g. humans: Hofman & Van Opstal, 1998; Hu et al., 2008; Hwang et al., 2008; Wenzel et al., 1993; owl: Hausmann et al., 2009; ferret: King et al., 2001; Mrsic-Flogel et al., 2001). It was shown that one of the main problems when using non-individualised HRTFs for stimulation was that the listeners committed front-back or back-front reversals, that is, they localised stimuli coming from the frontal hemisphere in the rear hemisphere or vice versa.

For many mammalian species, it was shown that in particular, notches in the high-frequency monaural spectra are relevant for sound localisation in the vertical plane (Carlile, 1990; Carlile et al., 1999; Koka & Tollin, 2008; Musicant et al., 1990; Tollin & Yin, 2003), and may help, together with ILD cues, to resolve front-back or back-front reversals as discussed in Hausmann et al. (2009). Whether this effect indeed occurs in the barn owl has yet to be proved.

In what concerns customisation of human HRTF-filtered signals, Middlebrooks (1999) proposed in his study how frequency-scaling of peaks and notches in directional transfer functions of human listeners allows generalisation of non-individualised HRTFs while preserving localisation characteristics. Such an approach may render extensive measurements for each individual unnecessary. Likewise, customisation of median-plane HRTFs is possible if the principal-component basis functions with largest inter-subject variations are tuned by one subject while the other functions are calculated as the mean for

all subjects in a database (Hwang et al., 2008). Since localisation accuracy is preserved even when HRTFs for human listeners account for only 30% of individual differences (Jin et al., 2003), slight customisation of measured HRTFs already yielded large improvements in localisation ability.

When individualised HRTF-filtered stimuli are used, the percepts in virtual auditory displays are identical to free-field percepts when the spatial resolution of HRTF-measurements is 6° or less (Langendijk & Bronkhorst, 2000). For 10 to 15° resolution, the percepts are still comparable (Langendijk & Bronkhorst, 2000), which implies that the spatial resolution for HRTF-measurements should not fall below 10°. This issue is of extreme importance in dynamic virtual auditory environments, because here it is required that transitions (switching) between HRTFs needed for the simulation of motion are inaudible to the listener. In other words, the listener should experience a smoothly moving sound image without disturbing clicks or jumps when the HRTF position is changed. Hoffman & Møller (2008) determined the minimum audible angles for spectral switching (MASS) to be 4-48° depending on the direction, and for temporal switching (minimum audible time switching MATS) to be 5-10 µs. That is, this resolution should not be under-run when switching between adjacent HRTF either temporally or spectrally. Interpolation of measured HRTFs is especially important if listeners are moving in the auditory world, to prevent leaps or gaps in the auditory percept. This interpolation has to be done carefully in order to preserve the natural auditory percept (Nishimura et al., 2009).

Standard sets of HRTFs are available on internet databases (e.g. on www.ais.riec.tohoku.ac.jp/lab/db-hrtf/). The availability of standard HRTFs recorded with artificial heads (reviewed in Paul, 2009) and of information and technology provided by head-acoustics companies allows scientists and private persons to benefit from sophisticated virtual auditory environments. Especially in what concerns users of cochlear implants, knowledge on the impact of individual HRTF features such as spectral holes (Garadat et al., 2008) on speech intelligibility has helped to improve hearing performance in those patients. Last but not least, much effort has been made to enhance the perceived "spaciousness" of virtual sounds for example to improve the impression of free-field sounds while listening to music (see Blauert, 1997).

7. Advantage, disadvantages and future prospects of virtual space techniques

There are still many challenges for the calculation of VASs. For instance, HRTFs have to be measured and interpolated very thoroughly for the various spatial positions in order to preserve the distributions of physical cues that occur in natural free-field sounds. This is to some extent easier for the largely frequency-independent ITDs, whereas slight mispositioning of the recording microphones can induce larger errors to the measured ILDs and spectral cues especially in the high-frequency range, which then may lead to mislocalisation of sound source elevation (Bronkhorst, 1995).

When measuring HRTFs, it is also important to carefully control the position of the recording microphone relative to the eardrum, since the transfer characteristics of the ear canal can vary throughout its length (Keller et al., 1998; Spezio et al., 2000; Wightman & Kistler, 1989a).

Another aspect is that the computational efforts for the complex and time-consuming creation of virtual stimuli may be reduced by reversing the positions of microphones and

sound source during HRTF measurements. The common approach, which has also been described in the present chapter, is placement of the microphone into the ear canal and subsequent application of sound from outside. In this case, the position of the sound source is varied systematically across representative spatial positions, in order to reflect the amplitude of each physical cue after filtering by the outer ear and ear canal.

However, it is also possible to take the reverse approach, that is, placing the sound source into the ear canal and record the signal that is arriving at a microphone after filtering by the ear canal and outer ear (e.g. Zotkin et al., 2006). The microphones that record the output signals are then positioned at the exact spatial locations where usually the loudspeaker would be. The latter approach has a huge advantage compared to the conventional way, because it saves an immense amount of time. Rather than placing the sound source sequentially to various locations in space, waiting until the signal has been replayed, reposition the sound source and repeat the measurement for another position, one single application of the sound suffices as long as an array of microphones is installed at each spatial location one wants to record an impulse response for. The time consuming conventional approach, however, has the advantage that only a single recording device is required. Furthermore, in the conventional approach, the loudspeaker is not as limited in size as is an in-ear loudspeaker. It may be difficult to build an in-ear loudspeaker with satisfying low-frequency sound emission.

Another possibility to save time when recording impulse responses is to use a microphone moving along a circle, which allows recording of impulse responses for each angle along the horizontal plane in less than one second (Ajdler et al., 2007). Also in this technique, the sound emitter is placed in the ear and the receiver microphone is placed outside the subject's ear canal.

Thus, depending on the purpose of an HRTF measurement, an experimenter has several choices and may simply decide which approach is more useful for his or her requirements.

Another important, but often neglected aspect of sound localisation that still awaits closer investigation is the role of auditory distance estimation. Kim et al. (2010) recently presented HRTFs for the rabbit, which show variances in HRTF characteristics for varying sound source distances. Overestimation of sound sources in the near field occur as commonly as underestimation of source distance in the far field (e.g. Loomis et al. 1998; Zahorik 2002), which again seems to be a phenomenon that is not due to headphone listening, but a common feature of sound localisation.

Loomis & Soule (1996) showed that distance cues are reproducible with virtual acoustic stimuli. The human listeners in their study experienced virtual sounds in considerable distance of several meters, even though the perceived distances were still subject to misjudgements. However, since the latter problem occurs also in free-field sounds (overestimation of near targets and underestimation of far targets), further efforts need to be spent to unravel distance perception in humans.

That is, it is possible to simulate auditory distance with stimuli provided via headphones. Noteworthy, virtual auditory stimuli may be scaled so that they simulate a specific distance, even if a corresponding free-field sound would be under- or overestimated, respectively. This is a considerable advantage of the virtual auditory space technique, because naturally occuring perceptional "errors" may be overcome by in- or decreasing the amplitude of virtual auditory stimuli according to the respective requirements. Fontana and coworkers (2002) developed a method to simulate the acoustics inside a tube in order to successfully provide distance cues in a virtual environment. It is also possible to calibrate distance

estimation using psychophysical rating methods, so to get a valid measure for distance cues (Martens, 2001).

How good distance cues, among which intensity, spectrum and direct-to-reverberant energy are especially important, are preserved with current HRTF recording techniques, i.e., how good they coincide with the natural distance cues, is still to be evaluated more closely.

In sum, the virtual space technique offers a wide range of powerful applications, not only for the general investigation of sound localisation properties, but also for its implementation in daily life. Once the cues that contribute to specific aspects of sound localisation are known, not only established techniques such as hearing aids may be improved, for example for the reduction of background noise or for better separation of several concurring sound sources, but the VAS also allows to introduce manipulations to sound stimuli that would naturally not occur. The latter possibility may be useful to create auditory illusions for various applications. Among these are auditory displays for navigational tasks for example during flight (Bronkhorst et al., 1996) or travel aids for both healthy and blind people (Loomis et al., 1998; Walker & Lindsay, 2006), as well as communicational applications such as telephone conferencing (see Martens, 2001).

However, it is indispensable to further evaluate if recording of HRTFs and creation of VASs indeed reflect all relevant aspects of sound localisation cues, in order to prevent unwanted artifacts that might confound the perceived spatial position.

Although a major goal of basic research has to be the long-time implementation of the gained knowledge for applications in humans, the extended use of animal models for the auditory system can yield valuable data on basic auditory processes, as was shown throughout this chapter.

8. References

Ajdler, T.; Sbaiz, L. & Vetterli, M. (2007). Dynamic measurement of room impulse responses using a moving microphone. *J Acoust Soc Am* 122, 1636-1645

Bala, A.D. ; Spitzer, M.W. & Takahashi, T.T. (2007). Auditory spatial acuity approximates the resolving power of space-specific neurons. *PLoS One,* 2, e675

Blauert, J. (1997). Spatial Hearing. The Psychophysics of Human Sound Localization, MIT Press, ISBN 3-7776-0738-X, Cambridge, Massachussetts.

Bronkhorst, A.W. ; Veltman, J.A. & Van Vreda, L. (1996). Application of a Three-Dimensional Auditory Display in a Flight Task. *Human Factors* 38, 23-33

Butts, D.A. & Goldman, M.S. (2006). Tuning Curves, Neuronal Variability, and Sensory Coding. *PLoS Biol* 4, e92

Calmes, L.; Lakemeyer, G. & Wagner, H. (2007). Azimuthal sound localization using coincidence of timing across frequency on a robotic platform. *J Acoust Soc Am*, 121, 2034-2048

Campenhausen, M. & Wagner, H. (2006). Influence of the facial ruff on the sound-receiving characteristics of the barn owl's ears. *J Comp Physiol A*, 192, 1073-1082

Carlile, S. (1990). The auditory periphery of the ferret. II: The spectral transformations of the external ear and their implications for sound localization. *J Acoust Soc Am* 88, 2195-2204

Carlile, S.; Leong, P. & Hyams, S. (1997). The nature and distribution of errors in sound localization by human listeners. *Hear Res* 114, 179-196

Carlile, S.; Delaney, S. & Corderoy, A. (1999). The localisation of spectrally restricted sounds by human listeners. *Hear Res* 128, 175-189

Coles, R.B. & Guppy, A. (1988). Directional hearing in the barn owl (Tyto alba). *J Comp Physiol A*, 163, 117-133

Delgutte, B. ; Joris P.X. ; Litovsky, R.Y. & Yin, T.C.T. (1999). Receptive fields and binaural interactions for virtual-space stimuli in the cat inferior colliculus. *J Neurophysiol* 81, 2833-2851

Dent, M.L.; Tollin, D.J. & Yin, T.C.T. (2009). Influence of Sound Source Location on the Behavior and Physiology of the Precedence Effect in Cats. *J Neurophysiol* 102, 724-734

Dietz, M. ; Ewert, S.D. & Hohmann, V. (2009). Lateralization of stimuli with independent fine-structure and envelope-based temporal disparities. *J Acoust Soc Am*, 125, 1622-1635

Drager, U. & Hubel, D. (1975). Physiology of visual cells in mouse superior colliculus and correlation with somatosensory and auditory input. *Nature* 253, 203-204

DuLac, S. & Knudsen, E.I. (1990). Neural maps of head movement vector and speed in the optic tectum of the barn owl. *J Neurophysiol*, 63, 131-146

Fontana, F. ; Rocchesso, D. & Ottaviani, L. (2002). A Structural Approach to Distance Rendering in Personal Auditory Displays. In : *Proceedings of the 4th IEEE International Conference on Multimodal Interfaces*, ISBN : 0-7695-1834-6, p. 33.

Garadat, S.N.; Litovsky, R.Y.; Yu, G. & Zeng, F.-G. (2009). Effects of simulated spectral holes on speech intelligibility and spatial release from masking under binaural and monaural listening. *J Acoust Soc Am* 127,2,977-989

Gardner, M.B. & Gardner, R.S. (1973). Problem of Localization in the Mediean Plane, Effect of Pinna Caity Occlusion. *J Acoust Soc Am* 53, 400-408

Harris, L. ; Blakemore, C. & Donaghy, M. (1980). Integration of visual and auditory space in the mammalian superior colliculus. *Nature* 5786, 56-59

Hartline P. ; Vimal, R. ; King, A. ; Kurylo, D. & Northmore, D. (1995). Effects of eye position on auditory localization and neural representation of space in superior colliculus of cats. *Exp Brain Res* 104, 402-408

Hartmann, W. & Wittenberg, A. (1996). On the externalization of sound images. *J Acoust Soc Am*, 99, 3678-3688

Hausmann, L.; von Campenhausen, M. ; Endler, F. ; Singheiser, M. & Wagner, H. (2009). Improvements of Sound Localization Abilities by the Facial Ruff of the Barn Owl (Tyto alba) as Demonstrated by Virtual Ruff Removal. *PLoS One*, 4, e7721

Hausmann, L.; von Campenhausen, M. & Wagner, H. (2010). Properties of low-frequency head-related transfer functions in the barn owl (tyto alba). *J Comp Physiol A*, epub ahead of print

Hebrank, J. & Wright, D. (1974). Are Two Ears Necessary for Localization of Sound Sources in the Median Plane? *J Acoust Soc Am* 56, 935-938

Hill, P. ; Nelson, P. ; Kirkeby, O. & Hamada, H. (2000). Resolution of front-back confusion in virtual acoustic imaging systems. *J Acoust Soc Am* 108, 2901-2910

Hoffman, P.F. & Møller, H. (2008). Audibility of Direct Switching Between Head-Related Transfer Functions. *Acta Acustica united with Acustica* 94, 955-964

Hofman, P.M.; Van Riswick, J.G.A. & Van Opstal, A.J. (1998). Relearning sound localization with new ears. *Nature Neuroscience* 1, 417-421

Hu, H.; Zhou, L.; Ma, H. & Wu, Z. (2007). HRTF personalization based on artificial neural network in individual virtual auditory space. *Applied Acoustics* 69, 163-172

Hwang, S.; Park, Y. & Park, Y. (2008). Modeling and Customization of Head-Related Impulse Responses Based on General Basis Functions in Time Domain. *Acta Acustica united with Acustica*, 94, 965-980

Jin, C.; Leong, P.; Leung, J.; Corderoy, A. & Carlile, S. (2000). Enabling individualized virtual auditory space using morphological measurements. Proceedings of the First IEEE Pacific-Rim Conference on Multimedia, pp. 235-238

Keller, C.; Hartung, K. & Takahashi, T. (1998). Head-related transfer functions of the barn owl : measurement and neural responses. *Hear Res* 118, 13-34

King, A. & Calvert, G. (2001). Multisensory integration : perceptual grouping by eye and ear. *Curr Biol* 11, R322-R325

King, A.; Kacelnik, O.; Mrsic-Flogel, T.; Schnupp, J.; Parsons, C. & Moore, D. (2001). How plastic is spatial hearing? *Audiol Neurootol* 6, 182-186

Krämer, T. (2008). Attempts to build an artificial facial ruff mimicking the barn owl (Tyto alba). *Diploma thesis*, RWTH Aachen, Aachen

Knudsen, E.I. & Konishi, M. (1979). Mechanisms of sound localisation in the barn owl (Tyto alba). *J Comp Physiol A*, 133, 13-21

Knudsen, E.I.; Blasdel, G.G. & Konishi, M. (1979). Sound localization by the barn owl (Tyto alba) measured with the search coil technique. *J Comp Physiol A* 133, 1-11

Knudsen, E.I. (1981). The Hearing of the Barn Owl. *Scientific American*, 245, 113-125

Koeppl, C. (1997). Phase locking to high frequencie in the auditory nerve and cochlear nucleus magnocellularis of the barn owl, Tyto alba. *J Neurosci* 17, 3312-3321

Koka, K. & Tollin, D. (2008). The acoustical cues to sound location in the rat : measurements of directional transfer functions. *J Acoust Soc Am* 123, 4297-4309

Lee, N.; Elias, D.O. & Mason, A.C. (2009). A precedence effect resolves phantom sound source illusions in the parasitoid fly *Ormia ochracea*. *Proc Natl Acad Sci USA*, 106(15), 6357-6362

Loebach, J.L. & Pisoni, D. (2008). Perceptual learning of spectrally degraded speech and environmental sounds. *J Acoust Soc Am* 123, 2, 1126-1139

Loomis, J.M. & Soule, J.I. (1996). Virtual acoustic displays for real and virtual environments. In: *Proceedings of the Society for Information Display 1996 International Symposium*, pp. 965-968. San Jose, CA : Society for Information Display.

Loomis, J.M.; Klatzky, R.L., Philbeck, J.W. & Golledge, R.G. (1998). Assessing auditory distance perception using perceptually directed action. *Perception & Psychophysics* 60, 6, 966-980

Loomis, J.M.; Golledge, R.G. & Klatzky, R.L. (1998). Navigation System for the Blind : Auditory Display Modes and Guidance. *Presence*, 7, 193-203

Makous, J. & Middlebrooks, J.C. (1990). Two-dimensional sound localization by human listeners. *J Acoust Soc Am* 87, 2188-2200

Martens, W.L. (2001). Psychophysical calibration for controlling the range of a virtual sound source: multidemensional complexity in spatial auditory display. Proceedings of the 2001 International Conference on Auditory Display, Espoo, Finland, July 29-August 1.

May, B.J. & Huang, A.Y. (1995). Sound orientation behavior in cats. I. Localization of broadband noise. *J Acoust Soc Am* 100, 2, 1059-1069

Mazer, J.A. (1998). How the owl resolves auditory coding ambiguity. *Proc Natl Acad Sci USA*, 95, 10932-10937

Meredith, M. & Stein, B. (1986). Visual, auditory, and somatosensory convergence on cells in superior colliculus results in multisensory integration. *J Neurophysiol* 56, 640-662

Middlebrooks, J. & Knudsen, E.I. (1984). A neural code for auditory space in the cat's superior colliculus. *J Neurosci* 4, 2621-2634

Moiseff, A. & Konishi, M. (1981). Neuronal and behavioral sensitivity to binaural time differences in the owl. *J Neurosci* 1, 1, 40-48

Mrsic-Flogel, T.; King, A.; Jenison, R. & Schnupp, J. (2001). Listening through different ears alters spatial response fields in ferret primary auditory cortex. *J Neurophysiol* 86, 1043-1046

Musicant, A.; Chan, J. & Hind, J. (1990). Direction-dependent spectral properties of cat external ear: New data and cross-species comparisons. *J Acoust Soc Am* 87, 757-781

Nishimura, R.; Kato, H. & Inoue, N. (2009). Interpolation of head-related transfer functions by spatial linear prediction. IEEE 1901-1904

Parsons, C.H.; Lanyon, R.G.; Schnupp, J.W.H. & King, A.J. (1999). Effects of Altering Spectral Cues in Infancy on Horizontal and Vertical Sound Localization by Adult Ferrets. *J Neurophysiol* 82, 2294-2309

Paul, S. (2009). Binaural Recording Technology: A Historical Review and Possible Future Developments. *Acta Acustica united with Acustica* 95, 767-788

Plenge, G. (1974). On the differences between localization and lateralization. *J Acoust Soc Am* 56, 944-951

Poganiatz, I. & Wagner, H. (2001). Sound-localization experiments with barn owls in virtual space : influence of broadband interaural level difference on head-turning behavior. *J Comp Physiol A*, 187, 225-233

Poganiatz, I.; Nelken, I. & Wagner, H. (2001). Sound-localization experiments with barn owls in virtual space : influence of interaural time difference on head-turning behavior. *JARO* 2, 1-21

Populin, L.C. (2006). Monkey Sound Localization: Head-Restrained versus Head-Unrestrained Orienting. *J Neurosci* 26, 38, 9820-9832

Populin, L.C. & Yin, T.C.T. (1998). Pinna movements of the cat during sound localization. *J Neurosci* 18, 4233-4243

Rayleigh, Lord (1907). On our perception of sound direction. *Philos Mag* 13, 214-232

Saberi, K.; Farahbod, H. & Konishi, M. (1998). How do owls localize interaurally phase-ambiguous signals? *PNAS* 95, 6465-6468

Saberi, K. ; Takahashi, Y. ; Farahbod, H. & Konishi, M. (1999). Neural bases of an auditory illusion and its elemination in owls. *Nat Neurosci* 2, 656-659

Searle, C.L.; Braida, L.D. ; Cuddy, D.R. & Davis, M.F. (1975). Binaural pinna disparity : another auditory localization cue. *J Acoust Soc Am* 57, 2, 448-455

Spezio, M.L.; Keller, C.H.; Marrocco, R.T. & Takahashi, T.T. (2000). Head-related transfer functions of the Rhesus monkey. *Hear Res* 144, 73-88

Steinbach, M. (1972). Eye movements of the owl. *Vision Research* 13, 889-891

Takahashi, T.T. & Konishi, M. (1986). Selectivity for interaural time difference in the owl's midbrain. *J Neurosci* 6, 3413-3422

Tollin, D.J. & Koka, K. (2009). Postnatal development of sound pressure transformation by the head and pinnae of the cat: Monaural characteristics. *J Acoust Soc Am* 125, 2, 980-994

Tollin, D.J. & Yin, T.C.T. (2002). The Coding of Spatial Location by Single Units in the Lateral Superior Olive of the Cat. I. Spatial Receptive Fields in Azimuth. *J Neuroscience*, 22, 4, 1454-1467

Tollin, D.J. & Yin, T.C.T. (2003). Spectral cues explain illusory elevation effects with stereo sounds in cats. *J Neurophysiol* 90, 525-530

Usher, J. & Martens, W.L. (2007). Naturalness of speech sounds presented using personalized versus non-personalized HRTFs. *Proceedings of the 13th International Conference on Auditory Display*, Montréal, Canada, June 26-29.

Van Wanrooij, M.M., Van Der Willigen, R.F. & Van Opstal, A.J. (2010). Learning Sound Localization with the Barn-owl's Ears. *Abstracts to FENS 2010*, Poster number 169.25, Amsterdam, Netherlands, July 2010

Wagner, H. ; Takahashi, T. & Konishi, M. (1987). Representation of interaural time difference in the central nucleus of the barn owl's inferior colliculus. *J Neurosci* 7, 3105-3116

Walker, B.N. & Lindsay, J. (2006). Navigation Performance With a Virtual Auditory Display : Effects of Beacon Sound, Capture Radius, and Practice. *Human Factors* 48, 2, 265-278

Wenzel, E. ; Arruda, M. Kistler, D. & Wightman, F. (1993). Localization using nonindividualized head-related transfer functions. *J Acoust Soc Am* 94, 111-123

Wightman, F.L. & Kistler, D.J. (1989a). Headphone simulation of free field listening. I : Stimulus synthesis. *J Acoust Soc Am* 85, 2, 858-867

Wightman, F.L. & Kistler, D.J. (1989b). Headphone simulation of free field listening. II : Psychophysical validation. *J Acoust Soc Am* 85, 2, 868-878

Zahorik, P. (2002). Assessing auditory distance perception using virtual acoustics. *J Acoust Soc Am* 111, 4, 1832-1846

Zahorik, P.; Bangayan, P.; Sundareswaran, V.; Wang, K. & Tam, C. (2006). Perceptual recalibration in human sound localization: learning to remediate front-back reversals. *J Acoust Soc Am* 120, 343-359

Zotkin, D.N.; Duraiswami, R.; Grassi, E. & Gumerov, N.A. (2006). Fast head-related transfer function mesurement via reciprocity. *J Acoust Soc Am* 120, 4, 2202-2215
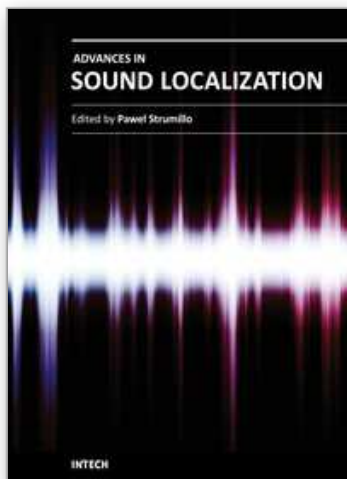
Advances in Sound Localization

Edited by Dr. Pawel Strumillo

Sound source localization is an important research field that has attracted researchers' efforts from many technical and biomedical sciences. Sound source localization (SSL) is defined as the determination of the direction from a receiver, but also includes the distance from it. Because of the wave nature of sound propagation, phenomena such as refraction, diffraction, diffusion, reflection, reverberation and interference occur. The wide spectrum of sound frequencies that range from infrasounds through acoustic sounds to ultrasounds, also introduces difficulties, as different spectrum components have different penetration properties through the medium. Consequently, SSL is a complex computation problem and development of robust sound localization techniques calls for different approaches, including multisensor schemes, null-steering beamforming and time-difference arrival techniques. The book offers a rich source of valuable material on advances on SSL techniques and their applications that should appeal to researches representing diverse engineering and scientific disciplines.

How to reference

In order to correctly reference this scholarly work, feel free to copy and paste the following:

Laura Hausmann and Hermann Wagner (2011). Using Virtual Acoustic Space to Investigate Sound Localisation, Advances in Sound Localization, Dr. Pawel Strumillo (Ed.), ISBN: 978-953-307-224-1, InTech, Available from: http://www.intechopen.com/books/advances-in-sound-localization/using-virtual-acoustic-space-to-investigate-sound-localisation

INTECH
open science | open minds