Neural responses to free field and virtual acoustic stimulation in the inferior colliculus of the guinea pig

Oliver Behrend¹, Benjamin Dickson¹, Elizabeth Clarke¹, Craig Jin¹ and Simon Carlile¹,²

¹Auditory Neuroscience Laboratory, Department of Physiology, ²School of Biomedical Sciences, University of Sydney, NSW 2006, Australia

Abstract: 255 words
Text pages: 24
Figures: 13

Corresponding author: Junior Professor Oliver Behrend, Ph.D.
Head, Aquatic Bioacoustics Laboratory
Institute of Biology
Humboldt University Berlin
Invalidenstr. 43
10115 Berlin
Germany
Fax: ++49-30-20938491
Phone: ++49-30-20938861
Email: behrend@rz.hu-berlin.de
Abstract

Virtual auditory space (VAS) stimuli based on outer ear transfer functions became increasingly important in spatial hearing research. However, few studies have investigated the match between responses of auditory neurons to VAS and free field (FF) stimulation. This study validates acoustic spatial receptive fields (SRFs) of 183 individual midbrain units using both VAS and FF stimuli.

The 1st spike latency, which varied systematically across SRFs, was 14.9 ± 8.3 ms in FF, and 15.1 ± 8.3 ms in VAS (mean ± SD). Spike count based SRFs measured 0-20 dB above the neural threshold covered on average 44.5 ± 18.0% of the recorded sphere in FF and 45.5 ± 18.7% in VAS. The average deviation of the centroid position of SRFs using FF and VAS stimuli was 7.4º azimuth and 3.3º elevation. The average spike rate remained unchanged. The SRF overlap recorded using FF and VAS stimuli (mean 71.3 ± 12.6%) or repeated FF stimuli (70.2 ± 14.2%) was high and strongly correlated (r=0.96; p<0.05). The SRF match observed with FF and VAS stimuli was not significantly altered over a range of stimulus levels (paired t-test p=0.51; N=6). Randomized VAS barely affected SRF sizes, centroids, or maximum spike count but decreased the average minimum response to 59% compared to sequential stimulation (paired t-test; p=0.05; N=26). SRF recordings in VAS excluding the acoustic distortions of the recording equipment differed from those in VAS incorporating the equipment (paired t-test p=0.01; N=5).

In conclusion, neurophysiological recordings demonstrate that individualized VAS stimuli provided a good simulation of a FF environment.
Introduction

In the vertebrate central nervous system, the nature of the representation of acoustic space and, more generally, the encoding and processing of auditory cues related to auditory localization have inspired neurophysiological investigations from the VIII nerve, brainstem, midbrain, thalamic to cortical areas (Middlebrooks and Pettigrew 1981; King and Palmer 1983; Ahissar et al. 1992; Young et al. 1992; Poon and Brugge 1993b; Brugge et al. 1994; King and Carlile 1994; Young et al. 1995; Brugge et al. 1996; Hartung and Sterbing 1997; Imig et al. 1997; Middlebrooks et al. 1998; Sutherland et al. 1998; Furukawa et al. 2000; Imig et al. 2000; May, 2000; McAlpine et al. 2001; Furukawa and Middlebrooks, 2002; Middlebrooks et al. 2002; Sterbing et al. 2003). Particularly in barn owls, studies have demonstrated that a map of acoustic space based on ITD and IID cues can be found in the midbrain (Knudsen and Konishi 1978a; Knudsen and Konishi 1978c; Knudsen 1982; Moiseff and Konishi 1983; Olsen et al. 1989; Knudsen et al. 1991). Also, a number of studies have demonstrated spatial maps in the mammalian midbrain, namely the external nucleus of the inferior colliculus (ICX), the nucleus of the brachium of the inferior colliculus (NBIC), and the superior colliculus (SC; King and Palmer 1983; Middlebrooks and Knudsen 1984; King and Hutchings 1987; Binns et al. 1992; Schnupp and King 1997). Yet in the mammalian SC, it had long been shown that ILDs and ITDs alone are insufficient to compute unambiguous spatial receptive fields (SRFs; Wise and Irvine 1983; King and Carlile 1994). A number of models had proposed that frequency integration helps to resolve these ambiguities (Brainard et al. 1992).

Indeed, reports on monaural map formation in the mammalian midbrain (Palmer and King 1985) indicated that sound localization can be performed with no interaural cues present at all. Physiologically, it was demonstrated that neural frequency tuning varies along the rostro-caudal axis of the spatially tuned SC (Carlile and Pettigrew 1987b). Thereby frequency tuning, and location dependent acoustic filtering of the outer ear (Carlile and Pettigrew 1987a), was linked to the spatial map of the SC. Perceptually, monaural filtering cues are crucial for, e.g., human auditory “externalized” localization (Angell and Fite 1901; Rayleigh 1907; Plenge 1974; Butler and Belendiuk 1977; Musicant and Butler 1984). In complementary animal experiments, encoding of monaural cues has been demonstrated physiologically at early levels such as the VIII nerve and the DCN (Young et al. 1992; Poon and Brugge


So a picture emerged in which binaural cues and the spectral variations produced by the auditory periphery are integrated for a proper representation of acoustic space (Shaw 1974; Middlebrooks and Pettigrew 1981; Carlile and Pettigrew 1987a; Middlebrooks et al. 1989; Carlile 1990a; Carlile 1990b; Musicant et al. 1990; Rice et al. 1992; Carlile and King 1994; Pralong and Carlile 1994).

The peripheral filters were described as head-related-transfer-functions (HRTF) or head-related-impulse-responses, HRIR (Mellert et al. 1974; Mehrgardt and Mellert 1977; Wightman and Kistler 1989a; Poon and Brugge 1993b; Brugge et al. 1994; Hartung and Sterbing 1997). Based on these filters, virtual auditory stimulation (VAS) proved to be an extremely useful tool to investigate both the neurophysiology and the psychophysics of auditory space perception and had been implemented by several research groups using humans, other mammals, and birds (Wightman and Kistler 1989b; Poon and Brugge 1993b; Brugge et al. 1994; Chen et al. 1995; Brugge et al. 1996; Reale and Chen 1996; Hartung and Sterbing 1997; Keller et al. 1998; Semple 1998; Schnupp et al. 2001). However, only one study in barn owls directly addressed the agreement between spatial neural responses to free field (FF) and virtual space stimuli at the level of single units (Keller et al. 1998). Most studies relied on acoustic measurements close to the eardrum in VAS and FF to match both stimulation regimes (Brugge et al. 1994; Sterbing et al. 2002). The general conclusion advanced is that on a population basis the spatial neural response characteristics observed within a brain structure were similar regardless of whether the units were tested using VAS or free field stimuli (Brugge et al. 1994; Brugge et al. 1996). However, at the level of the single neuron, the accuracy of VAS stimulation still had to be demonstrated in mammals. Our study directly assessed the fidelity of the VAS delivery system by comparing the neural responses of the same single unit to spatial acoustic stimuli using both FF and VAS stimuli. The system was comprised of two major components under remote control. Firstly a FF speaker mounted at a robot arm covering spherical positions between -45° and 90° elevation, and secondly a VAS earphone system mounted on retractable pressure-driven actuators. The combination of these two systems allowed the recording of the SRF of isolated single units using both FF and VAS stimuli. High fidelity VAS was generated using real-time control of the acoustics and online calibration of the delivery system. In the second part of the study, we examined the effects of randomisation of the spatial sequence of stimuli. In most
FF studies the mechanics of stimulus placement require an ordered sequence of replicate stimuli at each stimulus location. In the third part of this study the neurophysiological consequences of the acoustic distortion produced by all of the necessary recording and supporting apparatus surrounding the animal were investigated.

The guinea pig served as an animal model which features a widely overlapping hearing range with humans (Syka et al. 2000), and whose non-moveable pinna parallels the human ear to some extent (Hartung and Sterbing 1997). We have found that for auditory midbrain units VAS can mimic the acoustic free field quite realistically.

Parts of the data have been presented in abstract form (Behrend et al. 2003; 2004).
Material and Methods

Anaesthesia and surgery

A total of 31 adult guinea pigs with unobstructed ear canals and no sign of middle ear infection were anesthetized by inhalation of Isoflurane over a period of 2 minutes (0.5% in Carbogen; Forthane, Abbott, Australia; inhalator: Advanced Anaesthesia Specialists, Australia). Subsequently, an intramuscular injection of Ketamine™ and Xylazine™ ([2:1]; 0.1 ml/kg, Sigma-Aldrich, USA) was applied. Anaesthesia was maintained during the course of surgery and neural recordings by continuous intravenous administration via the jugular or femoral vein of Hypnorm™ (1 ml/kg/h; Janssen Animal Health, UK; A99 syringe pump, Razel, USA). Levels of anaesthesia were assessed by continuous monitoring of the cardiac rate (custom made device; Physics department of the University of Sydney) and frequent testing of the cornea and startle reflexes. Small quantities of local anaesthetic (Xylocaine™, Sigma-Aldrich) were applied subcutaneously at the body locations where surgery was performed. To expose the skull, an incision of the skin was made above the approximate location of the bregmoid and the lambdoid suture, and the tissue was reflected laterally. A metal holder was then attached to the skull by 1.5 mm diameter screws and dental cement (Palladur™, Kulzer, Germany). A hole of approximately 0.5 cm in diameter was then drilled (Moto-flex driller, Dremel, Germany) into the bone above the brain target region to expose the brain’s surface for electrode penetrations. Finally, the dura mater was locally opened according to landmarks on the brain’s surface. To enable acoustic recordings from within the external auditory canal a probe tube was implanted into the ear canal using an inside-out approach described elsewhere (Behrend et al. 2001).

Neural recordings

The animals’ position in the recording chamber was standardized by stereotactic landmarks on the surface of the skull (intersections of the bregmoid and lambdoid sutures with the sagittal suture in horizontal alignment; rostro-caudal axis aligned to setup midline). To properly adjust the electrode trajectory relative to the midbrain and to facilitate stable acoustic conditions, all animals were kept in this position throughout the experiment. Micromanipulators were used to position the recording electrode according to landmarks on the brain’s surface. Extracellular single and multi
unit responses were recorded by tungsten electrodes (~5 MΩ impedance; A-M Systems Inc.) and signals were processed via a head stage (MM-333, Narishige, Japan), then amplified and band pass filtered (0.7 to 3 kHz) by a microelectrode AC amplifier (model 1800, A-M Systems Inc., USA). Next the amplified signals were passed through a 50/60 Hz noise eliminator (Humbug, Quest Scientific, UK), and then fed into a Tucker-Davis-Technology System II (component DD1, TDT, USA) for digitisation and further analysis. Only action potentials from units featuring a signal to noise ratio greater than 3 were selected for analysis. Subsequently, action potentials were registered using an event timer (ET1, TDT) and a DSP-Board (System II, TDT) before storage for offline analysis. During the recordings the animal was wrapped in a heating blanket (custom built) in a sound attenuated anechoic chamber. Typical recording periods were 10-30 hours. Carbon coated electrodes (A-M Systems) were used to mark electrode tracts for subsequent reconstruction of recording sites.

**Histology and reconstruction**

At the end of the experiment the animals were euthanized by an overdose of barbiturate (Nembutal™ 2 ml/kg, Sigma-Aldrich) and intracardially perfused with heparinized 0.9% saline for 5 minutes followed by 4% paraformaldehyde and 25% glutaraldehyde for 40 minutes (Masterflex pump, Cole Parmer Instrument Co. Illinois, USA; 5 ml/min flow rate; Optiva 14G needle, 2.2 mm outer diameter, Terumo Medical Corporation, USA). The brain was then removed and stored in sucrose for kryoprotection (30% solution) and remained in the solution until it sank. Finally, the brain was embedded in Agar-agar (Merck, Germany) and mounted on a (Leica 1320, Germany) cryostat to obtain frontal sections of the midbrain (each 40 µm). For counterstaining the slices were treated with diaminobenzidine (DAB; Adams 1977). To verify approximate recording sites the sections were analyzed by light microscopy.

**Acoustic recordings and stimulus generation**

All bioacoustic and neurophysiological measurements took place in a darkened anechoic chamber. Gaussian noise stimuli (digitally generated; 80 kHz sampling rate; band-passed 0.3-30 kHz; 47.5-100 ms duration; 10 ms raised-cosine onset and offset ramp; 0-40 dB above neural threshold) were presented in two configurations, i.e. free field (FF) and virtual acoustic space (VAS). All stimuli were generated on a workstation (Pentium III 450 MHz, Intel, USA) using MatlabR12 DSP (The Maths
Work Inc., USA). Stimuli were delivered by a TDT System II with 16-bit D/A converters (DA3-2; sampling rate 80 kHz), anti-aliasing filters (FT-6; cut-off 30 kHz), and digital attenuators (PA4).

The FF stimuli were presented at 27 or 80 spherical positions covering the frontal hemisphere (above -30° elevation) or the full sphere (above -20° elevation; figure 1), respectively. Sounds were delivered at a distance of 1 m from the animal’s head, by means of a VIFA-D26TG-35 speaker (Danish Sound Technology, Denmark) mounted on a robot arm under remote control (custom made; Electrical and Information Engineering, University of Sydney). A digital inverse filter applied for FF sound delivery flattened out the speaker frequency characteristics from 0.3-30 kHz. VAS at corresponding virtual positions was presented via custom made earphones, for which a detailed description is given by Chan and colleagues (Chan et al. 1993). The earphone delivered VAS stimuli were matched spectrally and in overall level (± 1 dB) to the FF stimuli for each test location. To this end the difference of power spectral densities of sounds in VAS and subsequent FF were calculated across 760 frequency bins for stimuli given at 0° azimuth and elevation. The FF attenuation was then adjusted to VAS levels by minimizing the average FF-VAS difference observed at both ear canals. The power spectral densities of FF and VAS stimuli recorded in the ear canal were again compared on screen after each set of spatial stimuli. To best match the spatial features of FF acoustics, VAS was generated by convolving the identical noise stimuli used in FF stimulation with the recorded HRTFs.

HRTF measurements were made at 393 spherical positions evenly distributed around the animal (between -45° and 90° elevation) using a pair of complementary Golay sequences (Zhou et al. 1992) with radix 2 and length 1024 (12.8 ms duration; 12 repetitions; maximal 75 dB SPL at the ear canal). The impulse responses were recorded over 9.4 ms by the probe tube microphone implanted in the animal’s ear canal (KE4-211-2; Quad 306 amplifier, Sennheiser, Germany) and averaged across all repetitions for each position. To remove the effects of the sound delivery system for the calculation of HRTFs, the speaker-to-microphone transfer function was recorded and digitally de-convolved from FF signals recorded at all spatial locations. In general the HRTFs were recorded with all of the neurophysiological and other supporting apparatus in place to allow a direct comparison of the responses of the units to stimuli presented in the free field with stimuli presented in the VAS exactly matching the FF acoustic environment. For some animals, the HRTFs were also recorded in the
absence of supporting equipment. These latter recordings were intended as an accurate measure of the HRTFs of the animal in an acoustically uncluttered free field.

To match free field sounds as closely as possible it was also necessary to neutralize the filter effects of the sound delivery system in VAS by the implementation of a digital inverse filter. An inverse filter model corresponding to the sound driver was derived using the least mean square (LMS) adaptive filtering algorithm as described by Widrow (1985). The tap length for the LMS inverse filter was empirically chosen as 512 and the delay factor was set to 256 taps, corresponding to 3.2 ms. The LMS algorithm typically took a few hundred iterations (roughly 1-3 minutes) to converge with the learning rate gradually reduced online. The adaptive filtering algorithm was implemented as a C program which was linked to the MatlabR12 programming environment using the MEX function interface. There is a separate inverse filter for the left ear and right ear which together establish the flat frequency characteristic of the VAS sound delivery system (figure 2A). Interaural time differences (ITDs) captured in the HRTFs were preserved correctly. See figure 2B (in the frequency domain) and figure 2C (in the temporal domain) for typical ear canal recordings of stimuli delivered from corresponding positions in FF and VAS.

Spatial acoustic stimulation regimes and auditory responses

Earphones were positioned close (≤ 1 cm) to the aperture of the animal’s ear canal and calibrated as described above. VAS sounds were generated on-the-fly for all tested positions according to the individual earphone-to-microphone transfer functions whenever the earphones were repositioned (i.e. for each individual unit recording). To search for acoustic responses, 100 ms noise stimuli were delivered in VAS (repetition rate 4 Hz; 10-40 dB attenuation) at random positions and along the acoustic axis of the contralateral ear, respectively. Except for when searching for single unit activity, the stimulus amplitude was 0-20 dB above the threshold of a neuron when stimulated in the vicinity of the acoustic axis of the contralateral ear. Once isolated, single unit responses were recorded for virtual space stimuli presented across all tested positions (27 or 80 positions to N=77 and N=106 units, respectively). The neural responses were recorded on one channel, whereas the corresponding sound stimuli were recorded on a second channel at the animals’ right ear (sampling rate 80 kHz).

VAS stimuli were presented in both sequential order (to mimic stimulation in the FF setup; N=183) and in pseudo-randomized order (VASRAN; N=26) to assess
spatial adaptation processes. For the latter, the original position list for sequential stimulus presentation was separated into 8 sub-lists with evenly distributed position indices, which were then randomized and concatenated to create 10 new pseudo-random full position lists.

For a small number of units (N=5), VAS stimuli were also generated using HRTFs which had been recorded in the absence of all neurophysiological recording and supporting equipment. The aim of this manipulation was to examine if the presence of this equipment produced distortions of the acoustical environment that could affect the SRF responses of the auditory neurons recorded in the midbrain.

After completion of a VAS recording set, the earphones were retracted under remote control by pressure driven actuators (cylinder CJ2B6-60SR, SMC, Australia). Then FF stimuli were presented for corresponding positions. At least 10 repetitions per spatial position were presented in either configuration. The inter-stimulus interval was 250 ms or more, to minimize neural adaptation by repeated simulation.

**Data analysis**

Neural data and concurrent ear canal recordings were analyzed offline using individual spike times and peri-stimulus-time histograms (bin width 1 ms). Recording periods were 137.5 ms for 27 position recordings and 250 ms for 80 position recordings, respectively. Neural recordings were separated into single unit responses when necessary, by means of either threshold spike discrimination (N=202) and principal component analysis of the spike waveform (PCA; N=183). For each unit, the spike timing of all discharges was calculated relative to the start of the recording period. Secondly, the first spike latency was calculated relative to 50% of the rising envelope of the stimulus recorded at a frontal position (0° azimuth and elevation) where no interaural time difference occurs. The spike jitter was computed from the SD of the mean onset latency across 10 repetitions of an identical stimulus. Spike counts were evaluated for 10 stimulus repetitions over a time window (11-210 ms after recording onset) which covers twice the stimulus duration. The underlying spontaneous activity was computed using a silent 40 ms period at the end of each recording period, and was then subtracted from the spike count (27 position regime: For this subset the spike count was conducted over 11-110 ms after recording onset; the spontaneous activity was computed from the last 10 ms of the recording period). The optimal response area of a receptive field (SRF) of a unit was defined as those
locations at which the spatial stimulus response exceeded 50% of the maximal response of a unit. The number of positions which fulfil this condition was used to approximate the size of a receptive field and expressed as a % of the tested spherical positions. For each unit, the optimal response area overlap between a pair of receptive fields obtained in (i) free field and (ii) virtual acoustic space was calculated by:

\[
\frac{N}{T} \times 100
\]

\(N\) = Number of shared SRF positions exceeding 50% of the maximal response  
\(T\) = Total number of positions of the larger SRF of the pair exceeding 50% of the maximal response

The overlap of the optimal response areas of two receptive fields recorded in FF and VAS was used as a measure to indicate the SRF overlap. To test the statistical significance of a SRF match in FF and VAS, 1000 random pairs of SRFs obtained in FF and VAS were generated and their optimal response area overlap calculated. Our sample neurons were considered to have produced a significant SRF overlap in FF and VAS (p<0.05) if their match exceeded the 95\(^{th}\) percentile of the overlaps generated across 1000 random SRF combinations taken from the same sample. Secondly, the overall agreement of individual neural discharges in FF and VAS was explored by a bootstrap analysis. For each unit and each recorded position, two random sub-samples of 5 spike counts (recorded over 5 stimulus repetitions in FF) were drawn from 10 available FF recordings (with replacement), and their spike difference vectors calculated. The procedure was repeated 1000 times and the distribution of spike count differences observed across all bootstrapped FF stimulus repetition sub-samples was computed. A second distribution of spike count differences was calculated using the difference vectors observed between one of the bootstrapped FF samples and a VAS sample of 1000 x 5 spike counts generated in the same manner. The two distributions of spike count differences (i) being ‘expected’ for identical stimulus repetitions and (ii) being ‘observed’ across FF and VAS stimulus repetitions were tested for significant deviations by a standard chi-square test. For plotting and graphically analyzing the neural responses and acoustic properties, the data obtained at spatially distinct locations (27, 80, and 393 positions respectively) were interpolated using a spherical thin plate spline (Wahba 1981) across the whole
sphere of space. An estimate of the centre of gravity (centroid) was calculated for each SRF by treating interpolated spatial responses as vectors, indicating the strength of the discharges elicited. Vectors across 2601 interpolated data points on the sphere were summed up to determine the centroid vector. The interpolation algorithm estimated a neural response in the unrecorded spherical region at low elevations, and thereby eliminated inconsequential zero length response vectors to minimize the bias of centroid locations towards high elevations.

The recording and analysis of action potentials and acoustic signals, as well as the stimulus generation and delivery were executed by custom-written software (MatlabR12; supported by J. Leung, University of Sydney, NSW).

All experiments and procedures were approved by the Animal Ethics Committee of the University of Sydney.
Results

Responses were recorded from 183 midbrain units. The locations of the 116 neurons were stereotactically reconstructed based on microscopic assessment of frontal sections of the brain regions tested. Unit recordings were made from the central nucleus of the inferior colliculus (ICC; N=75), the border region between the rostral ICC and the superior colliculus (SC; N=17), and the anatomical transition between the ICX and the brachium of the inferior colliculus (BIC; N=24), respectively (taking into account an accuracy of ± 150 µm). Figure 3A depicts a frontal section illustrating electrode tracts (indicated by arrows) which indicate that recording sites along these penetrations were localized in the ICC and dorsal tegmental areas. The outline of the guinea pig IC is indicated in figure 3B. For all recordings, spatial receptive fields (SRFs) were computed based on a unit’s response rate or 1st spike latency, respectively (see methods).

Spontaneous activity

Roughly one third (29%) of the midbrain units showed low spontaneous activity (<1 spike/s). Some 62% of the units displayed moderate (1-10 spikes/s) and another 10% showed high spontaneous activity (>10 spikes/s).

Response patterns

The distribution of response patterns is shown in figure 3C. Of the 183 acoustically responsive cells isolated by PCA, 44% responded to the onset of the broadband noise stimulus and 34% displayed sustained discharges. Transient stimulus offset and onset-offset triggered spiking was observed for a total of 20 units or 10% of our sample. Primary-like spike patterns and complex discharges were recorded for 14 and 6 units, respectively.

Latencies in FF and VAS

The 1st spike latency has been shown to systematically vary across auditory SRFs (Brugge et al. 1996). To assess the temporal fidelity of the VAS stimulus system compared to FF stimulation the neural 1st spike latency and spike jitter were analysed in both regimes. At a stimulus location directly ahead of the animal (0º azimuth and elevation) the onset response latencies varied across all units from 3.3 ms to 38.9 ms.
in FF, and 2.1 ms to 40.8 ms in VAS (figure 4A; mean 1st spike latency over 10 stimulus repetitions; late responses >50 ms and offset responses excluded). The latencies were highly correlated between the two acoustically matched configurations FF and VAS (figure 4B; r=0.90; p<0.001; N=96). Over the sample, the mean response latency was 14.9 ± 8.3 (SD) ms in FF, and 15.1 ± 8.3 (SD) ms in VAS, and latency differences remained at an insignificant level (paired t-test; p=0.5). Thus, little or no systematic shift of response latencies was induced by acoustic stimulation in either FF or VAS. Temporally unstable recordings were excluded from the mean latency calculation. An individual average response jitter variation of less than 2 ms in FF and VAS was used as a criterion for temporal stability of a recording. This seemed appropriate since the latency match and the jitter match across FF and VAS were only weakly correlated (r=0.44; p<0.001; N=111). Averaged over all stable units the observed response jitter was unchanged in FF and VAS. Individual neurons displayed jitter values from 1.0 ms to 30 ms (figure 4C). The observed individual latency deviations between FF and VAS increase with the individual neural jitter (figure 4D; r=0.87; p<0.001; N=111). Units with large (but stable) jitter values displayed the largest deviations in latency between FF and VAS. For individual units showing very stable jitter characteristics over the course of a recording (FF-VAS jitter difference <1 ms; N=82), the absolute latency difference values were averaged over the sample, so as to estimate the mean individual latency deviation in FF and VAS. Such calculated differences of the response latency across the SRF overlap positions between FF and VAS yielded a mean absolute value of 2.65 ± 2.22 (SD) ms, and varied from 0.3 ms to >10 ms (the latter were considered outliers; N=6). Hence, the mean individual latency deviation across the two configurations was well within the range of the mean latency jitter in this sample (4.43 ± 3.29 (SD) ms in FF; 4.50 ± 3.22 (SD) ms in VAS). On a spherical projection plot the 1st spike latency pattern in FF was well preserved in VAS (right panels in figure 5). For spatial acoustic stimulation in both FF and VAS, the location-dependent neural latency plots (on a logarithmic scale) approximated those of (linearly scaled) spike response counts of the neurons in both size and shape (left panels in figure 5).

*Spikes count based spatial receptive fields in FF and VAS: Controls*

*Effects of the stimulus regime on spatial response properties*
Some intrinsic variation in the formation of SRFs could only be assessed by repeatedly recording units that responded to a set of unchanged spatial acoustic stimuli. For 6 units the SRF was tested repeatedly in both FF and VAS. The average overlap of SRFs was calculated for both FF versus VAS stimulation (figure 6A; open triangles) and repeated FF stimulation (1st versus 2nd recording; black squares). The SRF match for repetitive recordings in FF was strongly correlated to the SRF overlap observed in FF and VAS for this small sample of units (figure 6B; r=0.98; p<0.05). And despite the variation of SRF overlap across our units, the observed differences between the test sets (FF-VAS and FF-FF repeats, respectively) remained at an insignificant level (paired t-test; p=0.67). Since repeated FF recordings were obtained subsequently from individual units under constant stimulus conditions, the SRF overlap calculated across those repeats presumably reflects the stability of the recordings over the period of observation. Hence, a poor match of SRFs between FF and VAS also does not necessarily reflect low fidelity VAS. We have chosen random pairing of SRFs obtained in FF and VAS to explore the chance level of SRF overlaps. The random pairing of 1000 SRFs recorded in FF and VAS yields an average SRF overlap of 38 ± 18.6 (SD) %. Given a normal distribution of chance level SRF overlaps, two thirds of random pairs generate SRF overlaps in FF and VAS between 19.4% and 56.6%. To a first approximation, extremely unstable recordings will generate a “random” SRF overlap distribution not unlike the artificial random pairing. Consequently, a conservative cut-off criterion of 50% SRF overlap between FF and VAS seemed appropriate to eliminate the bulk of those units with an unstable recording history from further analysis. This selection process aimed to minimize the bias in the VAS performance assessment. Consequently, only selected units meeting the cut-off criterion (N=87) were used for further analysis of spike count based SRFs in FF and VAS. For the selected sample, 1000 random combinations of SRFs in FF and VAS yield an increased average overlap of 47.1 ± 19.4% (SD). Given that mostly unmatched, i.e. erratic and incomplete SRFs were excluded, the remaining SRFs recorded under stable conditions tend to be (i) more homogeneous and (ii) centred contralaterally, which favours random overlapping. In a second step, a bootstrap analysis was employed to individually verify neural responses in FF and VAS (see methods). For each unit, the spike count difference distribution observed over of two bootstrapped FF stimulus repetition samples (N=1000 at each recorded position) was tested against the spike count difference distribution computed across FF and VAS.
samples equally generated. Significant discharge deviations using FF and VAS stimuli were found in 26 units ($p \leq 0.05$; chi-square test). For the vast majority, 70% of 87 units, the spike count differences registered across FF and VAS stimulation regimes were within the spike count variation calculated from FF stimulation only. The actual mean overlap of SRFs obtained in FF and VAS (71.3 ± 12.6 (SD) %) closely approximated the mean SRF overlap across subsequent FF-FF recordings conducted in unchanged acoustic environments (70.2 ± 14.2 (SD) %).

**Effects of stimulus sound pressure levels on spatial response properties**

Figure 7 exemplifies a typical unit recorded in FF and VAS, firstly, close to response threshold levels and, secondly, at a sound pressure level increased by 20 dB. The neuron responds to an increased SPL with a larger SRF (e.g. in FF: now covering 60% of the recorded positions instead of 6% at threshold SPL levels) and only slightly increased spike count (see colour bars). However, some more complex units responded to an increase in the stimulus SPL with very little or no changes in the size of the SRF and concurrently even decreased spike rates could be observed (figure 8). Yet, for all cases these effects obtained using FF stimuli were faithfully replicated using stimuli presented in VAS. The SRF overlap obtained with FF and VAS stimulation was well correlated across different SPLs ($r=0.82$). 6 units were repeatedly recorded in FF and VAS, presenting spatial stimuli at least at two different sound pressure levels between 0 and 20 dB above neural response thresholds. The SRF overlap in FF and VAS changed by an average of 9.0 ± 2.6 (SD) % per any given SPL-change. This variation was well within the normal overlap variation for repeated recordings in constant conditions shown in the previous section (10.8 ± 6.1 (SD) %; N=5). Across the set of neurons tested, the SRF overlap between FF and VAS did not significantly change when stimulus levels were varied concurrently in both configurations (paired t-test $p=0.51$; N=6). The data indicate that the VAS system replicated FF largely independent of the sound pressure levels applied.

**Spike count based SRFs in FF and VAS: Validation**

**Spike count in FF and VAS**

The position dependent maximal discharge elicited in FF and VAS was computed for stable unit recordings and then averaged over the sample (N=87). In terms of spike
count, the elicited responses in both virtual and real configurations differed very little. The average maximal response was 90 spikes per 10 stimuli in FF and was strongly correlated to the mean maximum discharge in VAS over 10 stimulus repetitions (89 spikes; r=0.97; p<0.001). No significant differences were observed across the sample (paired t-test; p=0.53). The latter also holds for the minimal discharges in FF and VAS (paired t-test; p=0.38). The average minimal response, taken over the positions where the least spikes have been evoked by repetitive acoustic stimulation, were 18 spikes per 10 stimuli in FF and 17 spikes likewise in VAS (r=0.94; p<0.001).

**SRF overlap in FF and VAS**

Figure 9A displays the SRF overlap in FF and VAS across all PCA units selected for stable recording conditions (mean 71.3 ± 12.6 (SD) %; N=87; left bar). These data demonstrated that at the level of the auditory midbrain neurons VAS stimulation can mimic real free field stimulation to a high level of fidelity. N=35 of 87 units showed a SRF overlap of \( \geq \)75% when stimulated in FF and VAS, including two units featuring a 100% SRF overlap across stimulus modes.

To assess the statistical significance of the overlap observed in our sample we employed again a randomly combined sample of recorded SRFs in FF and VAS to create a chance level distribution of respective SRF overlaps. 1000 random pairs were computed from the sample of 87 units. The right bar in figure 9A shows the mean overlap of SRFs so randomly generated (47.1 ± 19.4 (SD) %). Based on units that were stimulated over 80 positions, 1000 random SRF pairs across FF and VAS generated a 95\(^{th}\) percentile of 72% SRF overlap. The 95\(^{th}\) percentile represented the hypothetical SRF overlap shown in only the best 5% of 1000 random SRF pairings. In our real sample, however, 41% of the units (36 of 87 recorded pairs) displayed an individual SRF overlap in FF and VAS which matches the best 5% of a hypothetical chance level distribution of SRF overlaps. The performance was independent of whether 27 or 80 positions were tested (for 27 positions tested 40% of the units exceeded the 95\(^{th}\) percentile of the respective random SRF overlap distribution). In statistical terms, 41% of units subsequently recorded in FF and VAS have produced an SRF overlap which was significant at a p<0.05 level. Across the sample of stable recordings (N=87), the observed average SRF match between FF and VAS closely approximated the 95\(^{th}\) percentile of a random SRF overlap distribution.
For FF stimulation 23% of the units had SRFs of less than a spherical quadrant (VAS 22%), 48% between a quadrant and a hemisphere (VAS 46%), and 29% greater than a hemisphere (VAS 32%). In FF the SRF covered an average of 44.5 ± 18.0 (SD) % of the recorded spatial positions compared to 45.5 ± 18.7 (SD) % in VAS. Figure 9B depicts the size of SRFs in both FF and VAS for each unit. The SRFs size was well correlated in FF and VAS (r=0.83; p<0.001). Across stimulus configurations SRF size differences remained at an insignificant level (paired t-test; p=0.38).

**SRF centre of gravity in FF and VAS**

A measure of the centre of gravity of the spatial response of a neuron is the centroid. The centroids of neural responses were estimated for each SRF by integrating the weighted neural discharges and the position from which they were elicited (see methods). Figure 10A depicts all centroids computed for recordings in FF (blue triangles) and VAS (purple crosses). The background of the plot exemplifies a typical individual distribution of broadband noise ILDs calculated from the HRTF recordings (post surgery). The bulk of the centroids were located in the contralateral hemisphere around -90° azimuth, corresponding to the high ILDs. The deviation of the average centroid position between FF and VAS was small (7.4° azimuth; 3.3° elevation; N=87).

**Spike count based SRFs in FF and VAS: Applications**

**Spatially randomized VASRAN**

In a first step exploiting the flexibility of our validated VAS system, we pseudo-randomized the spatial stimulation sequence (VASRAN) to test for possible spatial adaptation effects of sequentially presented stimuli. The average centroid position of receptive fields observed under sequential and spatially randomized stimulus regimes barely differed (N=26). Across this sample of units, figure 10B indicates the average centroid position for all tested stimulus configurations in blue (triangle; FF; -89.1° azimuth; 6.5° elevation), purple (cross; VAS; -90.0°; 6.7°), and red (dot; VASRAN; -92.1°; 10.2°).

Figure 11 also indicates that in terms of the SRF shape, little adaptation was observed for sequential stimulation when compared to randomized presentation.
Taken across this subset of neurons, the mean SRF overlap in FF and VAS was 75.6 ± 13.6 (SD) %. The average SRF overlap obtained in FF and spatially randomized VASRAN was almost identical (76.1 ± 14.7 (SD) %), indicating that the sequential ordering of the stimuli had no significant effect on the recorded SRFs. In addition, the mean SRF overlap obtained across the two virtual stimulus configurations VAS and VASRAN was very close (79.3 ± 12.9 (SD) %). Since in the latter case the stimuli were identical (and the altered sequence showed little effect before) much of the observed SRF deviation is probably the result of intrinsic changes in neural response or recording properties over a time course of 40-60 minutes. This holds true for the equally small SRF deviation between FF and VAS or VASRAN, respectively.

The average size of receptive fields in sequential (FF and VAS) and randomized stimulus configurations (VASRAN) differed only by 0.3% of the recorded sphere (41.0 ± 21.0 (SD) % and 40.7 ± 21.8 (SD) %, respectively). This supports the overall reliability of the sequential stimulus paradigm routinely used to obtain SRF characteristics in FF and VAS in this and many previous studies.

However, the range of spike discharges did differ across conditions (see colour code in figure 11). The average maximum response to random stimulation tended to increase the number of spikes relative to sequential presentation (by 7 spikes per 10 stimuli from 83 to 90 spikes for the sub-sample of N=26). This equals a rise in spike count of 8%. However, across the population tested the trend to increased maximal responses was not statistically significant (paired t-test; p=0.26). Only a set of units recorded in the border region between the rostral ICC and the SC (N=13) displayed significantly increased maximum responses to spatially randomized stimulation (i.e. a rise in spike count of 18%; paired t-test; p<0.05 re sequential FF and p<0.01 re sequential VAS). In contrast, the average minimal response, which was usually evoked in the ipsilateral hemisphere, decreased significantly over all tested units (paired t-test; p=0.05 re sequential FF and p<0.01 re sequential VAS). The spike count dropped from a mean of 9.3 spikes per 10 sequential stimuli to only 5.5 spikes per 10 randomized stimuli (59% of the sequential response). On possible explanation of this observation is that ipsilateral inhibitory inputs may be more susceptible to adaptation under sequential stimulation.

*Minimal surgery HRTF recordings: ‘Naturalized’ VASNAT*
To interpret how well the SRFs recorded in ‘standard’ FF and VAS reflected physiologically relevant response patterns, acoustic artefacts were minimized for a subset of unit recordings. Recall that for all neural recordings presented so far (i) the recording environment was acoustically distorted by the presence of electrophysiological recording equipment, and (ii) the animals’ HRTFs were distorted by the skull surgery performed to access the brain, for instance by changes of the pinna position. To eliminate these effects VAS was generated from HRTFs recorded in the absence of any physiological recording apparatus, and before skull surgery was performed. In this somewhat ‘naturalized’ (VASNAT) environment the SRFs of 5 units were recorded and tested against SRFs obtained in FF and corresponding ‘standard’ VAS. Figure 12 displays the responses of an ICC unit to VASNAT versus the discharges recorded subsequently in the three standard configurations FF, VAS, and VASRAN. As discussed above, the receptive fields in all three standard stimulus regimes were very well matched, regardless whether presented in real or virtual field and sequential or random order (76-88% overlap). Against all standard configurations the SRF recorded in VASNAT stood out for its poor SRF match (32-41% overlap), showing a substantially altered shape which was split into two response areas. Its SRF shape suggests that under ‘naturalized’ conditions the neural representation of space differs from the one observed using a standard neurophysiological recording environment. Across the five units tested the parameters describing SRFs only partially reflected these observations (figure 13A; N=5): The average centroid position of SRFs obtained in ‘standard’ VAS was near the average centroid calculated from VASNAT stimulation (mean shift: 1° azimuth; 3.8° elevation). The average size of SRFs recorded in VAS and in VASNAT was also stable (28 ± 12.6 (SD) % and 25.4 ± 11.8 (SD) % of the sphere, respectively). However, figure 13B depicts that the average overlap of SRFs obtained in FF and its corresponding ‘standard’ VAS (both incorporating equipment and surgery effects) was far greater than the overlap of SRFs in ‘standard’ VAS and VASNAT. Apparently, receptive fields obtained in an acoustically ‘uncluttered’ and surgically undistorted VASNAT environment generally differed in shape from SRFs observed in ‘standard’ conditions (paired t-test; p<0.01; N=5).
Discussion

While much has been published concerning the nature of the neural representation of acoustic space, the limitation of the traditional free field delivery systems have hampered investigations using more naturalistic stimulus paradigms. These involve multiple stationary and/or moving sound sources and varying acoustic environments. The advent of virtual auditory space (VAS) stimulation techniques provides a very promising approach to addressing these shortcomings. Yet, in general, it remains an open question as to what extent the conclusions derived from current VAS experiments on the neurophysiology of spatial hearing sufficiently apply to the real world. Here we have presented the first systematic study addressing the matter in mammals, aided by real-time control of the acoustics to accurately match the free field recording environment and VAS. Neural responses of individual midbrain units to stimuli presented in both FF and VAS could well be affected by the subtle acoustic differences between stimuli presented in either mode. Therefore, it was critical to systematically assess the impact of either stimulation regime on the spatial receptive fields (SRFs) recorded from the same unit under both conditions.

The analysis of the correspondence of neural responses addressing spike timing and spike rate in both stimulus configurations provides a quantification of the fidelity of the VAS system in use. To optimize the VAS presentation used in our study, the bandwidth of VAS stimuli (0.3-30 kHz) was significantly greater when compared to previous studies (e.g. Hartung and Sterbing 1997; Sterbing et al. 2002; 0.2-16 kHz), and consequently covered most of the guinea pig audiometric range. A default stimulus SPL of 20 dB above threshold determined using a contralateral stimulus location, was chosen to ensure good binaural interactions in the recorded SRFs in this study (King and Palmer 1983; Palmer and King 1985). Earphone calibration has been aided previously by precise positioning of the earphone and measurement probe (Poon and Brugge 1993b; Hartung and Sterbing 1997), however, we have experienced dramatic effects of small position deviations on the transfer functions of the VAS delivery system. We have therefore tried to improve on the overall stability of the system by carrying out online calibration of the drivers prior to each unit recording. Most importantly, the use of the embedded microphone for both recording the HRTFs and calibrating the VAS stimuli ensured that the pattern of sound waves at the point of
the recording microphone, and hence at the ear drum, was identical in both FF and VAS stimulus conditions.

We chose to examine the SRFs of units recorded from the ICC, ICX, and BIC as these regions are known to either provide essential acoustic input to spatially tuned nuclei like the SC or to be spatially tuned themselves, or both (Kudo et al. 1984; Fuzessery et al. 1985; Wenstrup et al. 1988; Binns et al. 1992; Schnupp and King 1997; Doubell et al. 2000). In our study response patterns, latencies, and spontaneous activity were found to be typical for the mammalian IC (Aitkin et al. 1975; Hartung and Sterbing 1997; McAlpine et al. 2001; Sterbing et al. 2003). SRFs at this level are thought to be mainly based on directional amplifying effects of the outer ear (Calford and Pettigrew 1984; Moore et al. 1984; Fuzessery et al. 1985; Middlebrooks et al. 1989; Aitkin and Martin 1990; Musicant et al. 1990). Consequently, spatial responses reflect a convolution of these frequency dependent directional pinna effects and a unit’s frequency sensitivity (Semple et al. 1983). While our data were recorded using broad band noise stimuli, the overall directional properties of the SRFs recorded are consistent with this view in that the centroids of neural responses were mainly found at azimuth angles close to that of the acoustic axis of the pinna. Average neural SRF centroids were tuned to positions near the audio-visual plane which is in accord with a strong neural representation of this area demonstrated earlier (Sterbing et al. 2003).

**Variability and stability observed in FF and VAS**

Generally, we have observed a fair amount of variability in the neural responses to spatial acoustic stimulation across both the FF and VAS conditions, and also for repeat recordings in any one condition. Earlier studies have reported similarly, yet did not focus on the variability (King and Palmer 1983; Middlebrooks and Knudsen 1984; Palmer and King 1985). Possible contributions to such variations could include intrinsic neural processes or ‘neural noise’, e.g. trial-by-trial variability of neural responses (Eggermont 1992; Ringach, 2003) and external factors such as VAS fidelity and/or stability of the recording conditions and physiological changes. We have endeavoured to control as carefully as possible physiological changes, e.g., by the use of a continuous anaesthetic regime and of Hypnorm™ instead of barbiturates so as to avoid previously observed depressive affects (Poon and Brugge 1993b; Brugge et al. 1994; Brugge et al. 1996; Imig et al. 2000).
For animals with moveable pinnae, a strong effect on HRTFs and neural responses of spatially tuned units has been observed for different pinna positions (Calford and Pettigrew 1984; Jen and Sun 1984; Middlebrooks and Knudsen 1987; Young et al. 1996). Therefore, even though the pinna of guinea pigs only have a small range of active movement, accidental changes in their position could have contributed to changed neural SRFs.

Most importantly, our study has shown that the differences observed across repeated recordings in unchanged stimulus conditions were very similar to those deviations of SRFs observed in FF and VAS. This suggests that the bulk of the variation observed between the stimulus conditions simply represents intrinsic response variation over time rather than differences in the acoustics of the stimulation. Also, the VAS stimulus system behaved linearly with respect to stimulus amplitude and the SRFs obtained at different sound levels with VAS stimuli matched well the SRFs obtained using free field stimuli at comparable sound levels. Overall, different response patterns to increased SPLs were observed in the population of neurons tested at more than one sound level. Some SRFs resembled so-called ‘bounded’ neurons for which inhibitory input had been predicted or even demonstrated to be responsible in mammals and specialized birds, namely barn owls (Knudsen 1982; King and Palmer 1983; Middlebrooks and Knudsen 1984; Palmer and King 1985; Brugge et al. 1994; Brugge et al. 1996). These neurons feature a relatively constant SRF across different stimulus levels but in our sample only represented a small subset of neurons. The majority of units in our sample demonstrated expanded SRFs for increased stimulus levels (see also King and Hutchings 1987).

In the owl auditory midbrain, which is homologous to the mammalian IC, neural SRFs might be organized in a centre-surround fashion different from SRFs found in mammals (Knudsen and Konishi 1978b). Still, the only other study systematically investigating subsequent single unit recordings in FF and VAS was conducted in barn owls. Keller and colleagues observed a remarkable stability of receptive fields recorded in FF and VAS (Keller et al. 1998), which was assessed by the calculation of a correlation coefficient derived from aligned SRFs. Their report supports the view that in barn owls VAS stimuli could largely simulate an acoustic FF environment on a neural level. Due to the different data analysis applied to measure the stability of SRFs across stimulus configurations a quantitative comparison between Keller’s and our study is hampered, however.
An important question is whether a VAS system not only delivers proper spectrally filtered stimuli, but also does so with sufficient temporal accuracy to support neural codes based on first spike latencies or spike pattern analysis (Brugge et al. 1996; Middlebrooks et al. 1998; Furukawa et al. 2000; Furukawa and Middlebrooks, 2002; Middlebrooks et al. 2002). The performance of the VAS system tested yielded an average temporal match of 1st spike timing in FF and VAS which was well below the trial-to-trial jitter of the recorded units. No systematic shift was observed for response latencies in FF and VAS, and across both configurations the difference in the average latency remained small and at an insignificant level. Whether the temporal acuity of VAS stimuli supports the investigation of neuron populations specialized in the processing of very small interaural time differences (McAlpine et al. 2001; Brand et al. 2002) remains to be investigated. The spatial map analysis of first spike timing revealed a picture of decreasing latency values approaching the centre of a receptive field as was observed in earlier studies on cats (Brugge et al. 1996). Just as for the spike rate based SRFs, the maps of SRFs based on 1st spike timing were well matched for both FF and VAS stimulus configurations. Thus, our results indicated that VAS stimulation properly supports the spatially dependent temporal pattern formation incorporated into the SRFs observed in FF.

**Spatially randomized VASRAN and naturalized VASNAT**

For a smaller sample of neurons VAS stimuli were not only presented in sequential repetitions at each position (to match sequential FF stimulation) but also in a spatially randomized fashion (VASRAN). The overall stimulus repetition rate and the VAS stimuli were identical for both conditions. A long inter-stimulus interval was chosen for the sequential presentations to avoid adaptation effects, and accordingly randomized VASRAN yielded no dramatic effect on the formation of SRF in terms of size or degree of overlap with FF or VAS. The lack of strong adaptation effects supports the overall reliability of the sequential stimulus paradigm for an assessment of the VAS system over a long time period. We have observed, however, a trend to larger maximal responses using VASRAN stimuli and have demonstrated a significantly smaller minimal response in SRFs when compared to sequential stimuli presentation. This results in an expanded dynamic range of neural discharge across a receptive field, and is in line with an increased sensitivity for novel spatial stimuli. Our data is consistent with the idea that the decrease of the minimal response may
result from an adaptation of ipsilateral inhibition, since in sequential presentation the ipsilaterally elicited responses tended to strengthen over the course of repeated stimuli, whereas in a random stimulus regime ipsilateral responses remained weak over all stimulus repeats. With respect to spatial hearing, such ipsilateral inhibitory input to the IC has, for instance, been demonstrated to underlie apparent motion detection, and the formation of a directional best response in the guinea pig (Ingham et al. 2001). As a principle, it has long been shown in the spatially tuned SC that ipsilateral inhibitory inputs were pivotal for the formation of a spatial map for supra threshold stimulation (Wise and Irvine 1983; Palmer and King 1985; Wise and Irvine 1985).

We also recorded the VASRAN sample with a new set of tungsten electrodes, which yielded more stable recording conditions. As a by-product of these recordings a small sample of FF and VAS recordings were obtained which showed an even stronger overlap of SRFs between FF and VAS up to an average of 76% (N=26). Consequently, this smaller sample leads us to speculate that under stable conditions, a VAS system has little limitations and can almost perfectly reconstruct a FF setting. Interestingly the average overlap of SRF across virtual acoustic configurations (VAS and VASRAN) was only marginally greater than the overlap of SRF obtained in FF and virtual space. This margin of ± 3% SRF overlap might, on a neural level, reflect the true technical limit of the VAS sound delivery system in reconstructing a FF environment.

In a further step, HRTFs were obtained at different stages of surgery and in a reduced setup that was stripped down to the essential equipment for acoustic recordings. The virtual stimuli based on these HRTFs were called VASNAT due to the minimized incorporation of acoustic artefacts introduced by the neurophysiological recording apparatus. The SRFs elicited in this environment were tested against those elicited in the standardized VAS, which was designed to closely match FF conditions. It has been shown earlier that in mammals, but not so much in barn owls (Keller et al. 1998), that the use of non-individual HRTFs to generate VAS results in profoundly different SRFs (Mrsic-Flogel et al. 2001; Schnupp et al. 2003). Comparable effects were observed in regard of the maturation of ferrets when using HRTFs obtained from adult and juvenile animals for neural stimulation (Mrsic-Flogel et al. 2003). Thus it was not surprising to see that the SRFs recorded in VASNAT differed significantly in shape from those obtained in standardized VAS. It was
surprising, however, that the SRFs in VASNAT did occasionally split up in a manner similar to that which has been described for stimulation with VAS based on non-individualized or foreign HRTFs (Mrsic-Flogel et al. 2001; Sterbing et al. 2003). These were anecdotal observations, however, supported only by a small number of neurons. Still, the deviating responses in VASNAT (compared to both FF and standard VAS), underline the importance of virtual stimulation for probing the neural representation of acoustic space. Only in ‘natural’ virtual space can the acoustic distortions induced by surgery and the free field equipment be eliminated.

**Conclusion**

Our study supports the assumption that VAS largely resembles a free field acoustic environment, and that this similarity is reflected by neural responses in the mammalian auditory midbrain. Yet it also emphasized the need for careful calibration of a VAS system with respect to variations in the equipment configuration, such as those, for instance, induced by changes in earphone position. In contrast to free field, studies in virtual acoustic space can suffer from both low recording stability and low fidelity spatial stimulation, either of which limits the conclusions that can be reliably drawn. To optimize the spatial fidelity of a VAS sound delivery system, feedback obtained from neurophysiological recordings could also help to match FF and VAS responses at certain test positions.
Acknowledgements

We are indebted to Johan Leung for his continuous support, Jan Schnupp for fruitful discussions, and to Anandhi Anandan for technical help.

Current address of C. Jin: School of Elect. and Inform. Eng., Bld. J03, University of Sydney, Sydney, NSW 2006, Australia. E.Clarke: Prince of Wales MRI, Barker St., Randwick, NSW 2031, Australia. O. Behrend: Institute of Biology, Humboldt University Berlin, Invalidenstr. 43, 10115 Berlin, Germany

Grants

This work was supported by the National Health and Medical Research Council (NHMRC) and the German Academy of Natural Scientists Leopoldina (Deutsche Akademie der Naturforscher Leopoldina).
Figure legends

Figure 1
The 80 standard positions for spatial stimulation around the animal (black dots).

Figure 2
A A spectrally flat noise stimulus (0.3-30 kHz) delivered through the LMS-calibrated earphone system and recorded at the left ear canal (left panel) and the right ear canal, respectively (right panel). B Superimposed power spectral densities of two corresponding FF and VAS stimuli (90° azimuth; 30° elevation), subsequently recorded in the ipsilateral ear canal. C The same two stimuli superimposed in the temporal domain.

Figure 3
A Frontal section showing penetration tracts left by the recording electrode in the ICC (arrows). B The principal structures are indicated schematically. IC – inferior colliculus; LL – lateral lemniscus; PAG – periaqueductal grey. C Observed auditory response types (N=183).

Figure 4
A Mean 1st spike latency of units stimulated at a frontal position (0° azimuth and elevation) in FF (grey bars) and VAS (black bars; likewise in C). B Correlation plot of 1st spike latencies in FF and VAS (r=0.90; p<0.001; N=96). Dotted line indicates r=1.0, solid line indicates data trend (likewise in D). C Mean 1st spike jitter of units across their shared receptive field in FF and VAS. D Correlation plot of 1st spike
latency deviation and average 1st spike jitter in FF and VAS \((r=0.87; \ p<0.001; N=111)\). All plots based on mean spike timing over 10 stimulus repetitions.

**Figure 5**

**Upper panels:** SRFs under FF stimulation computed from spike count (**left panel**), and 1st spike latency (**right panel**). **Lower panels:** The same unit’s SRFs in calibrated VAS. The 1st spike latency pattern observed in FF was clearly preserved. Note that the outline of the 1st spike latency SRF fairly resembles the SRF based on the spike count in both FF and VAS. Colour bars indicate the number of spikes elicited, or the average 1st spike latency respectively, across 10 stimulus repetitions. The size of the spike count based receptive field (% of recorded sphere) and the SRF overlap between FF and VAS is indicated below the respective maps.

**Figure 6**

**A** Average overlap of the SRFs of 6 midbrain units for repeated recordings in FF (black squares) versus the average overlap of the same units’ SRFs derived from FF and VAS stimulation (open triangles). **B** Correlation of the SRF overlap of the two combinations in **A**. The correlation coefficient is indicated above the graph.

**Figure 7**

**Upper panels:** SRF of a unit recorded in FF stimulation close to the neural threshold (**left**) and 20 dB above threshold (**right**). **Lower panels:** The same unit’s SRF in calibrated VAS. The size of the receptive field (% of recorded sphere) and the SRF overlap between FF and VAS is indicated below the maps. Colour bar indicates the number of spikes elicited for 10 stimulus repetitions.
**Figure 8**

**Upper panels:** SRF of a complex unit for FF stimulation close to the neural threshold (*left*), and 10 dB above threshold (*right*). **Lower panels:** The same unit’s SRF in calibrated VAS. Note that for lower stimulus SPL the SRF size is only slightly smaller (compare figure 7), and the spike count increased. The size of the receptive field (% of recorded sphere) and the SRF overlap between FF and VAS is indicated below the maps. Colour bar indicates the number of spikes elicited for 10 stimulus repetitions.

**Figure 9**

**A** The average SRF overlap in FF and VAS across two samples: Firstly, over all units recorded in FF and VAS (in stable recording conditions; N=87), secondly, across a hypothetical sample of units generated by random combination of the SRFs observed in the first sample (see text; N=1000). **B** Correlation plot of the SRF size in FF and VAS observed over the 87 units in **A** (r=0.83; p<0.001). Dotted line indicates r=1.0, solid line indicates data trend.

**Figure 10**

**A** The centroid positions of the SRFs of auditory midbrain units stimulated in FF (blue triangles) and VAS (purple crosses; N=87). **B** Average centroid positions across units tested in FF (blue triangle), VAS (purple cross) and spatially randomized VASRAN (red dot; N=26). The positions are almost identical. The background depicts a typical individual ILD-distribution after surgery, computed across 393 spherical positions. Colour bar indicates the ILD.
**Figure 11**

**Left panel** A unit’s SRF under sequential FF stimulation. **Right panel** The same unit’s SRFs stimulated in calibrated VAS under a spatially pseudo-randomized regime (VASRAN). The SRFs are very well matched in FF, VAS (not shown) and VASRAN. In VASRAN a tendency to higher maximal and lower minimal responses was observed compared to sequential stimulus presentation in FF and VAS. The size of the receptive field (% of recorded sphere) and the SRF overlap between stimulus configurations is indicated for this unit below the maps. Colour bars indicate the number of spikes elicited across 10 stimulus repetitions.

**Figure 12**

**Left panels:** SRFs of a unit recorded under three different virtual stimulation regimes, ‘naturalized’ sequential VASNAT (see text; **top**), sequential standard VAS (**middle**), and spatially randomized VASRAN (**low**). **Right panel:** The same unit’s SRF in FF. For all but VASNAT stimulation the SRFs display great similarity. The size of the receptive field (% of recorded sphere) and the SRF overlap between stimulus configurations is indicated below the maps. Colour bar indicates the number of spikes elicited for 10 stimulus repetitions.

**Figure 13**

**A** Average size of the SRFs of 5 midbrain units stimulated in both ‘standard’ VAS (**black bar**) and ‘naturalized’ VASNAT (**grey bar**). No significant difference was observed. Also, the average SRF centroid position in VAS and VASNAT displayed very little change (positions indicated above the data bars, respectively). **B** However,
the average SRF overlap between ‘standard’ VAS and VASNAT recordings was significantly smaller than between VAS and FF recordings. Error bars show SD.
References


Figure 1  VAS in the guinea pig IC
Figure 2  
VAS in the guinea pig IC

A  
[b]  
Frequency [kHz]
0 10 20 30 40 50 60 70 80 90 100
-20 0 20 40 60 80

B  
[dB]  
Frequency [kHz]
0 5 10 15 20 25 30 35 40
0 10 20 30 40

C  
Stimulus magnitude
$1 \times 10^4$ $1$ $0.5$ $0$ $-0.5$ $-1$ $-1.5$

Time [sampling points/80kHz s. rate]
Figure 3  VAS in the guinea pig IC

A

B

C

<table>
<thead>
<tr>
<th>discharge pattern</th>
<th>% of units</th>
</tr>
</thead>
<tbody>
<tr>
<td>on</td>
<td>40</td>
</tr>
<tr>
<td>on-off</td>
<td>5</td>
</tr>
<tr>
<td>off</td>
<td>1</td>
</tr>
<tr>
<td>sustained</td>
<td>30</td>
</tr>
<tr>
<td>pri</td>
<td>10</td>
</tr>
<tr>
<td>complex</td>
<td>5</td>
</tr>
</tbody>
</table>
Figure 4  VAS in the guinea pig IC

A

![Histogram showing the distribution of latency for FF and VAS units.]

B

![Scatter plot showing latency VAS vs latency FF. Regression line: y = 0.9x + 1.8.]

C

![Histogram showing the distribution of jitter for FF and VAS units.]

D

![Scatter plot showing FF-VAS latency deviation vs average jitter. Regression line: y = 1.1x + 1.5.]

y = 0.9x + 1.8  
y = 1.1x + 1.5

Figure 4  VAS in the guinea pig IC
Figure 5  
**VAS in the guinea pig IC**

**SPIKE COUNT**

- **FF**
  - Gp230102cn3peak1
  - FF sphere cover: 50%

- **VAS**
  - VAS sphere cover: 57%
  - FF-VAS overlap: 80%

**LATENCY**

- [# of spikes/10 repetitions]
- [log₁₀ ms]

- FF sphere cover: 50%
- VAS sphere cover: 57%
- FF-VAS overlap: 80%
Figure 6  
VAS in the guinea pig IC

A

B

FF / FF overlap of SRF in %

FF / VAS overlap of SRF in %

r=0.98
Figure 7  VAS in the guinea pig IC

AT THRESHOLD

FF

contra  -180°  0°  +180°  ipsi
+60°
+30°
0°

Gp230102cn4pca2
FF sphere cover:  6 %

[ 0°  180°  360°  0°  180°  360° ]

[# of spikes/10 repetitions]

20dB ABOVE THRESHOLD

FF sphere cover:  60 %

VAS

+60°  -180°  0°  +180°
+30°
0°

VAS sphere cover:  5 %
FF-VAS overlap:  80 %

VAS sphere cover:  60 %
FF-VAS overlap:  92 %
Figure 8  

VAS in the guinea pig IC

AT THRESHOLD

FF sphere cover: 21%

10dB ABOVE THRESHOLD

FF sphere cover: 28%

[AT THRESHOLD]

contra  
ipsi

ipsicontra

frontal

Gp210202surgn2peak1

FF sphere cover: 21 %

[10dB ABOVE THRESHOLD]

FF sphere cover: 28 %

[AT THRESHOLD]

FF sphere cover: 21 %

0° -180° +180° 0° +60° +30°

FF-VAS overlap: 76 %

[10dB ABOVE THRESHOLD]

FF-VAS overlap: 64 %

VAS sphere cover: 19 %

VAS sphere cover: 23 %

[AT THRESHOLD]

0° -180° +180° 0° +60° +30°

FF-VAS overlap: 76 %

[10dB ABOVE THRESHOLD]

FF-VAS overlap: 64 %

[AT THRESHOLD]

ipsi  
contra

ipsicontra

frontal

Gp210202surgn2peak1

FF sphere cover: 21 %

[10dB ABOVE THRESHOLD]

FF sphere cover: 28 %

[AT THRESHOLD]

FF sphere cover: 21 %

0° -180° +180° 0° +60° +30°

FF-VAS overlap: 76 %

[10dB ABOVE THRESHOLD]

FF-VAS overlap: 64 %

[AT THRESHOLD]

ipsi  
contra

ipsicontra

frontal

Gp210202surgn2peak1

FF sphere cover: 21 %

[10dB ABOVE THRESHOLD]

FF sphere cover: 28 %

[AT THRESHOLD]

FF sphere cover: 21 %

0° -180° +180° 0° +60° +30°

FF-VAS overlap: 76 %

[10dB ABOVE THRESHOLD]

FF-VAS overlap: 64 %

[AT THRESHOLD]

ipsi  
contra

ipsicontra

frontal

Gp210202surgn2peak1

FF sphere cover: 21 %

[10dB ABOVE THRESHOLD]

FF sphere cover: 28 %

[AT THRESHOLD]

FF sphere cover: 21 %

0° -180° +180° 0° +60° +30°

FF-VAS overlap: 76 %

[10dB ABOVE THRESHOLD]

FF-VAS overlap: 64 %

[AT THRESHOLD]

ipsi  
contra

ipsicontra

frontal

Gp210202surgn2peak1

FF sphere cover: 21 %

[10dB ABOVE THRESHOLD]

FF sphere cover: 28 %

[AT THRESHOLD]

FF sphere cover: 21 %

0° -180° +180° 0° +60° +30°

FF-VAS overlap: 76 %

[10dB ABOVE THRESHOLD]

FF-VAS overlap: 64 %

[AT THRESHOLD]

ipsi  
contra

ipsicontra

frontal

Gp210202surgn2peak1

FF sphere cover: 21 %

[10dB ABOVE THRESHOLD]

FF sphere cover: 28 %

[AT THRESHOLD]

FF sphere cover: 21 %

0° -180° +180° 0° +60° +30°

FF-VAS overlap: 76 %

[10dB ABOVE THRESHOLD]

FF-VAS overlap: 64 %

[AT THRESHOLD]

ipsi  
contra

ipsicontra

frontal

Gp210202surgn2peak1

FF sphere cover: 21 %

[10dB ABOVE THRESHOLD]

FF sphere cover: 28 %

[AT THRESHOLD]

FF sphere cover: 21 %

0° -180° +180° 0° +60° +30°

FF-VAS overlap: 76 %

[10dB ABOVE THRESHOLD]

FF-VAS overlap: 64 %

[AT THRESHOLD]

ipsi  
contra

ipsicontra

frontal

Gp210202surgn2peak1

FF sphere cover: 21 %

[10dB ABOVE THRESHOLD]

FF sphere cover: 28 %

[AT THRESHOLD]

FF sphere cover: 21 %

0° -180° +180° 0° +60° +30°

FF-VAS overlap: 76 %

[10dB ABOVE THRESHOLD]

FF-VAS overlap: 64 %

[AT THRESHOLD]

ipsi  
contra

ipsicontra

frontal

Gp210202surgn2peak1

FF sphere cover: 21 %

[10dB ABOVE THRESHOLD]

FF sphere cover: 28 %

[AT THRESHOLD]

FF sphere cover: 21 %

0° -180° +180° 0° +60° +30°

FF-VAS overlap: 76 %

[10dB ABOVE THRESHOLD]

FF-VAS overlap: 64 %
Figure 9  
VAS in the guinea pig IC

A

<table>
<thead>
<tr>
<th></th>
<th>real SRF pairing (N=87)</th>
<th>random SRF pairing (N=1000)</th>
</tr>
</thead>
<tbody>
<tr>
<td>avg. SRF overlap FF / VAS in %</td>
<td>75</td>
<td>35</td>
</tr>
</tbody>
</table>

B

$y = 0.868x + 6.886$
Figure 10  VAS in the guinea pig IC
Figure 11  VAS in the guinea pig IC

FF

> contra -180° 0° +180° ipsi

Gp230102cn3peak2

FF sphere cover: 63%

VAS sphere cover (not shown): 70%

FF-VAS overlap (not shown): 89%

VASRAN

> -180° 0° +180°

VASRAN sphere cover: 70%

FF-VASRAN overlap: 86%

VAS-VASRAN overlap: 96%

[# of spikes/10 repetitions]
Figure 12  VAS in the guinea pig IC

VIRTUAL FIELD

VASNAT

Gp210202n8peak1
VASNAT sphere cover: 11%

VAS

VAS sphere cover: 19%
VAS-VASNAT overlap: 33%

VASRAN

VASRAN sphere cover: 19%
VASRAN-VAS overlap: 87%
VASRAN-VASNAT overlap: 40%

FREE FIELD

[# of spikes/10 repetitions]
Figure 13  VAS in the guinea pig IC

A

avg. centroid position (deg):
az.       -71.6       -72.6
el.       -9.8         -6.0

B

p=0.01