Neural Responses to Free Field and Virtual Acoustic Stimulation in the Inferior Colliculus of the Guinea Pig

Oliver Behrend,1 Benjamin Dickson,1 Elizabeth Clarke,1 Craig Jin,1 and Simon Carlile1,2

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

Submitted 20 April 2004; accepted in final form 21 June 2004

INTRODUCTION

In the vertebrate CNS, 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, brain stem, midbrain, thalamic, to cortical areas (Ahissar et al. 1992; Brugge et al. 1994, 1996; Furukawa and Middlebrooks 2002; Furukawa et al. 2000; Hartung and Sterbing 1997; Imig et al. 1997, 2000; King and Carlile 1994; King and Palmer 1983; May 2000; McAlpine et al. 2001; Middlebrooks and Pettigrew 1981; Middlebrooks et al. 1998, 2002; Poon and Brugge 1993b; Sterbing et al. 2003; Sutherland et al. 1998; Young et al. 1992, 1995). Particularly in barn owls, studies have demonstrated that a map of acoustic space based on interaural time difference (ITD) and interaural intensity difference (IID) cues can be found in the midbrain (Knudsen 1982; Knudsen and Konishi 1978a,c; Knudsen et al. 1991; Moiseff and Konishi 1983; Olsen et al. 1989). 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) (Binns et al. 1992; King and Hutchings 1987; King and Palmer 1983; Middlebrooks and Knudsen 1984; Schnupp and King 1997). Yet in the mammalian SC, it had long been shown that IIDs and ITDs alone are insufficient to compute unambiguous spatial receptive fields (SRFs) (King and Carlile 1994; Wise and Irvine 1983). 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 rostrocaudal 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 “externalization” localization (Angell and Fite 1901; Butler and Belendiuk 1977; Musicant and Butler 1984; Plenge 1974; Rayleigh 1907). In complementary animal experiments, encoding of monaural spatial cues has been demonstrated neurophysiologically at early levels such as the VIII nerve and the DCN (Imig et al. 2000; Poon and Brugge 1993a; Young et al. 1992). 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 (Carlile 1990a,b; Carlile and King 1994; Carlile and Pettigrew 1987a; Middlebrooks and Pettigrew 1981; Middlebrooks et al. 1989; Musicant et al. 1990; Pralong and Carlile 1994; Rice et al. 1992; Shaw 1974).

The peripheral filters were described as head-related transfer functions (HRTF) or head-related impulse responses, HRIR (Brugge et al. 1994; Hartung and Sterbing 1997; Mehrgardt and Mellert 1977; Mellert et al. 1974; Poon and Brugge 1993b; Wightman and Kistler 1989a). 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 (Brugge et al. 1994, 1996; Chen et al. 1995; Fite 1901; Butler and Belendiuk 1977; Musicant and Butler 1984; Plenge 1974; Rayleigh 1907). In complementary animal experiments, encoding of monaural spatial cues has been demonstrated neurophysiologically at early levels such as the VIII nerve and the DCN (Imig et al. 2000; Poon and Brugge 1993a; Young et al. 1992). 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 (Carlile 1990a,b; Carlile and King 1994; Carlile and Pettigrew 1987a; Middlebrooks and Pettigrew 1981; Middlebrooks et al. 1989; Musicant et al. 1990; Pralong and Carlile 1994; Rice et al. 1992; Shaw 1974).

The peripheral filters were described as head-related transfer functions (HRTF) or head-related impulse responses, HRIR (Brugge et al. 1994; Hartung and Sterbing 1997; Mehrgardt and Mellert 1977; Mellert et al. 1974; Poon and Brugge 1993b; Wightman and Kistler 1989a). 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 (Brugge et al. 1994, 1996; Chen et al. 1995;
Hartung and Sterbing 1997; Keller et al. 1998; Poon and Brugge 1993b; Reale and Chen 1996; Schnupp et al. 2001; Semple 1998; Wightman and Kistler 1989b). 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 FF stimuli (Brugge et al. 1994, 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 composed of two major components under remote control. First a FF speaker mounted at a robot arm covering spherical positions between −45 and 90° elevation, and second 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 on-line calibration of the delivery system. In the second part of the study, we examined the effects of randomization 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 that features a widely overlapping hearing range with humans (Syka et al. 2000) and the nonmovable pinna of which parallels the human ear to some extent (Hartung and Sterbing 1997). We have found that for auditory midbrain units VAS can mimic the acoustic FF quite realistically.

Parts of the data have been presented in abstract form (Behrend et al. 2003, 2004).

METHODOLOGY

Anesthesia 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 min (0.5% in Carbogen; Forthane, Abbott; inhalator: Advanced Anesthesia Specialists). Subsequently, an intramuscular injection of Ketamine and Xylazine ([2:1]; 0.1 ml/kg, Sigma-Aldrich) was applied. Anesthesia 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 per h; Janssen Animal Health, UK; A99 syringe pump, Razel). Levels of anesthesia 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 anesthetic (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-diam screws and dental cement (Palladur, Kulzer, Germany). A hole of ~0.5 cm in diameter was then drilled (Moto-flex driller, Dremel) 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; rostrocaudal 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 multiunit responses were recorded by tungsten electrodes (~5 MΩ impedance; A-M Systems), and signals were processed via a head stage (MM-333, Narishige, Japan), then amplified and band-pass filtered (0.7–3 kHz) by a microelectrode AC amplifier (model No. 1800, A-M Systems). The amplified signals were passed through a 50/60 Hz noise eliminator (Humbug, Quest Scientific) and then fed into a Tucker-Davis-Technology System II (component DD1, TDT) for digitization and further analysis. Only action potentials from units featuring a signal-to-noise ratio >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 off-line 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 h. 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 killed by an overdose of barbiturate (Nembutal 2 ml/kg, Sigma-Aldrich) and intracardially perfused with heparinized 0.9% saline for 5 min followed by 4% paraformaldehyde and 25% glutaraldehyde for 40 min (Masterflex pump, Cole Parmer Instrument; 5 ml/min flow rate; Optiva 14G needle, 2.2 mm OD, Terumo Medical). 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) and mounted on a (Leica 1320) 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 to 100 ms duration; 100 ms raised-cosine onset and offset ramp; 0–40 dB above neural threshold) were presented in two configurations, i.e., FF and VAS. All stimuli were generated on a workstation (Pentium III 450 MHz, Intel) using MatlabR12 DSP (The Math Works). 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).
covering the frontal hemisphere (above \( \pm 30^\circ \) elevation) or the full sphere (above \( -20^\circ \) elevation; Fig. 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) 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 to 30 kHz. VAS at corresponding virtual positions was presented via custom-made earphones, for which a detailed description is given by Chan and colleagues (1993). The earphone delivered VAS stimuli were matched spectrally and in overall level (\( \pm 1 \text{ dB} \)) 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 \( 90^\circ \) elevation) using a pair of complementary Golay sequences (Zhou et al. 1992) with radix 2 and length 1,024 (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) 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 FF 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 FF.

To match FF 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 (\( \sim 1-3 \text{ min} \)) to converge with the learning rate gradually reduced on-line. The adaptive filtering algorithm was implemented as a C program that was linked to the MatlabR12 programming environment using the MEX function interface. There is a separate inverse filter for the left and right ears, which together establish the flat frequency characteristic of the VAS sound-delivery system (Fig. 2A). Interaural time differences (ITDs) captured in the HRTFs were preserved correctly. See Fig. 2, B (in the frequency domain) and C (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 (\( \pm 1 \text{ cm} \)) to the aperture of the animal’s ear canal and calibrated as described in the preceding text. 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; \( 10-40 \text{ dB} \) 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 \( 0-20 \text{ dB} \) 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 pseudorandomized order (VASRAN; \( n = 26 \)) to assess spatial adaptation processes. For the latter, the original position list for sequential stimulus presentation was separated into eight sublists with evenly distributed position indices, which were then randomized and concatenated to create 10 new pseudorandom full position lists.

For a small number of units (\( n = 5 \)), VAS stimuli were also generated using HRTFs that 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 CJ286-60SR, SMC). Then FF stimuli were presented for corresponding positions. At least 10 repetitions per spatial position were presented in either configuration. The interstimulus interval was \( \geq 250 \text{ ms} \) to minimize neural adaptation by repeated stimulation.

Data analysis

Neural data and concurrent ear canal recordings were analyzed off-line using individual spike times and peristimulus 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 (PCA) of the spike waveform (\( n = 183 \)). For each unit, the spike timing of all discharges was calculated relative to the start of the recording period. Second, 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 fulfill this condition was used to approximate the size of a receptive field and expressed as a percentage of the tested spherical positions. For each unit, the optimal response area overlap between a pair of receptive fields obtained in FF and VAS was calculated by

$$\frac{N}{T} \times 100$$

$N =$ number of shared SRF positions exceeding 50% of the maximal response and $T =$ 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, 1,000 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 95th percentile of the overlaps generated across 1,000 random SRF combinations taken from the same sample. Second, 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 subsamples of five 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 1,000 times, and the distribution of spike count differences observed across all bootstrapped FF stimulus repetition subsamples 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 1,000 spike counts generated in the same manner. The two distributions of spike count differences being “expected” for identical stimulus repetitions and being “observed” across FF and VAS stimulus repetitions were tested for significant deviations by a standard \( \chi^2 \) 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 center of gravity (centroid) was calculated for each SRF by treating interpolated spatial responses as vectors, indicating the strength of the discharges elicited. Vectors across 2,601 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 toward 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 \( \rightarrow \)), 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 Fig. 3B. For all recordings, SRFs were computed based on a unit’s response rate or first-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 Fig. 3C. Of the 183 acoustically responsive cells isolated by PCA, 44%

---

**Figure 3.** A: frontal section showing penetration tracts left by the recording electrode in the IC (\( \rightarrow \)). B: the principal structures are indicated schematically. IC, inferior colliculus; LL, lateral lemniscus; PAG, periaqueductal gray. C: observed auditory response types (\( n = 183 \)).
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 first-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 with FF stimulation, the neural first-spike latency and spike jitter were analyzed in both regimes. At a stimulus location directly ahead of the animal (0° azimuth and elevation), the onset response latencies varied across all units from 3.3 to 38.9 ms in FF, and 2.1 to 40.8 ms in VAS (Fig. 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 (Fig. 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 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 <2 ms in FF and VAS was used as a criterion for temporal stability of a recording. This seemed appropriate because 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 to 30 ms (Fig. 4C). The observed individual latency deviations between FF and VAS increase with the individual neural jitter (Fig. 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 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 ms in FF; 4.50 ± 3.22 ms in VAS). On a spherical projection plot the first-spike latency pattern in FF was well preserved in VAS (Fig. 5, right). 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 (Fig. 5, left).

**FIG. 4.** A: mean 1st-spike latency of units stimulated at a frontal position (0° azimuth and elevation) in FF (■) and VAS (□; likewise in C). B: correlation plot of 1st-spike latencies in FF and VAS (\( r = 0.90; P < 0.001; n = 96 \)); ---, data trend (like wise in D). C: mean 1st-spike jitter of units across their shared receptive field in FF and VAS (\( r = 0.87; P < 0.001; n = 111 \)). All plots based on mean spike timing over 10 stimulus repetitions.
Spike-count-based SRFs 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 six units, the SRF was tested repeatedly in both FF and VAS. The average overlap of SRFs was calculated for both FF versus VAS stimulation (Fig. 6A; \(\triangle\)) and repeated FF stimulation (1st vs. 2nd recording; ■). 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 (Fig. 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; \(P = 0.67\)). Because 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 1,000 SRFs recorded in FF

---

**Figure 5**. Top: spatial receptive fields (SRFs) under FF stimulation computed from spike count (left) and 1st-spike latency (right). Bottom: 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. Color 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 (■) vs. the average overlap of the same units’ SRFs derived from FF and VAS stimulation (\(\triangle\)). B: correlation of the SRF overlap of the 2 combinations in A. The correlation coefficient is indicated above the graph.
and VAS yields an average SRF overlap of $38 \pm 18.6\%$ (mean ± 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, 1,000 random combinations of SRFs in FF and VAS yielded an average overlap of 47.1 ± 19.4%. Given that mostly unmatched, i.e., erratic and incomplete SRFs were excluded, the remaining SRFs recorded under stable conditions tend to be more homogeneous and centered contralaterally; this favors 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 obtained over two bootstrapped FF stimulus repetition samples ($n = 1,000$ 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 ($P \leq 0.05$; $\chi^2$ 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%) closely approximated the mean SRF overlap across subsequent FF-FF recordings conducted in unchanged acoustic environments ($70.2 \pm 14.2\%$).

**Effects of stimulus sound pressure levels on spatial response properties.** Figure 7 exemplifies a typical unit recorded in FF and VAS, first, close to response threshold levels and, second, 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 color 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 (Fig. 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 ($r = 0.82$). Six 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 $9.0 \pm 2.6\%$ 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 ($10.8 \pm 6.1\%; n = 5$). 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.

**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%; \( n = 87 \); left bar). These data demonstrated that at the level of the auditory midbrain neurons, VAS stimulation can mimic real FF stimulation to a high level of fidelity; \( n = 35 \) of 87 units showed a SRF overlap of ≥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. One thousand random pairs were computed from the sample of 87 units. The right bar in Fig. 9A shows the mean overlap of SRFs so randomly generated (47.1 ± 19.4%). Based on units that were stimulated over 80 positions, 1,000 random SRF pairs across FF and VAS generated a 95th percentile of 72% SRF overlap.

**FIG. 9.** A: the average SRF overlap in FF and VAS across 2 samples: 1st over all units recorded in FF and VAS (in stable recording conditions; \( n = 87 \)); 2nd across a hypothetical sample of units generated by random combination of the SRFs observed in the 1st sample (see text; \( n = 1,000 \)). B: correlation plot of the SRF size in FF and VAS observed over the 87 units in A (\( r = 0.83; P < 0.001 \)). – – – , \( r = 1.0 \); ---, data trend.
The 95th percentile represented the hypothetical SRF overlap shown in only the best 5% of 1,000 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 95th percentile of the respective random SRF overlap distribution). In statistical terms, 41% of units subsequently recorded in FF and VAS have produced a SRF overlap that 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 95th percentile of a random SRF overlap distribution.

**SRF size in FF and VAS.** 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 $44.5 \pm 18.0\%$ of the recorded spatial positions compared with $45.5 \pm 18.7\%$ 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 center of gravity in FF and VAS.** A measure of the center 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 interaural level differences (ILDs) calculated from the HRTF recordings (post surgery). The bulk of the centroids were located in the contralateral hemisphere around $-90^\circ$ azimuth, corresponding to the high ILDs. The deviation of the average centroid position between FF and VAS was small ($7.4^\circ$ azimuth; $3.3^\circ$ elevation; $n = 87$).

**APPLICATIONS.** Spatially randomized VASRAN. In a first step exploiting the flexibility of our validated VAS system, we pseudorandomized 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, Fig. 10B indicates the average centroid position for all tested stimulus configurations in blue (triangle; FF; \(-89.1^\circ\) azimuth; \(6.5^\circ\) elevation), purple (cross; VAS; \(-90.0^\circ; 6.7^\circ\)), and red (dot; VASRAN; \(-92.1^\circ; 10.2^\circ\)).

Figure 11 also indicates that in terms of the SRF shape, little adaptation was observed for sequential stimulation when compared with randomized presentation. Taken across this subset of neurons, the mean SRF overlap in FF and VAS was 75.6 ± 13.6%. The average SRF overlap obtained in FF and spatially randomized VASRAN was almost identical (76.1 ± 14.7%), 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%). Because 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 min. 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 and 40.7 ± 21.8%, 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 color code in Fig. 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 subsample 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). One 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 artifacts were minimized for a subset of unit recordings. Recall that for all neural recordings presented so far, the recording environment was acoustically distorted by the presence of electrophysiological recording equipment, and 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 five 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 in the preceding text, 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 that 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 (Fig. 13A; \(n = 5\)): The average centroid position of SRFs obtained in standard VAS was near the average centroid calculated from VASNat stimulation (mean shift: \(1^\circ\) azimuth; 3.8° elevation). The average size of SRFs recorded in VAS and in VASNat was also stable (28 ± 12.6 and 25.4 ± 11.8% of the sphere.

![FIG. 11](image-url) **Left:** a unit’s SRF under sequential FF stimulation. **Right:** the same unit’s SRFs stimulated in calibrated VAS under a spatially pseudorandomized 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 with 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. Color bars indicate the number of spikes elicited across 10 stimulus repetitions.
respectively). However, Fig. 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 VAS NAT. Apparently, receptive fields obtained in an acoustically “uncluttered” and surgically undistorted VAS NAT environment generally differed in shape from SRFs observed in standard conditions (paired t-test; \( P < 0.01; n = 5 \)).

**DISCUSSION**

Although much has been published concerning the nature of the neural representation of acoustic space, the limitation of the traditional FF 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 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 FF 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 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 with 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 (Hartung and Sterbing 1997; Poon and Brugge 1993b); 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 on-line calibration of the drivers prior to each unit recording. Most importantly, the use of the embedded
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 endeavored to control as carefully as possible physiological changes, e.g., by the use of a continuous anesthetic regime and of Hypnorm instead of barbiturates so as to avoid previously observed depressive affects (Brugge et al. 1994, 1996; Imig et al. 2000; Poon and Brugge 1993b).

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 FF 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 (Brugge et al. 1994, 1996; King and Palmer 1983; Knudsen et al. 1983; Middlebrooks and Knudsen 1984; Palmer and King 1985). 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 homologue to the mammalian IC, neural SRFs might be organized in a center-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 (1998) observed a remarkable stability of receptive fields recorded in FF and VAS that 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; Furukawa and Middlebrooks 2002; Furukawa et al. 2000; Middlebrooks et al. 1998, 2002). The performance of the VAS system tested yielded an average temporal match of first-spike timing in FF and VAS that 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 (Brand et al. 2002; McAlpine et al. 2001) remains to be investigated. The spatial map analysis of first-spike timing revealed a picture of decreasing latency values approaching the center 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 first-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 in 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 interstimulus 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 with sequential stimuli presentation. This result 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 are consistent with the idea that the decrease of the minimal response may result from an adaptation of ipsilateral inhibition because 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 suprathreshold stimulation (Palmer and King 1985; Wise and Irvine 1983, 1985).

We also recorded the VASRAN sample with a new set of tungsten electrodes, which yielded more stable recording conditions. As a byproduct 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 artifacts 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 nonindividual 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 nonindividual 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 with 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 FF equipment be eliminated.

Conclusion

Our study supports the assumption that VAS largely resembles a FF 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 FF, 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 help to match FF and VAS responses at certain test positions.

Acknowledgments

We are indebted to J. Leung for continuous support, J. Schnupp for fruitful discussions, and A. Anandan for technical help.
REFERENCES


