Echo-acoustic evaluation of real and phantom objects in phyllostomid bats

Jan-Eric Grunwald

SIGILL: UNIVERSITAT: LUDOVIC: MAXIM:

München 2004

Echo-acoustic evaluation of real and phantom objects in phyllostomid bats

Jan-Eric Grunwald

Dissertation
an der Fakultät für Biologie
der Ludwig–Maximilians–Universität
München

vorgelegt von
Jan-Eric Grunwald
aus München

München, den 18. Februar 2004

Erstgutachter: Prof. Dr. Dr. Gerhard Neuweiler

Zweitgutachter: Prof. Dr. Gerd Schuller

Tag der mündlichen Prüfung: 18. Mai 2004

Contents

Zusammenfassung

Mithilfe ihrer Echoabbildung vollbringen Fledermäuse erstaunliche Leistungen bei der Erkennung von Objekten in völliger Dunkelheit. Eine wachsende Anzahl von Studien belegt, dass für frugivore und nectarivore Fledermäuse der Familie der Phyllostomidae die Echoabbildung eine zentrale Rolle bei der Erkennung von unbewegten Objekten spielt. Die Grundvorraussetzung für die Erkennung von Objekten ist die sensorische Analyse von physikalischen Objekteigenschaften. Die akustischen Eigenschaften eines Objekts werden durch dessen Impulsantwort beschrieben. Die Impulsantwort ist die Summe aller Reflektionen, die ein Objekt zurückwirft, wenn es mit einem akustischen Impuls beschallt wird. Das spezifische Echo, das eine Fledermaus von einem Objekt empfängt, entsteht, indem die Impulsantwort des Objekts auf den ausgesandten Ortungslaut „aufgeprägt“ wird.

Bis heute ist es nicht klar, ob und wie eine akustische Objektepräsentation aus den Impulsantworten im Gehirn der Fledermaus gebildet wird. Die vorliegende Arbeit beschäftigt sich mit drei verschiedenen Aspekten der echoakustischen Objektverarbeitung bei Phyllostomiden: Der Detektion von gefilterten Lauten in einem Rückspielexperiment, der Objekterkennung über verschiedene Modalitäten hinweg und der Klassifikation komplexer Phantomziele.

Im ersten Teil dieser Arbeit sollte die Empfindlichkeit des Gehörs einer blütenbesuchenden Fledermaus, *Glossophaga soricina* bestimmt werden. Im ersten Teilversuch wurden in einem Rückspielexperiment die Detektionsschwellen für Echos des eigenen Echoortungslautes als Funktion des Frequenzgehalts und der Echoverzögerung gemessen. Bis zu vier Fledermäuse wurden dafür gleichzeitig in einem voll automatisierten Versuchsaufbau dressiert. Während des Versuchs wurden Ortungslaute von *G. soricina* zur Quantifizierung verschiedener Lautparameter aufgenommen und analysiert. Die Empfindlichkeit für spektrale Lautkomponenten war eng an die spektrale Lautzusammensetzung gekoppelt. Außerdem wird gezeigt, dass Nachverdeckung bei der Echoabbildung von Objekten in mehr als 41 cm Entfernung eine untergeordnete Rolle spielt. In einem zweiten Teilversuch wurden die relative Empfindlichkeit und die Abstimmschärfe der auditorischen Peripherie anhand von Verzerrungsprodukt otoakustischen Emissionen (distortion-product otoacoustic emissions, DPOAEs) und Suppressionstuningkurven (STCs) gemessen. Die DPOAE Schwellenkurven zeigten eine relativ gleichbleibend hohe Empfindlichkeit für Frequenzen bis zu 135 kHz. Die Abstimmschärfe der cochleären Filter entspricht weitestgehend der anderer unspezialisierter Säuger.

Neben der Echoabbildung kann bei Fledermäusen natürlich auch der Gesichtssinn zur Wahrnehmung von Objekten beitragen. Für eine zuverlässige Objekterkennung wäre es von Vorteil, wenn die Repräsentation eines Objektes einer Modalität auch der anderen zur Verfügung stünde. Mit diesem Thema befasst sich der zweite Teil dieser Arbeit. Mehrere *G. soricina* wurden darauf andressiert, zwei verschiedene reale Objekte nur mit Hilfe der Echoabbildung zu unterscheiden. Hatten sie diese Aufgabe erlernt, wurde getestet, ob sie dieselben Objekte auch nur mithilfe visueller Wahrnehmung unterscheiden konnten. Keines der Tiere zeigte sich in der Lage, die Testaufgabe zu lösen. Dies zeigt, dass Flermäuse nicht ad-hoc eine akustische Objektrepräsentation bilden, die für andere Modalitäten zugänglich ist.

Große, komplexe Objekte wie z.B. Bäume besitzen sehr viele Reflektoren, wodurch chaotische Impulsantworten zustandekommen. Daher können solche Objekte nur anhand statistischer Eigenschaften der Impulsantworten, wie z.B. der Rauhigkeit, unterschieden und klassifiziert werden. Im dritten Teil der vorliegenden Arbeit sollte die Fähigkeit der frugivoren Art *Phyllostomus discolor* untersucht werden, komplexe, chaotische Echos, wie sie z.B. von verschiedenen Vegetationstypen zurückgeworfen würden, zu unterscheiden und zu klassifizieren. In einem Dressurexperiment lernten die Tiere, eine bestimmte rauhe Impulsantwort von einer bestimmten weniger rauhen Impulsantwort zu unterscheiden. Darüber hinaus waren die Tiere in der Lage, unbekannte Impulsantworten anhand ihrer Rauhigkeit zu klassifizieren. Diese Fähigkeit gibt Fledermäusen die Möglichkeit, große, komplexe Objekte trotz ihrer chaotischen Impulsantworten in Klassen einzuordnen. Basierend auf einem Modell der auditorischen Peripherie von *P. discolor* konnte in einer Computersimulation der Klassifikationsaufgabe anhand von auditorischen Repräsentationen der Echos gezeigt werden, dass modulationssensitive Neurone für die Kodierung der Rauhigkeit chaotischer Echos eine zentrale Rolle spielen könnten.

Summary

Their echo-imaging system has enabled microchiropteran bats to perform extraordinary object recognition tasks in complete darkness. A growing number of studies provide evidence that echo-imaging plays an important role for the recognition of silent and motionless objects in nectivorous and frugivorous species of the phyllostomid family. The basic prerequisite for object recognition is the perceptual analysis of an object's physical properties. The acoustic properties of an object are encoded in its impulse response (IR). The IR is defined as the sum of reflections recorded when an object is ensonified with an acoustic impulse. Whenever a bat ensonifies an object it receives a specific echo, which is the object's IR imprinted on the bat's sonar emission. Up to now, it is not clear if and how an auditory representation of an object derives from the analysed IRs in the bat's brain.

The current thesis addresses three different aspects of echo-acoustic object evaluation in two species of phyllostomid bats: detection of filtered calls in a playback experiment, cross-modal object recognition and classification of complex phantom objects.

The aim of the first part of the present thesis was to characterise the auditory sensitivity of the nectar feeding bat, *Glossophaga soricina*. In a first experiment, the detection thresholds for echoes of the own emitted calls were determined as a function of echo spectral content and echo delay. To this end, up to four bats were trained simultaneously in a fully automatic two-alternative, forced choice (2-AFC) playback paradigm. For a quantification of call parameters, sonar emissions of the bats were recorded and analysed during this experiment. The sensitivity for spectral components of the echo was closely related to the distribution of energy in the echolocation call. The playback of unfiltered calls revealed that forward masking plays a minor role for echo imaging in distances to the target of more than 41 cm. In a second experiment, distortion-product otoacoustic emission (DPOAE) thresholds and suppression tuning curves of *G. soricina* were measured to investigate the sensitivity and tuning of the auditory periphery in this species. The obtained DPOAE thresholds showed no significant deterioration of sensitivity for frequencies up to 135 kHz. The tuning sharpness of the cochlear filters is in the range of other unspecialised mammals.

Of course, bats can also perceive objects visually. In terms of object recognition, it would be advantageous if a representation of an object in one modality would also be available to the other. This topic was addressed in the second part of the thesis in *G. soricina*. The bats were trained to discriminate two different objects only with echolocation cues available. Subsequently, it was tested whether the bats were still able to accomplish this discrimination when the objects were presented to vision alone. The fact that none of the bats was able to solve this task in the test situation indicates that echolocating bats do not ad-hoc form an acoustic object representation that is transferrable across modalities.

Large, complex objects such as trees have very many reflective surfaces which result in a chaotic impulse response. Thus, such objects can only be discriminated and classified by statistical properties of their impulse responses, e.g. their degree of roughness. The aim of the third part was to investigate the ability of the frugivorous species *Phyllostomus discolor* to discriminate and classify complex chaotic echoes as they would be reflected from different foliage types, for example. In a 2-AFC playback paradigm, the bats learned to discriminate a specific rough IR from a specific smooth IR. Further, it was shown that the bats were able to classify unknown IRs according to their roughness. This capability allows bats to classify large, complex objects in spite of their chaotic echoes. In a computer simulation based on a detailed model of the auditory periphery of *P. discolor*, the behavioural performance of the bats was related to possible physiological representations of the perceived echoes. This simulation revealed that modulation sensitive neurons may play a role in encoding signal roughness.

1 Echo-detection thresholds and their relationship to the hearing threshold as determined by otoacoustic emissions in *Glossophaga soricina*

1.1 Introduction

Echolocating bats ensonify their surroundings with ultrasonic emissions and analyse the reflections from objects. In the course of evolution, their sonar system has become the bats' primary perceptual window to their outside world, and a large variety of echolocation strategies have developed in association with newly exploited niches (reviews: Kalko and Schnitzler, 1998; Neuweiler, 1989; Schnitzler and Kalko, 1998).
But the capabilities of bat sonar go far beyond a precise location of an ensonified object in space: bats can also extract information about the size and three-dimensional shape of an object from the reflected echo (e.g. Habersetzer and Vogler, 1983; Neuweiler, 1989; Schmidt, 1988, 1992; Schmidt et al., 2000; Simmons et al., 1974). This capability constitutes the basis for the discrimination and recognition of specific objects or object classes by echolocation. Rhinolophid bats, for example, can discriminate and even classify insects by their wingbeat rate: the fluttering of the insect-wings imprints different amplitude and frequency modulations on the long constant-frequency portion of the echolocation call (v. d. Emde and Menne, 1989; v. d. Emde and Schnitzler, 1990).
A growing number of studies provide evidence for the significance of echolocation for the detection and identification of silent, motionless prey by bats (Kalko and Condon, 1998; Schmidt et al., 2000; Thies et al., 1998; v. Helversen et al., 2003; v. Helversen and v. Helversen, 1999, 2003). These findings show that bats, especially nectivorous and frugivorous species, can detect, localise, and identify objects like flowers or fruit which are often buried in clutter. The major drawback of echolocation is that it provides a good resolution only along one dimension, the depth, of an object. Information about the object's height and width has to be extracted from successive echoes from different "viewing"-angles or positions. Thus, for a primarily one-dimensional sense it is a highly challenging task to detect and identify a target echo among a

multitude of echoes or a composite echo reflected from the target and background. Moreover, the recognition of objects independent of conditions like "viewing"-angle or distance to and quality of the background would require an image-like representation of acoustic object properties extracted from the echoes.

Up to now, the basic psychophysical principles underlying echo-imaging are poorly understood. For example, it is still in dispute whether information about object structure is represented in the time- or in the frequency domain (Schmidt, 1988; Simmons et al., 1990; Weissenbacher and Wiegrebe, 2003). However, these questions can hardly be answered in psychophysical experiments employing real objects as targets, because these do not allow the selective manipulation of echo parameters.

Recent advances in digital-signal processing (DSP) technology now provide new opportunities to address these questions. Real time FIR-filtering (Finite-Impulse Response-filtering; see below) at high sampling rates enables us to create acoustic images of objects in the ultrasonic range, and to present them in playback experiments, similar to the presentation of visual images on a screen (cf. Aubauer and Au, 1998; Aubauer et al., 2000). As in visual imaging, these acoustic images can be manipulated specifically to assess information about the parameters important for object recognition.

The species studied here, the common long-tongued bat, *Glossophaga soricina* (Pallas 1766, Phyllostomidae, Glossophaginae), is an ideal model animal to approach this topic: it forages for nectar, pollen, fruit, and also insects in subtropical and tropical forests and savannas in Central and South America (Alvarez et al., 1991). Glossophagine bats serve as important pollinators in the neotropics (v. Helversen, 1993; v. Helversen and Winter, 2003), and the results of recent studies indicate that echolocation plays an important role for the detection and recognition of their target flowers (v. Helversen et al., 2003; v. Helversen and v. Helversen, 1999, 2003). Furthermore, this species has been successfully trained in psychoacoustic experiments (López, 2002; v. Helversen and v. Helversen, 2003).

The aim of the present study was to lay the foundations for subsequent playback experiments focusing on object recognition in this nectar-feeding bat in two respects: first, to characterize the auditory sensitivity of the model animal, especially for stimulation in a playback-paradigm. Second, to establish a fully computer-controlled psychoacoustic training set-up for playback experiments in *G. soricina* which allows simultaneous training of several individuals.

According to López (2002), the sonar emissions of *G. soricina* are short (< 2 ms), multiharmonic frequency sweeps covering a frequency range from 140 to 60 kHz, with their intensity peak near 105 kHz in the second harmonic. These short, broadband calls with high upper frequency limits yield high accuracy in localization and characterization of objects (Schnitzler and Henson, 1980; Simmons et al., 1995a; Simmons and Stein, 1980).

Interestingly, the predominance of the second harmonic of the call is not reflected in auditory threshold curves measured so far: recordings of cochlear potentials indicated a sensitivity maximum at 60 kHz in the auditory system of *G. soricina* (Howell, 1974). Unfortunately, these potentials were only recorded up to 100 kHz. A more recent behavioural audiogram showed lowest hearing thresholds at 80 kHz (López, 2002). Hence, the auditory sensitivity maxima are far below the intensity peak of the echolocation calls of this species. Most bats studied show relatively high sensitivity in the area of their most prominent echolocation call component (e.g. Heffner et al., 2003; Koay et al., 2002; Neuweiler et al., 1984).

In which frequency range is *G. soricina* most sensitive for echoes of the own emitted call? This question is of special interest for the generation and presentation of stimuli in playback experiments: key parameters such as sampling rate and frequency response of the system have to be adjusted to the effective frequency range used by the experimental bat. These parameters, of course, determine the choice of equipment such as microphone, speakers, and DSP-unit.

A further critical aspect in acoustic playback experiments is forward masking. Forward masking occurs when two acoustic signals are separated by a short time interval: the preceding signal affects the threshold of the following one (Moore, 1997). In echolocation, either the emitted call or echoes arriving earlier than the target echo might provide forward masking, e.g., echoes from objects in the set-up like the speakers or the microphone. Hence, in playback experiments echo stimulation should be presented in a time window free of masking effects.

These topics were addressed in two different experiments.

In the first experiment of this study, the role of echo spectral content and of forward masking for the detection of echoes from emitted sonar calls was characterized.

Up to four individual *G. soricina* were trained simultaneously in a fully automatized,

two-alternative forced-choice (2-AFC) playback paradigm. The sonar emissions of the bats were band-pass filtered on-line with different centre frequencies and played back with different delay times. In addition, unfiltered calls were presented at different delays to measure the influence of forward masking on detection of echoes of the entire call.
To quantify parameters of calls relevant for the bats' responses, sonar emissions were recorded in the course of the experiment.

In a second experiment, distortion-product otoacoustic emissions (DPOAEs) were measured to investigate the sensitivity of the auditory periphery of *G. soricina*. This non-invasive method is based on the fact that the mammalian cochlea is a mechano-electrical transducer with nonlinear mechanical properties. As a consequence of this nonlinearity, mechanical distortion can be evoked by stimulation with two tones of different frequency ($f_1 < f_2$). Among other cubic distortion products at frequencies $nf_1 - (n-1)f_2$ and $nf_2 - (n-1)f_1$, the most prominent distortion is produced at the frequency $2f_1$-f_2. These distortion products can be recorded from the outer ear canal (reviews: Kössl, 1997; Probst et al., 1991). As the $2f_1$-f_2 distortion product is produced near the f_2-place in the cochlea, threshold curves can be obtained by plotting the f_2-level sufficient to elicit a distortion of a defined level as a function of f_2-frequency (Brown and Kemp, 1984; Foeller and Kössl, 2000; Frank and Kössl, 1995). These threshold curves describe the mechanical sensitivity of the cochlea and often run parallel to neuronal and behavioural threshold curves (Gaskill and Brown, 1990; Kössl, 1997).
Furthermore, cochlear tuning properties, i.e., the frequency selectivity of cochlear filters, were investigated by means of suppression-tuning curves (STCs): the DPOAE is suppressed by a certain level criterion with a third tone, f_3, of variable frequency. Plotting the f_3-level sufficient to suppress the emission as a function of f_3-frequency results in STCs which describe the tuning of cochlear filters (Brown and Kemp, 1984). Since the cochlea is the gateway between the physical stimulus and its representation along the auditory pathway, these measurements of cochlear sensitivity can provide information about whether high frequency hearing is limited already by the characteristics of the auditory periphery.

The findings of this study constitute the basis for the design of further playback experiments with their focus on object recognition in this species.

Impulse responses and FIR-filtering

FIR-filtering of echolocation calls is a powerful tool to create and manipulate phantom targets. The next paragraph is intended to give some information required for the understanding of the basic principle underlying this method.

An acoustic image of an object is defined by the object's impulse response. This is the echo an object produces when ensonified with an acoustic impulse. An impulse, or a click, is an indefinitely short impulse with an indefinitely high amplitude. In such an impulse, all frequencies are represented equally (cf. Fig. 1.1 a). The impulse response of an object is defined by its reflective surfaces. For example, a flat, totally smooth surface will reflect an exact copy of the impinging sound, thus, its impulse response is again an impulse. When we imagine ensonifying a two-front target with a raised and a recessed surface, its impulse response will consist of two clicks, one reflection from each surface. The time interval between these clicks is defined as

$$\frac{sound\ velocity}{2 \times distance\ between\ surfaces}$$

because the sound has to travel back and forth between these surfaces. The magnitude spectrum of such an impulse response shows a ripple. The distance between two peaks of this ripple represents the inverse of the time interval between the two clicks. Thus, this two-front target acts as a comb filter in the frequency domain. It is easily conceivable that more complex objects with more surfaces will have more complex impulse responses with more complex frequency spectra. Consequently, every object acts as a filter, i.e. it imprints its impulse response on the impinging sound and thus produces a characteristic echo. Mathematically, this physical composition of an echo corresponds to either a convolution of the impulse response with the sound in the time domain, or to a multiplication of the complex spectra of both. This convolution is the basic principle of FIR-filtering.

1.2 Materials and methods

1.2.1 Animals

For the present study, four male individuals of *G. soricina* were kept and trained in an anechoic chamber ($w \times d \times h$: 3 x 3 x 2.2 m). Inside the chamber, the temperature and humidity were kept constant at approx. 25°C and 65%, respectively. The animals were kindly provided by Prof. O. von Helversen from his breeding colony kept at the Institut für Zoologie der Universität Erlangen.

The bats were subject to an artificial diurnal rhythm of 12 hours light (1300 to 0100) and 12 hours darkness (0100 to 1300). All animals were trained simultaneously, usually 24 hours per day from Monday to Saturday. During training, the bats were rewarded with an 18% sugar-water solution (1/3 glucose, 1/3 fructose, and 1/3 saccharose). The amount of food intake per animal was calculated after each training session. If one of the animals did not receive enough reward (less than 8 ml/day on three consecutive days), the bats were given additional food at free disposal for approx. 2 h per day.

In order to maintain an adequate supply with nutrients, the bats were regularly provided with instant nectar (Nektar Plus, *Nekton*) or human medical diet for tube feeding (Nutricomp, *B.Braun Pharma*) diluted in a 20 % honey-water solution, and pollen in the non-training periods. Flies (*Calliphora sp.*) and Banana were given less regularly.

For individual identification, each bat carried a necklace made of silicone tubing with a transponder (GlasTag micro, *IQ Automation*, size 12 x 2.12 mm) fixed to the necklace with heat-shrink tubing.

Because of interference between individuals, one bat had to be removed from training. Only two animals learned the task and were further trained for data acquisition. Two other male *G. soricina* were used for the DPOAE measurements. These animals were kept with conspecifics in a basement room under identical climatic conditions.

1.2.2 Echo-detection thresholds

Stimulus (echo) generation and presentation

To manipulate the calls for playback, FIR-filtering was used: the impulse response of a filter was convolved with the bat call in the time domain. To quantify the effect of forward masking on the echo detection, the echoes were played back with different digital delays: 0, 0.5, 1, 2, and 4 ms. As the delay caused by signal manipulation (0.2 ms) and the sound travel time from the speakers back to the presentation feeder (2 ms) have to be taken into account, the effective delays amounted up to 2.4, 2.9, 3.4, 4.4, and 6.4 ms. These delays correspond to target distances between 41 and 109 cm.

Impulse responses For the playback of the unfiltered call, the echolocation call was convolved with a click-like IR (Fig. 1.1 a). This IR has a flat frequency response and therefore the echo is an exact copy of the call.

To vary the spectral content of the echoes, four different IRs with band-pass characteristics were used (Fig. 1.1 b – e). All IRs were symmetrical and 111 samples long (corresponding to a duration of 278 μs at a sampling rate of 400 kHz). At the given sampling rate, the IRs had centre frequencies of 63, 80, 100, and 126 kHz and bandwidths of ± 10% of the centre frequency. All IRs of the bandpass filters had the same root-mean squared amplitude (RMS). Thus, all FIR-filters had the same amplification factor (5 dB). Figure 1.1 shows the impulse responses of the FIR-filters used.

Echo generation In each trial of the experiment, every echolocation call that was picked up by the microphone was manipulated online, and the resulting echo was played back to the bat. The bats' sonar emissions were picked up by a 1/4 inch condenser microphone (*Brüel & Kjær* 4135) connected to a measuring amplifier (*Brüel & Kjær* 2610). A high-pass filter (30 kHz, *Krohn Hite* 3550) was connected to the filter loop of the measuring amplifier to remove low-frequency noise. The amplified signals were again filtered for anti-aliasing (200 kHz low-pass, *Wavetek* model 442), and then digitized by a DSP-board (*Microstar* DAP 5200a) with a sampling rate of 400 kHz. On the DSP-board, each call was convolved online with a specific IR in the time domain. After a specific delay of 0 to 4 ms, the resulting echo was converted from digital to analogue and again low-pass filtered (200 kHz, *Wavetek* model 442) for anti-aliasing. Then the signal was split into two channels, left and right. Each channel was fed into a computer-controlled attenuator (custom built), amplified (*Harman/Kardon* 6150) and sent to a speaker (*Matsushita* EAS10 TH1000D). Fig. 1.2 shows a block diagram of the playback set-up.

In each trial, only one channel was used for playback. Therefore, one channel, either left or right, was set silent by setting its attenuator to full attenuation (128 dB). The attenuation of the active channel was varied depending on the individual bat's training status.

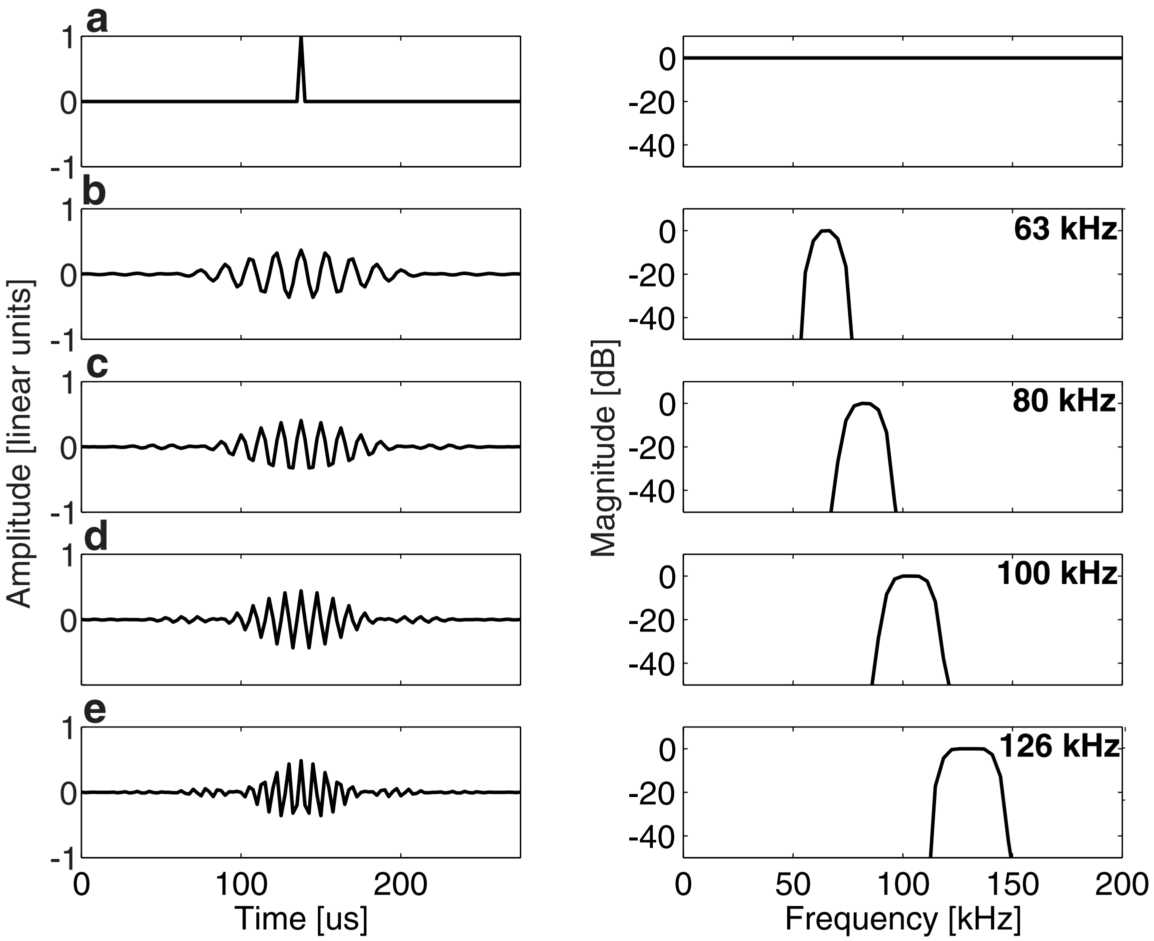

Figure 1.1: Impulse responses (left) and frequency responses (right) of the FIR-filters used for measuring echo-detection thresholds. For unfiltered playback, the filter with the click-like impulse response (**a**) was used. Note the flat frequency response of this filter. The bandpass filters **b** – **e** were used to vary the echo spectral content. The filter centre frequencies are given in the right panel column. Note that filter bandwidth changes proportional to filter centre frequency.

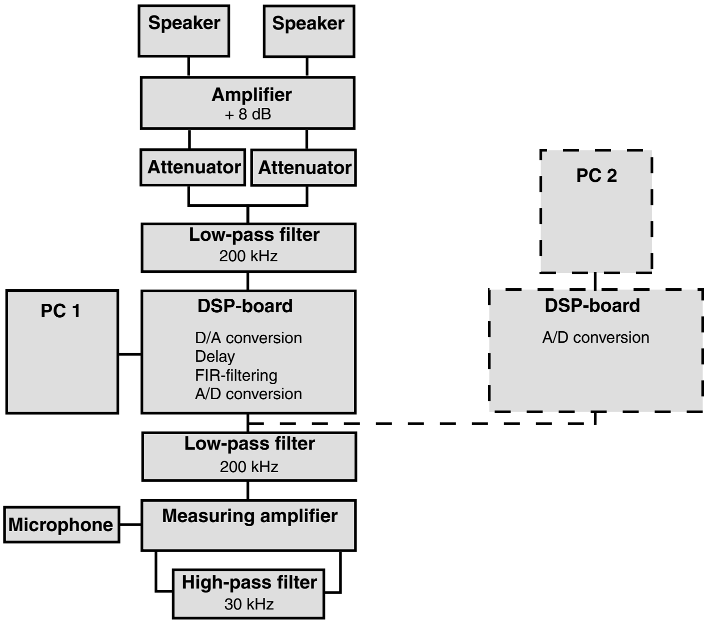

Figure 1.2: Block diagram of the playback and sound-recording set-up. The solid boxes show the playback section. The dashed lines and boxes illustrate the sound-recording section.

Experimental set-up

The training and playback set-up was mounted in a box of the dimensions : $1.4 \times 1.1 \times 0.5\,\mathrm{m}$ $(w \times d \times h)$, outlined as shown in Fig. 1.3, located in the chamber. The inside of the box was lined with sound absorbing foam.

Inside the box, three computer-controlled feeders were suspended from the ceiling to about mid-height of the box: one "presentation feeder" in the centre, and two "reward feeders" left and right of the presentation feeder. The feeders consisted of a PVC ring with a shortened syringe as spout. The feeders were custom made by a design communicated by Dr. Y. Winter. In addition, each feeder was equipped with an infrared light barrier. This light barrier was interrupted by the snout of a bat hovering in front of the feeder and indicated the presence of a bat to the computer. For the identification of the visiting bat, a custom built circular antenna connected to a transponder reader (easykey /R, *IQ Automation*) was mounted in front of each feeder. The access to the reward feeders could be denied by computer controlled flaps in front of the feeders.

The sugar solution was conducted in silicone tubing from the reservoirs (50 ml syringes) to the feeders. The reward was dosed by computer-controlled squeeze-

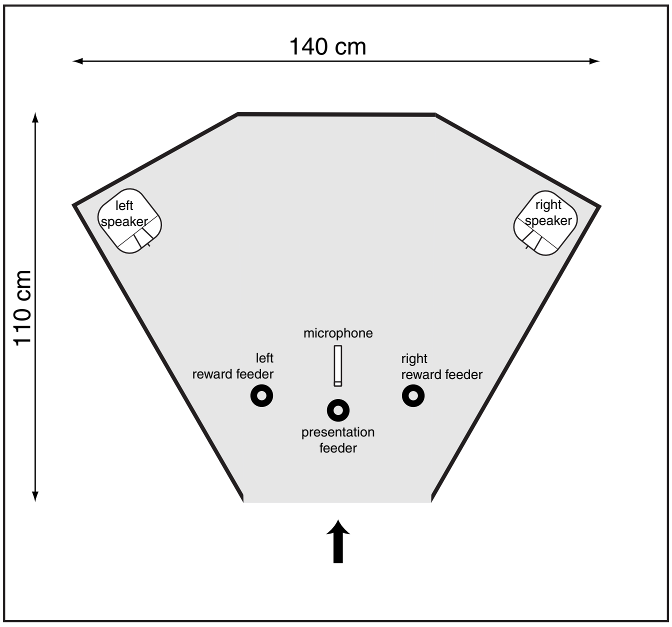

Figure 1.3: Experimental set-up used for the measurement of echo-detection thresholds. The filled box illustrates the outline of the set-up. The arrow indicates the entrance to the box.

valves (*Asco Joucomatic* W295A112-12VDC).

The playback section of the set-up consisted of a microphone and two speakers. The microphone was mounted centrally in front of the presentation feeder in 12 cm distance from the feeder. The speakers were positioned in the far left and far right corner of the box, facing towards the presentation feeder. The distance from each speaker's membrane to the presentation feeder was 68 cm.

The bats' behaviour could be observed from the outside of the chamber on a Monitor connected to an infrared sensitive camera mounted on the set-up box. Further, the bats' vocalisations and the echoes for playback were monitored with a storage oszilloscope (*Hameg* HM 407). The training routines were programmed in *MATLAB* v5.3 (*The MathWorks*).

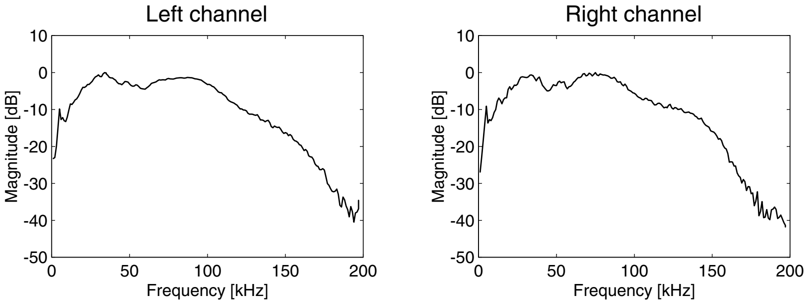

Figure 1.4: Frequency response of the left and right playback channel.

Playback system calibration

To determine the frequency response and the overall amplification of the playback system, the set-up was calibrated before starting the data acquisition. The frequency response of the output section, i.e. the amplifier and speakers, was determined by recording its impulse response. Therefore, a 2.5 μs impulse was emitted from the DSP-board at a sampling rate of 400 kHz and sent through the output section as it was used in the experiment. The response was recorded with a 1/8 inch microphone (*Brüel & Kjær* 4138) mounted in front of the speaker's membrane at a distance of 10 cm. This measurement was done separately for each channel. The frequency responses for the two channels are shown in Fig. 1.4.

The sensitivity of the speakers was measured with a 63 kHz sinus with 10 ms duration and an amplitude of ±2 V (4 V peak to peak) at the output of the amplifier. The speaker output was recorded with a 1/8 inch microphone (*Brüel & Kjær* 4138) mounted in front of the speaker's membrane at a distance of 70 cm, corresponding to the distance of the speakers to the presentation feeder. The intensity of the speakers' output was read from the measuring amplifier (*Brüel & Kjær* 2610). The speakers produced an output of 75 dB SPL at 63 kHz.

The effective gain of the FIR-filtering was measured by FIR-filtering broad band noise created by a frequency analyser on the DAP-board. The output of the board was fed into the direct input of the measuring amplifier and the measured level was compared to the level of unfiltered noise. The resulting gain factors were 0 dB for the "click"-filter and 5 dB for the bandpass filters.

The maximal total gain of the playback system amounted to 1.7 dB for unfiltered playback calculated for a distance of approx. 70 cm to the speaker membrane. For bandpass filtered playback, the maximal total gain amounted to 4.9, 6.9, 6.9, and 3.4 dB for bandpass filtering at 63, 80, 100, and 126 kHz, respectively, due to the frequency response of the system.

Procedure

The bats were trained in a two-alternative, forced-choice (2-AFC) paradigm: a food reward at the left or right feeder had to be associated with an echo from the left or right speaker, respectively.

To start a trial, a bat had to visit the presentation feeder. Here, it was identified by the transponder reader and received a small reward (5 μl). As soon as the bat was identified, a specific IR for FIR-filtering, and a specific playback delay were chosen by the training routine, dependent on the individual bat's training status. The playback was active for five seconds. The flaps in front of the reward feeders were opened for six seconds from the start of the playback, and the bat had to indicate its decision by visiting either reward feeder within this time span. Correct choices were rewarded with sugar solution (between 15 and 25 μl). A reward was only given when the bat at the reward feeder was identified as the one that started the trial. Otherwise the trial was aborted.

The choice of the playback channel was randomized by the training routine. However, the randomized choice was limited by setting a maximum of three successive echo presentations from the same side.

For the initial training of the bats, unfiltered calls were played back with a fixed delay of 3.4 ms. Data acquisition was started, as soon as the bats had reached a stable performance above 70 % correct choices in this task. To obtain psychometric detection functions (Fig. 1.5) for each filter-delay combination, the playback attenuation was increased successively in 5 dB steps from 0 to a maximum of 40 dB. For each attenuation level and filter-delay combination 64 data points were recorded: first, 32 data points were recorded for each attenuation level successively until the bats reached a performance of 50 % correct choices. Then, the next 32 data points were recorded for the same filter-delay combination to the determined maximum attenuation level.

For subsequent analysis, all the bats' decisions were time stamped and stored in a log file automatically after every trial.

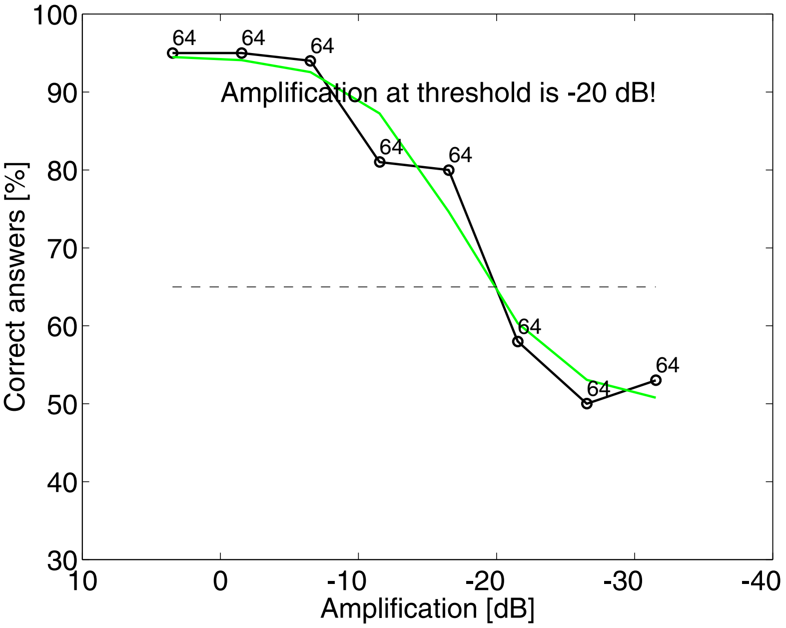

Figure 1.5: Example of a psychometric function (green) fitted to experimental data (black) of bat2 when presented with unfiltered calls with a delay of 6.4 ms. The data are plotted as percent correct choices of 64 trials as function of the system amplification. The dashed line represents the threshold criterion of 65 % correct choices.

Data analysis

To determine the echo-detection thresholds, a sigmoidal function was fitted to the data recorded for each filter-delay combination. The equation for the function F dependent on the system amplification is as follows:

$$F(p) = b/(exp^{-\lambda \times (p-a)} + 1) + c$$

where a is the expected value for $F = 65$; b is the bandwidth, c is the starting point, and λ is the slope of the function F. An example of the function fitted to data obtained in the experiment is shown in Fig. 1.5.

In a 2-AFC paradigm, a performance of 50 % correct choices corresponds to random choice, i.e., the echo intensity was too low to be detected. For 64 recorded data points per condition, a performance of 65 % correct choices is significantly different from a 50 %-performance (Miller, 1996). This 65 % criterion was used to extract the system amplification at the echo-detection thresholds from the psychometric functions. All resulting thresholds were corrected for the frequency response of the playback system.

1.2.3 Sound recording and analysis

To assess information about the characteristics of the echolocation calls of *G. soricina*, the bats' sonar emissions during approx. 9,000 trials in the experiment were recorded. Therefore, the output of the measuring amplifier was split into two channels after anti-aliasing filtering. One channel was used for playback, the other was fed to a second DAP-board in a second PC. Here, the signals were digitized with a sampling rate of 380 kHz and stored in the DAP-board's memory buffer.
To trigger onset and offset of the sound recording, the recording set-up was also connected to the light barriers of presentation and reward feeders. Whenever the presentation feeder was visited, the recording started automatically. The recording was stopped by a bat's visit to a reward feeder or after five seconds, in case that no decision was made. Each recorded call sequence was retrieved from the DAP-board and then saved to disk as a .wav file by a *MATLAB* routine. The time of recording was logged for each sequence. In addition, the system clocks of the recording and the playback set-up were synchronized with an atomic-clock server by a network time client (*NTPC*). Thus, the recorded sequences could be assigned to a specific trial before analysis.

The analysis of the recorded sequences was programmed in *MATLAB*. To quantify the background noise level of the recordings, the RMS of a 50 ms interval without bat calls was determined in each sequence. To measure the onset and offset of the echolocation calls, the hilbert envelope of the waveform of each sequence was calculated. The hilbert envelope was lowpass-filtered at a cut-off frequency of 3,500 Hz to prevent short-term amplitude modulations from being interpreted as on- or offset of a call. The on- and offset of the signals were measured at those points, where the amplitude of the hilbert envelope exceeded the background RMS by a factor of 1.5. Afterwards, signals of less than 0.1 ms duration were removed, to exclude transient noise pulses from analysis. This duration, as well as the filter cut-off frequency and the amplitude threshold for on- and offset measurement were chosen, because these parameter settings yielded the most accurate results in test analyses.
From the remaining on- and offsets, the call durations and the inter-pulse intervals

(from one onset of a call to the onset of the next) were calculated.

For the spectral analysis of the calls, the magnitude spectrum of each call was calculated using a 512-point rectangular FFT-window centred around each bat call. At the given sampling rate, the FFT-window had a duration of 0.67 ms and a frequency resolution of 750 Hz. The duration of this FFT-window was longer than that of the longest call recorded (0.46 ms). From this magnitude spectrum, the peak frequency of each signal was measured as the most intense frequency. The bandwidth of the call was determined as the frequency band covered by the most prominent harmonic 20 dB below the peak frequency. Echolocation calls with a peak level that was less than 22 dB above the noise level of the magnitude spectrum (mean intensity between 170 and 190 kHz) were excluded from the bandwidth analysis.

The intensity of each call was calculated as its peak-equivalent sound-pressure level (peSPL):

$$peSPL = 20 \times \log \frac{\frac{x_{max}}{\sqrt{2}} / Sens_{mic}}{20 \mu Pa}$$

where x_{max} is the maximum amplitude of the call in Volts and $\frac{x_{max}}{\sqrt{2}}$ represents the RMS of a sinus with the amplitude x_{max}, and $Sens_{mic}$ is the sensitivity of the microphone multiplied with the gain factor of the measuring amplifier (here: 9.9 V/Pa). Additionally, the energy content of the calls in the frequency bands of the filters used for playback were determined. Therefore, each recorded call was convolved with each of the filter-IRs. After compensating for the frequency response of the microphone, the intensity of each filtered portion was calculated as above.

After all parameters were calculated by the *MATLAB* program, each analysed signal was displayed on the screen as waveform, spectrogram, and magnitude spectrum (cf. Fig. 1.8). All signals were visually inspected to identify echoes or noise pulses that the program may have misinterpreted as bat echolocation calls. The signal parameters were recalculated for all sequences containing such signals.

The results are based on a total of 82 sequences (41 per individual bat). Values are given as mean ± SD of mean of sequences, unless stated otherwise.

1.2.4 Distortion-product otoacoustic emissions

Two male individuals of *G. soricina* were used to measure DPOAEs and STCs of DPOAEs. The head of the awake bat was fixed with a mouth holder made of dental cement. A closed coupler system was introduced into the animal's outer ear canal, close to the bat's eardrum. This coupler consisted of two conical tubes. One tube was connected to three 1/2-inch microphones (*Brüel & Kjær* 4133) serving as loudspeakers for the primary (f_1, f_2) and suppressor (f_3) tones. The second tube was connected to a 1/4-inch microphone (*Brüel & Kjær* 4135) powered by a measuring amplifier (*Brüel & Kjær* 2636). The coupler system and recording set-up is described in detail by Kössl (1994) and Kössl et al. (1999).

The primary tones were generated and converted from digital to analogue on a DSP-board (*Microstar* DAP 3200/440i), the suppressor tone was produced by a synthesizer (DS335, *Stanford Research Systems*). Each tone was attenuated by a programmable attenuator (PA5, *Tucker-Davis Technologies*) and amplified (custom built by M. Kössl) before being broadcast from a speaker. The resulting emissions recorded by the microphone were fed into an input of the DSP-board and sampled at a rate of 333 kHz. In order to measure the DPOAEs in relation to primary-tone phase, the input and output channels of the DSP-board were synchronised in phase.

Sound generation and recording was controlled on a portable PC with Testpoint (*Keithley*) programs written by M. Kössl.

The system was calibrated in situ, before each measurement session and, if necessary, during the experiment after movement of the animal. Each speaker was calibrated separately with broadband noise between 1 and 166 kHz. Subsequently, the frequency dependent gain factors were determined with a reference pure-tone emitted from one of the speakers. The frequency response of the recording microphone was incorporated in the calibration routine beforehand.

DPOAEs were measured between 10 and 140 kHz (f_2). For each f_2, the best f_2/f_1 ratio was determined before measuring growth functions by keeping f_2 constant while shifting f_1. During this process, the level of f_2 was maintained 10 dB below that of f_1. This level ratio elicits highest DPOAE levels in mammals (Faulstich and Kössl, 2000; Foeller and Kössl, 2000; Probst et al., 1991). Growth functions were then measured with the best f_2/f_1 ratio by increasing the stimulus levels in steps of 5 dB and recording the levels of the resulting $2f_1$-f_2 emissions. To obtain DPOAE threshold curves, the f_2 level required to produce a $2f_1$-f_2 emission of -10 dB SPL was interpolated and plotted as a function of f_2.

Suppression tuning curves were measured using a third tone, f_3, to suppress the $2f_1$-f_2 emissions. Therefore, $2f_1$-f_2 emissions at levels more than 6 dB above noise level were elicited using the best f_2/f_1 ratio for the given f_2. Then f_3 with a frequency distance of at least 200 Hz to the DPOAE was broadcast simultaneously with the speakers. The level of f_3 was increased in 5 dB steps and the level sufficient to suppress the DPOAE by 6 dB was recorded and plotted as a function of f_3 frequency (cf. Brown and Kemp, 1984; Foeller and Kössl, 2000; Frank and Kössl, 1995). This procedure was repeated for a range of frequencies around each f_2 to obtain a tuning curve for each f_2 place on the basilar membrane.

1.3 Results

1.3.1 Echo-detection thresholds

General behaviour in the experiment

Once acquainted with the location and function of the feeders, the bats learned quickly to visit the presentation feeder to activate the reward feeders, usually within 100 trials. Generally, all bats left the box for a few seconds after visiting the presentation feeder and then came back to make their decision.

On average, the animals made between 300 and 2,000 decisions per training session. The number of decisions was correlated to the number of bats trained simultaneously; the lower the number of individuals in training, the more decisions per individual were made.

The association of a stimulus presentation with a reward at a certain feeder took the bats between approx. 1500 (bat2) to approx. 7,000 visits (bat1).

Inter-individual interference varied considerably. During initial training with four bats, One individual had to be removed from training because it was constantly pursued by another whenever visiting the feeders, and was thus in danger of being starved. Two and three individuals, however, were trained together successfully.

Detection thresholds for unfiltered calls

The detection thresholds for echoes of unfiltered calls show the influence of forward masking on the detection of the entire call.

The results illustrated in Fig. 1.6 show that the detection of only one of two bats, bat1, was affected by forward masking. At the shortest delay (2.4 ms) both bats were equally sensitive, but for longer delays, the thresholds of bat1 decreased considerably by up to 8 dB at the longest delay. In contrast, the thresholds of bat2 increased slightly

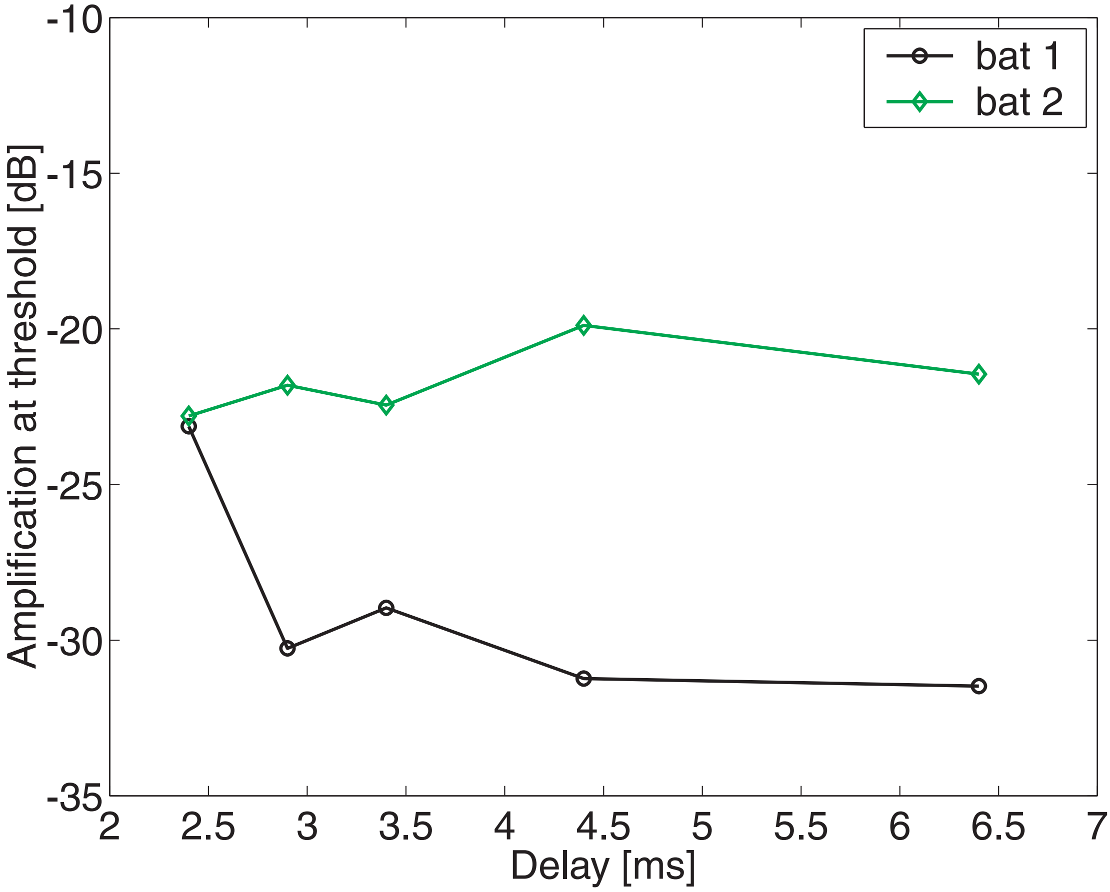

Figure 1.6: Detection thresholds for unfiltered calls plotted as amplification at threshold against playback-delay time for two bats. The lower the amplification at threshold, the higher is the sensitivity. Note that only the performance of bat1 showed influence of forward masking.

with delay time, up to 3 dB at 4.4 ms.

Detection thresholds for filtered calls

The detection thresholds of filtered calls are illustrated in Fig. 1.7 and the values are given in Tab. 1.1 for the individual bats.

Along the delay axis, the data represent the influence of forward masking in each frequency band. Along the frequency axis, the data show the sensitivity for different frequency components of the call for each delay. The performance of bat1 (Fig. 1.7 a) showed an influence of forward masking at the shortest delay time, especially in the frequency band around 126 kHz. In this frequency band, the sensitivity improved by

Table 1.1: Detection thresholds for filtered calls as function of delay and filter centre frequency for bat1 (top) and bat2 (bottom). Thresholds are given as amplification at threshold in dB (cf. Fig. 1.7).

bat1

Frequency \ Delay	**2.4 ms**	**2.9 ms**	**3.4 ms**	**4.4 ms**	**6.4 ms**
126 kHz	-15.39	-27.58	-28.25	-27.84	-33.88
100 kHz	-9.81	-8.34	-11.24	-14.2	-16.48
80 kHz	-0.17	-4.1	n.n.	-3.64	-4.75
63 kHz	2.43	-3.76	n.n.	-11.28	n.n.

bat2

Frequency \ Delay	**2.4 ms**	**2.9 ms**	**3.4 ms**	**4.4 ms**	**6.4 ms**
126 kHz	-21.02	-18.7	-25.63	-21.01	-22.76
100 kHz	-4.48	-6.67	-8.25	-12.59	-16.85
80 kHz	1.75	9.2	1.67	-1.69	2.42
63 kHz	-11.45	-12.28	0.78	-10.92	n.n.

more than 12 dB at a delay of 2.9 ms. In the frequency band around 100 kHz, the influence of forward masking was not as distinct. In the lower frequency bands (around 63 and 80 kHz), the changes in sensitivity along the delay axis were non-monotonic, and no data could be recorded at a delay of 3.4 ms.

Especially in the frequency bands around 126 and 100 kHz, bat1 exhibited lower thresholds than bat2. The performance of bat2 (Fig. 1.7 b) showed a slight sensitivity increase only in the 100 kHz band. All other changes along the delay axis were non-monotonic.

For frequencies around 63 kHz, both bats showed a high sensitivity at a delay of 4.4 ms. None of the bats was able to detect echoes in this frequency band at a delay of 6.4 ms. Both bats exhibited highest sensitivity for frequencies around 126 kHz, independent of the delay. In general, the sensitivity dropped with filter centre frequency.

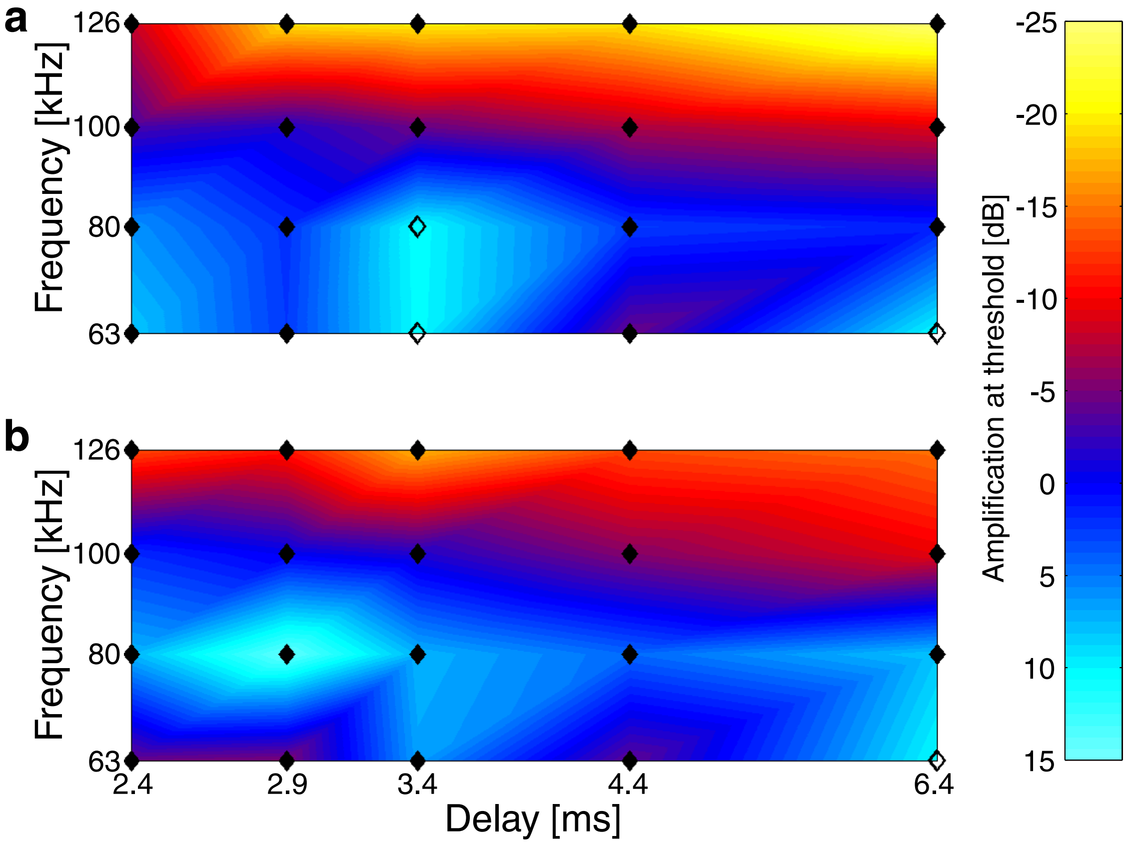

Figure 1.7: Detection thresholds for filtered calls plotted as function of delay and filter centre frequency for two bats, bat1 (**a**) and bat2 (**b**). The amplification at threshold is colour coded: low amplification (yellow) represents low thresholds, high amplification (blue) represents high thresholds. The colours are interpolated between the data points (filled symbols). The open symbols represent filter-delay combinations where the bats were not able to detect any echoes, thus, no thresholds could be measured at these points. The amplification at threshold at such points represent the maximal amplification of the system in the specific frequency range.

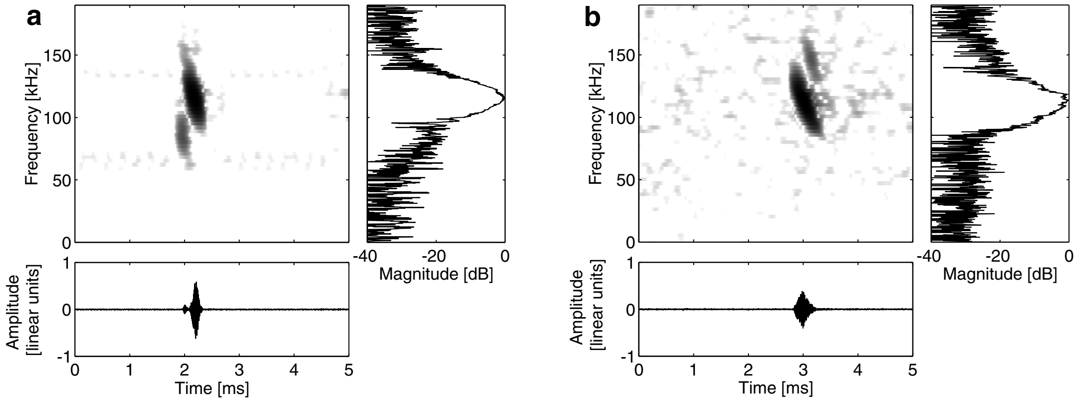

Figure 1.8: Two examples of echolocation calls of *G. soricina* recorded in the experimental set-up. Both calls are displayed as sonagram, waveform, and magnitude spectrum. Note the presence of the first harmonic in call **a** and the presence of the third harmonic in call **b**.

1.3.2 Echolocation calls

All echolocation calls emitted by *G. soricina* in the behavioural experiment were short, multiharmonic, downward modulated sweeps as illustrated in Fig. 1.8.

The duration of the echolocation calls was very short (0.33 ± 0.06 ms; mean ± SD), the longest call recorded was 0.46 ms long. The average inter-pulse interval was 43.0 ± 24.5 ms, the shortest inter-pulse interval was 24.5 ms.

The most energy was always contained in the second harmonic, the average peak frequency was measured at 116.6 ± 1.9 kHz. The second harmonic covered a frequency range from 144.6 ± 1.9 kHz to 93.0 ± 2.9 kHz, 20 dB below the peak. Thus, the average bandwidth of the second harmonic amounted to 51.5 ± 2.9 kHz. Nevertheless, the harmonic structure of the recorded echolocation calls varied: the second harmonic was always present, but the presence of first and third harmonic altered between calls. If present, the first harmonic covered a frequency range from 87.0 ± 4.1 kHz to 72.1 ± 5.6 kHz, 20 dB below the calls' intensity peak.

The average sound-pressure level of the recorded calls was 76.9 ± 5.4 dB peSPL. The highest call intensity measured amounted to 91.0 dB peSPL for both bats. The calls recorded from bat1 were by 3.8 dB louder than those recorded from bat2 (78.7 ± 4.8 versus 74.9 ± 5.4 dB peSPL). This difference was significant ($p < 0.01$, Student's t-test; Zar, 1999).

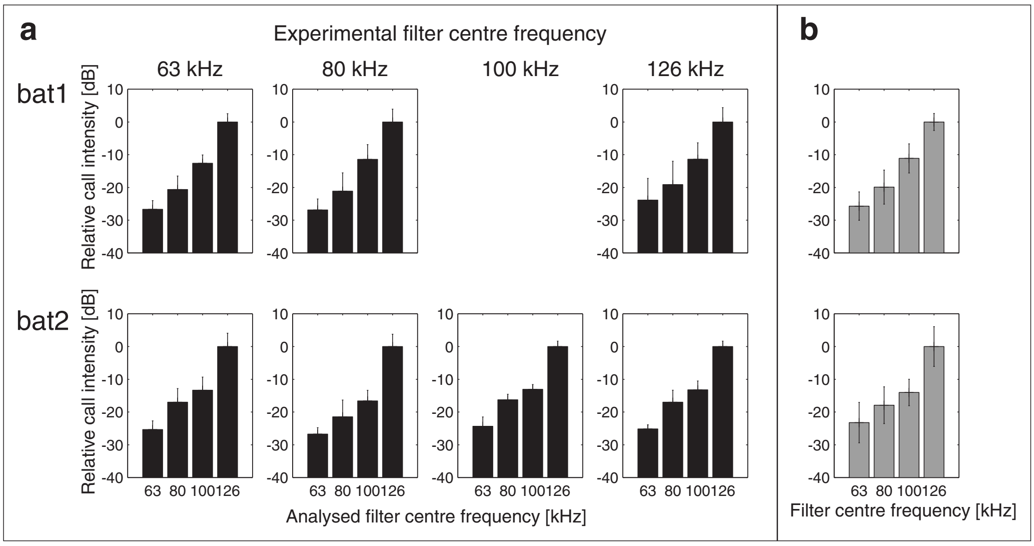

Figure 1.9: Relative intensity of echoes from filtered calls recorded during the experiment. The bars in panel **a** illustrate the intensities of the filtered call portions as function of the analysed filter centre frequency. The plot columns represent the experimental condition under which the calls were recorded, given as the experimental filter centre frequency. The rows in the plot represent recordings from the different individuals. The grey bars in panel **b** show the average data for all sequences analysed. All data are given as mean of mean of sequences, the error-bars show the SD across sequences. The data was normalised to the peak intensity.
No analysable recordings of bat1 were obtained during filtering with 100 kHz.

Relative intensity of echoes from filtered calls

The relative intensity of the filtered call portions recorded under different experimental conditions are plotted in Fig. 1.9. The energy maximum was always located in the frequency band around 126 kHz, independent of the filter used in the experiment. No systematic changes in the spectral composition of calls dependent on the experimental condition performance were detectable.

1.3.3 DPOAEs and STCs

The DPOAE threshold curves obtained from two bats are shown in Fig. 1.10. Both individuals had a sensitivity minimum at 55 kHz and another at 85 (bat3) or 80 kHz (bat4) respectively. The cochlear sensitivity in the range of the second harmonic of the echolocation calls (90 – 120 kHz) is rather high, but deteriorates above 130 kHz. The absolute sensitivity maxima were located at 80 kHz (bat3) and 100 kHz (bat4).

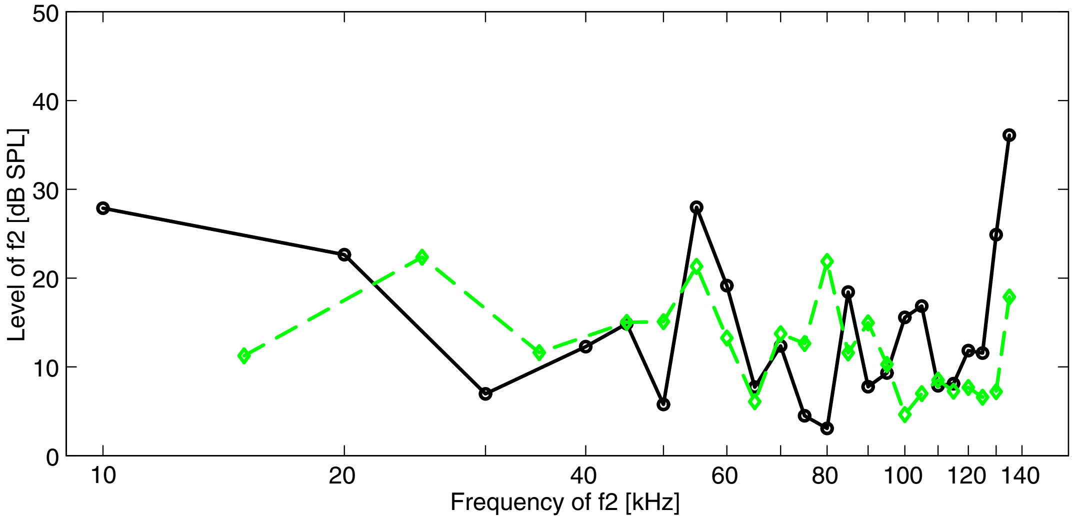

Figure 1.10: DPOAE threshold curves measured from two *G. soricina*, bat3 (circles, black line) and bat4 (diamonds, green line). The plots show the level of f_2 sufficient to elicit a $2f_1 - f_2$ DPOAE as a function of f_2 frequency.

A total of nine STCs were obtained from the two bats. Six examples are shown in Fig. 1.11. The BFs (minima of the STCs) are always located close to f_2, which indicates that the DPOAE is generated close to the place of f_2 in the cochlea on the basilar membrane. Whereas the STCs recorded from bat3 at frequencies up to 80 kHz show a relatively steep low frequency slope (Fig. 1.11 a – c), the STC recorded at 100 kHz shows a distinct low frequency tail (Fig. 1.11e). The Q_{10dB} value varied between 1.5 (at 63 kHz) and 5.0 (at 80 kHz), which indicates a rather broad tuning of the cochlear filters. The variation of the Q_{10dB} values was not correlated with the frequency of f_2.

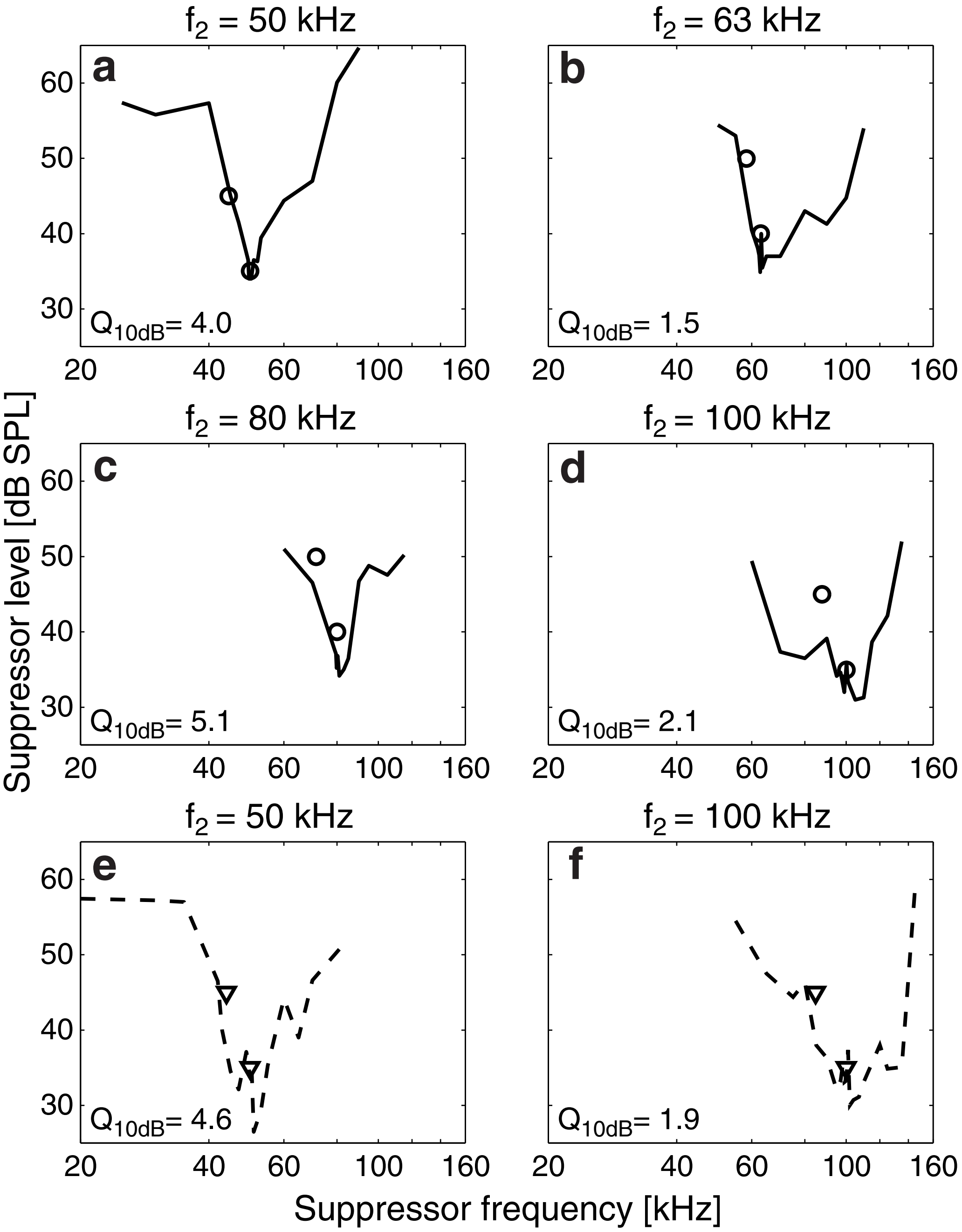

Figure 1.11: Six examples of suppression tuning curves (STCs) obtained from two bats, bat3 (solid lines; **a** – **d**) and bat4 (dashed lines; **e** – **f**). The curves are plotted as the level of f_3 sufficient to suppress f_1-f_2 by 6 dB as a function of the frequency of f_3. Q_{10dB} values and frequencies of f_2 are given in the lower left and above each panel, respectively. The frequencies and levels of the primary stimuli (f_1 and f_2) are plotted as symbols.

1.4 Discussion

1.4.1 Echo-detection thresholds

Detection of filtered echoes

The current echo-detection thresholds obtained in the playback experiment largely reflect the spectral composition of the echolocation call: lowest thresholds were measured for echoes of bandpass-filtered calls with a centre frequency of 126 kHz. In this frequency band, the relative intensity of the echoes was highest (cf. section 1.4.2). On average, the detection thresholds decreased with filter centre frequency and, correspondingly, with relative energy content in the echo. To compare these results with the behavioural audiogram (López, 2002), an estimate of thresholds to absolute sound pressure has to be calculated. The calculations plotted in Fig. 1.12 give rough approximations: the calculations are based on hypothetical calls with a relative distribution of energy across frequency bands as shown in Fig. 1.9 b and a maximal peSPL of 91 dB in the 126 kHz band. To obtain absolute thresholds, the measured amplification at the threshold averaged across playback delays was subtracted from the calculated call peSPL in each frequency band. The results plotted in Fig. 1.12 show that the thresholds for calls filtered at 126 kHz lie close to those of the behavioural audiogram by López (2002). The sensitivity for calls filtered at 100 and 80 kHz, however, seems extraordinary low. This is surprising, especially since this is the frequency range of best hearing in *G. soricina*, according to López (2002). The reason for this sensitivity drop is not clear. The comparison with the DPOAE thresholds reveals no correlation to cochlear sensitivity (Fig. 1.13). Thus, it is conceivable that higher order echo processing, e.g. by combination sensitive neurons in the auditory cortex (e.g. Maekawa et al., 1992), requires the representation of prominent call components in the echo.

Detection of unfiltered echoes

According to the detection thresholds of *G. soricina* reported here, an unfiltered echolocation call can be attenuated by 31 dB, in the best case, to be still detectable after playback. Assuming an 'ideal' call with a level of 91 dB peSPL (corresponding to the loudest calls recorded here), this would result in a detection threshold to an absolute sound pressure of 60 dB peSPL. Interestingly, this estimate matches the hearing threshold in the range of the calls' intensity peak in the behavioural audiogram

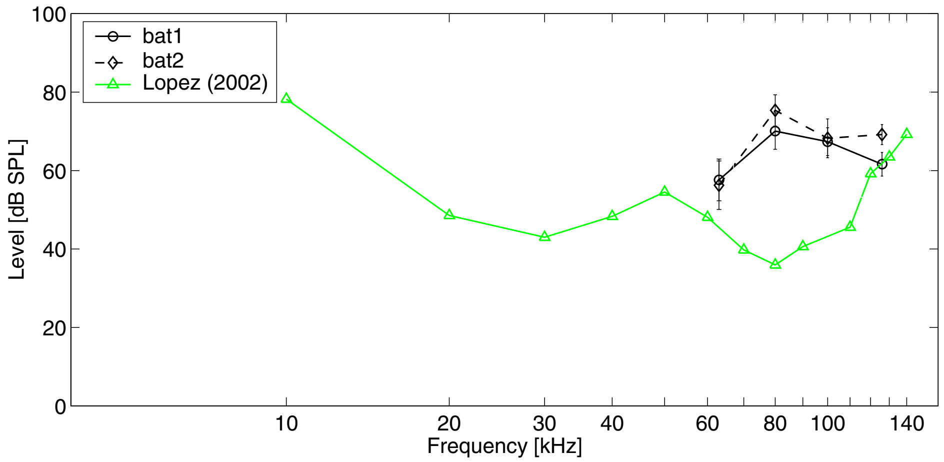

Figure 1.12: Comparison of echo-detection thresholds with the behavioural audiogram of *G. soricina*. The data plotted in black illustrate rough estimates of detection thresholds of the two individual bats. Data points are calculated based on a hypothetical call with an peSPL of 91 dB in the 126 kHz band and on the amplifications at threshold averaged across playback delays. The error-bars show the SD across playback delays and thus represent potential influence of forward masking. The green curve shows the behavioural audiogram by López (2002)

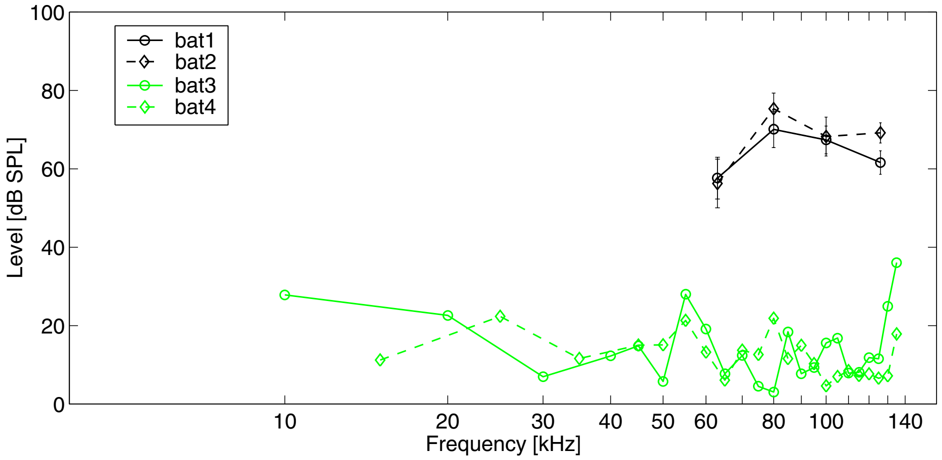

Figure 1.13: Comparison of the echo-detection thresholds with the DPOAE thresholds measured in the present study. The data plotted in black illustrate rough estimates of detection thresholds of the two individual bats. Data points are calculated as in described in Fig. 1.12. The green curves show the DPOAE threshold curves obtained from bat3 and bat4. Note that the DPOAE curve does only represent relative thresholds. Thus, only the progressions of the curves can be compared here.

by López (2002).

However, compared to echo-detection thresholds of an insectivorous bat, *Eptesicus fuscus* (Kick, 1982; Kick and Simmons, 1984; Thomson et al., 1985), these thresholds seem rather high: *E. fuscus* showed thresholds between 30 and below 10 dB SPL in the same delay range (Kick and Simmons, 1984; Thomson et al., 1985). Furthermore, the thresholds of *G. soricina* were not as strongly related to playback delay as those of *E. fuscus*. However, in contrast to *G. soricina* that forages in surroundings with high echo clutter, *E. fuscus* hunts for insects in open space habitat and thus depends on the detection of weak echoes over long distances. The detection sensitivity of the species might be adapted to these differences in habitat use. Moreover, whereas *E. fuscus* was maintained in a fixed position during the experiments, the detection thresholds of *G. soricina* were obtained from free-flying bats which complicates comparability, as discussed below.

However, also the behavioural audiogram of *G. soricina* was measured in free flying bats (López, 2002) and the data reveal high thresholds in comparison to other Phyllostomid bats (Esser and Daucher, 1996; Heffner et al., 2003; Koay et al., 2002, 2003), as well as to *E. fuscus* (Koay et al., 1997). The agreement of the data reported here with the behavioural audiogram might lead to the conclusion that *G. soricina* has lower auditory sensitivity than other species. It is more likely, though, that the method employed leads to underestimation of sensitivity. However, this problem could be overcome by using an acoustic tracking system that allows the reconstruction of beam patterns as presented by Ghose and Moss (2003). Nevertheless, the agreement of the current results with the behavioural audiogram, although based on rough estimates, indicates the significance of the current data for psychoacoustic experiments in this species under laboratory conditions.

The effect of forward masking on echo detection

Only in one of two bats (bat1), an effect of forward masking on echo detection could be measured: the sensitivity to echoes from unfiltered calls increased by approx. 7 dB for playback delays longer than 2.4 ms (cf. Fig. 1.6). As the bats' position relative to the microphone and speakers at the time of decision was not defined (see below), bat2, for example, could have traded in a lesser effect of forward masking with a lower echo intensity by making its decision farther away from the speakers and the microphone. In contrast, bat1 might have made its decision at a closer distance, which would have had the inverse effect. This example might explain the different performance of the two bats (cf. section 1.3.1). This difference in performance is consistent with the data obtained during filtered playback: the thresholds of bat1 in the 126 kHz frequency band show influence of forward masking at the shortest delay (cf. Fig. 1.7, Table 1.1). The sensitivity of bat2 changes rather erratically with increasing delay, only in the 100 kHz frequency band a slight increase of sensitivity is observable.

Hence it can be concluded that forward masking does not affect the detection of echoes for delays between 2.9 and 6.4 ms, indicating that forward masking plays a minor role for echo imaging in distances to the target of more than 41 cm.

The reason for the erratic thresholds of both bats across delays in the 63 kHz band is not clear.

1.4.2 Echolocation calls

The echolocation calls recorded in the course of the experiment show some considerable differences to those described by López (2002): first, they are much shorter (0.3 vs. 1.3 ms). Most probably, the bats have emitted shorter sounds due to the confined space conditions in the anechoic box to avoid pulse-echo overlap. However, the differences in analysis methods might also account for these large differences: López (2002) analysed call parameters exclusively in the spectrogram which can cause artifacts, especially in duration measurement (Beecher, 1988; Parsons et al., 2000).

Second, the calls' intensity peak is located at a higher frequency (117 vs. 105 kHz), but still in the second harmonic. Again, the methods of analysis might be involved here (magnitude spectrum vs. spectrogram), but the frequency might also have been shifted actively, as discussed e.g. for *Megaderma lyra* (Schmidt et al., 2000) or as a by-product of shorter call production.

Third, the occurrence and intensity of the first harmonic. López (2002) reports that the first harmonic is always present and the author considers it the most important call component, with respect to the behavioural hearing thresholds. The recordings obtained in the present study show that this is not the case: whereas the second harmonic was always present and the most prominent component of all echolocation calls of *G. soricina*, the first was, in many cases, not even emitted at a recordable level. These findings strongly indicate that the second harmonic is the most important call component of this species. This is corroborated by the echo-detection thresholds and cochlear sensitivity measured.

Although the spectral composition of the calls varied, no systematic changes dependent on the experimental condition in favour of a better detection performance were detected. For example, a systematic emphasis of the first harmonic would possibly have improved the bats' detection performance for echolocation calls filtered with 63 or 80 kHz bandpass. The systematic variation of call spectral composition has recently been described for the Indian false vampire bat: when gleaning prey, *M. lyra* varied the emphasis of call components dependent on the different stages of its hunting flight or even on different prey types (Leippert et al., 2002; Schmidt et al., 2000).

1.4.3 DPOAEs and STCs

The sonar emissions of *G. soricina* are certainly located in the higher frequency range of echolocation, not only for Phyllostomids (e.g. Kalko and Condon, 1998; Rother and Schmidt, 1982; Thies et al., 1998), but for bats using FM-calls in general: FM-calls rarely extend this far beyond frequencies of 100 kHz (e.g. Barclay, 1986; Schmidt, 1992; Schnitzler and Kalko, 1998). As already mentioned in the introduction, previous measurements of auditory sensitivity in this species do not show a specifically high sensitivity in this frequency range (Howell, 1974; López, 2002, Fig. 1.14). The results reported here, however, demonstrate a relatively high sensitivity in the range of the second harmonic of the echolocation call. The DPOAE threshold curves measured in this study attest this species excellent high frequency sensitivity up to 130 kHz in the auditory periphery.

Relative to lower frequencies, the thresholds are quite low at frequencies between 90 and 130 kHz. The DPOAE curve of bat4, for example, runs almost parallel to the abscissa in this range before sensitivity deteriorates steeply. In contrast, the behavioural audiogram by López (2002) and the recordings of cochlear microphonics by Howell (1974, Fig. 1.14) show a more or less dramatic sensitivity drop above 80 kHz. All curves show a decreased sensitivity in the same area at 50 kHz (Howell, 1974; López, 2002) or 55 kHz (present study), respectively. Also the relatively low thresholds at frequencies around 30 kHz are relatively consistent with the earlier measurements.

As in other species using FM-echolocation (e.g. *Megaderma lyra*: Weissenbacher et al., 2002), the STCs of *G. soricina* indicate that the tuning sharpness of cochlear filters (Q_{10dB} values between 1 and 5) is well within the range of those of other unspe-

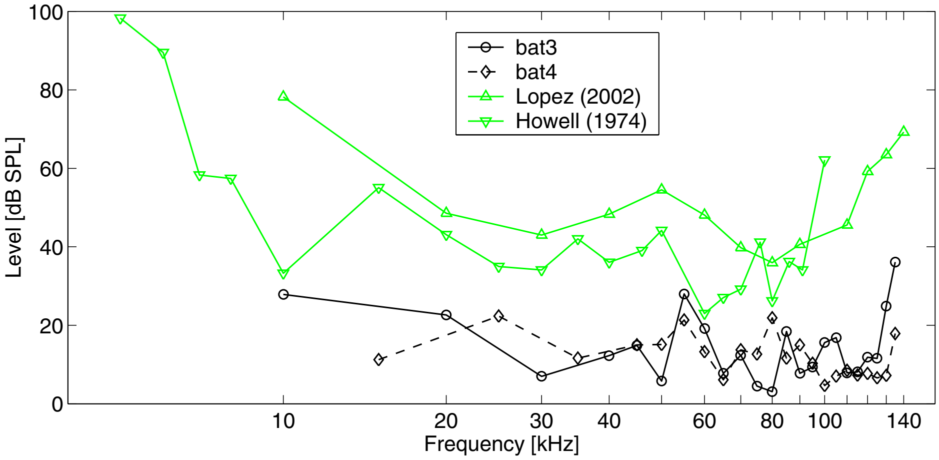

Figure 1.14: Comparison of auditory threshold measurements in *G. soricina*. The data plotted in black shows the DPOAE threshold curves (present study). The data plotted in green shows the results of earlier studies: cochlear sensitivity based on microphonic potentials (dashed; Howell, 1974) and the behavioural audiogram (solid; López, 2002).

cialised mammals (Brown and Kemp, 1984; Drexl et al., 2003). No changes in tuning width such as in bats with cochleae specialised on CF-echolocation could be observed (Foeller and Kössl, 2000; Frank and Kössl, 1995).

1.4.4 Methodological aspects

DPOAE measurements

DPOAEs were measured from awake bats. Therefore, slight movements of the head or ears unobserved by the experimenter might have caused minimal changes in coupler position. These changes, however, can affect the DPOAE intensity. This is a possible explanation for the erratic progression and the high inter-individual differences in parts of the threshold curves.

Simultaneous training of the bats

Although two to three individual bats were trained successfully together in one set-up, severe inter-individual interference was observed relatively often. The pursuit of conspecifics and colliding with them resembles the behaviour of wild *G. soricina* defending food sources (Alvarez et al., 1991). Furthermore, the number of decisions per individual was higher when less animals were trained simultaneously. This had to be compensated with frequent additional feeding. Hence, it is likely that inter-individual interference does affect the bats' performance in the experiment by causing additional stress. For further experiments, it is therefore suggested to train the animals individually.

Experimental set-up and procedure

The echo-detection thresholds were obtained from free flying bats. Therefore, their exact position relative to the microphone and speakers at the time they made their decision was not exactly defined. Even video observations of the bats' behaviour could not provide clear information in this respect. This affects the data interpretation in two important aspects: first, the playback delay. Total playback delay, for example, depends, apart from the system-innate delay, on the sound travel time between bat and microphone and between speaker and bat. Thus, the playback delays given here can only be regarded as minimal playback delays.

Second, the playback gain. Sound energy decreases with travelled distance. Thus, the gain of the playback system can only be seen as maximal gain and the thresholds obtained may be lower.

1.4.5 Summary and conclusions

The aim of the present study was to provide the basis for further playback experiments on object recognition in *G. soricina*. A fully automatic set-up for behavioural psychoacoustic playback experiments was successfully established. Although this set-up is laid out for simultaneous training of up to four individual bats, the experience gained from this study shows that individual, successive training of the animals is to be preferred.

This paradigm was successfully used to characterize the sensitivity of *G. soricina* for echoes of its own sonar emissions as a function of echo spectral content and echo delay. The data show that forward masking is not relevant for echo-delays between 2.9 and 6.4 ms. The sensitivity for echoes is highest in the frequency band between 113 and 139 kHz which contains the most energy of the call. In lower frequency bands, the sensitivity deteriorated more than it was expected considering the behavioural audiogram of this species (López, 2002). These findings indicate that the predominant second harmonic of the call is also the most relevant component for echo processing. Although the sensitivity of the auditory periphery does not deteriorate up to frequencies of 130 kHz, neither the current measurements of cochlear sensitivity and tuning nor the echo-detection thresholds revealed obvious specialisations of the auditory system to an enhanced processing of this call component. This leads to the assumption that, in the course of evolution, *G. soricina* might have traded detectability for an improved accuracy in resolution of object fine structure by shifting either the emphasis of the entire call or only of its processing to such high frequencies.

The fact that behavioural auditory thresholds of *G. soricina* are high compared to those of other species suggests that auditory sensitivity is likely to be underestimated in free-flying bats. Thus, to ensure a reliable detection of echoes by *G. soricina* in playback experiments, the stimuli should be presented with a sound pressure of at least 60 dB SPL and with delays longer than 2.9 ms. Furthermore, a calibration of the playback system up to frequencies of at least 150 kHz has to be accomplished.

If these parameters are taken into account, *G. soricina* offers a high potential for the investigation of object recognition by echolocation, based on its feeding ecology and the reduced procedural learning due to the unhesitant acceptance of automatic feeding which facilitates training in computer-controlled paradigms.

2 Cross-modal object recognition in *Glossophaga soricina*

2.1 Introduction

The ability to integrate information across senses maximizes the chances for the recognition of important objects, also under conditions in which the predominant sense is impaired. Especially in terms of foraging, this ability would be highly advantageous for the recognition of food items or landmarks for orientation under varying conditions. The most familiar example for cross-modal transfer is the interrelatedness of the visual and haptic sense in humans, which has been studied extensively in the last decades (e.g. Gottfried et al., 1977; Rose et al., 1998; Rose and Ruff, 1987): we can easily recognise objects by vision that we have only experienced before tactually and vice versa. A recent study has shown that even newborn children are able to fulfil this generalization (Streri, 2003). An information transfer between visual and haptic sense has also been found in apes (Davenport, 1976; Savage-Rumbaugh et al., 1988) and monkeys (Blakeslee and Gunter, 1966; Cowey and Weiskrantz, 1975; DiMattia et al., 1990; Tolan et al., 1981).

In echolocating dolphins, vision is interlinked with acoustic imaging: in a striking experiment, Pack and Herman (1995) have shown that bottlenose dolphins easily match familiar objects presented to echolocation to samples presented to vision and vice versa. A following study revealed that these animals spontaneously recognise unfamiliar objects across modalities, as well (Herman et al., 1998). More recently, the experiments of Harley et al. (2003) indicated that bottlenose dolphins extract the features of objects directly from the perceived echoes. These findings suggest, that, at least in dolphin echolocation, the percept of an object is not only formed by the proximal stimuli, i.e. the sensation of receptor cells, but by the distal stimuli, the object properties themselves. These results are of utmost importance, as they prove that the acoustic percepts of echoes can actually form an 'auditory object' in the animal brain and that the recognition of a certain object is not only based on the recognition of a recurring time- or spectral pattern in the echo.

Based on these findings, it may be hypothesized that also echolocating bats perform information transfer across echolocation and vision.

Apart from their echo-imaging system, all michrochiropteran bat species possess functional vision to perceive their environment. Although vision in bats has not been studied in as much detail as echolocation, the few existent studies revealed a well developed object recognition capabilities: it has been shown that bats can discriminate different orientations of stripe patterns or of forms (Neuweiler, 2000). *Anoura geoffroyi*, a glossophagine bat, can distinguish different forms, like circles and squares, from one another and is also able to discriminate between erect and inverted triangles (Suthers et al., 1969). Very recently, Winter et al. (2003) have demonstrated that the experimental animal of this study, *G. soricina* has a sensitivity maximum for ultraviolet light. These findings are of special interest with regard to the fact that some bat-pollinated, neotropical flowers reflect UV-light (Winter and v. Helversen, 2001). Can bats transfer the information about object characteristics assessed in one modality to the other?

This question was addressed in a computer-controlled 2-AFC training experiment. *G. soricina* had to discriminate a rewarded target from another one, having no cues but those accessible through echolocation. After the bats had learned this task, the objects were presented to the visual sense alone, to test whether the animals were still able to discriminate these objects. If so, this would indicate that an acoustic percept of an object can be transferred to visual sense in this bat.

2.2 Materials and methods

2.2.1 Animals

For the present experiment, four male individuals of *G. soricina* were trained one after another in a basement room of the dimensions 3.8 x 2 x 2 m ($d \times w \times h$). The temperature and humidity were kept constant at approx. 25°C and 65%, respectively. The bats were subject to an artificial diurnal rhythm of 12 hours light (1400 to 0200) and 12 hours darkness (0200 to 1400). Training and data acquisition was carried out exclusively in the dark phase, usually six days per week. The animals had no access to food during the light phase. In the experiment, the bats were rewarded with sugar water.

For further details about maintenance and feeding of the bats see chapter 1.

Bat1 and bat2 were trained in a behavioural experiment before and bat4 was used for

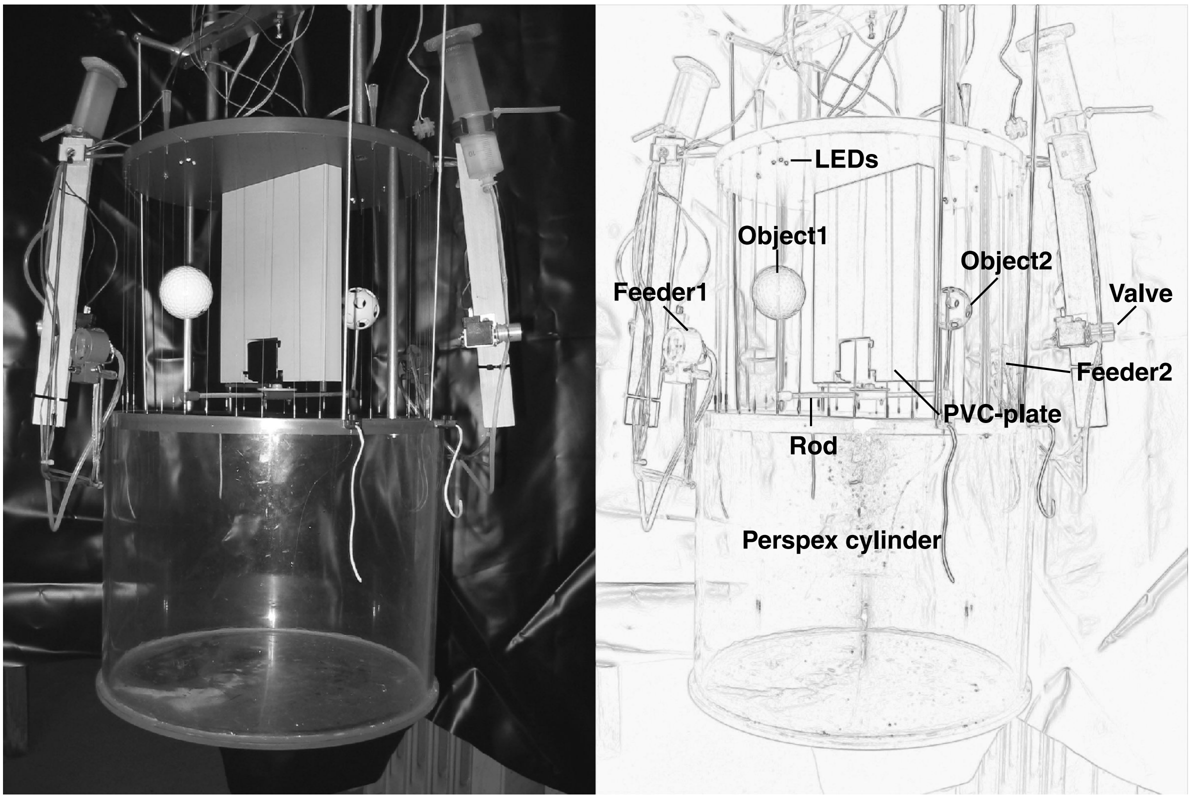

Figure 2.1: Experimental set-up for cross-modal object discrimination. The left panel shows a photograph of the set-up as it was used during training (object discrimination by echolocation cues). The legends to important items are given in the right panel. The position of the objects was altered by turning the rod, so that object1 was located behind feeder2 and vice versa. Note that the perspex cylinder is in the lowered position (cf Fig. 2.2).

DPOAE measurements (see chapter 1).

2.2.2 Experimental set-up

The aim of this training set-up was to present two objects in two different positions, perceivable only by echolocation or only by vision. A solid golf ball and a hollow, perforated training-golf ball (see Fig. 2.1, 2.2) were chosen as objects for discrimination.

The experimental set-up was suspended in the centre of the room from an iron u-profile mounted horizontally under the ceiling.

The apparatus had a cylindrical shape (Fig. 2.1). The cylinder (27 cm in diameter) was bisected vertically by a PVC plate. An object was presented in front of each face of the plate. Therefore, the two different objects were mounted on either end of a rod made of aluminum tubing. This rod could be rotated with a computer controlled servo motor, in order to alter the positions of the objects. Outside of the cylinder, a feeder (cf. section 1.2.2) was mounted in front of each object location. Each feeder was equipped with an infrared light barrier to indicate the visit of the bat to the computer. The centre of each object was located 10 cm behind and 6 cm above the feeder.

Above each object location, three white light-emitting diodes (LEDs) were installed in the top cover of the cylinder to illuminate the objects when presented in visual mode. To prevent the bats from flying inside the cylinder, fishing line (0.2 mm diameter) was tensed between top and bottom of the cylinder to form a 'fence' with 1 cm distance between the lines.

To shield the objects from echolocation during visual presentation, the entire apparatus could be encased with a perspex cylinder. This perspex cylinder was suspended on strings from an electro motor and could be moved to a raised and a lowered position to switch between visual and acoustic presentation modes, respectively. Of course, the feeders were accessible independent of the cylinder position.

The training of the animals was observed with an infrared-sensitive video-surveillance system from the outside of the experimental room.

The routines for the control of the experiment and for data acquisition and storage were self-programmed in *MATLAB* v. 5.3.

2.2.3 Procedure

Generally, the bats had to associate one of the two objects with a food reward (approx. 10 μl of sugar water). To receive a reward, the bat had to visit the feeder located in front of the rewarded object. The positions of the objects were altered pseudorandomly by the training routine. After a maximum of three successive trials with the same object positioning, the positions were altered forcedly by the computer. If the wrong feeder was visited, the bat was reinforced with a timeout of ten seconds. During this timeout period, the objects were moved away from the feeders, next to the PVC plate, and no visit was registered at either feeder.

Initially, the bats had to learn this task by using echolocation. Therefore, this training phase was carried out in complete darkness with lowered perspex cylinder, thus, the objects were only perceivable by echolocation ("acoustic mode"). As each individ-

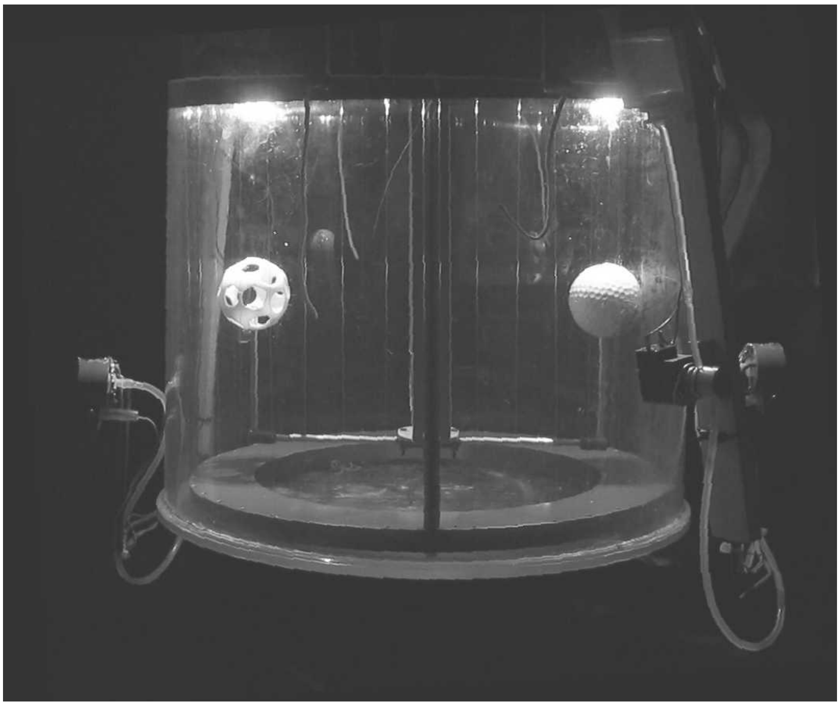

Figure 2.2: Experimental set-up as it was used for testing (object discrimination by visual cues). Note that the perspex cylinder is in raised position and that the LEDs illuminate the objects (cf Fig. 2.1).

ual bat lived in the experimental room during the experiment, the objects had to be shielded from both echolocation and vision during the 12 h of the light period, at least in the training phase. Therefore, the perspex cylinder was covered with paper. Each training session ended with the raising of the cylinder 10 min before the light went on.

After the bat showed a stable performance of at least 65 % correct choices per training session in this task, the procedure was changed to visual presentation, in order to test whether the representation of an object is stable across modalities: the objects were shielded from echolocation by the perspex cylinder and the LEDs illuminated the objects. Thus, the objects could only be perceived by vision (“visual mode”).

For data analysis the visiting time, feeder-number, object position, and correct/wrong answer were logged by the computer for each visit.

2.2.4 Data analysis

To obtain the bats' performance as a function of training time, learning curves were calculated as follows: initially, the data of each bat were normalised to the onset of visual object presentation, i.e., the first number of the first trial in visual mode was set to zero. This was done for an easier comparison between individuals. Subsequently, the discrimination performance was calculated as percent correct choices for portions of 100 successive trials. For 100 trials, a performance of 65 % correct choices is significantly different from chance level (Miller, 1996).

2.3 Results

Three of four bats learned to discriminate the two objects by echolocation after a training period of 20,000 (bat5) to 41,000 trials (bat1). Bat2 was excluded from training after a period of approx. 45,000 trials without any detectable learning effect.

All three bats were not able to solve the discrimination task when the objects were presented to vision afterwards. The results are shown as plots of the learning curves in Fig. 2.3. During visual presentation of the objects, the bats' performance dropped to chance level.

To control whether the bats were still able to solve the discrimination task by echolocation after testing in visual mode, the objects were again presented to bat4 in the acoustic mode. The results plotted in Fig. 2.3 show that bat4 could accomplish the discrimination again immediately.

2.4 Discussion

All three bats successfully trained in this experiment were not able to associate the visual impression of the objects with the template learned by echolocation. Are echolocating bats, in contrast to dolphins, unable to transfer the perception of objects across modalities? At first glance, the current results indicate that the answer to this question is "yes". But several aspects have to be taken into account before finally answering this question.

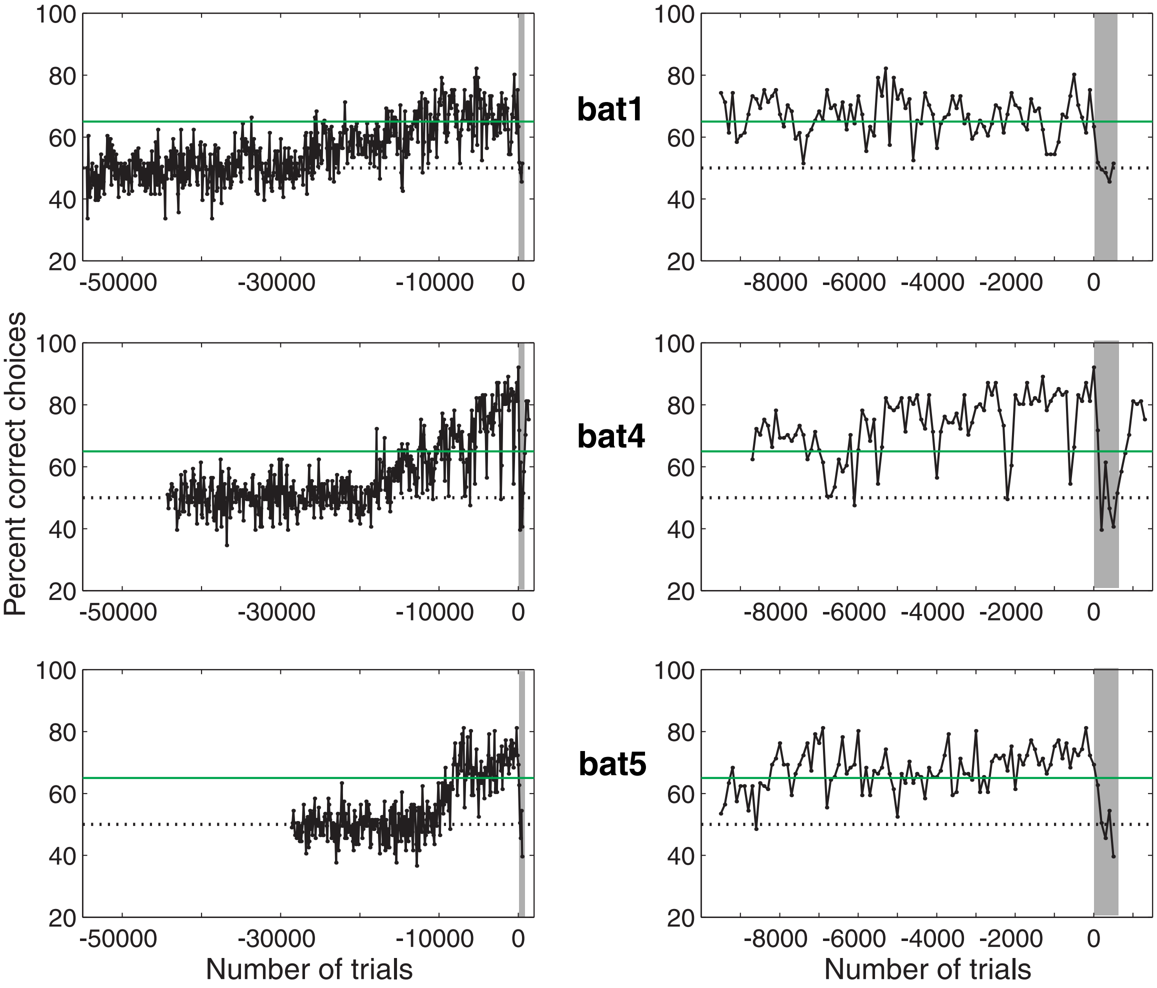

Figure 2.3: Discrimination performance of *G. soricina* plotted as the learning curves of the three individuals (plot rows). The left plot column shows the entire learning curve for each bat, the right column shows a more detailed part, starting from 9,500 trials before switching to visual object presentation. Each data point represents the percentage of correct choices of 100 trials. The data are normalised to the onset of visual presentation of the objects (zero-point on the x-axis). Negative trial numbers illustrate the period of training to discrimination by echolocation. The grey boxes show the time window in which the objects were exposed to vision only. The dotted horizontal lines represent the random choice level (50 %), the solid horizontal lines represent the 65 % criterion. Note that the performance of all bats dropped to random choice level when the objects were presented in visual mode. Also note that bat4 was still able to solve the task acoustically after testing in visual mode.

First of all, the bats needed a minimum of 20,000 trials to learn the discrimination of the two objects by echolocation. This is an extremely long learning period for *G. soricina*, compared to other studies with similar tasks (cf. v. Helversen and v. Helversen, 2003). This might have been caused by the relatively large distance between target and feeder due to the perspex cylinder that had to be moved up and down between both. The performance of *G. soricina* in object recognition tasks deteriorates with increasing distance of feeder and target (Y. Winter, personal communication). Thus, the association of an object with a reward was possibly hampered by the experimental design.

Furthermore, the test for a recognition of the objects by vision was based on the assumption that the bats had previously formed a cognition of the object characteristics which are encoded in the objects' IRs. However, the discrimination by sonar could have been accomplished simply by the detection of differences in a single echo parameter. Hence, the bats had not been forced to memorise IRs characterising each object as such.

In addition, in visual mode the bats perceived diverging information from echolocation and vision due to the perspex screen. Supposedly, this is a natural situation for dolphins that most likely cannot echolocate in air (Au, 1993) but nevertheless have to see and identify objects above water. For bats, however, that have to orient in one medium only, this situation must be quite unusual. It is thus conceivable that *G. soricina* is not able to extract information necessary to identify an object from two diverging sources which requires a certain level of abstraction.

These problems can be addressed in "matching to sample"paradigms as those used in the dolphin experiments (Harley et al., 2003; Herman et al., 1998; Pack and Herman, 1995): the dolphins had to match one of two or three objects presented to one sense to a sample object presented to the other. However, since the high density of water and the resulting impedance mismatch between water and air facilitate the exposure of objects to one sense alone (cf. Au, 1993; Pack and Herman, 1995), a comparable experimental design is hard to apply to airborne sonar.

Although Weissenbacher and Wiegrebe (2003) have shown that bats do indeed develop an internal representation of object IRs extracted from the echoes, it is still in question whether they are capable of assembling different IRs of the same object to a three-dimensional representation of the object. Whereas echolocation provides excellent depth resolution (Schmidt, 1988; Simmons et al., 1974) but will provide comparably less information about the contours along the dimensions of height and width of an object, the reverse will be the case in vision. In fact, height and width are probably the most important cues for visual object perception, not only for bats (review: Neuweiler, 2000). This could be of particular significance in the present experiment, where the outline of the two objects was basically the same.

Consequently, it is conceivable that, at least in bats, a three dimensional object representation can not be accomplished by one sense alone but by concurring perceptions of both, echolocation and vision. An example for such an interaction of both senses is the visual calibration of the three-dimensional auditory space map in the barn owl (Knudsen and Brainard, 1991). Regarding the well developed object recognition capabilities of glossphagine bats in both vision and echolocation (Suthers et al., 1969; v. Helversen et al., 2003), it appears valid to hypothesize that both modalities contribute to a representation of an object's dimensions. This hypothesis will be tested in the current paradigm by exposing the objects to both senses during training and subsequently testing in either mode, whether the bats are still able to accomplish the discrimination task.

To conclude, the question whether bats are able to transfer the perception of objects from echolocation to vision can not be answered entirely at this stage. Nevertheless, the present results show, that echolocating bats do not generally form a representation of an object perceived by echolocation that can be matched with a visual percept of the same object. These results, however, might be biased by an overestimation of the bats' abstraction capabilities. Thus, this interesting field has to be further investigated before satisfactory conclusions can be drawn about the information transfer between acoustic and visual imaging in bats.

3 Classification of natural textures in echolocation

This chapter has been published in 2004 under the title 'Classification of natural textures in echolocation' by Jan-Eric Grunwald, Sven Schörnich, and Lutz Wiegrebe in PNAS.

3.1 Abstract

Through echolocation, a bat can not only perceive the position of an object in the dark, it can also recognize its three-dimensional structure. A tree, however, is a very complex object; it has thousands of reflective surfaces which result in a chaotic acoustic image of the tree. Technically, the acoustic image of an object is its impulse response (IR), i.e., the sum of the reflections recorded when the object is ensonified with an acoustic impulse. The extraction of the acoustic IR from the ultrasonic echo and the detailed IR analysis underlies the bats' extraordinary object-recognition capabilities. Here, a phantom-object, playback experiment is developed to demonstrate that the bat *Phyllostomus discolor* can evaluate a statistical property of chaotic IRs, the IR roughness. The IRs of the phantom objects consisted of up to 4,000 stochastically distributed reflections. It is shown that *P. discolor* spontaneously classifies echoes generated with these IRs according to IR roughness. This capability enables the bats to evaluate complex natural textures, such as foliage types, in a meaningful manner. The present behavioral results and their simulations in a computer model of the bats' ascending auditory system indicate the involvement of modulation-sensitive neurons in echo analysis.

3.2 Introduction

The neural interpretation of sensory input into an object-based, sensory scenery is a main focus in neuroscience. The echolocation of bats and dolphins is an ideal model system because echolocating mammals have perfect control over their sensory data acquisition due to the active nature of echolocation. A useful analysis of the acoustic scenes, as they are represented in sequences of echoes, requires the identification of the acoustically complex objects surrounding the animals in their natural habitat. Many studies have provided insights into the extraordinary capabilities of echolocating animals in object recognition and classification (Dror et al., 1995; Harley et al., 2003, 1996; Helweg et al., 1996; Herman et al., 1998; Moss and Simmons, 1993; Saillant et al., 1993; Simmons, 1979; Simmons et al., 1990, 1995b; v. Helversen and v. Helversen, 2003; Weissenbacher and Wiegrebe, 2003).

In their natural nocturnal habitat, bats are forced to orient in and navigate through a highly structured environment. How can echolocation serve these tasks? The echoes produced by potential landmarks for orientation, such as trees or bushes, are highly chaotic: the ultrasonic emission of a bat is reflected from a multitude of surfaces, the leaves, which are chaotically distributed in space and angle to the sound source and receiver. Thus, the echoes reflected from such an object will have a chaotic waveform and no systematic spectral interference pattern (cf. Fig. 3.1). Moreover, the echoes are highly unstable over time, as they are susceptible to both changes of the bat's observation angle and e.g. wind-induced movement of the object. Thus, a bat will rarely receive the same echo of an individual object twice.

Up to now, object recognition in echolocation has been studied only with deterministic echoes from small objects with very few reflections. The echoes from such objects can be evaluated according to their characteristic waveforms and/or frequency patterns (Schmidt, 1988; Simmons et al., 1990; Weissenbacher and Wiegrebe, 2003). However, these concepts appear insufficient to describe the analysis of the chaotic echoes a bat has to cope with in its natural habitat.

An echo as it is perceived by a bat consists of its ultrasonic emission convolved with the acoustic impulse response (IR) of the ensonified object. The IR is the sum of the reflections when the object is ensonified with an acoustic impulse of theoretically infinite shortness and infinite amplitude. Thus, the IR is a physical object property whereas the echo as it is perceived by a bat also depends on the structure of the emitted sound.

What are the typical characteristics of the IRs of large natural objects? A conifer, for example, has needle-shaped, densely distributed leaves, i.e., many surfaces, each of them producing only a faint reflection. Thus the IR will consist of many chaotically distributed reflections, each with a relatively low amplitude. A synthetic IR with these characteristics is shown in Fig. 3.2 a. In contrast, a broad-leafed tree has fewer surfaces each of them producing a stronger reflection. Thus the IR will consist of fewer chaotically distributed reflections each with a relatively larger amplitude

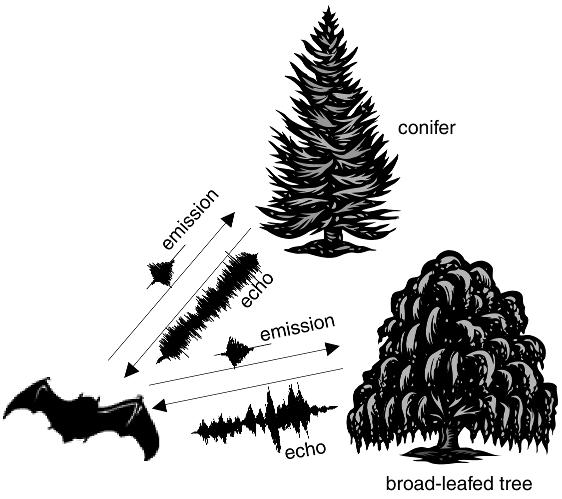

Figure 3.1: Illustration of sonar emissions of a bat and the echoes it may receive from different foliage types. Note that a conifer produces a smoother echo than a broad-leafed tree.

(cf. Fig. 3.2c). Thus, while both IRs are chaotic, they will differ in the statistical description of their envelopes: a conifer has a smoother IR than a broad-leafed tree (cf. Fig. 3.1). Recent theoretical work has confirmed the power of a statistical echo analysis for the classification of large natural objects (Müller and Kuc, 2000).

This study investigates whether bats are able to evaluate statistical properties of complex IRs in a behavioral experiment. The fruit-eating bat, *Phyllostomus discolor* was trained to evaluate echoes digitally generated from their ultrasonic emissions and IRs with up to 4,000 reflections.

3.3 Methods

3.3.1 Animals

The experimental animal, the lesser spear-nosed bat, *P. discolor*, forages for fruit, nectar, pollen, and insects in a neotropical forest habitat. Hence, this species has to navigate through highly structured surroundings. *P. discolor* emits brief (<3 ms), broadband, multi-harmonic echolocation calls covering the frequency range between 45 and 100 kHz (Rother and Schmidt, 1982). Four female individuals took part in the experiments.

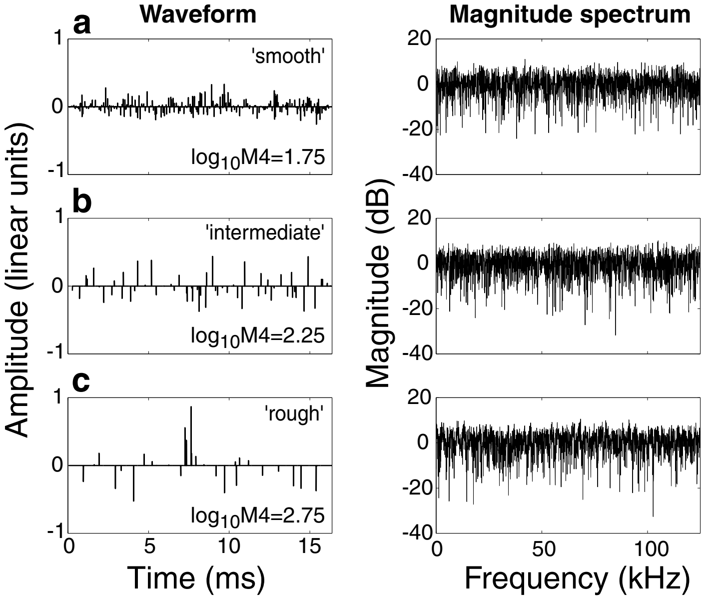

Figure 3.2: Three examples of complex IRs with increasing roughness from **a** to **c** as used in the experiment. IR roughness is quantified as the base-ten logarithm of the IR 4th moment ($\log_{10} M4$, see Methods). The IRs are plotted as waveform (left) and magnitude spectrum (right). Every non-zero amplitude value in the waveform represents a single reflection from a surface of a complex object. Note the frequency independent magnitude spectrum of all three IRs despite the large waveform differences. Each IR had a duration of 16.4 ms and equal root-mean-squared amplitude.

3.3.2 Impulse responses

We created complex IRs with different degrees of roughness: each IR consisted of a 4,096-sample portion of sparse noise (Hübner and Wiegrebe, 2003). Sparse noise is a Gaussian noise with random-width temporal gaps (nulls) between the amplitude values. The different degrees of roughness were achieved by varying the average width of the temporal gaps. All the resulting IRs had chaotic waveforms and frequency-independent magnitude spectra (cf. Fig. 3.2). We quantified the IR roughness by calculating the base-ten logarithm of the fourth moment ($\log_{10} M4$) of the IRs (Hartmann and Pumplin, 1988). The fourth moment is calculated as

$$M4 = \frac{\frac{1}{T}\int_0^T x^4(t)dt}{[\frac{1}{T}\int_0^T x^2(t)dt]^2}$$

where $x(t)$ is the time-domain representation and T is the duration of the IR.

For the initial training of the animals, two specific training-IRs were used, a smooth one ($\log_{10} M4$ = 1.75, cf. Fig. 3.2 a), and a rough one ($\log_{10} M4$ = 2.75, cf. Fig. 3.2 c). For the stimulation in the test trials, we used 50 test-IRs in five groups defined by their average $\log_{10} M4$. The five groups of IRs had a roughness of 1.75 ± 0.005, 2.0 ± 0.005, 2.25 ± 0.006, 2.5 ± 0.016 and 2.75 ± 0.026. Error values represent standard deviations. Thus, each group contained 10 individual IRs with a similar roughness. Fig. 3.2 shows three examples of the IRs used. All IRs had the same root-mean-squared amplitude. At the given sampling rate (250 kHz), the IRs had a duration of 16.4 ms, corresponding to an object depth of approx. 2.8 m.

3.3.3 Experimental setup

The bats were trained in a 2-alternative, forced-choice, playback setup, consisting of a horizontal Y-shaped maze (45×30 cm; wire mesh) in an echo-attenuated chamber. A starting perch was located at the bottom leg of the Y, a reward feeder was mounted at the end of each upper leg. The inner width of each leg was 10 cm. To indicate its decision, the bat had to crawl to the reward feeder in either the left or the right upper leg of the maze. For the playback of the echoes, an ultrasonic speaker (*Matsushita* EAS10 TH800D, Osaka, Japan) was mounted centrally between the upper legs of the Y-maze, directed towards the starting perch.

Further, a 1/4-inch ultrasonic microphone (*Brüel & Kjær* 4135, Nærum, Denmark) was located on top of the speaker to pick up the sonar emissions of the bat. The microphone-speaker unit was located at 25 cm distance from the perch.

During the experiment, the amplified and band-pass filtered (20–100 kHz, 24 dB/oct, *Krohn Hite* 3550, Brockton, Massachusetts, USA) echolocation calls were digitized by a data-acquisition board (*Microstar* DAP 5200a, Bellevue, Washington, USA) at a sampling rate of 250 kHz. Each recorded call was convolved with a specific IR on the DAP-board by multiplication of the complex spectra of the recorded emission, zero-padded to 4,096 samples, and the IR. This mathematical operation corresponds to the physical formation of the echo from a real object. Thus, any change of the bat's ultrasonic emission resulted in an immediate change of the perceived echo. The resulting echo was converted from digital to analogue and played back to the bat after a total delay of 18 ms relative to emission. This delay corresponds to a target distance of approx. 3 m.

3.3.4 Procedure

First, four individuals were trained in a two-alternative, forced-choice paradigm to discriminate two specific IRs. A smooth IR (Fig. 3.2 a) was associated with a food reward at the left feeder. A rough IR (Fig. 3.2 c) was associated with a reward at the right feeder.

Data acquisition started when the bats had achieved a performance of at least 85 % correct choices in this discrimination task. Then, test trials were randomly interspersed with a probability of 25 %. In these test trials, one of the 50 unknown IRs was presented and the bats were rewarded independently of their decision. Behavioural results are based on at least 40 test trials per animal and experimental condition.

3.3.5 Simulations

The classification of the IRs was simulated based on one of two representations of the perceived echoes: the auditory spectrograms or the output of a modulation-filterbank. These representations were obtained from a detailed computer model of the auditory peripheral processing in *P. discolor*. This model was fed with echoes, i.e., with the experimental IRs convolved with a typical *P. discolor* echolocation call. We have simulated the manipulation of acoustic stimuli applied both by the animals' outer- and middle ear (Esser and Daucher, 1996), and by cochlear processing based on distortion-product otoacoustic emission suppression tuning curves (Wittekind, 2003). Inner-hair cell processing was simulated by half-wave rectification, exponential compression (exponent = 0.4) and filtering with a second-order low-pass filter at 1 kHz (Palmer and Russell, 1986) to simulate the loss of phase locking. These manipulations resulted in an auditory-spectrogram representation of the perceived echoes. In the first simulation, these auditory spectrograms, arranged along the time- and auditory-frequency axes, were considered as the model output.

For the second simulation, the auditory spectrograms were fed into a modulation filterbank model (Dau et al., 1997a,b) with ten modulation filters with center frequencies (CFs) logarithmically spaced between 30 and 500 kHz. The model output for the second simulation was the AC-coupled root-mean-square of the modulation-filterbank output (Ewert and Dau, 2000). The model decisions were based on the similarity of the model output computed with a test IR relative to the model outputs computed with the two training IRs. The similarity was quantified as the root-mean-squared distance (Euclidean distance, ED) between the different model outputs. For each test IR, classification performance in percent was calculated according to the following equation:

$$judged\ as\ smooth(\%) = 100 \times \left(1 - \frac{EDsmooth}{EDsmooth + EDrough}\right)$$

where $EDsmooth$ is the Euclidean distance between the model outputs computed with a test IR and the smooth training IR and $EDrough$ is the Euclidean distance between the model outputs computed with a test IR and the rough training IR.

3.4 Results

3.4.1 Behavioral performance

In a two-alternative, forced-choice paradigm, the four individuals were successfully trained to discriminate a single smooth IR (Fig. 3.2 a) from a single rough one (Fig. 3.2c). The horizontal lines in Fig. 3.3 show the bats' discrimination performance recorded in the subsequent test phase.

In this test phase, we investigated to which extent the bats can generalize roughness to test IRs which the bats had not experienced before. The spontaneous responses to these test IRs are shown in Fig. 3.3 as a function of the IR roughness. The five bars represent the bats' spontaneous classification of unknown test IRs from the five groups with an IR roughness as specified on the abscissa. The bars show that spontaneous classification is monotonically related to the IR roughness: Unknown IRs with low roughness were spontaneously judged 'smooth' in a high percentage of test trials; unknown IRs with high roughness were only rarely judged 'smooth'. IRs with intermediate roughness resulted in a similar amount of 'smooth' or 'rough' judgements.

Human psychophysical studies have shown that stimuli with the same sound pressure level and long-term spectrum can produce different degrees of masking (Carlyon and Datta, 1997; Kohlrausch and Sander, 1995) and loudness (Gockel et al., 2003) dependent on their degree of envelope fluctuation. To investigate whether the bats may have based their decisions on differences in perceived echo loudness, we repeated the classification experiment with a roving-level paradigm: the echoes of the training IRs were attenuated by 6 dB compared to the level in the original paradigm; the echoes of the test IRs were presented at levels roving by ± 6 dB around that of the training

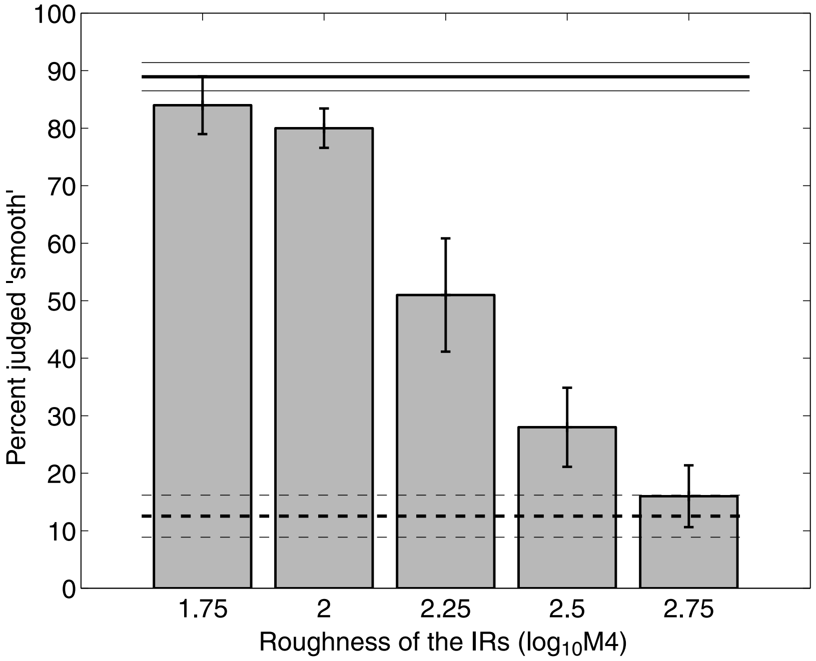

Figure 3.3: Discrimination and classification of chaotic IRs by *P. discolor*. Performance is quantified as percent of trials judged as 'smooth'. The solid and dashed strong horizontal lines show the discrimination of the smooth and rough IR in the training trials, respectively. The thin horizontal lines represent standard errors. The bars show the spontaneous classification of unknown chaotic IRs as a function of the IR roughness. The bats' spontaneous classification is monotonically related to the IR roughness. Error bars represent inter-individual standard errors.

IRs. The results of this control experiment, performed with two of the four bats, are shown in Fig. 3.4. The IRs were classified in the same manner as in the original experiment. This control experiment shows that the classification performance of the bats was not based on differences in perceived echo loudness.

How may the bats' auditory system evaluate echo roughness? As stated above, the bat does not perceive the IR as such but the IR imprinted on its own sonar emission, an echo. A spectral analysis of the echo envelope shows that the envelope spectrum is monotonically related to the echo roughness. In the mammalian auditory system, properties of the envelope spectrum can be encoded by modulation-sensitive neurons. Hence, it is conceivable that modulation-sensitive neurons described in the bats' auditory brainstem (Grothe et al., 2001; Huffman et al., 1998a,b) can encode the roughness of perceived echoes.

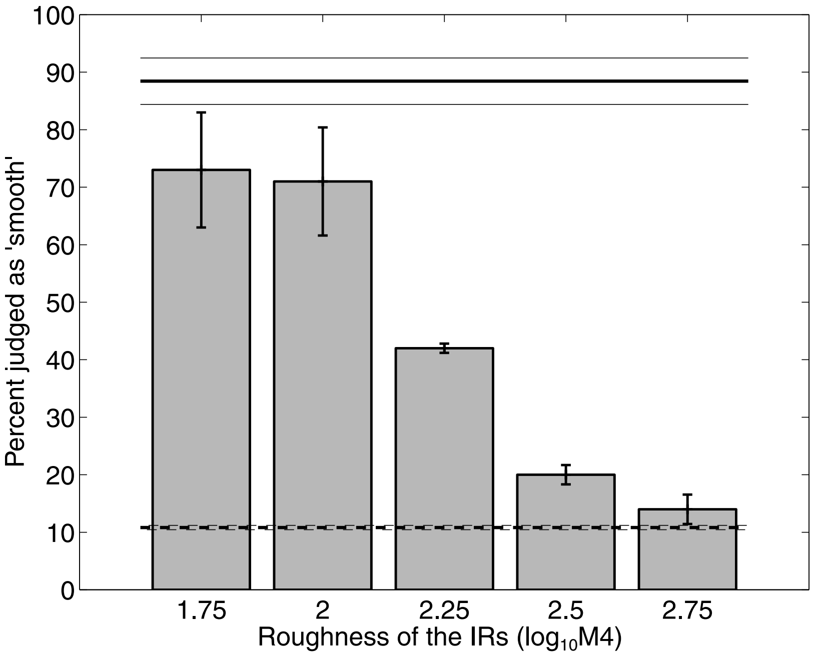

Figure 3.4: Classification of chaotic IRs by two *P. discolor* in a roving-level paradigm. Classification performance is plotted in the same format as in Fig. 3.3. The echoes in the training trials (horizontal lines) were attenuated by 6 dB relative to the main experiment. The echoes in the test trials were played back at randomized levels (± 6 dB) roving around that of the training echoes. Note that the bats spontaneously classified the IRs in a similar way as shown in Fig. 3.3.

3.4.2 Simulation results

As outlined in the methods, simulations of the behavioral performance were based on either auditory-spectrogram representations of the perceived echoes or on the outputs of a hypothetical modulation filterbank as a functional implementation of neural envelope analysis. Exemplary representations of the echo waveforms, the generated auditory spectrograms and the modulation filterbank outputs are shown in Fig. 3.5. The simulation results are shown in Fig. 3.6. The simulation based on the similarities in the auditory spectrograms (Fig. 3.6 a) show only a weak correlation to the bats' performance in the experiment.

The simulation based on the similarities in the modulation filterbank outputs (Fig. 3.6 b) generates a better fit to the behavioral data. This is true despite the fact that the modulation filterbank outputs appear not to vary very much with IR roughness (cf. Fig. 3.5, bottom row).

3.5 Discussion

The present results show for the first time that echolocating bats spontaneously evaluate and generalize the roughness of chaotic IRs. Such chaotic IRs arise from large natural objects like trees and bushes and they are thus abundant in the animals' natural habitat.

Which neural processing strategies may underlie the analysis of IR roughness? The bats do not perceive the IR of an object but the object's IR convolved with their sonar emission, i.e., an echo. Earlier studies have indicated that bats may be able to reconstruct the IR from the detailed comparison of their sonar emission and the echo (Simmons et al., 1990; Weissenbacher and Wiegrebe, 2003). However, it is not clear whether *P. discolor* can do so with such complicated IRs consisting of thousands of reflections. Thus, at present we reside to simulate the auditory analysis of chaotic echoes based on the echoes themselves, not on the IRs.

The simulation results show that an auditory spectrogram representation of the perceived echoes, as it would exist in the bats' auditory nerve, does not provide a reliable estimate of IR roughness. The deterministic encoding of the echo temporal structure in the auditory spectrograms precludes a successful evaluation of the statistical echo properties.

When the information from the auditory nerve is subjected to a modulation-filterbank analysis, the modulation-filterbank output provides an improved fit to the experi-

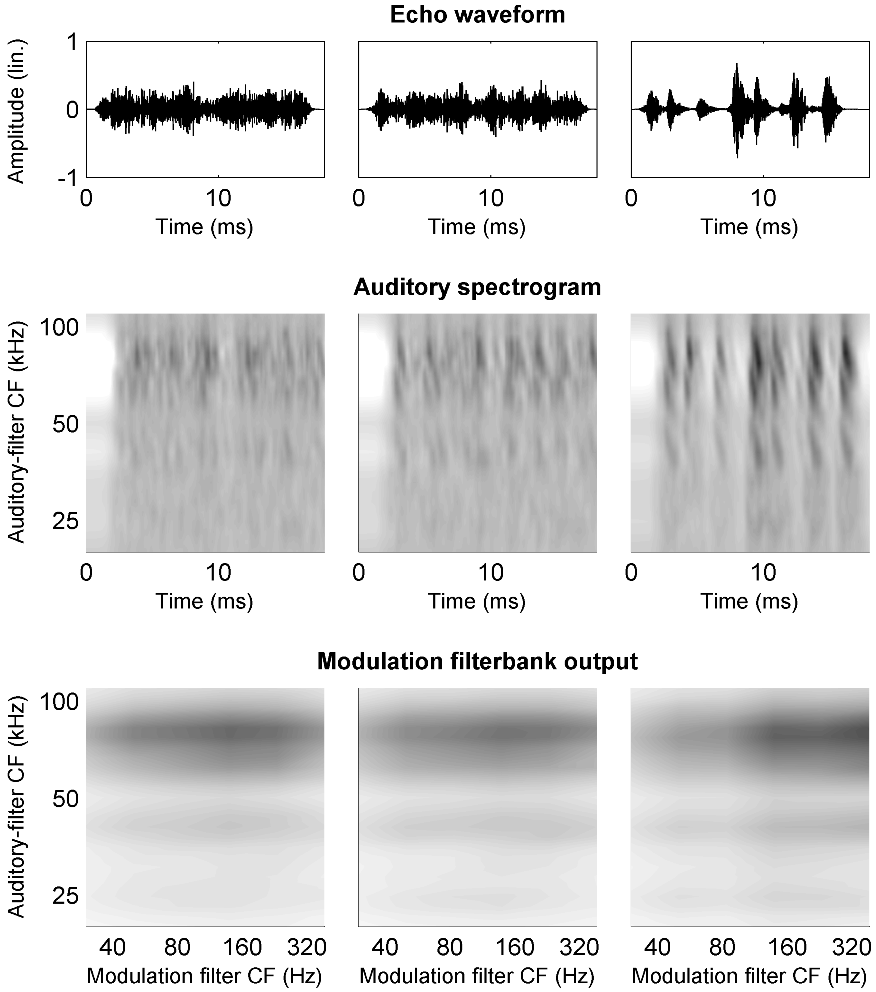

Figure 3.5: Simulated echoes generated with IRs as used in the behavioral experiments and their auditory representations as simulated in a computer model of the peripheral auditory system of *P. discolor*. The left and right columns show the representations for the two IRs used in training; the middle column shows the representation for a single test IR from the IR group with a $\log_{10} M4$ of 2.0. The top row shows the echo waveforms, the middle row shows the simulated auditory spectrograms, and the bottom row shows the simulated modulation filterbank outputs. The grayscale in the middle and bottom row encodes the simulated neural activation in arbitrary units.

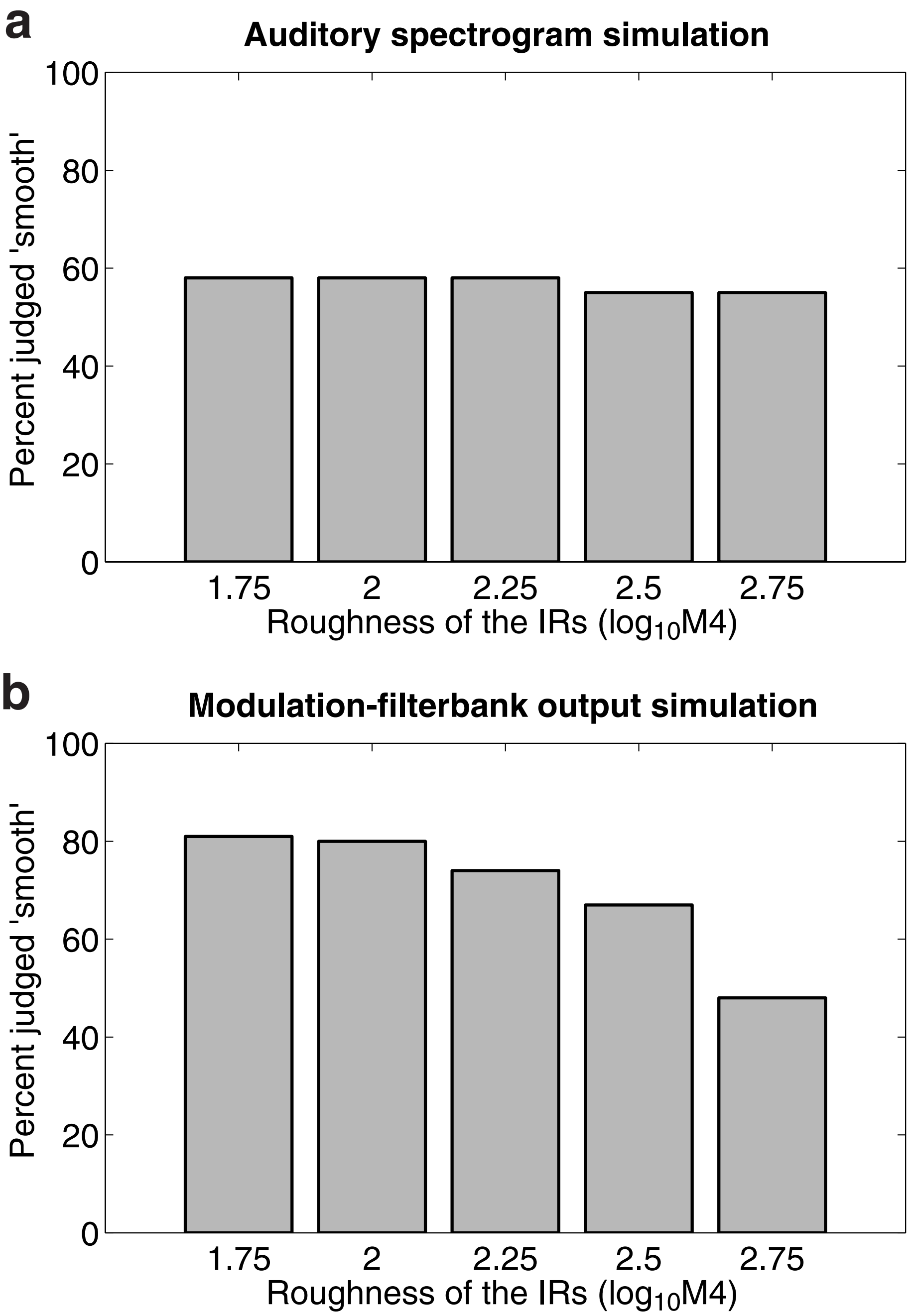

Figure 3.6: Simulation results of the behavioral performance based on two different auditory representations of the generated echoes. The simulation based on auditory-spectrograms (**a**) provides a poor fit to the experimental data (cf. Fig. 3.3) because, in the auditory spectrogram, the temporal structure of the echo waveform is encoded in a deterministic fashion. A subsequent analysis of the auditory spectrograms in a modulation filterbank (**b**) provides a better representation of echo roughness and consequently results in an improved fit to the experimental data.

mental data. The success of the modulation-filterbank simulation results from the stability of the filterbank output across different realizations of stochastic IRs with similar roughness.

The modulation filterbank analysis revealed that modulation magnitude in the modulation frequency range around 80 to 200 Hz can be used to evaluate the roughness of the experimental echoes. Modulation sensitive units covering this range have been characterized physiologically in bats (Grothe et al., 2001; Huffman et al., 1998a). Thus, the simulations support the hypothesis that modulation sensitive neurons, e.g. in the auditory midbrain may play an important role in the processing of stochastic echoes. However, even with a modulation-filterbank analysis, the fit to the experimental data is not fully satisfactory. Better fits to the experimental data can be obtained if a simulation was based on the evaluation of the IR itself, not on the echo. This, however, requires the preceding reconstruction of the IR from the echo. Future studies will reveal to which extent *P. discolor* can reconstruct the IR of complex objects having thousands of reflective surfaces.

In previous research, chaotic echoes from natural textures have mostly been regarded as disturbing 'clutter'. In the light of the current data, these chaotic echoes should be regarded as a contribution to a meaningful acoustic image of the bat's surroundings. It is conceivable that for flying bats, the perception of an acoustic stream on the basis of changes of echo roughness facilitates navigation guided by echolocation. The spontaneous classification of unknown chaotic IRs along an ecologically meaningful parameter indicates the significance of a statistical evaluation of echo properties for the natural behavior of bats.

Bibliography

Alvarez, J., Willig, M. R., Jones, J. K., and Webster, W. D. (1991). *Glossophaga soricina*. *Mammalian Species*, 379:1–7.

Au, W. W. L. (1993). *The sonar of dolphins*. Springer Verlag, New York.

Aubauer, R. and Au, W. W. (1998). Phantom echo generation: a new technique for investigating dolphin echolocation. *J Acoust Soc Am*, 104(3):1165–1170.

Aubauer, R., Au, W. W., Nachtigall, P. E., Pawloski, D. A., and DeLong, C. M. (2000). Classification of electronically generated phantom targets by an Atlantic bottlenose dolphin (*Tursiops truncatus*). *J Acoust Soc Am*, 107(5):2750–2754.

Barclay, R. M. R. (1986). The echolocation calls of hoary (*Lasiurus cinereus*) and silver-haired (*Lasionycteris noctivagans*) bats as adaptations for long versus short-range foraging strategies and the consequences for prey selection. *Can J Zool*, 64(12):2700–2705.

Beecher, M. D. (1988). Spectrographic analysis of animal vocalizations: implications of the “uncertainty principle”. *Bioacoustics*, 1:187–208.

Blakeslee, P. and Gunter, R. (1966). Cross-modal transfer of discrimination learing in Cebus monkeys. *Behaviour*, 26(1):76–90.

Brown, A. M. and Kemp, D. T. (1984). Suppressibility of the 2f1-f2 stimulated acoustic emissions in gerbil and man. *Hear Res*, 13(1):29–37.

Carlyon, R. P. and Datta, A. J. (1997). Masking period patterns of Schroeder-phase complexes: effects of level, number of components, and phase of flanking components. *J Acoust Soc Am*, 101(6):3648–3657.

Cowey, A. and Weiskrantz, L. (1975). Demonstration of cross-modal matching in rhesus monkeys, *Macaca mulatta*. *Neuropsychologia*, 13(1):117–120.

Dau, T., Kollmeier, B., and Kohlrausch, A. (1997a). Modeling auditory processing of amplitude modulation. I. Detection and masking with narrow-band carriers. *J Acoust Soc Am*, 102(5 Pt 1):2892–2905.

Dau, T., Kollmeier, B., and Kohlrausch, A. (1997b). Modeling auditory processing of amplitude modulation. II. Spectral and temporal integration. *J Acoust Soc Am*, 102(5 Pt 1):2906–2919.

Davenport, R. K. (1976). Cross-modal perception in apes. *Ann. N. Y. Acad. Sci.*, 280:143–149.

DiMattia, B. V., Posley, K. A., and Fuster, J. M. (1990). Crossmodal short-term memory of haptic and visual information. *Neuropsychologia*, 28(1):17–33.

Drexl, M., Faulstich, M., v. Stebut, B., Radtke-Schuller, S., and Kössl, M. (2003). Distortion Product Otoacoustic Emissions and Auditory Evoked Potentials in the Hedgehog Tenrec, *Echinops telfairi*. *J Assoc Res Otolaryngol*, 4:555–564.

Dror, I. E., Zagaeski, M., and Moss, C. F. (1995). Three-Dimensional Target Recognition via Sonar: A Neural Network Model. *Neural Networks*, 8(1):149–160.

Esser, K. H. and Daucher, A. (1996). Hearing in the FM-bat *Phyllostomus discolor*: A behavioral audiogram. *J Comp Physiol A*, 178(6):779–785.

Ewert, S. D. and Dau, T. (2000). Characterizing frequency selectivity for envelope fluctuations. *J Acoust Soc Am*, 108(3 Pt 1):1181–1196.

Faulstich, M. and Kössl, M. (2000). Evidence for multiple DPOAE components based upon group delay of the 2f(1)-f(2) distortion in the gerbil. *Hear Res*, 140(1-2):99–110.

Foeller, E. and Kössl, M. (2000). Mechanical adaptations for echolocation in the cochlea of the bat *Hipposideros lankadiva*. *J Comp Physiol A*, 186(9):859–870.

Frank, G. and Kössl, M. (1995). The shape of 2f-1-f-2 suppression tuning curves reflects basilar membrane specializations in the mustached bat, *Pteronotus parnellii*. *Hear Res*, 83(1-2):151–160.

Gaskill, S. A. and Brown, A. M. (1990). The behavior of the acoustic distortion product, 2f1-f2, from the human ear and its relation to auditory sensitivity. *J Acoust Soc Am*, 88(2):821–839.

Ghose, K. and Moss, C. F. (2003). The sonar beam pattern of a flying bat as it tracks tethered insects. *J Acoust Soc Am*, 114(2):1120–1131.

Gockel, H., Moore, B. C., Patterson, R. D., and Meddis, R. (2003). Louder sounds can produce less forward masking: effects of component phase in complex tones. *J Acoust Soc Am*, 114(2):978–990.

Gottfried, A. W., Rose, S. A., and Bridger, W. H. (1977). Cross-modal transfer in human infants. *Child Dev.*, 48(1):118–123.

Grothe, B., Covey, E., and Casseday, J. H. (2001). Medial superior olive of the big brown bat: neuronal responses to pure tones, amplitude modulations, and pulse trains. *J Neurophysiol*, 86(5):2219–2230.

Habersetzer, J. and Vogler, B. (1983). Discrimination of surface-structured targets by the echolocating bat *Myotis myotis* during flight. *J Comp Physiol A*, 152:275–282.

Harley, H. E., Putman, E. A., and Roitblat, H. L. (2003). Bottlenose dolphins perceive object features through echolocation. *Nature*, 424(6949):667–669.

Harley, H. E., Roitblat, H. L., and Nachtigall, P. E. (1996). Object representation in the bottlenose dolphin (*Tursiops truncatus*): integration of visual and echoic information. *J Exp Psychol Anim Behav Process*, 22(2):164–174.

Hartmann, W. M. and Pumplin, J. (1988). Noise power fluctuations and the masking of sine signals. *J Acoust Soc Am*, 83(6):2277–2289.

Hübner, M. and Wiegrebe, L. (2003). The effect of temporal structure on rustling-sound detection in the gleaning bat, *Megaderma lyra*. *J. Comp. Physiol. A*, 189(5):337–346.

Heffner, R. S., Koay, G., and Heffner, H. E. (2003). Hearing in American leaf-nosed bats. III: *Artibeus jamaicensis*. *Hear Res*, 184(1-2):113–122.

Helweg, D. A., Roitblat, H. L., Nachtigall, P. E., and Hautus, M. J. (1996). Recognition of aspect-dependent three-dimensional objects by an echolocating Atlantic bottlenose dolphin. *J Exp Psychol Anim Behav Process*, 22(1):19–31.

Herman, L. M., Pack, A. A., and Hoffmann-Kuhnt, M. (1998). Seeing through sound: dolphins (*Tursiops truncatus*) perceive the spatial structure of objects through echolocation. *J Comp Psychol*, 112(3):292–305.

Howell, D. J. (1974). Acoustic behavior and feeding in glossophagine bats. *J Mammal*, 55(2):293–308.

Huffman, R. F., Argeles, P. C., and Covey, E. (1998a). Processing of sinusoidally amplitude modulated signals in the nuclei of the lateral lemniscus of the big brown bat, *Eptesicus fuscus*. *Hear Res*, 126(1-2):181–200.

Huffman, R. F., Argeles, P. C., and Covey, E. (1998b). Processing of sinusoidally frequency modulated signals in the nuclei of the lateral lemniscus of the big brown bat, *Eptesicus fuscus*. *Hear Res*, 126(1-2):161–180.

Kalko, E. K. V. and Condon, M. A. (1998). Echolocation, olfaction and fruit display: how bats find fruit of flagellichorous cucurbits. *Functional Ecology*, 12(3):364–372.

Kalko, E. K. V. and Schnitzler, H. U. (1998). How echolocating bats approach and acquire food. In Kunz, T. H. and Racey, P. A., editors, *Bat biology and conservation*, chapter 13, pages 197–204. Smithsonian Institution Press, Washington, London.

Kick, S. A. (1982). Target detection by the echolocating bat, *Eptesicus fuscus*. *J Comp Physiol A*, 145:431–435.

Kick, S. A. and Simmons, J. A. (1984). Automatic gain control in the bat's sonar receiver and the neuroethology of echolocation. *J Neurosci*, 4(11):2725–2737.

Knudsen, E. I. and Brainard, M. S. (1991). Visual instruction of the neural map of auditory space in the developing optic tectum. *Science*, 253(5015):85–87.

Koay, G., Bitter, K. S., Heffner, H. E., and Heffner, R. S. (2002). Hearing in American leaf-nosed bats. I: *Phyllostomus hastatus*. *Hear Res*, 171(1-2):96–102.

Koay, G., Heffner, H. E., and Heffner, R. S. (1997). Audiogram of the big brown bat (*Eptesicus fuscus*). *Hear Res*, 105(1):1–2.

Koay, G., Heffner, R. S., Bitter, K. S., and Heffner, H. E. (2003). Hearing in American leaf-nosed bats. II: *Carollia perspicillata*. *Hear Res*, 178(1-2):27–34.

Kohlrausch, A. and Sander, A. (1995). Phase effects in masking related to dispersion in the inner ear. II. Masking period patterns of short targets. *J Acoust Soc Am*, 97(3):1817–1829.

Kössl, M. (1994). Otoacoustic emissions from the cochlea of the 'constant frequency' bats, *Pteronotus parnellii* and *Rhinolophus rouxi*. *Hear Res*, 72(1-2):59–72.

Kössl, M. (1997). Sound emission from cochlear filters and foveae–does the auditory sense organ make sense? *Naturwissenschaften*, 84(1):9–16.

Kössl, M., Mayer, F., Frank, G., Faulstich, M., and Russell, I. J. (1999). Evolutionary adaptations of cochlear function in Jamaican mormoopid bats. *J Comp Physiol A*, 185(3):217–228.

Leippert, D., Frank, E., Gabriel, P., Kutter, S., Scheidemann, K. D., von Stillfried, N., and Weller, F. (2002). Prey-correlated spectral changes in echolocation sounds of the Indian false Vampire *Megaderma lyra*. *Ethology*, 108:139–156.

López, J. E. (2002). *Verhaltensphysiologische Bestimmung der optischen und akustischen Empfindlichkeit der Blumenfledermaus Glossophaga soricina (Phyllostomidae)*. Dissertation, Universität Erlangen.

Maekawa, M., Wong, D., and Paschal, W. G. (1992). Spectral selectivity of FM-FM neurons in the auditory cortex of the echolocating bat, *Myotis lucifugus*. *J Comp Physiol A*, 171(4):513–522.

Miller, J. (1996). The sampling distribution of d'. *Percept. Psychophys.*, 58(1):65–72.

Müller, R. and Kuc, R. (2000). Foliage echoes: a probe into the ecological acoustics of bat echolocation. *J Acoust Soc Am*, 108(2):836–845.

Moore, B. C. J. (1997). *An introduction to the psychology of hearing*. Academic Press, London, San Diego, 4th edition.

Moss, C. F. and Simmons, J. A. (1993). Acoustic image representation of a point target in the bat *Eptesicus fuscus*: Evidence for sensitivity of echo phase in bat sonar. *J Acoust Soc Am*, 93(3):1553–1562.

Neuweiler, G. (1989). Foraging ecology and audition in echolocating bats. *Trends in Ecology and Evolution*, 4(6):160–166.

Neuweiler, G. (2000). *The biology of bats*. Oxford University Press, New York.

Neuweiler, G., Singh, S., and Sripathi, K. (1984). Audiograms of a south indian bat community. *J. Comp. Physiol. A*, 154(1):133–142.

Pack, A. A. and Herman, L. M. (1995). Sensory integration in the bottlenosed dolphin: immediate recognition of complex shapes across the senses of echolocation and vision. *J Acoust Soc Am*, 98(2):722–733.

Palmer, A. R. and Russell, I. J. (1986). Phase-locking in the cochlear nerve of the guinea-pig and its relation to the receptor potential of inner hair-cells. *Hear Res*, 24(1):1–15.

Parsons, S., Boonman, A. M., and Obrist, M. K. (2000). Advantages and disadvantages of techniques for transforming and analyzing chiropteran echolocation calls. *J Mammal*, 81(4):927–938.

Probst, R., Lonsbury-Martin, B. L., and Martin, G. K. (1991). A review of otoacoustic emissions. *J Acoust Soc Am*, 89(5):2027–2067.

Rose, S. A., Feldman, J. F., Futterweit, L. R., and Jankowski, J. J. (1998). Continuity in tactual-visual cross-modal transfer: infancy to 11 years. *Dev. Psychol.*, 34(3):435–440.

Rose, S. A. and Ruff, H. A. (1987). Cross-modal abilities in human infants. In Osofsky, J., editor, *Handbook of infant development*, pages 318–362. Wiley, New York, 2nd edition.

Rother, G. and Schmidt, U. (1982). The influence of visual information on echolocation in Phyllostomus discolor (Chiroptera). *Z. Säugetierk.*, 47(6):324–334.

Saillant, P. A., Simmons, J. A., Dear, S. P., and McMullen, T. A. (1993). A computational model of echo processing and acoustic imaging in frequency-modulated echolocating bats: the spectrogram correlation and transformation receiver. *J Acoust Soc Am*, 94(5):2691–2712.

Savage-Rumbaugh, S., Sevcik, R. A., and Hopkins, W. D. (1988). Symbolic cross-modal transfer in two species of chimpanzees. *Child Dev.*, 59(3):617–625.

Schmidt, S. (1988). Evidence for a spectral basis of texture perception in bat sonar. *Nature*, 331:617–619.

Schmidt, S. (1992). Perception of structured phantom targets in the echolocating bat, *Megaderma lyra*. *J Acoust Soc Am*, 91(4):2203–2223.

Schmidt, S., Hanke, S., and Pillat, J. (2000). The role of echolocation in the hunting of terrestrial prey- new evidence for an underestimated strategy in the gleaning bat, *Megaderma lyra*. *J Comp Physiol A*, 186(10):975–988.

Schnitzler, H. U. and Henson, O. W. (1980). Performance of animal sonar systems: I. Microchiroptera. In Busnel, R. G. and Fish, J. F., editors, *Animal Sonar Systems*, pages 109–181. Plenum Press, New York.

Schnitzler, H. U. and Kalko, E. K. V. (1998). How echolocating bats search and find food. In Kunz, T. H. and Racey, P. A., editors, *Bat biology and conservation*, chapter 12, pages 183–196. Smithsonian Institution Press, Washington, London.

Simmons, J. A. (1979). Perception of Echo Phase Information in Bat Sonar. *Science*, 204(4399):1336–1338.

Simmons, J. A., Ferragamo, M. J., Saillant, P. A., Haresign, T., Wotton, J. M., Dear, S. P., and Lee, D. N. (1995a). Auditory dimensions of acoustic images in echolocation. In Popper, A. N. and Fay, R. R., editors, *Hearing by bats*, Springer Handbook of Auditory Research. Springer- Verlag, New York, 5th edition.

Simmons, J. A., Lavender, W. A., Lavender, B. A., Doroshow, C. A., Kiefer, S. W., Livingston, R., Scallet, A. C., and E., C. D. (1974). Target Structure and Echo Spectral Discrimination By Echolocating Bats. *Science*, 186:1130–1132.

Simmons, J. A., Moss, C. F., and Ferragamo, M. (1990). Convergence of temporal and spectral information into acoustic images of complex sonar targets perceived by the echolocating bat, *Eptesicus fuscus*. *J Comp Physiol A*, 166(4):449–470.

Simmons, J. A., Saillant, P. A., Wotton, J. M., Haresign, T., Ferragamo, M. J., and Moss, C. F. (1995b). Composition of Biosonar Images for Target Recognition by Echolocating Bats. *Neural Networks*, 8(7-8):1239–1261.

Simmons, J. A. and Stein, R. A. (1980). Acoustic imaging in bat sonar echolocation signals and the evolution of echolocation. *J Comp Physiol A*, 35(1):61–84.

Streri, A. (2003). Cross-modal recognition of shape from hand to eyes in human newborns. *Somatosens Mot Res*, 20(1):13–18.

Suthers, R., Chase, J., and Braford, B. (1969). Visual form discrimination by echolocating bats. *Biol Bull*, 137(3):535–546.

Thies, W., Kalko, E. K. V., and Schnitzler, H. U. (1998). The roles of echolocation and olfaction in two Neotropical fruit-eating bats, *Carollia perspicillata* and *C. castanea*, feeding on Piper. *Behav Ecol Sociobiol*, 42(6):397–409.

Thomson, C. E., Fenton, M. B., and Barclay, R.-M. R. (1985). The role of infant isolation calls in mother-infant reunions in the little brown bat *Myotis lucifugus* chiroptera vespertilionidae. *Can J Zool*, 63(8):1982–1988.

Tolan, J. C., Rogers, C. M., and Malone, D. R. (1981). Cross-modal matching in monkeys: altered visual cues and delay. *Neuropsychologia*, 19(2):289–300.

v. d. Emde, G. and Menne, D. (1989). Discrimination of insect wingbeat frequencies by the bat *Rhinolophus ferrumequinum*. *J Comp Physiol A*, 164(5):663–672.

v. d. Emde, G. and Schnitzler, H. U. (1990). Classification of insects by echolocating greater horseshoe bats. *J Comp Physiol A*, 167(3):423–430.

v. Helversen, D., Holderied, M. W., and v. Helversen, O. (2003). Echoes of bat-pollinated bell-shaped flowers: conspicuous for nectar-feeding bats? *J Exp Biol*, 206(6):1025–1034.

v. Helversen, D. and v. Helversen, O. (1999). Acoustic guide in bat-pollinated flower. *Nature*, 398:759–760.

v. Helversen, D. and v. Helversen, O. (2003). Object recognition by echolocation: a nectar-feeding bat exploiting the flowers of a rain forest vine. *J Comp Physiol A*, 189(5):327–336.

v. Helversen, O. (1993). Adaptations of flowers to the pollination by glossophagine bats. In Barthlott, W., editor, *Animal-plant interactions in tropical environments*, pages 41–59. Museum Koenig, Bonn.

v. Helversen, O. and Winter, Y. (2003). Glossophagine bats and their flowers: costs and benefits for plants and pollinators. In Kunz, T. H. and Fenton, M. B., editors, *Ecology of bats*, pages 346–397. University of Chicago Press, Chicago.

Weissenbacher, P. and Wiegrebe, L. (2003). Classification of virtual objects in the echolocating bat, *Megaderma lyra*. *Behav Neurosci*, 117(4):833–839.

Weissenbacher, P., Wiegrebe, L., and Kössl, M. (2002). The effect of preceding sonar emission on temporal integration in the bat, *Megaderma lyra*. *J Comp Physiol A*, 188(2):147–155.

Winter, Y., López, J., and v. Helversen, O. (2003). Ultraviolet vision in a bat. *Nature*, 425:612–614.

Winter, Y. and v. Helversen, O. (2001). Bats as pollinators: foraging energetics and floral adaptations. In Chittka, L. and Thomson, J. D., editors, *Cognitive ecology of pollination*, chapter 8, pages 149–170. Cambridge University Press, Cambridge.

Wittekind, A. (2003). *Periphere Hörverarbeitung bei Phyllostomus discolor*. Diplom thesis, Universität Frankfurt.

Zar, J. H. (1999). *Biostatistical analysis*. Prentice Hall, Upper Saddle River, 4th edition.

List of Figures

Abbreviations

2-AFC	Two-alternative, forced choice
$2f_1 - f_2$	Cubic distortion product
AC	Alternating current
CF	Centre frequency
DPOAE	Distortion product otoacoustic emission
DSP	Digital signal processing
ED	Euclidean distance
f_1	Frequency of primary tone f_1
f_2	Frequency of primary tone f_2
f_3	Frequency of primary tone f_3
FFT	Fast Fourier Transformation
FIR	Finite impulse response
IR	Impulse response
LED	Light-emitting diode
$log_{10}M4$	base-ten logarithm of the fourth moment
peSPL	peak-equivalent sound-pressure level
$Q_{10}dB$ value	Ratio of filter best frequency and filter bandwidth 10 dB above threshold minimum
RMS	Root-mean squared amplitude

SD	Standard deviation
SPL	Sound-pressure level
STC	Suppression tuning curve

Curriculum vitae

Jan-Eric Grunwald

Bucher Weg 4
82211 Herrsching

Tel.: (08152) 399 609
e-mail: j.grunwald@lmu.de

Date of birth: 20. 08. 1970
Place of birth: Munich

Education

1976–1978	Primary school in Hongkong
1978–1980	Primary school in Herrsching
1980–1989	Secondary school: Luisengymnasium in Munich
1989	Abitur

Community service

11/1989–01/1991	Community service at the Kinderkrankenhaus an der Lachnerstr. in Munich

Vocational school

09/1992–03/1993	Berufsgrundschuljahr Holztechnik at the Kerschensteiner BBZ in Munich

University studies

WS 1991/92–SS1992	Economic geography at the LMU Munich
WS1993–SS1995	Biology at the TU Munich; Vordiplom

WS1995/96–SS1999	Biology at the LMU Munich; Diplom. Major subject: Zoology (systematics). Minor subjects: botany (systematics), ecology, neurobiology. Thesis work for Diplom on 'Untersuchungen zur Echoortung von *Megaderma lyra* im Freiland und deren Einsatz bei der Jagd auf fliegende Beute' at the Department of Neurobiology, LMU Munich
10/1999–2004	Doctoral thesis on 'Echo-acoustic evaluation of real and phantom objects in phyllostomid bats' completed 02/2004 at the Department of Neurobiology, LMU Munich

Work experience abroad

10/1997–03/1998	Research on 'The Indian false vampire and its acoustic environment' in Sri Lanka
03/1999–04/1999	Research on the foraging ecology of *Hipposideros speoris* in Sri Lanka
12/1999	Research on the acoustic properties of bat pollinated flowers at La Selva Biological Station, Costa Rica
10/2001	Workshop on biology and conservation of bats held at the National Wildlife Training Centre, Giritale, Sri Lanka

Publications

Chris R. Pavey, Jan-Eric Grunwald, and Gerhard Neuweiler (2001). Foraging habitat and echolocation behaviour of Schneider's leafnosed bat, *Hipposideros speoris*, in a vegetation mosaic in Sri Lanka. Behav. Ecol. Sociobiol. 50:209-218.

Chris R. Pavey, Chris J. Burwell, Jan-Eric Grunwald, Chris J. Marshall, and Gerhard Neuweiler (2001). Dietary benefits of twilight foraging by the insectivorous bat *Hipposideros speoris*. Biotropica 33(4):670-681.

Jan-Eric Grunwald and Sabine Schmidt. Echolocation on the wing in the Indian false vampire bat, *Megaderma lyra*: a field study. (In prep.).

Jan-Eric Grunwald, Sven Schörnich, and Lutz Wiegrebe (2004). Classification of natural textures in echolocation. Proc. Natl. Acad. Sci. U.S.A. 101(15):5670-5674.

Acknowledgements

I want to express my gratitude to Prof. Gerhard Neuweiler for giving me the opportunity to do this thesis, and his constant support. My special thanks belongs to PD Dr. Lutz Wiegrebe for his persistent, always enthusiastic and motivating guidance and assistance in every aspect throughout the entire progress of this thesis.

I want to thank PD Dr. York Winter for his cooperation and his invaluable technical help concerning the training of *G. soricina*. Prof. Otto von Helversen kindly provided me with animals, and he and his wife Dagmar took me on an interesting, unforgettable field-trip to Costa Rica.

Markus Drexl helped me a lot with the otoacoustic emissions and Prof. Manfred Kössl kindly supplied all the gear.

Many thanks to Sven Schörnich for the training of the *P. discolor* and Maike Schuchmann for her support. I also like to thank Prof. Gerhard Neuweiler, Prof. Benedikt Grothe, Prof. John Casseday, and Prof. Gerd Schuller for helpful comments on the manuscript of chapter 3. This work was supported by the 'Deutsche Forschungsgemeinschaft', Wi1518/6.

Jenny Holzhaider was always great help with catching and marking the bats and with solving animal care problems. Thanks to her and Udo Gröger, Michael Sala, Petra Weißenbacher, and Ariane Santoso my bats were always well looked after in times of my absence.

Without the excellent service of Mr. Zaschka and his workshop staff some set-up ideas would not have been realisable. Erika Young solved all sorts of administration problems.

The bat workshop in Sri Lanka with Gerd, Chris, and Andreas and the annual field trips with Lutz and Andreas were a welcome and interesting diversion from labwork, and great fun.

I enjoyed working at the department very much over the last couple of years, owing to all the members of the neurobiology-group creating an exceptionally nice working atmosphere, especially Alex, Andi, Ariane, Chris, Claudia, Dieter, Elisabeth, Frank, Jenny, Lutz, Maike, Manfred, Markus, Mathias, Petra, Sven, Tom, Rudi, Udo, and

Uwe. Thanks to all of you for discussions, motivation, support, interest, and friendship.

Finally, I want to thank my family, friends, and especially my wife Ingrid for their patience, support, and interest in what I'm doing.

Ehrenwörtliche Erklärung

Die Arbeit wurde von mir selbständig und nur unter Verwendung der angegebenen Hilfsmittel angefertigt.