Erstgutachter: PD Dr. Lutz Wiegrebe

Zweitgutachter: Prof. Dr. Benedikt Grothe

Tag der mündlichen Prüfung: 08. April 2008

DYNAMIC OBJECT ANALYSIS IN ECHO IMAGING

DISSERTATION
AM DEPARTMENT BIOLOGIE II
DER FAKULTÄT FÜR BIOLOGIE

LUDWIG-MAXIMILIAN-UNIVERSITÄT MÜNCHEN

DARIA GENZEL
NOVEMBER 2007
MÜNCHEN

To grizzly, madl, pancake and an evil monkey

INDEX

FIGURE INDEX

ABBREVIATIONS

AM	amplitude modulated
2-AFC	two-alternative, forced-choice
CF	centre frequency
cpd	cycles per degree
2D	two-dimensional
3D	three-dimensional
DSP	digital signal processor
IR	impulse response
MR	modulation rate
SAM	sinusoidal amplitude modulated
Scf	scaling factor
SPL	sound pressure level
VB	virtual bat
VO unit	virtual object unit

ZUSAMMENFASSUNG

Fledermäuse sind in der Lage, dreidimensionale Objekte in vollkommener Dunkelheit zu erkennen und zu unterscheiden, indem sie die Echos ihrer ausgesandten Ultraschalllaute analysieren. Diese Dissertation bearbeitet die dynamische Objektanalyse in der Echoabbildung. Der erste Teil befasst sich mit der passiven und aktiven Echoanalyse von virtuellen Objekten. Es wird untersucht, ob Fledermäuse einer schnellen Aufeinanderfolge von Echos folgen können und wie empfindlich sie sind für spektrale oder zeitliche Veränderungen von passiv präsentierten Signalen oder von aktiv akquirierten Echos. Der zweite Teil behandelt die echo-akustische Analyse von Objektgröße und die Reaktion von Fledermäusen auf unterschiedlich große virtuelle Objekte in ihrer Flugbahn.

Fledermäuse gewinnen detaillierte Information über ein Objekt, indem sie sich um das Objekt herumbewegen und es mit einer Reihe von Echoortungsrufen beschallen. Das dreidimensionale, echoakustische Abbild des Objekts wird durch das Aneinanderreihen relevanter Information gebildet, die von Echos gewonnen wird, die aus verschiedenen Beschallungswinkeln aus zurückkehren. Das Ziel in dem ersten Teil der Dissertation ist es, die Verhaltensstrategien der Fledermaus, *Megaderma lyra*, zu analysieren, mit denen sie die dreidimensionale Form eines komplexen Objekts aus Echosequenzen rekonstruiert.

Frühere Studien weisen darauf hin, dass Fledermäuse zeitliche und spektrale Muster von Echos auswerten, um ihre Umwelt echo-akustisch wahrzunehmen. Wir haben in zwei getrennten Studien mit Hilfe des so genannten ‚two-alternative, forced-choice' Paradigma die Detektionsschwellen der Fledermaus, *M. lyra,* für zeitliche Variationen von Signalen gemessen. In der ersten Studie wurden Detektionsschwellen für spektrale Modulationen gemessen. Hierzu wurde die Modulationstiefe von sich zeitlich verändernden, synthetischen Echoortungsrufsequenzen variiert. Es wurden Modulationsraten von 2 bis 16 Hz verwendet. Die Schwellen lagen ungefähr bei 11 % der Mittenfrequenz. Interessanterweise waren sie relativ unabhängig von der Modulationsrate. Die wirksame Empfindlichkeit und die Modulationsratenunabhängigkeit der erhaltenen Daten zeigen, dass Fledermäuse in der Lage sind, der spektralen Zusammensetzung von Echos zu folgen, die von komplexen Objekten von verschiedenen Winkeln reflektiert werden. Zu beachten ist jedoch, dass unterschiedliche Meinungen existieren, ob man Vergleiche zwischen aktiv- und passiv-akustischen Verarbeitungen ziehen kann.

In der zweiten echo-akustischen Studie haben wir die Korrelation zwischen der Bewegung der Fledermaus im Raum und der Wahrnehmung von Echos, die sich in Abhängigkeit von ihrer Position verändern untersucht. Wir haben eine Echtzeit-Technik zur Generation holographischer, virtueller, echoakustischer Objekte anhand eines omnidirektionalen Phantomziels entwickelt. Wir können zeigen, dass Fledermäuse, die diese holographischen Objekte erkunden, eine Amplitudenmodulation von mindestens 10 dB benötigen, um ein holographisches Objekt, das sich in Korrelation mit der Position der Fledermaus verändert, von einem invarianten Objekt zu unterscheiden.

Die Ergebnisse erlauben erste Einblicke in die Ultraschallempfindlichkeit und die Koordination von Flug- und Ultraschallaktivität echoortender Fledermäuse, während sie komplexe Objekte im Raum erkunden. Wir konnten bestätigen, dass Fledermäuse sowohl passiv als auch aktiv echo-akustisch imstande sind, Veränderungen der spektralen Komposition von aufeinander folgenden Echos zu folgen, die von einem komplexen Objekt reflektiert werden.

Nachdem wir die Mechanismen der dynamischen Objektanalyse untersucht hatten, haben wir in dem zweiten Teil dieser Arbeit die Objektgröße in der Echoanalyse untersucht. Gefragt wurde zu einem, ob Fledermäuse anhand der Echoortung Unterschiede in der Größe gleich geformter Objekte kompensieren können. Als zweites interessierte uns, wie die Objektgröße die Flugbahn frei fliegender Fledermäuse beeinflusst.

In der ersten Studie wurde untersucht, ob die Fledermaus, *Phyllostomus discolor,* Größen-skalierte Versionen (Testobjekte) vorher erlernter Objekte (Standardobjekte) korrekt klassifizieren kann. Dazu mussten die Fledermäuse eine interne Repräsentation für jedes Standardobjekt bilden. Drei Fledermäuse konnten mindestens vier der sechs Testobjekte den entsprechenden Standardobjekten zuordnen. In einem weiteren psychophysikalischen Experiment haben wir getestet, ob Menschen dies auch können. Drei der Versuchspersonen konnten zum Großteil die Größenunterschiede kompensieren. Die Ergebnisse bestätigen, dass das auditorische System der Fledermaus einen festen Mechanismus besitzt, um mit Veränderungen der Größe von echo-akustischen Objekten umzugehen. Weiterhin ist zu vermuten, dass Information, die von aufeinander folgenden Echos aus verschiedenen Beschallungswinkeln gewonnen wird, die Normalisierung von Objekten fördert. Wir schlagen vor, dass die Klangfarbe der Signale von den Menschen und vermutlich von den Fledermäusen für die Klassifizierung verwendet wurde.

In einer zweiten Studie, die aufgrund der Tatsache, dass sie noch nicht abgeschlossen ist, in dieser Dissertation im Anhang zu finden ist, wollen wir die Korrelation zwischen der Flugbahn einer Fledermaus und der Größe eines echo-akustischen Objekts untersuchen. Die Tamanahöhle in Trinidad bietet die einzigartige Möglichkeit, die Phantomzieltechnik auf untrainierte Fledermäuse anzuwenden. Die hohe Anzahl der Tiere in dieser Gegend von mehreren Tausend und die parallele Flugbahn der Tiere aus der Höhle erleichtern die Datenaufnahme. Gleichzeitig dient dieses Projekt als kritischer Test der Phantomzieltechnik, die in unserem und anderen Laboren verwendet wird, da wir die Reaktion untrainierter Fledermäuse, die nur mit reellen Zielen vertraut sind, auf die Präsentation von Echos virtueller Objekte studieren können. Da die Datenaufnahme erst im Dezember stattfinden wird, kann sie in dieser Dissertation nicht mit aufgenommen werden. Ich werde eine Einleitung geben und die Methoden beschreiben, die wir anwenden werden. Wir hoffen, die erhaltenen Daten in einer späteren Veröffentlichung zu präsentieren und zu diskutieren.

SUMMARY

Bats are able to recognize and discriminate three-dimensional objects in complete darkness by analyzing the echoes of their ultrasonic emissions. This thesis addresses the topic of dynamic object analysis in echo imaging. The first part concerns passive and active echo analysis of virtual objects. It investigates, whether bats can track fast successions of echoes and how sensitive they are for spectral or temporal changes in presented signals in passive listening and active echolocation. The second part deals with the echo-acoustic analysis of object size. We are interested, if bats can compensate for size-induced variations of objects and how bats react to differently sized virtual objects in their flight path.

Bats acquire detailed information about an object by moving around the object and scanning it with a series of echolocation calls. Stringing together object related information from echoes obtained from different observation angles will produce a three-dimensional echo-acoustic image of the object. Our goal in the first part of the thesis is to analyze the behavioural strategies with which the bat, *Megaderma lyra*, is able to reconstruct the three-dimensional shape of complex objects by perceptual integration of the information acquired through sequences of echoes.

Previous work suggests that bats rely on both temporal and spectral cues for the echo-acoustic analysis of their environment. In two separate studies we measured in two-alternative, forced-choice procedures the detection thresholds for temporal variations in returning signals in the echolocating bat, *M. lyra*. In the first passive-acoustic study, detection thresholds of spectral modulations were measured by varying the modulation depth of time-variant synthetic echolocation-call sequences for modulation rates ranging from 2 to 16 Hz. The thresholds of about 11 % of the centre frequency were interestingly relatively independent of modulation rate. Acknowledging reservations about direct comparisons of active-acoustic and passive-acoustic auditory processing, the effectual sensitivity and modulation-rate independency of the obtained results indicate that the bats are well capable of tracking changes in the spectral composition of echoes reflected by complex objects from different angles.

In the second active echo-acoustic study we investigated the correlation between the bat's movement in space and the perception of position-variant echoes with an omni-directional phantom target. We realized a real-time technique for the generation of holographic echo-acoustic objects. We show that the bats exploring these holographic objects required an amplitude modulation of at least 10 dB to be able to discriminate a space-variant, holographic object from an invariant object.

The data provide first insight into the sonar sensitivity and into the coordination of flight and sonar activity of echolocating bats exploring complex objects in space and corroborate that bats are well capable of tracking changes in the spectral composition of consecutive echoes reflected by complex objects from different angles.

In the second part of the thesis we wanted to investigate object size in echo imaging. For one, can bats normalize for objects which differ in size, but not in shape, through echolocation alone, and secondly, how does object size affect a bat's flight path?

In the first study, the bat, *Phyllostomus discolor,* was tested whether it can classify scaled versions (test-objects) of previously learned objects (standard-objects). In contrast to an earlier study, we forced the bats to generate an internal representation of each standard-object. Three bats correctly classified at least four of six scaled objects. In a second psychophysical experiment we tested whether humans have the same ability. Three listeners were able to normalize almost all size variations. The results corroborate that the bat's auditory system has dedicated mechanisms like humans to deal with size-induced variations of echo-acoustic objects and that information from successive echoes from different ensonification angles probably facilitate object normalization. We propose that timbre may have been employed as a cue for classification by the human listeners and possibly by the bats.

In a second study, which, due to the fact that the project is still in progress, is included as an appendix in this thesis, we want to investigate the correlation between a bat's flight path and the size of echo-acoustic objects. The Tamana-Cave in Trinidad provides the unique opportunity of testing this correlation on untrained bats, as the number of bats active in this area mounts up to thousands and the bats exhibit parallel flight paths out of the cave. Simultaneously this project acts as a critical test for the phantom target technique used in our and other labs, as we can study how untrained bats familiar only with real targets react to the presentation of echoes of virtual objects. Data acquisition will take place in December and can therefore not be included in this thesis. I will give an introduction and describe the methods we will be using. We hope to present and discuss the obtained results in a later publication.

GENERAL INTRODUCTION

All living beings need to orient in their environment. Vision, hearing, touch and scent are used to generate an internal spatial representation of the surroundings. During evolution different species were forced due to their natural environment or lifestyle to specialise sensory systems. Some fish species discriminate and identify three-dimensional (3D) objects through electrolocation alone (Schwarz and von der Emde, 2000a; Schwarz and von der Emde, 2000b; von der Emde and Schwarz, 2000; von der Emde, 2004; Graff *et al.*, 2004; von der Emde, 2006). Cave dwelling spiders lost their 'eye sight' and developed an advanced vibration sensitivity (Krajick, 2007). A further specialization of cave-dwelling animals is the sophisticated olfactory system found for salamander species allowing the evaluation of their surrounding environment (Uiblein *et al.*, 1992). Long distance travelling animals deprived of distinct landmarks orient with the help of the earth's magnetic field (birds, turtles) (Edwards *et al.*, 1992; Rodda and Phillips, 1992). A further advanced sensory modality in the auditory system is echolocation, through which bats and dolphins orient in their environment with sparse or no visual feedback (Neuweiler, 1990; Kalko *et al.*, 1998; Kalko and Condon, 1998; Neuweiler, 2000; von Helversen and von Helversen, 2003; von Helversen, 2004). Lazzaro Spallanzani, the Bishop of Padua in the late eighteenth century, was the first person to demonstrate that bats do not need visual input to orient in space. By conducting a few experiments he assumed the bats were orienting by hearing, but as the bats did not produce sounds he himself could perceive, he was unable to determine how. Not until 1938 and 1943 did two scientists, Donald Griffin and Sven Dijkgraf, independently discover how the bats were orienting acoustically: by producing short ultrasonic calls through their mouth or nose bats employ the echoes produced by reflective surfaces. This auditory analysis enables them to orient in space, localize objects and measure distances and even evaluate complex shapes and structures for food localization and identification (e.g., fruit, pollen, nectar, insects, mice, fish, frogs) (Simmons *et al.*, 1974; Schmidt, 1988; Kalko *et al.*, 1998; Kalko and Condon, 1998; von Helversen and von Helversen, 1999; Neuweiler, 2000; Weissenbacher and Wiegrebe, 2003; von Helversen and von Helversen, 2003; von Helversen *et al.*, 2003; Korine and Kalko, 2005). Bat species have adapted their echolocation calls to their surrounding and the context in which they are calling. Bats for example living and foraging near vegetation for flying insects predominately call with longer (10 to 100 ms), constant-frequency signals (Neuweiler, 2000). Gleaning bats on the other hand, which take their prey from surfaces, employ very short (0.5-

3 ms), downward frequency-modulated pulses as their call (Neuweiler, 2000). In both cases the bats facilitate the separation of their prey from the background. Many studies concerning echolocation in bats have already investigated different parameters of echo imaging. In general, the intensity, temporal structure, and spectral composition of an echo provide information about the object's size, shape and structure (Schmidt, 1988; Grunwald *et al.*, 2004; Simon *et al.*, 2006). How a bat's environment is internally represented echo-acoustically though is still not fully understood.

When an object is ensonified, it reflects the ensonification signal in an object-specific manner, producing an echo. Every echo contains the acoustic image of the ensonified object. The acoustic image or impulse response (IR) is defined by the reflection characteristics of an object ensonified with an impulse with theoretically infinite duration and amplitude, containing every frequency with the same amplitude (Fig. 1).

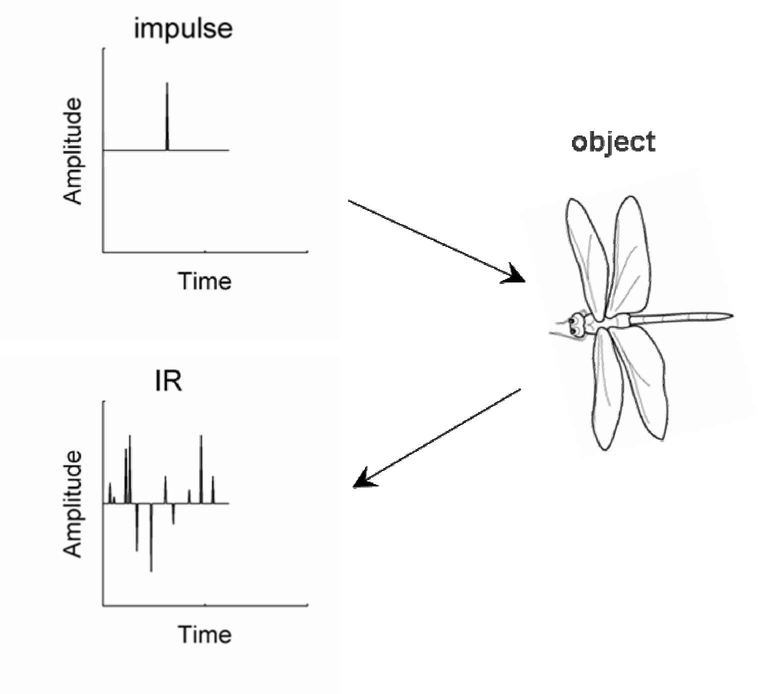


Figure 1: IR generation
When an object is ensonified with an impulse, its surfaces reflect in a characteristic manner, generating the object's impulse response, IR.

When such an impulse is reflected by an object, frequency minima and maxima are generated by destructive and constructive interference of the reflections, generating object specific interference patterns which are visible in the IR's magnitude spectrum. For example the interfering waveforms add up and generate a spectral peak when the difference in depth between these surfaces is 1/2 of the wavelength or a multiple of this relationship. A cancellation or notch within a spectrum will occur when the distance between the high and low surface points on an object is 1/4th of the wavelength or an uneven multiple of this relationship. Consequently temporal and spectral reflection patterns depend on an object's shape, size and material. Schmidt (1992) showed that bats are very sensitive for such spectral features. The bat *Megaderma lyra* can detect a 7 – 9 kHz difference in the position of a spectral notch, corresponding to a depth difference of 0.2 mm. Figure 2 depicts the time signal of an IR and its magnitude spectrum sampled with a frequency of 10 kHz. The different reflections generate frequency cancellations and positive summations which are visible in the spectrum.

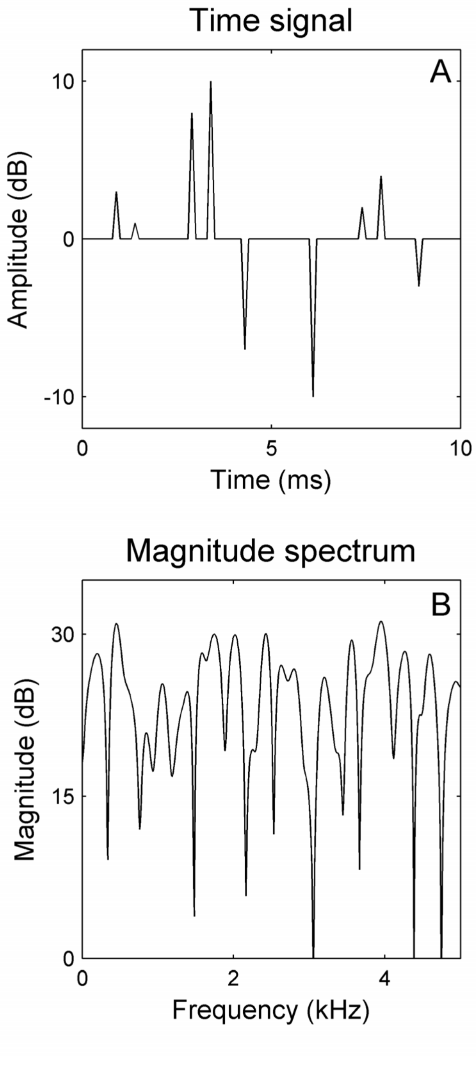


Figure 2: Time signal and magnitude spectrum of an IR
Panel A depicts the time signal of the IR. Plotted is the amplitude in dB as a function of time in s. Panel B shows the corresponding magnitude spectrum of the IR. Plotted is the magnitude in dB against the frequency in kHz. Visible are the frequency notches and peaks generated through the destructive and constructive interferences of the single reflections.

The echo a bat perceives, when ensonifying an object, is the result of the convolution of the bat's echolocation call with the IR of the object (Fig. 3). The convolution of two signals is defined as the integral of the product of the two functions after one is reversed and shifted sample by sample. Note that when two signals are convolved temporally their magnitude spectra are multiplied. Consequently only frequencies existent in the ensonification signal are present in the reflected signal. A bat's echolocation call is not as broadband as an impulse; therefore a bat can only receive spectral information about an object's IR in the frequency range covered by the echolocation call. This means the object's IR is imprinted on the bat's echolocation call.

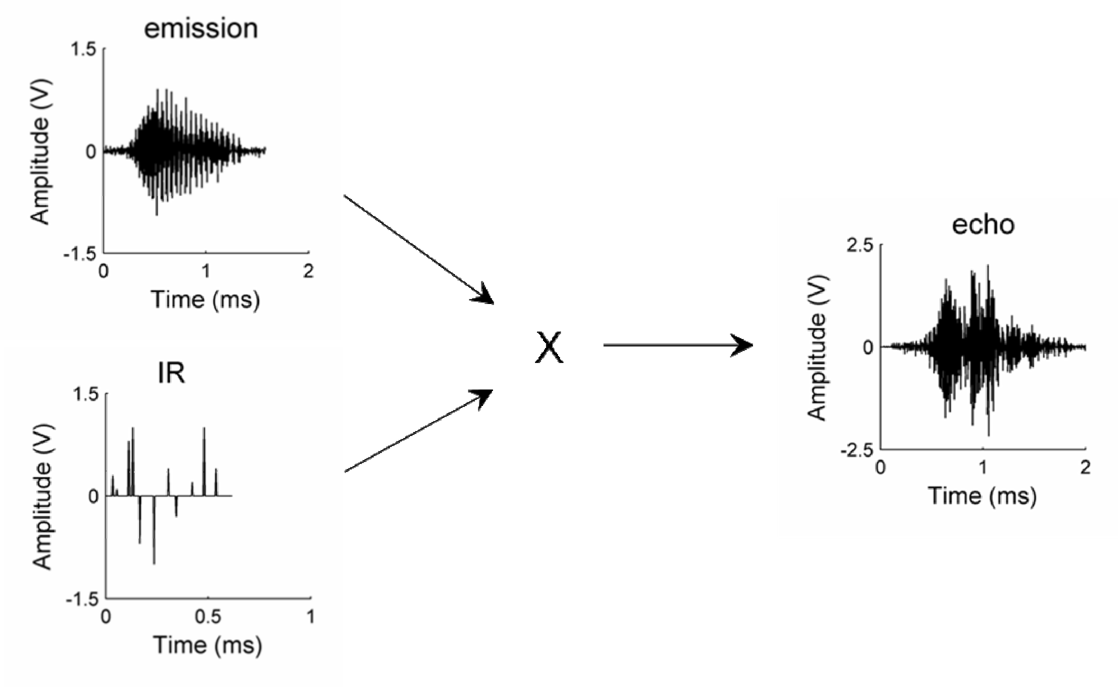


Figure 3: Echo generation
The bat's echolocation call, emission, is convolved with the object's IR, producing the object's echo.

When receiving an echo of an object, the bat still needs to reconstruct the information about the object contained in the echo. As the bat has full information about its echolocation call and an echo is dependent on the IR and this echolocation call, the bat can extract the IR from the reflected echo, by comparing its call with the returning echo (Weissenbacher and Wiegrebe, 2003).

We have elaborated how bats deal with echoes reflected by ensonified objects. How are such topics experimentally approached? For the investigation of how bats evaluate objects echo-acoustically, many studies have presented real objects and had bats evaluate them according to certain behavioural tasks. However, using real objects limits presentable object types (object size, design and quantity) and complicates the method of presentation and especially exchange. Furthermore, not only echo-acoustic, but other sensory cues might be presented. An alternative is the employment of virtual objects. A virtual visual object for example can be generated by displaying an image of an object on a computer monitor. Generating virtual acoustic objects is not quite as simple, but has already been established in echo-acoustic experiments with dolphins (Aubauer *et al.*, 2000) and bats (Schmidt, 1988; Weissenbacher and Wiegrebe, 2003; Grunwald, 2004; Grunwald *et al*., 2004; Firzlaff *et al.*, 2006; Firzlaff *et al*., 2007). It is realized by a so called real-time playback, phantom-target technique. A virtual echo-acoustic object is presented by recording the animal’s echolocation calls with ultrasonic microphones, convolving the calls with IRs of virtual objects and playing back these computer generated echoes to the same animal over ultrasonic speakers. The IRs are constructed according to the echo-acoustic parameters of interest. The animals evaluate these ‘virtual’ echoes corresponding to a certain task. Using virtual objects instead of real objects eliminates confounding visual, olfactory or tactile cues. The technique allows complete control and easy manipulation over the object’s reflection characteristics. A further advantage is a quick exchange of and the potential of presenting numerous virtual objects without mechanical investments. Within this thesis, the technique of virtual-object presentation will be exploited and extended to address two different research areas in bat sonar: in the first part of this thesis, the virtual-object technique will be extended to a presentation of holographic objects whose echo-acoustic appearance changes dependent on the position of the bat relative to the virtual object. In the appendix of the second part of this thesis, a first attempt is made to carry the virtual-object technique into the field and study spontaneous responses of untrained bats to virtual objects.

PART I: DYNAMIC OBJECT ANALYSIS

1 INTRODUCTION

For the generation of a visual representation of an object, our visual system compares dynamic changes in height and width of two-dimensional (2D) visual images between consecutive angles to extract the depth dimension (Bulthoff and Edelman, 1992). To achieve a 3D object representation from echo-acoustic information, however, is more difficult. An echo only reflects an object's distance (depth dimension) to the ensonifier unambiguously. The time delay, when a bat receives the echo of its emitted call, depends on the distance of the ensonified object to the bat. The bat knowing when it had echolocated can derive the time and measure the distance the sound wave needed to travel to the object and back again. The height and width of an object, its shape, are not imaged distinctly. Therefore not only one, but two dimensions need to be extracted through sequential echo analysis. Figure 4 schematically depicts such an example. Shown are two objects with the same depth dimensions and surface area, but with different shapes. When the objects are ensonified once, they reflect unambiguously their depth dimensions in the same manner. Their shapes are not imaged distinctly, complicating object discrimination.

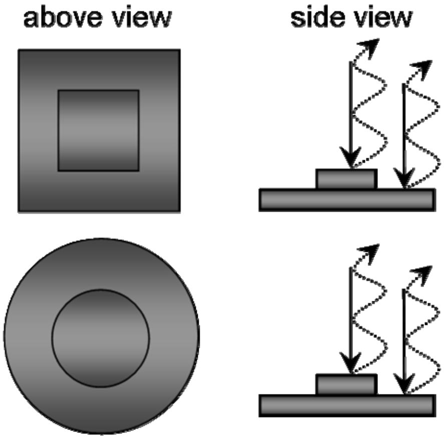


Figure 4: Echo-acoustic object depth
Two objects with the same depth dimensions and surface area, but with different shapes. When ensonified once, they only reflect their depth dimension unambiguously, not their shape.

It has been shown in many behavioural studies that bats can echo-acoustically discriminate 2D and 3D forms and objects (Kalko and Condon, 1998; von Helversen and von Helversen, 1999; von Helversen and von Helversen, 2003; von Helversen *et al.*, 2003; von Helversen, 2004; Korine and Kalko, 2005; Simon *et al.*, 2006; Stich and Winter, 2006; Firzlaff *et al.*, 2007). Dolphins are able to recognize simple 3D geometric shapes, i.e. cubes or tetrahedrons, perceived from unknown angles (Helweg *et al.*, 1996a; Helweg *et al.*, 1996b) and even generate a modality (echo-acoustical or visual) independent internal representation of the objects (Harley *et al.*, 1996; Harley *et al.*, 2003). This is not an easy accomplishment as objects feature differences in shape dependent on the observation angle. So, how is this dealt with echo-acoustically?

Visually, object analysis and recognition is aided by translating an object's shape from different view points (Logothetis and Sheinberg, 1996). Humans and monkeys recognize objects viewed from unknown angles, with mental translation and rotation of the object's shape playing an important role (Murray *et al.*, 1993; Logothetis and Sheinberg, 1996; Hamm and McMullen, 1998; Willems and Wagemans, 2001; Lloyd-Jones and Luckhurst, 2002a). As in the visual system, an echo is dependent on the observation angle. Subsequently for obtaining echoes from different view points and for broadening the echo-acoustic image, ensonification with multiple successive calls is required. Exactly this behaviour can be seen for echolocating animals and echolocation-trained humans: they produce a series of short pulses, when echo-acoustically analysing an object (Fish *et al.*, 1976; Au and Martin, 1989; Helweg *et al*., 1996a; Helweg *et al*., 1996b; Schaub and Schnitzler, 2007b). Bats typically scan an object with a series of echolocation calls with varying ensonification angles by moving around it. This produces amplitude and frequency modulations in the echoes' spectral envelopes depending on the angle from which the object is ensonified. Stich and Winter (2006) described this echo-acoustic perceptual experience as resembling a visual experience caused by so-called physical or metallic colours: due to spectral interferences, these colours change their appearance with the angle of illumination and observation. It is plausible that bats extract object height and width by sequentially analyzing the dynamically changing echoes. By stringing together one-dimensional echoes a 3D acoustic image might be generated, which contains complete information about an object's shape. Consequently the construction of an internal representation of a 3D object might be possible (Fig. 5).

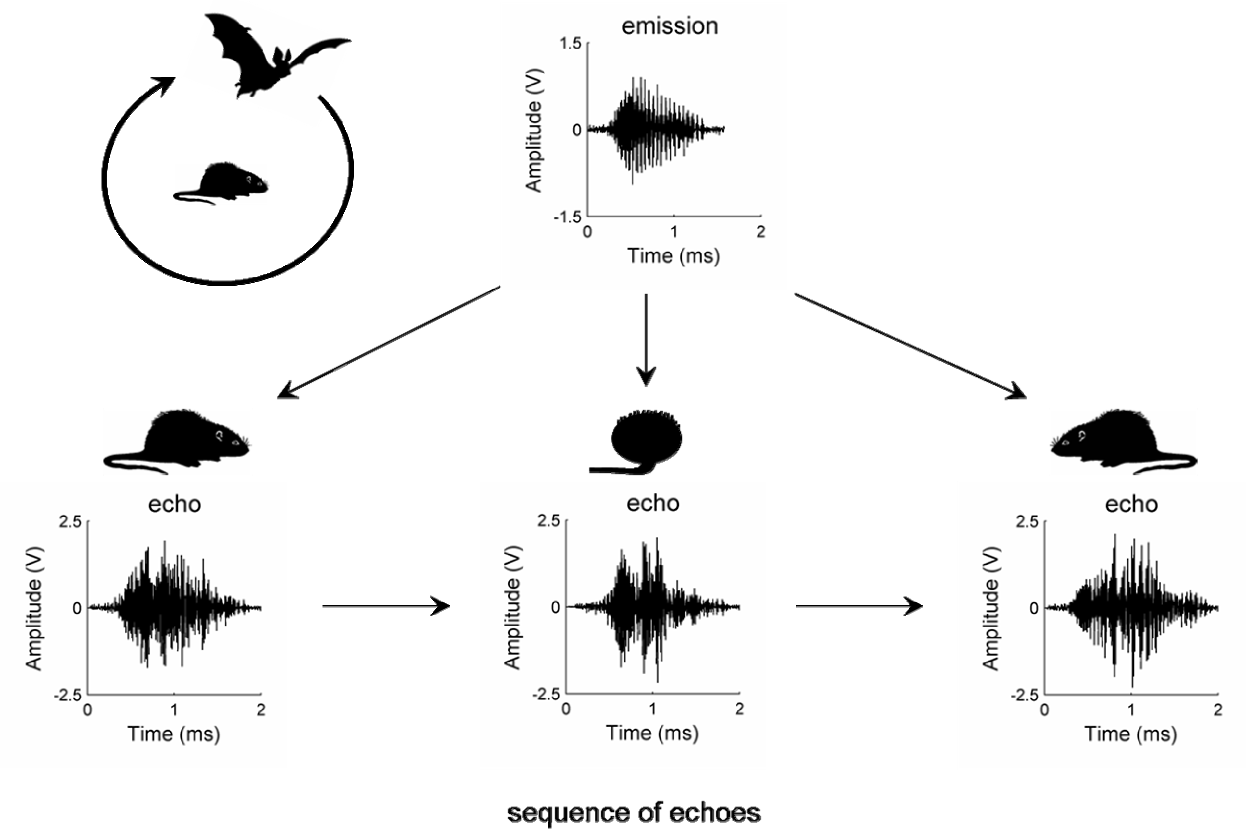


Figure 5: Echo sequence
By ensonifying an object, e.g. a mouse, from different angles and stringing the reflected echoes together, a 3D acoustic image of the object may be generated.

The significance of sequential data acquisition and analysis of dynamically changing echoes has been shown in earlier studies about echo imaging for dolphins (Helweg *et al.*, 1996a; Helweg *et al.*, 1996b; Delong *et al.*, 2006). Further studies showed that human divers were able to discriminate metal plates with different structures and thicknesses with the aid of ultrasonic transmitters and receivers (Fish *et al.*, 1976; Au and Martin, 1989). Both dolphins and divers moved around the objects enabling an ensonification from different acoustic perspectives. This strategy allowed the deduction of object information from dynamic changes of the resulting echoes. The question addressed here is, how fast can bats track temporal changes in sequential echoes and how sensitive are they for these variations? This will be the topic of the following studies.

Experimental animal

Behavioural studies investigating echolocation with real or virtual targets have already been undertaken for many different bat species. An experimental animal which has proven itself of value in many behavioural studies is the tropical, gleaning bat, *Megaderma lyra* (Geoffroy

1810), the great false vampire bat (e.g. (Schmidt, 1988; Schmidt, 1992; Sedlmeier, 1992; Wiegrebe and Schmidt, 1996; Preisler and Schmidt, 1998; Krumbholz and Schmidt, 1999; Weissenbacher and Wiegrebe, 2003)). An image of this bat and its typical echolocation call can be seen in Figure 6.

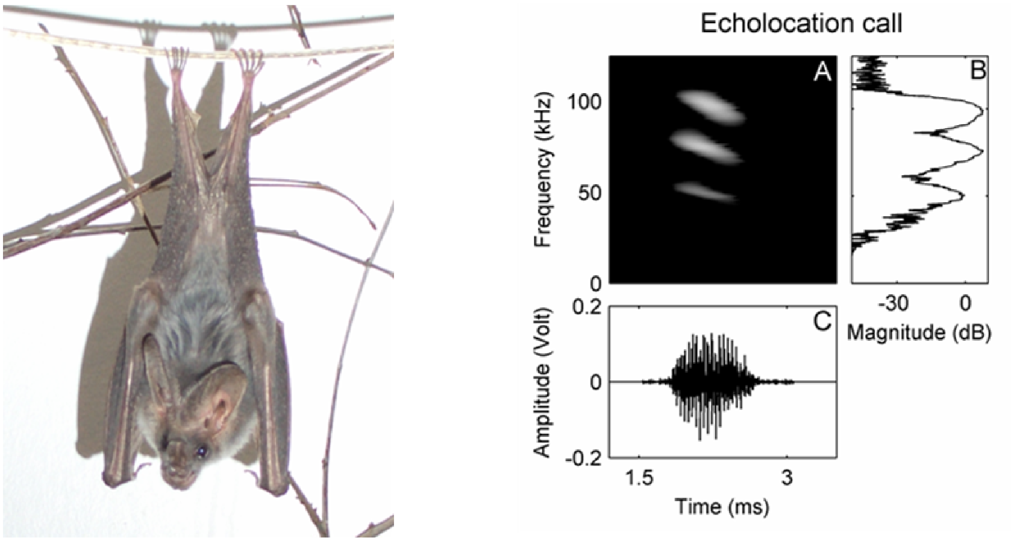


Figure 6: Megaderma lyra
This bat's echolocation call is illustrated on the left hand. The echolocation call is typically about 1 ms long and features multi-harmonic, downward modulated frequency sweeps with three dominant harmonics. Panel A displays the spectrogram of the call. Plotted is the frequency in kHz as a function of the time signal in ms. Panel B shows the magnitude spectrum, plotted is again the frequency in kHz as a function of the magnitude in dB. Panel C shows the time signal of the call; plotted is the amplitude in Volt as a function of the time signal in ms.

M. lyra can be found on the Asian continent, from Sri Lanka to Northern Malaysia (Lekagul and McNeely, 1988). They belong to the *Megadermatidae* family. These bats are carnivorous and pick their prey up from the ground (Audet *et al.*, 1991). For prey detection, *M. lyra* relies on prey generated rustling noises (Neuweiler, 1990), featuring very low passive hearing thresholds in the range of 1- 130 kHz (Schmidt *et al.*, 1983). For the further analysis of its prey and to facilitate separation of prey objects from background, *M. lyra* employs short (0.4 – 1.2 ms), multi-harmonic (up to 6), downward modulated frequency sweeps as their echolocation calls, all sweeps in the frequency range between 120 to 18 kHz (Schmidt *et al.*, 2000). Objects of interest are then typically ensonified from different aspects in flight. Furthermore, *M. lyra* exhibits prey specific changes of its echolocation call's spectral content, showing quick echo-acoustic adaptations from prey detection to capture (Leippert *et al.*, 2002). Therefore in terms of auditory processing, the bat's behaviour requires tracking changes of spectral interference patterns over time. These criterion and the bat's flight ability to manoeuvre in close spaces, a prerequisite for the experimental setup, justifies the use of this bat species as an experimental animal in our behavioural paradigms.

Experimental goal

Our goal is to analyze the behavioural strategies with which the bat, *M. lyra*, is able to reconstruct the 3D shape of complex objects by perceptual integration of the information acquired through sequences of echoes. We measured in two behavioural studies, during passive listening and active echolocation, *M. lyra's* capability for detection of fast changes in consecutive signals and sensitivity for changes in the spectral and temporal content of echoes, as they would be reflected by an object, when it is scanned echo-acoustically.

1.1 Time-variant spectral peak and notch detection in echolocation-call sequences in bats

This chapter has been accepted in the Journal of Experimental Biology (October the 17th, 2007) under the same title by Daria Genzel and Lutz Wiegrebe.
A few supplementary figures have been added.

In a real-target paradigm von Helversen (2004) showed that the bat *Glossophaga soricina* was able to discriminate two hollow forms, a hemisphere and a paraboloid with the same diameter and depth. The extracted IRs of each object generated a spectral interference pattern with frequency peaks and notches which varied systematically with ensonification angle. This variation was highly specific to the object. It is conceivable that the bats solved this task by evaluating the changes in the peak and notch patterns in correlation with their movement around the objects (Moss and Surlykke, 2001; von Helversen and von Helversen, 2003). The speed with which the bats auditory system can follow time-variant spectral interference patterns is unknown. It was therefore interesting to investigate the effect of time-variant spectral changes on sequential echo-acoustic object analysis.

We decided to design the psychophysical study as a classical, passive acoustic task, which is nevertheless adequate as an experimental approach for questions concerning echolocation. In a two-alternative, forced-choice (2-AFC) experiment we investigated the auditory sensitivity of the bat, *M. lyra*, to changes in the position of spectral peaks and notches across a sequence of synthesized echolocation calls. These call sequences were generated to mimic the echoes as they would return from a 3D object whose reflection characteristics change with ensonification angle. Thereby, we wanted to analyze the importance of these spectral features for echo-acoustic object recognition. Unlike in previous studies, the changes of the peak and notch centre frequencies were time-variant, varying sinusoidally with a certain modulation frequency. The bats' detection threshold for variations in the spectral envelope was measured by presenting a synthesized echolocation-call sequence filtered with time-variant filters.

1.1.1 Material and Methods

1.1.1.1 Experiment 1: time-variant peak detection

Animals

Four adult *M. lyra*, one male and three females, took part in the training. One of the female bats died within the data-acquisition period, thus most of the data presented is from the remaining two females and one male. They were kept in a 12 m² room with free access to water. During training periods consisting of five consecutive days the bats were fed with mealworms as a reward. Apart from the training rewards the animals were fed one mouse per week.

Experimental setup

All experiments were performed in an echo attenuated chamber (3.5 m x 2.2 m x 2.2 m) with a wall foam coating. The setup consisted of a starting perch on one side of the room, ensuring a precise positioning of the bat, and two ultrasonic speakers, one in the left and one in the right hemi field. The two ultrasonic speakers were placed at the same distance and angle in each hemi field to the bat's starting position: the distance from the speakers to the bat's head was 1.2 m; the angle between the speakers and the bat's head was 90°. A feeding dish was placed below each speaker. The setup is depicted in Figure 7.

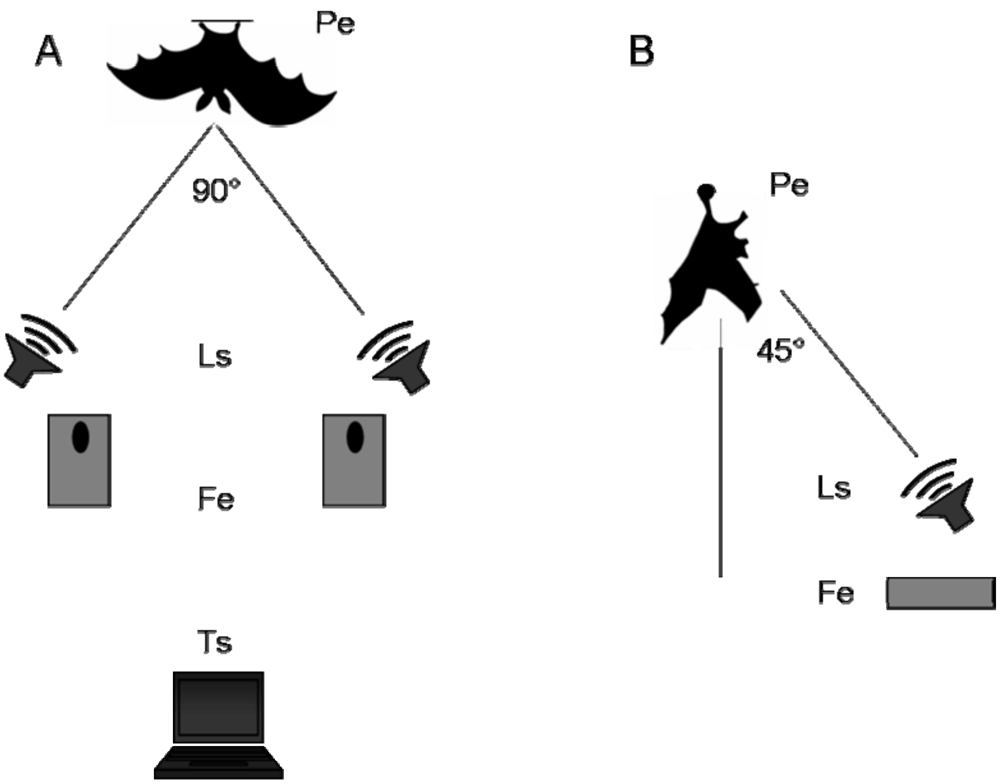


Figure 7: Illustration of the setup for the peak and notch detection
Panel A depicts the setup from the front, panel B from the side. Feeder (Fe), loudspeaker (Ls), starting perch (Pe) and touch screen (Ts) are depicted. The angle relations are indicated.

Stimuli

The source signal was a sequence of 17 synthesized echolocation calls. Each call was a multi-harmonic frequency sweep with a duration of 1.5 ms. The fundamental frequency swept from 23 to 19 kHz. Five harmonics were generated with attenuations of 30, 10, 5, 0, and 5 dB for harmonics one to five, respectively. The call was windowed with a raised-cosine window with a 0.2 ms rise time, 1.1 ms steady state and 0.2 ms decay time.

For the time-invariant echolocation-call sequence, a band-pass filter with a reference centre frequency (CF) of 60 kHz and a bandwidth of ± 10% of the CF was applied to all 17 calls in the sequence. For the generation of the time-variant echolocation-call sequences, the CF of the band-pass (peak) filter was sinusoidally modulated around the reference CF along a log-frequency axis. The filter was designed as a finite-impulse-response, band-pass filter of order 62. The detection threshold for variations of spectral peaks was measured by varying the modulation depth (in % of the CF) of the time-variant filtered echolocation-call sequence. To measure the bats' sensitivity to the CF modulation, we presented modulation depths of 100, 52, 40, 30, 24, 18, 14, 11, and 9 % of the CF. A modulation depth of 100% defined a frequency range of ± one octave around the CF and produced filter CFs between 30 and 120 kHz. The modulation rate of the CF modulation was 2, 4, 8, or 16 Hz. One echolocation-call sequence always contained two modulation periods. In consequence, the overall duration of the echolocation-call sequence and the temporal separation between the echolocation calls in the sequence decreased with increasing modulation rate. For a modulation rate of 2 Hz, the echolocation-call sequence was 1 s long and the temporal separation between the echolocation calls was about 61 ms; for a modulation rate of 16 Hz, the echolocation-call sequence was 125 ms long and the temporal separation between the echolocation calls was about 6 ms. Spectrograms of an unfiltered call and time-variant and time-invariant echolocation-call sequences are shown in Figure 8 A and B (page 25). These echolocation-call sequences simulate a bat moving twice around an abstract virtual acoustic object and ensonifying it from eight different angles. Different flight speeds are represented by modulation rates between 2 and 16 Hz. While this range of modulation rates is low compared to many auditory studies on the perception and encoding of time-variant signals, the rates are certainly high enough to include the speed of spectral or temporal modulations encountered by a bat when it moves around an object ensonifying it from different angles.

To preclude the bats' use of overall presentation level or absolute-frequency cues (Krumbholz and Schmidt, 1999), the presentation level was roved by ± 6 dB and the reference CF was roved by ± 10 % over trials. Moreover, the phase of the sinusoidal frequency modulation was roved over trials.

The echolocation-call sequences were computer generated (Matlab 5.3, Mathworks, Natick, MA) and digital-analog converted (RX6, sampling rate 260 kHz, Tucker Davis Technologies, Gainesville, FI). The echolocation-call sequences were amplified (Rotel RB 976 MK II, Worthing, England) and presented over the ultrasonic loudspeakers (Matsushita EAS 10 TH 800D, Osaka, Japan) at a level of 65 dB SPL (preceding the roving level). The frequency response of all setup components including speakers was flat within +/- 5 dB between 5 and 100 kHz. The echolocation-call sequences were heterodyned by two real-time digital signal processors, DSPs, (RP2, sampling rate 200 kHz, Tucker Davis Technologies, Gainesville, FI) allowing the experimenter to follow the presentation acoustically via headphones.

Procedure

In a 2-AFC experiment, psychometric functions were obtained for variations in the spectral content of synthesized echolocation calls. The time-variant filtered echolocation-call sequence was played back by one speaker and the time-invariant filtered echolocation-call sequence by the other. While hanging on the perch, the bat perceived the echolocation-call sequences alternately from each speaker. There was a fixed inter-stimulus interval of 500 ms between successive echolocation-call sequence presentations. The echolocation-call sequence presentations stopped as soon as the bat left the perch. The bat had to therefore make its decision at the starting position. On the other side of the room, opposite to the perch, the experimenter was seated, controlling the procedure and the data storage via touch screen (WES TS, ELT121C-7SWA-1, Nidderau-Heldenbergen, Germany). The experimental program was written in Matlab 5.3.

The bats were trained to fly to the speaker from where they perceived the time-variant filtered echolocation-call sequence. For the initial training, the modulation depth was set to 40 % of the CF. As a control, one bat was trained to fly to the time-invariant filtered echolocation-call sequence. Whether the time-variant echolocation-call sequence was presented at the left or right position was determined by a pseudo-random sequence, with the same echolocation-call sequence never occurring more than three times in a row at the same position. As soon as the bats were able to solve this task with a stable performance of > 85 % correct choices over several days, the modulation depth of the time-variant filtered echolocation-call sequence was decreased and increased. 30 trials for each modulation depth were collected. The performance was calculated as decisions for the side of the time-variant echolocation-call sequence in percent correct as a function of the modulation depth. The significance level was set to 75 % correct choices. After evaluating the threshold modulation depth for a specific modulation rate, the bats were trained to the next modulation rate and the corresponding threshold was measured.

1.1.1.2 Experiment 2: time-variant notch detection

The animals, the experimental setup, and the procedure were the same as in Experiment 1.

Stimuli

The source signals were the same synthetic call sequences as in Experiment 1. The filter was designed as a finite-impulse-response, band-stop (notch) filter of order 64, a reference CF of 60 kHz and a bandwidth of ± 10 % of the CF. For the time-invariant echolocation-call sequence, this filter was applied to all 17 calls in the echolocation-call sequence.

For the generation of the time-variant echolocation-call sequences, the CF of the band-stop filter was sinusoidally modulated around the reference CF along a log-frequency axis. As in Experiment 1, the detection threshold for variations of spectral notches was measured by varying the modulation depth (in % of the CF) of the time-variant filtered echolocation-call sequence. The stimuli are illustrated in Figure 8 C and D.

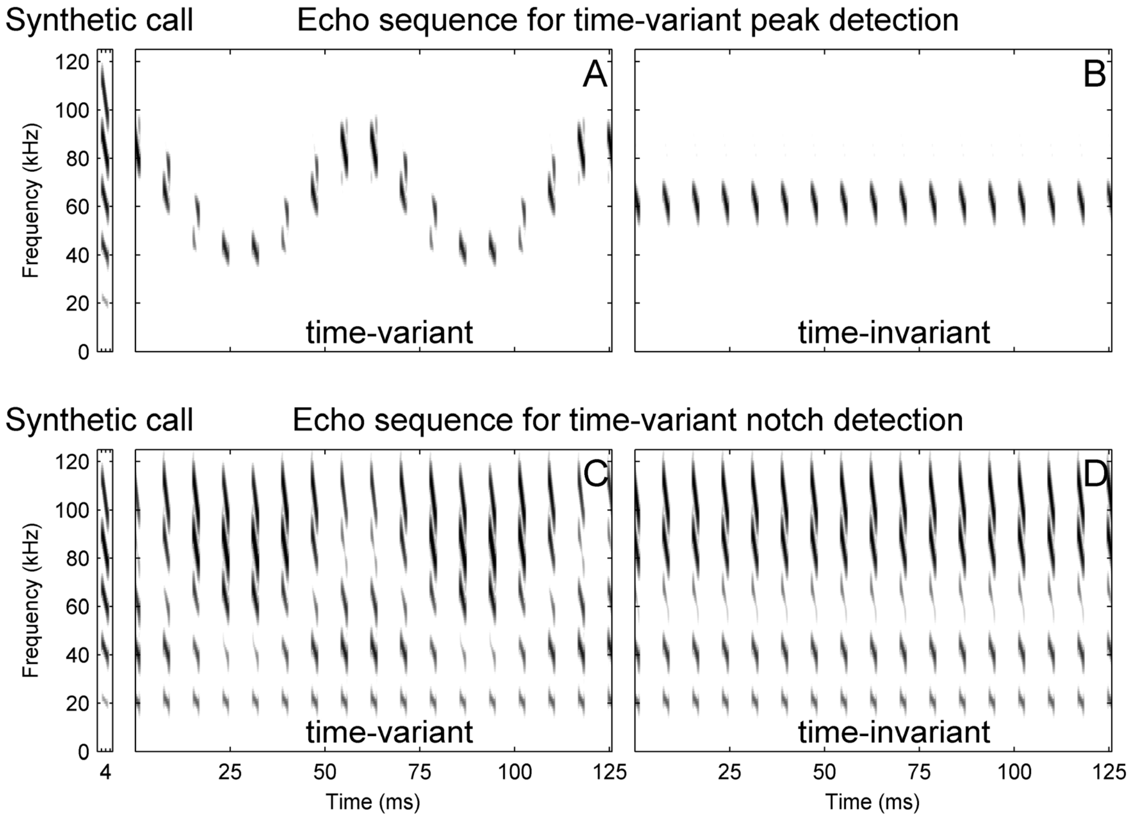


Figure 8: Spectrograms of the echolocation-call sequences with the time-variant peaks (A, B) and notches (C, D)

To the left of A and C is the spectrogram of the unfiltered FM sweep consisting of F0 (23 to 19 kHz) and 4 harmonics with a duration of 1.5 ms. Panels A and B show the echolocation-call sequences to be discriminated for the time-variant peak detection. Panels C and D show the echolocation-call sequences to be discriminated for the time-variant notch detection. All echolocation-call sequences have a modulation depth of 40 % of the CF and a modulation rate of 16 Hz which results in an overall duration of 125 ms.

Figure 9 depicts modulation depth examples of 100, 40, 11 and 5 % of the CF (panel A – D, respectively) for the time-variant peak detection task. Plotted is the frequency in kHz as a function of time in ms. The applied modulation rate is 16 Hz, resulting in an overall duration of 125 ms per echolocation-call sequence.

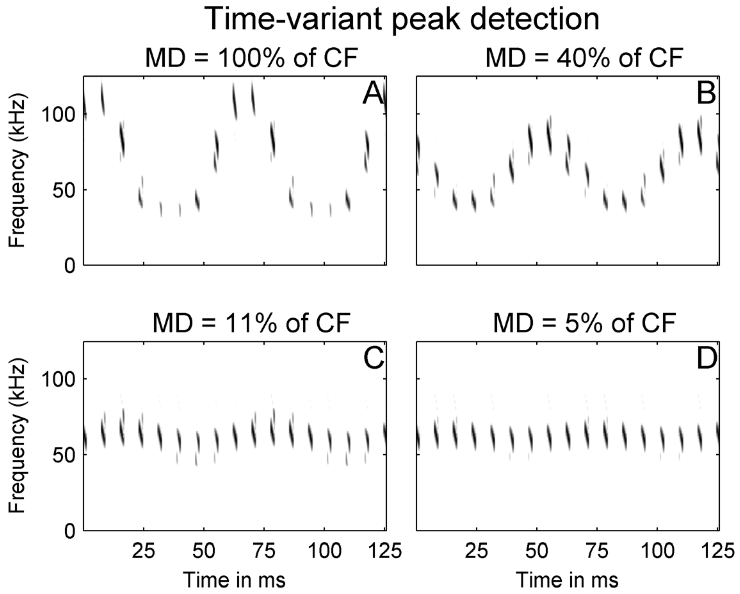


Figure 9: Spectrograms of the echolocation-call sequences with the time-variant peaks with different modulation depths

Panels A, B, C and D show the echolocation-call sequences with a modulation depth of 100, 40, 11 and 9 % of the CF, respectively. Applied is a modulation rate of 16 Hz which results in an overall duration of 125 ms.

1.1.2 Results

1.1.2.1 Experiment 1: time-variant peak detection

The bats' performance in the time-variant peak detection task was very similar and thus, the threshold was calculated as a mean value for individuals. Psychometric functions for the modulation rates of 2, 4, 8, and 16 Hz for all bats are shown in the four panels of Figure 10.

At a modulation rate of 2 Hz, the four bats were able to detect a frequency modulation depth of 10.9 % of the CF, on average (Fig. 10 A). At a modulation rate of 4 Hz, the four bats could detect a modulation depth of 10.9 % of the CF, on average (Fig. 10 B). At a modulation rate of 8 and 16 Hz, the three remaining bats could detect a modulation depth of 11.2 and 11 % of the CF, respectively (Fig. 10 C, D). For a CF of 60 kHz 11 % of the CF corresponds to a frequency bandwidth of 13 kHz.

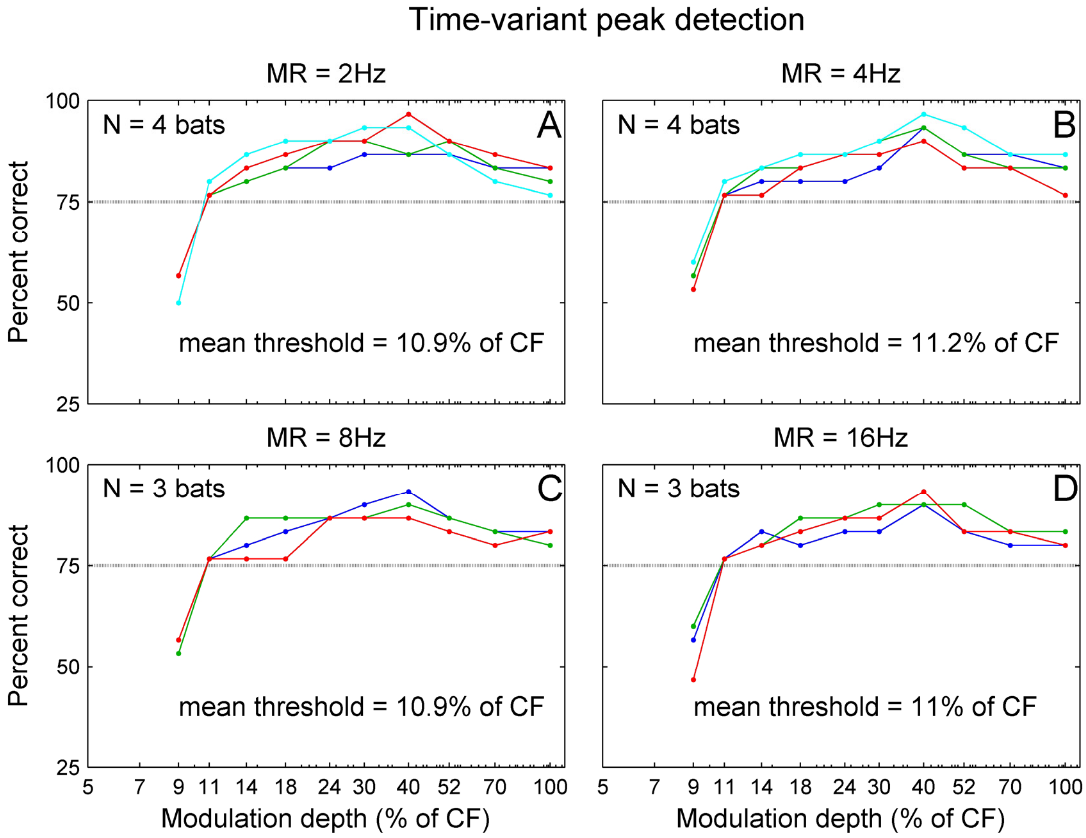


Figure 10: Mean psychometric functions for all bats for the time-variant peak detection for the tested modulation rates (MR)
Plotted are percent correct decisions as a function of the modulation depths, the significant threshold level was set to 75 %. Each panel depicts the psychometrics curves for all animals as a function of modulation depth. The threshold is calculated as a mean value for all bats. Panels A to D show data for MRs of 2, 4, 8, and 16 Hz, respectively.

Surprisingly all animals readily transferred the discrimination task from one modulation rate to the next, although not only the modulation rate but also the overall echolocation-call sequence duration changed. This is shown in an exemplary training curve for one bat in Figure 11.

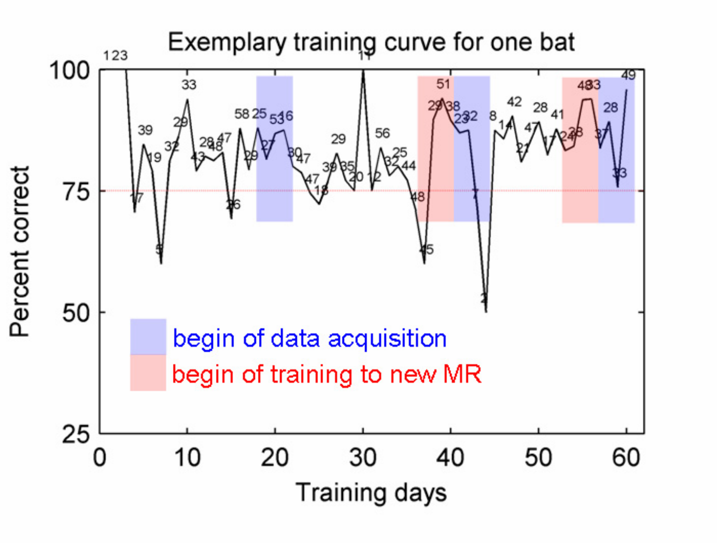


Figure 11: Exemplary training curve for one bat during part of the data acquisition
Plotted are percent correct decisions as a function of the number of training days, the significant threshold level was set to 75 %. The numbers above the curve indicate the number of trials per day. The blue fields highlight the beginning of threshold data acquisition, the red fields emphasize when a new MR was initiated. Note that with initiation of a new MR the bat's performance is above the 75 % level, denoting that the bat readily transferred the discrimination task from one modulation rate to the next.

The slight decrease of the bats' performance when the modulation depth was increased from 40 % of the CF to 100 % can be attributed to the bats being trained on a modulation depth of 40 % and they seemed slightly irritated by the high modulation depths, allowing the assumption that these signals may have sounded different than the initially trained condition (40 %).

However, the modulation depth used for the training did not affect the threshold value: as a control, one female bat was retrained at a modulation rate of 4 Hz to a modulation depth of 100 % after the data acquisition for all other experimental conditions was finished. Figure 12 depicts the psychometric function obtained after initial training to a 40 % modulation depth and the function obtained after initial training to a 100 % modulation depth. Although the above-threshold performance differs somewhat between these data acquisitions, near-threshold performance is very similar ensuring the validity of the obtained threshold values for all animals.

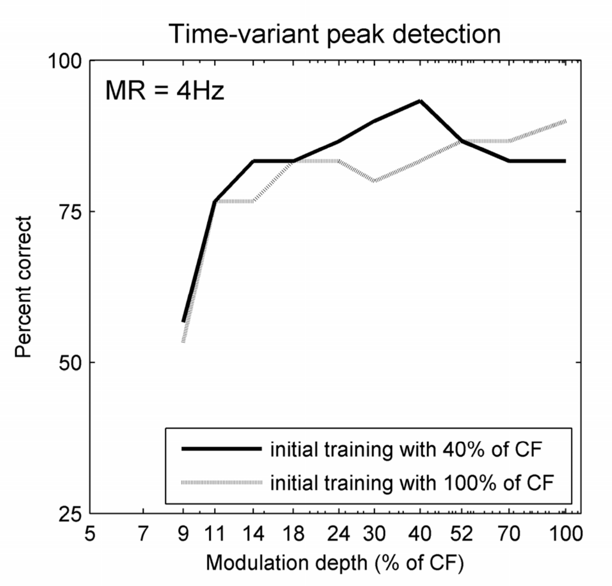


Figure 12: Psychometric functions for the peak detection for one bat (MR = 4 Hz)
The solid line represents the performance with an initial training at a modulation depth of 40 % of CF. The dotted line displays the same bat's performance for the same experimental condition, only with an initial training at a modulation depth of 100 % of CF.

In general we were able to observe that the spectral peaks have to vary by about 11 % of the CF to be discriminated from the time-invariant peaks. Furthermore, this threshold seems to be independent of the modulation rate in the tested range.

1.1.2.2 Experiment 2: time-variant notch detection

Psychometric functions for the detection of a time-variant spectral notch are shown in Figure 13 in the same format as for Experiment 1. Again, the threshold was calculated as a mean value for three bats. For a modulation rate of 2 Hz, the threshold was 11.3 % of the CF. For modulation rates of 4, 8, and 16 Hz, the thresholds were 11.4, 10.9, and 11.8 % of the CF, respectively. The general performance was slightly worse, but all in all did not differ from that of the first experiment.

As it was the case for the time-variant peak detection, the transfer to a new modulation rate did not require retraining and thresholds were largely independent of modulation rate.

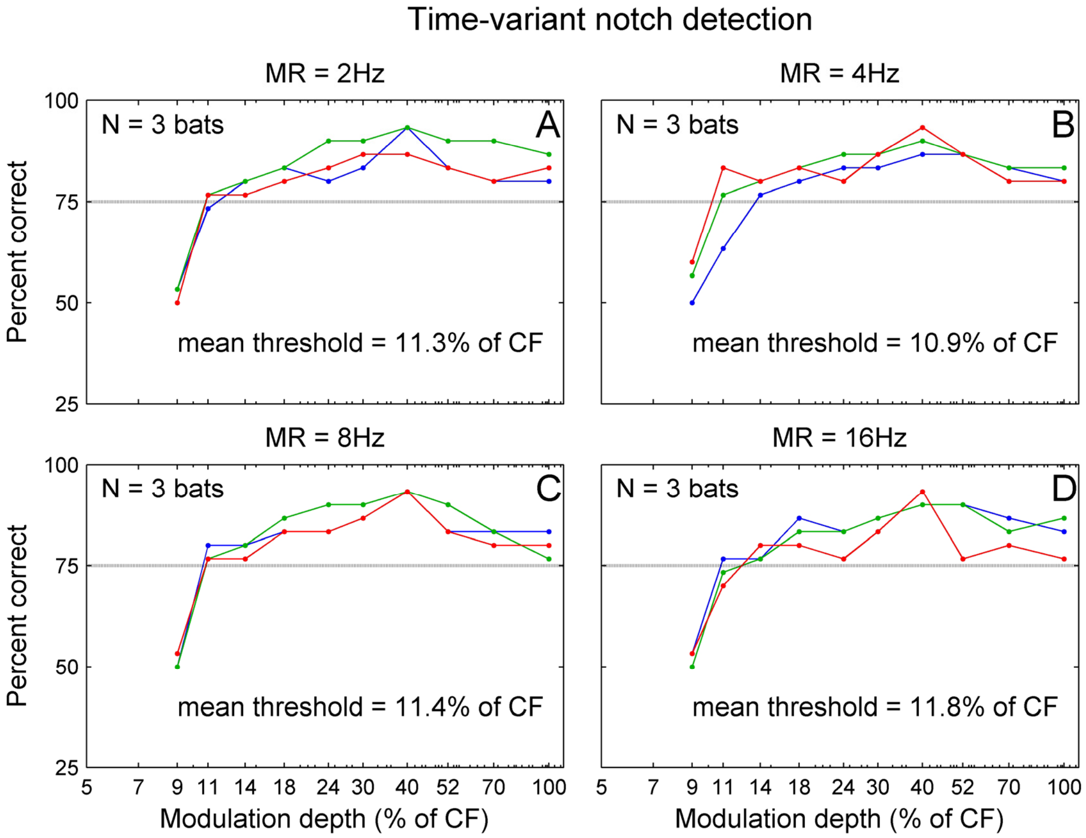


Figure 13: Mean psychometric functions for all bats for the time-variant notch detection for the tested MRs Plotted are percent correct decisions as a function of the modulation depths, the significant threshold level was set to 75 %. Each panel depicts the psychometrics curves for all animals as a function of modulation depth. The threshold is calculated as a mean value for all bats. Panels A to D show data for MRs of 2, 4, 8, and 16 Hz, respectively.

1.1.3 Discussion

The current psychoacoustical study was designed to investigate the auditory sensitivity of the bat, *M. lyra,* to time-variant spectral peaks and notches imposed on sequences of synthesized echolocation calls. We found that *M. lyra* is well able to discriminate a time-variant echolocation-call sequence from a time-invariant echolocation-call sequence. The detection threshold for the time-variant echolocation-call sequence in the tested range lies at 11 % of the CF independent of whether a spectral peak or notch was modulated. Furthermore the detection threshold seems to be unaffected by the modulation rate across the tested range from 2 to 16 Hz. In the following, we will discuss the obtained data in regard to these three points: the obtained threshold values in general, the apparent modulation rate independency and the threshold values of the peak and notch signals in comparison.

Threshold value

The detection threshold for changes in the spectral domain lies at about 11 % of the CF, independent of whether the CF of a peak- or notch filter was varied. This threshold is comparable to frequency modulations (7 – 21 %) occurring in the active-acoustic object-discrimination experiment of Simon et. al (2006) based on the assumption that the bats exploited spectral-notch changes in that experiment. In an earlier two-front, phantom-target study, Schmidt (1992) obtained similar threshold values for *M. lyra* of 6 – 13 % for spectral-notch centre frequency changes. Note, however, that again, these thresholds were obtained in an active-acoustic paradigm where the bats evaluated the spectral content of echoes of their own calls. The frequency differences, on the other hand, were static within a trial. The current data obtained in a passive-acoustic paradigm with time-variant filtering corroborate these findings.

M. lyra is a gleaning bat; it rarely actively hunts for flying insects and therefore does not have to detect wing flutter. Nevertheless, the current thresholds, obtained in a passive-acoustic paradigm, are comparable to values obtained for other bat species in active-acoustic paradigms (Mogdans and Schnitzler, 1990; Bartsch and Schmidt, 1993; Esser and Kiefer, 1996). Typically, *M. lyra* catches its prey from the ground, and first detects it by listening to prey-generated rustling noises. By first relying on passive rustling noises and then moving in to evaluate and catch possible prey it might not need to analyse fine modulation differences.

Lyzenga and Carlyon (1999) measured in humans the detection of just noticeable differences for a sinusoidal modulation of the CF of a synthetic formant with a fixed fundamental. The thresholds they obtained were larger, by a factor of 2, than thresholds for the discrimination of (static) formant frequencies (Lyzenga and Horst, 1997). This seems to hold for starlings as well, which also show 2-3 times larger threshold depths for low modulation frequencies than for just noticeable frequency differences between pure tones (Langemann, 1991). This difference might explain our slightly increased thresholds in comparison to other studies, where frequency differences were static (Schmidt, 1992; Simon *et al*., 2006).

Modulation-rate independency

The current detection thresholds for spectral changes in the envelope were apparently independent of modulation rate. Temporal processing therefore does not seem to be a critical factor for the discrimination task. In contrary, the current data are consistent with an analysis of place cues along a tonotopic frequency axis. Moore and Sek (1995; 1996) and Sek and Moore (1995; 2000) claim that the detection threshold for frequency modulations of low-frequency pure tones for humans is modulation-rate dependent, as the low-frequency tones are encoded by phase-locked, temporal cues. In mammals, phase locking is limited to frequencies below about 5 kHz (Rose *et al.*, 1968; Palmer and Russell, 1986; Oertel, 1999). Higher frequencies are encoded exclusively by place cues. In humans, spectral place cues provide worse frequency accuracy than phase-locked, temporal cues (Moore and Sek, 1995).

In the bat, each of the presented ultrasonic calls can only be encoded by auditory place cues. Thus, no phase-locked, temporal information concerning the current frequency composition of the call is available. Consequently, the frequency acuity is limited. The phase-locking, low-pass filter does not impair the peripheral auditory representation of the modulation rate as all tested modulation rates were considerably lower than the phase-locking filter cut-off frequency; meaning that the fluctuations of the spectral envelope can easily be encoded through phase-locking. In summary, the current data are consistent with the hypothesis that the spectral peaks and notches are encoded via place cues in the peripheral auditory system and that the bats' central auditory system is fast enough to follow the changes of these place cues over time for the range of modulation frequencies tested.

Several electrophysiological studies on temporal encoding in the mammalian auditory cortex have revealed a low-pass characteristics of synchronous cortical discharges with a cut-off frequency around 20 Hz (Schulze and Langner, 1997; Lu *et al.*, 2001; Liang *et al.*, 2002). In an electrophysiological study with rising and falling FM stimuli, responses of neurons in the primary auditory cortex of the gerbil were recorded (Ohl *et al.*, 2000). Across the range of tested modulation frequencies (1 to 24 Hz), the neurons' responses did not vary with modulation rates. This again fits with the modulation rate independent thresholds we obtained in this study for modulation rates lying in a similar range.

In a psychophysical study in the bat *Tadarida brasiliensis*, Bartsch and Schmidt (1993) tested perceptual sensitivity to sinusoidal frequency modulation at much higher rates (10-2000 Hz, CF = 40 kHz). They found that threshold modulation depths deteriorated with increasing modulation rate. As we only tested modulation rates between 2 and 16 Hz, we are not able to comment on whether the bats may have showed increased thresholds for even higher modulation rates. Note that our stimulus trains were intended to simulate a stationary complex object ensonified by a bat surrounding the object twice and ensonifying it from eight different angles. In this context, modulation rates above 16 Hz would have represented a highly unnatural situation, 16 Hz already representing an extreme.

Comparison of peak- and notch thresholds

In the current study, the bats were equally sensitive for time-variant peaks and notches. In the following, we discuss this finding in regard to the question, whether the bat *M. lyra* extracts pitch information from its harmonically structured echolocation calls or whether echo analysis is based on the auditory processing of spectral place profiles.

Sedlmeier (1992) was able to show that *M. lyra* categorizes ultrasonic pure tones and complex harmonic structures with attenuated or missing fundamentals almost identically. This was interpreted as that the bat perceives the 'missing fundamental', enabling it to integrate different acoustic qualities to a complex perception. Sedlmeier (1992) suggested that the bats perceive a pitch corresponding to the fundamental frequency of a sound and categorize sounds with different spectral features according to their pitches. Preisler and Schmidt (1998) further investigated this topic, and examined whether the bat *M. lyra* evaluates complex harmonic structures according to their pitch or on the basis of overall spectral similarity. They observed that the tested bats differed in which of the strategies they applied to solve the task. Krumbholz and Schmidt (1999) showed that *M. lyra* spontaneously classified test signals according to their broadband spectral similarity, using trained signals as spectral templates, not pitch cues.

As the slope of the filters used in the current study was rather steep (filter-order 62), an echolocation call filtered with a band-pass (peak) filter centred at 60 kHz will cause a pitch percept corresponding to 60 kHz. Due to the time-variant filtering, the bats would hear a time-variant pitch. When the notch filters are applied, on the other hand, the bats always hear all harmonics except the one filtered out by the notch filter. Thus, the pitch would always correspond to the calls' fundamental frequency of about 21 kHz. As pitch extraction is rather insensitive to amplitude modulations of higher harmonics, this percept would not be strongly affected by the time-variant filtering. In summary, if the bats had applied a pitch-based analysis, one would expect a better performance with the band-pass (peak) filters than with the band-stop (notch) filters. The finding that this is not the case corroborates the conclusions of Krumbholz and Schmidt (1999) that in most cases the bats recruit a spectral profile- rather than a pitch analysis for echo imaging.

In summary, the current data show that the bat *M. lyra* can discriminate time-variant from time invariant echolocation-call sequences with good accuracy. In the range of modulation rates tested (2 to 16 Hz), the discrimination performance was constant. The fact that sensitivity to time-variant spectral peaks and notches was similar argues in favour of a spectral profile analysis rather than a pitch-based analysis of the harmonic echolocation-call sequences. With the reservation of comparing passive-acoustical and active-acoustical auditory processing, the current data indicate that the bats' central auditory system is fast enough to track the changes in the spectral composition of returning echoes when the bat ensonifies an object while flying around it.

1.1.4 Conclusion

M. lyra is able to follow changes in the spectral composition of echolocation-call sequences in a passive acoustic task. The question now arises, whether the bat is able to detect changes in the composition of returning echoes, while exploring and actively ensonifying an object in space. This leads us to the following study, where we analyzed the correlation between the bat's movement in space and the perception of position-variant echoes. We wanted to investigate whether *M. lyra* is able to acquire information about an object through perceptual integration of echoes and its sensitivity to temporal changes of reflected echoes in an active echo-acoustic task.

1.2 Evaluation of holographic echo-acoustic images by echolocating bats

We have developed an experimental setup which can generate holographic echo-acoustic objects, i.e., the echo-acoustic image perceived by the bats depends on their echo-acoustic azimuthal observation angle. Previously the bat had received a static acoustic image independent from where the bat had echolocated. With this technical extension the real-time presentation of space-variant acoustic images is possible. Each echolocation call is convolved with a certain IR, which is determined by the bat's position in space. The echo structure is therefore dependent on the echolocation call and the ensonification angle.

Our goal is to investigate *Megaderma lyra's* sensitivity for dynamic changes of the spectral and temporal content of IRs, while it is scanning virtual objects echo-acoustically. To analyze this echolocation performance we created two different virtual 3D objects defined exclusively in terms of their IRs. One object, made up of 36 different IRs, changes in dependence of the observation angle and the other, consisting of 36 identical IRs, is independent from the bat's observation angle. In a 2-AFC paradigm bats are trained to fly around two units each consisting of a microphone, loudspeaker and feeder. A high-speed video camera above these units determines the bat's position in space. The bat's echolocation calls are recorded with the microphone, filtered with one of the 36 IRs (dependent on the bat's position) and played back in real-time over the loudspeaker. While flying around both units the bat experiences echoes, as it would experience when moving around two real objects. The units therefore substitute real objects. The IR of one virtual object changes in dependence of the bat's position in space, the IR of the second virtual object is independent of the ensonification angle.

1.2.1 Material and Methods

Animals

Five adult *M. lyra*, two males and three females, took part in the training. They were kept in a 12 m² room with free access to water. During training periods of five consecutive days the bats were fed with mealworms as a reward. Apart from the training rewards the animals were fed one mouse and two crickets per week.

Experimental setup

All experiments were performed in an echo attenuated chamber (3.5 m x 2.2 m x 2.2 m) with a wall foam coating. The setup consisted of a starting perch on one side of the room, ensuring a precise positioning of the bat, two virtual-object units (VO unit), depicted in Figure 14 A, and a high-speed video camera (Basler A602f 1394 camera, Basler, Ahrensburg, Germany) above the VO units. The dimension of the camera's observation field was 1x2 m. The VO units were placed in the centre of each hemi-field of the observation field and consisted of an ultrasonic ¼ inch microphone (B&K 4135 microphone, B&K 2807 power supply, Brüel & Kjer Instruments, Nearum, Denmark), an omni-directional speaker (Elac 4piplus.2, Elac Electroacustic GmbH, Kiel, Germany), a landing platform and a feeding dish (Fig. 14 B). The bats were rewarded by opening and closing an iris diaphragm (Linos Photonics GmbH & Co.KG, Göttingen, Germany) above the feeding dish. The chamber was illuminated by a 40 watt red light bulb positioned above the camera.

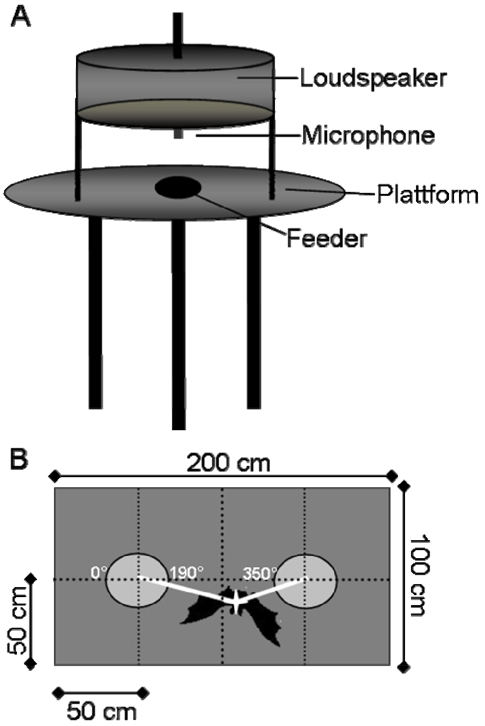


Figure 14: Illustration of an omni-directional virtual-object (VO unit) (A) and the spatial arrangement of the two VO units in the experimental setup (B)
The VO unit consists of a down-facing ultrasonic microphone centred in an omni-directional ultrasonic speaker. This combination is mounted above a circular landing platform with a centred feeder. The feeder could be opened and closed with a motorized iris diaphragm. The setup consisting of two such VO units (light-grey circles in B) was monitored by a high-speed video camera with an observation area as outlined by the grey area in B. The angle of the bat relative to the VO units is calculated by the image-processing software in real-time every 10 ms.

Stimuli

We created two different holographic 3D objects defined exclusively in terms of their IRs. All IRs were band-pass filtered impulses with cut-off frequencies of 20 and 120 kHz and had a filter order of 164. One object, made up of 36 different IRs (space-variant), changed in dependence of the observation angle and the other, consisting of 36 identical IRs (space-invariant), was independent from the bat's observation angle. The space-variant IRs were generated by generating a block of six IRs. The first four had an attenuation of 70 dB, the last two of 30 dB, thereby having a level difference of 40 dB. This spatial reflection pattern was repeated six times creating a cylindrical virtual object, which had 12 strong and 24 weak reflections. The space-invariant IRs all had the same attenuation level of 35 dB (level difference of 0 dB), leading to a virtual cylinder. The periodicity was defined as spatial frequency cpd (cycles per degree) with a cpd of 1/60. For threshold measurements, seven interpolation levels were calculated. This led to nine spatial amplitude modulated (AM) depths of 40 – 0 dB in 5 dB steps. Both virtual objects are shown in Figure 15 (panels A and D). Two interpolations (20 and 10 dB difference) are depicted in panels B and C.

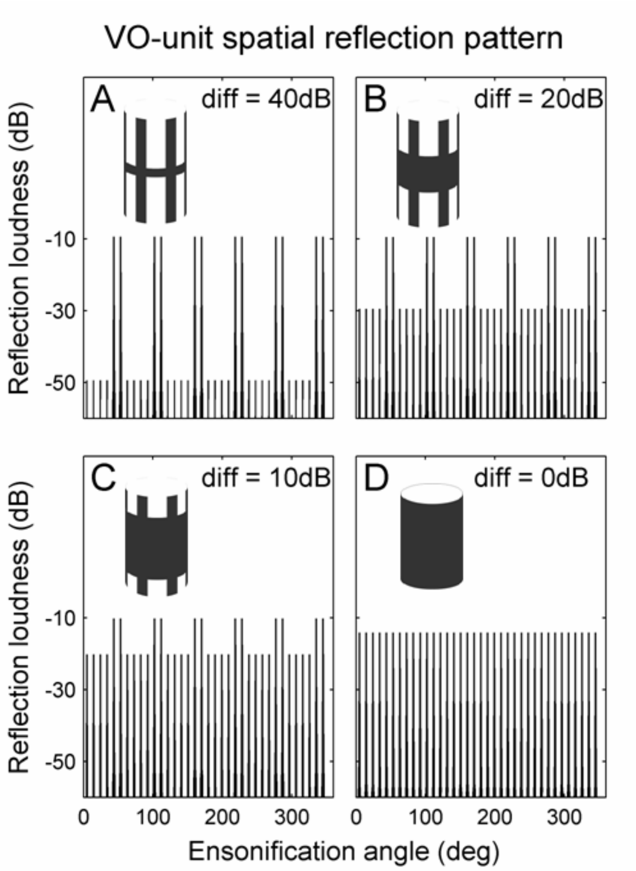


Figure 15: Illustration of the virtual objects presented via the VO units
The loudness of the reflections of the rewarded virtual object (A to C) changed periodically with ensonification angle while the reflection loudness of the unrewarded virtual objects is independent of ensonification angle (D). The insets of the panels depict a visual illustration of how such objects might look like. Note, however, that due to the roving-level paradigm, the overall size of the object and consequently the absolute target strength could not be used to solve the task. Moreover, the random rotations of the virtual objects around the vertical axis prevented the bats from homing onto singular reflections.

To preclude the bats' use of overall presentation level or absolute-frequency cues (Krumbholz and Schmidt, 1999), the presentation level was roved by ± 6 dB over trials. To eliminate absolute position cues and to force the bats to fly around the VO units, the space-variant virtual object was randomly rotated around the vertical axis, meaning the peak positions were rotated, but the overall periodicity was kept intact.

Recorded echolocation calls were band-pass filtered (20 – 100 kHz), amplified by 80 dB (PM 5171, Philips, Hamburg, Germany), analog-digital converted and convolved with an IR by a real-time processor (RX6, sampling rate 260 kHz, Tucker Davis Technologies, Gainesville, FI). The space-variant and –invariant signals were computer generated (Matlab 5.3, Mathworks, Natick, MA). The bat's position in space was determined by a second computer running a customized program version of EyeSeeCam (EyeSeeCam, T. Dera and E. Schneider, Technical University Munich, Germany). This customized program converts the camera signals (100 frames per second) by subtracting each frame from a 'quiet' background acquired when the program is initiated, calculating the centroid of the resulting darkest pixel group and converting this centroid into degrees. As each virtual object is defined by 36 IRs, each VO unit is divided into 36 10° angles, thus encoding the bat's position to 36 possibilities around each VO unit. This degree calculation was sent bit-encoded to the real-time processor via an I/O-input. The outgoing digital-analog converted echoes were amplified (RB 976 MK II, Rotel, Worthing, England) and presented over the ultrasonic loudspeakers. A delay of 113 samples is generated by the analog-digital and digital-analog conversions. The echoes were additionally heterodyned by two further real-time processors (RP2, sampling rate 200 kHz, Tucker Davis Technologies, Gainesville, FI) allowing the experimenter to follow the presentation acoustically via headphones. The technical setup is schematically depicted in Figure 16.

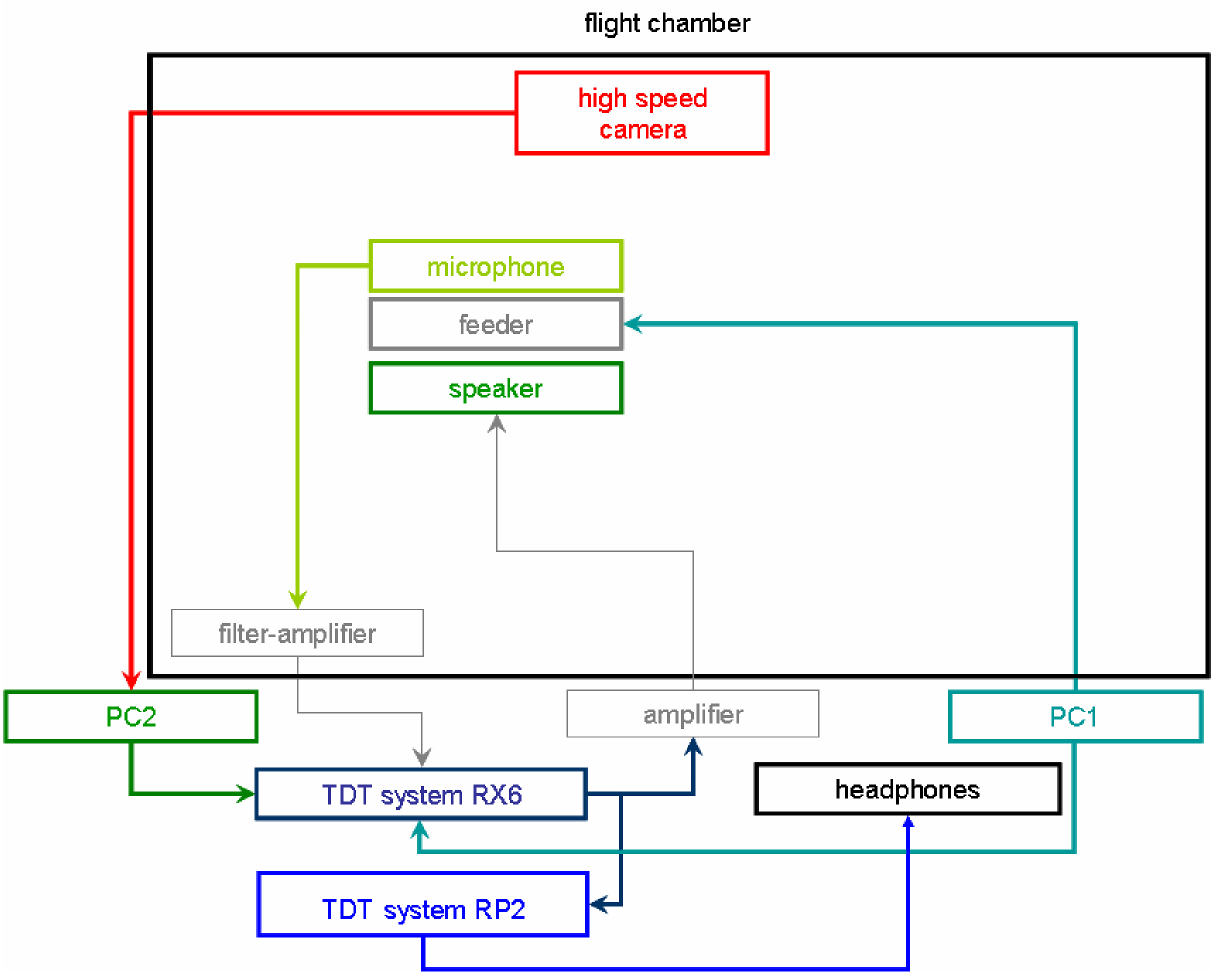


Figure 16: Technical setup of the VO units
Schematic for one VO-unit: the high speed video camera sends its input to the second computer (PC2), which calculates the current IR-number and sends this to the real-time processor RX6. The RX6 filters the band-pass filtered (20 to 100 kHz) and amplified (80 dB) microphone-input with the current IR and sends this via an amplifier to the speaker, which plays back the echo. The RX6 output is also connected to the second real-time processor RP2, which heterodynes the signal and sends it to headphones for the experimenter as an acoustic control. The first computer (PC1) controls the two DSPs through Matlab5.3. The feeder can be opened and closed via PC1. The camera, microphone, feeder, speaker and filter-amplifier are positioned inside the flight chamber; all other units are outside of the chamber.

Procedure

In a 2-AFC experiment, psychometric functions were obtained for space-variant echoes. The bats were trained to fly around the VO units. Each echolocation call was filtered with the current IR. The camera tracking the bat's position in space determined the IR for filtering. The space-variant echoes were played back by one VO unit and the space-invariant echoes by the other. Playback did not start until the bat left the perch. On the other side of the room, opposite to the perch, the experimenter was seated, controlling the procedure and the data storage via touch screen (WES

TS, ELT121C-7SWA-1, Nidderau-Heldenbergen, Germany). The experimental program was written in Matlab 5.3.

Three conditions were determined. In the first condition, 'ACT', an echo a bat received was a convolution of the bat's own echolocation call with the current IR. Playback was only triggered by call detection. In the second condition, 'TRIG', the bats had to echolocate as in the ACT condition for playback triggering, but an echo was generated by convolving a synthetic echolocation call with the current IR. In the third condition, 'PAS', playback was not triggered by the bats' echolocation calls. An echo consisted of a convolution of a synthetic echolocation call with the current IR as in the TRIG condition, but the echoes were presented with a fixed interpulse interval of 100 ms.

For conditions TRIG and PAS, the synthetic echolocation call was a multi-harmonic frequency sweep with a duration of 1.5 ms. The fundamental frequency swept from 23 to 19 kHz. Five harmonics were generated with attenuations of 30, 10, 5, 0, and 5 dB. The chirp was windowed with a raised-cosine window with a 0.2 ms rise time, 1.1 ms steady state and 0.2 ms decay time.

For all three conditions the bats were trained to fly to the VO unit from where they perceived the space-variant echoes. The bat made a choice by landing on the platform. For correct choices the feeding dish was opened and the bat was rewarded with mealworms which were always present in both VO units. Whether the space-variant echoes were presented at the left or right VO unit was determined by a pseudo-random sequence, with the same signal never occurring more than three times in a row at the same VO unit. As soon as the bats were able to solve this task with a stable performance of > 80 % correct choices over several days, the absolute level difference between peaks and baseline was decreased. 30 trials for nine absolute level differences were collected. The performance was calculated as decisions for the side of the space-variant echoes in percent correct as a function of absolute level difference. The significance level was set to 75 % correct choices. After evaluating the threshold absolute level for a specific condition, the bats were trained to the next condition and the corresponding threshold was measured.

Echo-acoustic Calibration

To ensure, that the VO unit's physical echo did not mask the echo of the virtual object, we additionally ensonified one VO unit with a virtual bat (VB). This enabled a visualization of the echoes the bats really perceived. The VB consisted of an ultrasonic ¼ inch microphone (B&K 4135 microphone, B&K 2610 preamplifier, Brüel & Kjer Instruments, Nearum, Denmark) a two-channel spectrum analyzer (SR780, Stanford Research Systems, Sunnyvale, CA), an integrated amplifier (RB 960 BX, Rotel, Worthing, England), and an ultrasonic speaker (EAS 10 TH 800D, Matsushita, Osaka, Japan). The VB, positioned 20 cm away from the VO unit, was tracked by the high-speed camera as implemented above. The analyzer sent out over the VB's speaker 16 synthetic *M. lyra* echolocation calls per second. The echolocation calls were dealt with in the same manner as the real bats' echolocation calls. They were recorded, filtered and played back by the VO unit. The VB's microphone recorded the returning echoes, which the analyzer cross-correlated with the echolocation calls. The cross-correlation produces the IR of the VO unit, weighted with the typical frequency content of an echolocation call. The IR of the VO unit reveals both the physical reflections generated by the VO unit and the virtual-object reflections which depend on the position of the VB unit relative to the VO unit. Figure 17 illustrates the echoes extracted from the VB's recording. In all panels the first signal responds to the crosstalk from the VB's chirp, the second signal, labeled with one asterisk, corresponds to the real echo of the VO-unit and the last signal marked by two asterisks is the echo of the virtual object. Panel A and B depict a weak and strong reflection of the space-variant object with a spatial AM depth of 40 dB, respectively. Panel C shows the stronger reflection of the space-variant object with a spatial AM depth of 25 dB and panel D the reflection of the space-invariant object (spatial AM depth 0 dB). As can be seen in these panels an interfering masking effect of the virtual-object reflections by the physical reflection of the VO unit were not present as both reflections are separated in time, and the virtual object produces in terms of loudness the more dominant reflection.

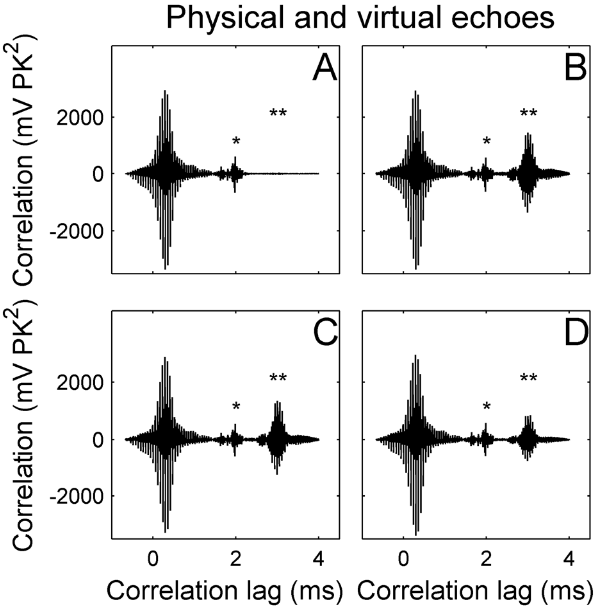


Figure 17: Physical and virtual echoes reflected by a VO unit obtained in the calibration
In all panels the first signal responds to the crosstalk from the VB's chirp, the second signal, labelled with one asterisk, corresponds to the real echo of the VO-unit and the last signal marked by two asterisks is the echo of the virtual object. Panel A depicts a weak and panel B and C depict a strong reflection of the space-variant object with a spatial AM depths of 40 and 25 dB, respectively. Panel D shows the reflection of the space-invariant object (spatial AM depth 0 dB).

1.2.2 Results

Psychoacoustic results

Five bats were successfully trained to discriminate the space-variant echoes from invariant echoes as they would be radiated by a plain cylinder with different sized surfaces in at least one of the three experimental conditions. In the ACT condition, the bats had to evaluate the real-time generated echoes of their own calls while flying around the VO units. Psychometric functions obtained from four bats in the ACT condition are shown in Figure 18. At a spatial AM depth of 40 dB, all animals perform significantly correct. At this high AM depth, the loudness of the reflections of the rewarded virtual object fluctuated strongly when the bat flew around it. Note that due to the roving-level paradigm, the rewarded virtual object was not necessarily the one producing the louder reflections. With decreasing modulation depth, the spatial distribution of echoes from the two VO units becomes more similar and consequently, performance decreases. Threshold spatial AM depths, derived from a sigmoidal fit to the psychometric functions, are given in the legend of Figure 18. Whereas one of the bats can detect a spatial AM depth of only 9 dB and one bat a depth of 12 dB, the other two bats need a spatial AM depth of 21 to 23 dB to discriminate the space-variant echoes from space-invariant echoes.

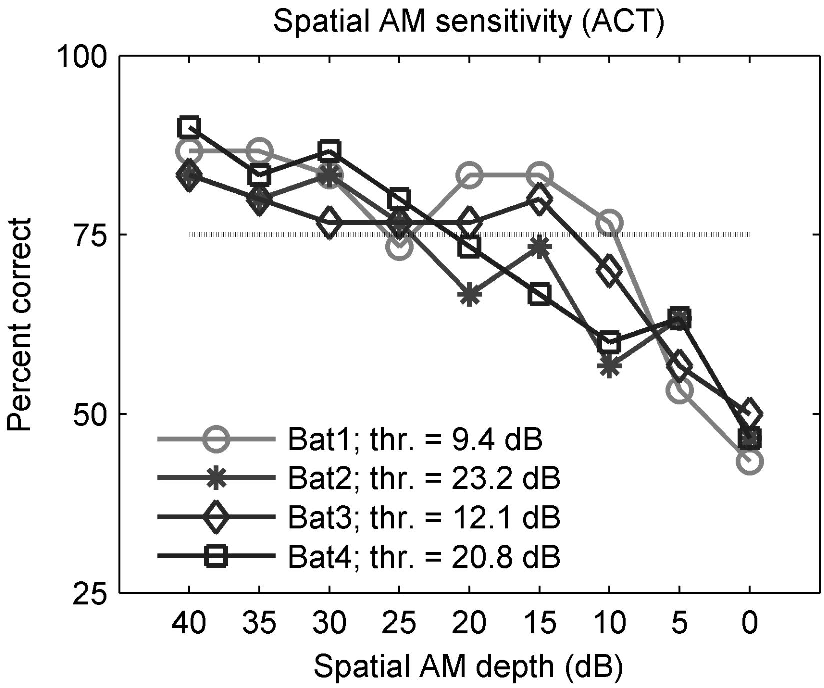


Figure 18: Psychometric functions for the detection of space-variant virtual objects in the ACT condition
Here the echoes generated by the VO units corresponded to exact copies of the sonar emissions. While all bats could reliably discriminate between a virtual object with a spatial AM depth of 40 dB, none of the bats could reliable detect the virtual object with a spatial AM depth of 5 dB. Individual spatial AM depth thresholds are given in the legend.

Four of the five bats learned to discriminate echoes from the VO units in condition TRIG. Here, the animals' vocalizations triggered the presentation of a synthetic echo generated by the convolution of a synthetic call with the same space-variant IRs. As in the ACT condition, the loudness of the IR depended on the position of bats relative to each VO unit. Psychometric functions for the bats' performance in the TRIG condition are shown in Figure 19 in the same format as Figure 18. Spatial AM depth thresholds amount to 10 and 14 dB for Bats 1 and 3, respectively, and amount to 18 and 23 dB for Bats 2 and 4, respectively.

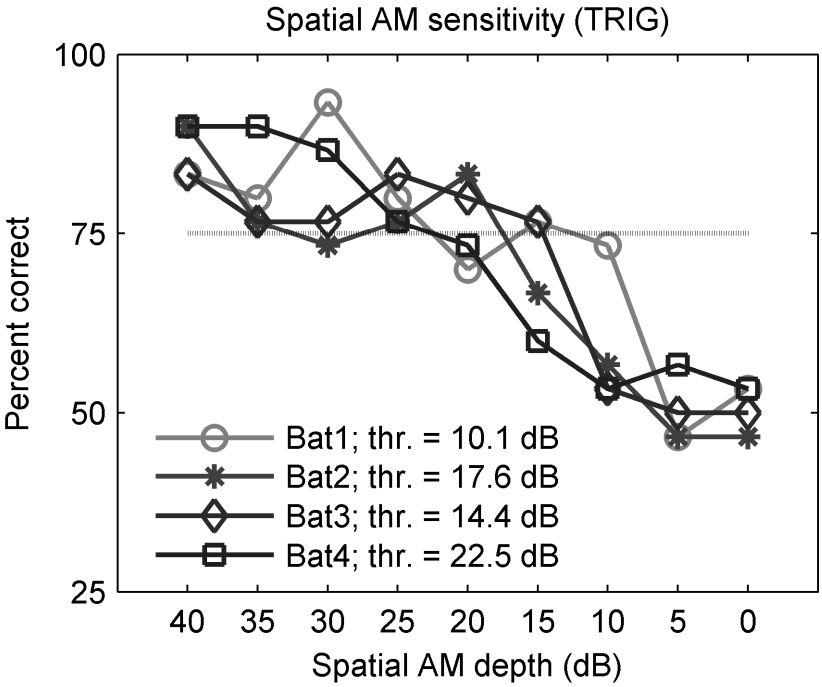


Figure 19: Psychometric functions for the detection of space-variant virtual objects in the TRIG condition
Here the echoes generated by the VO units were generated with a synthesized, stereotyped echolocation call but the same impulse responses as in the ACT condition. Echo playback, however, was triggered by the bats' sonar emissions. Again, individual spatial AM depth thresholds are given in the legend.

Finally, three of the five bats were successfully trained in the PAS condition. Here, the bats were not required to trigger the playback of synthetic echoes but the synthetic echoes were presented at a fixed repetition rate of 9.85 Hz from both VO units. Only the echo amplitude depended on the bats' position relative to the rewarded VO unit. Psychometric functions obtained in the PAS condition are shown in Figure 20. Thresholds for Bats 1 and 5 amounted to spatial AM depths of 12 and 15 dB, respectively. Bat 2 achieved a threshold of 11 dB.

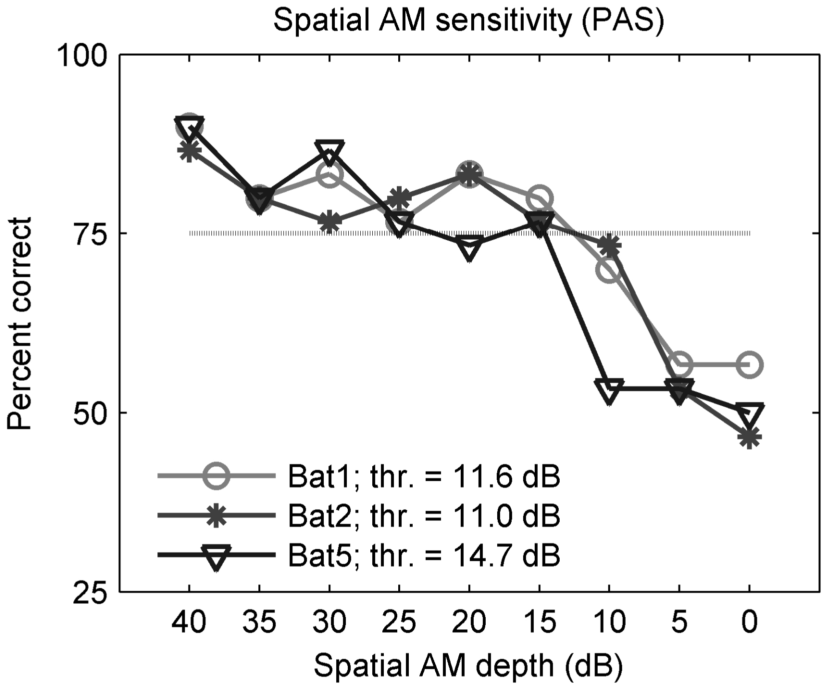


Figure 20: Psychometric functions for the detection of space-variant virtual objects in the PAS condition
Here the echoes generated by the VO units were generated with a synthesized, stereotyped echolocation call and the same impulse responses as in the ACT condition. In the PAS condition, echo playback was independent of the bats' sonar emissions at a rate of 9.85 Hz. Again, individual spatial AM depth thresholds are given in the legend.

Analysis of flight patterns of the bats around the VO units

Typical flight paths are depicted in two examples for each animal in Figure 21. Shown is a flight path during one trial from the moment the bat left the starting perch until it made a decision by landing on one of the two VO units. Bats 1 (row 1) and 3 (row 3) typically flew around the VO units before deciding for one object. Bat 4 (row 4) displayed relative long flight paths, as it usually flew back and forth between both VO units.

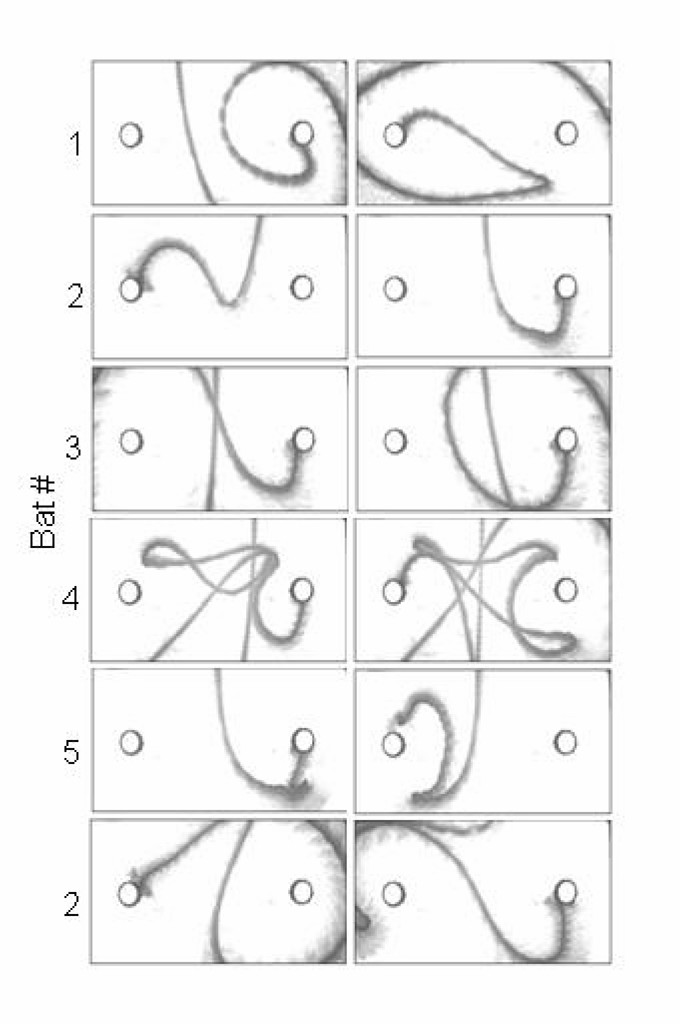


Figure 21: Exemplary flight paths for each bat
Two representative flight paths were extracted for each animal. The row numbers correspond to the bat number. Each flight path, grey sinuous line, was extracted during one trial; the two VO units are depicted as dark circular contours. Highly different flight strategies are evident. Additional flight paths for Bat 2 extracted towards the end of data acquisition are depicted in the last row, displaying a slight change in flight strategy in comparison to the whole duration of training (row 2).

In a subset of recorded trials, the sequence of radial segments which were visited by the bats around each VO unit were recorded and analysed. The position of the bat was sampled at a rate of 100 Hz. As each VO unit is divided into 36 10° angles the smallest resolvable position is 10°. A spatial modulation period always span six such segments, i.e., an angle of 60 degrees around a VO unit. The average number of different segments visited by a bat for at least 25 ms is shown in Figure 22 for the rewarded and unrewarded VO unit. In general, the bats visited on average 10 to 20 different segments in each trial and thus, they were well able to acquire enough spatial information from both the rewarded and unrewarded VO units. While the number of visited segments did not differ significantly between the rewarded and the unrewarded VO unit for each bat, individual differences of the overall flight patterns are apparent: Specifically, Bat 4 visited the highest number of different segments for at least 25 ms and Bat 2 the least number of segments. This result confirmed the experimenter's subjective observations.

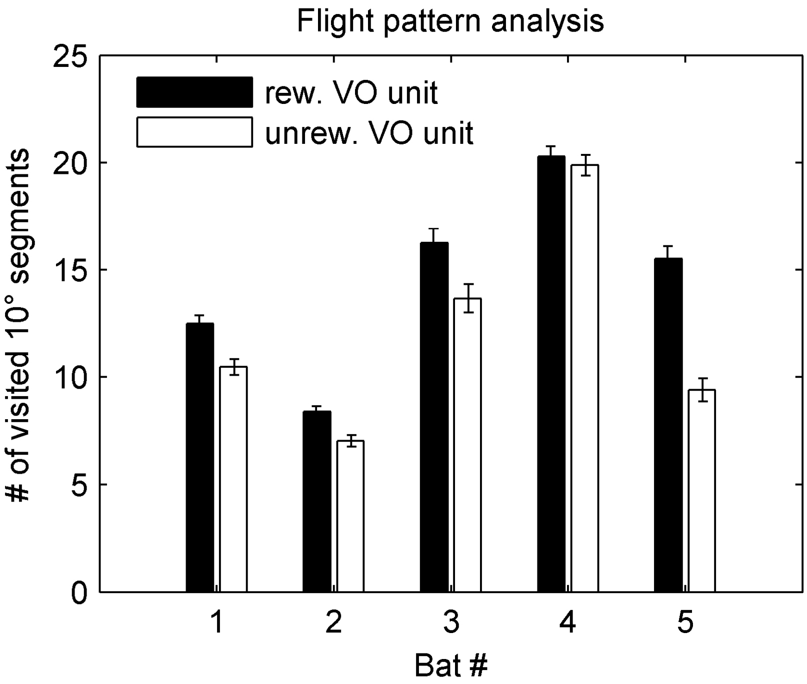


Figure 22: Analysis of individual flight patterns around the VO units
The bars show for each bat the average number of 10-degree segments visited for at least 25 ms around each VO unit in a trial. The analysis is motivated by the assumption that the higher this number is, the higher is the probability that a bat can acquire information about the angle-dependent properties of a virtual object. Most bats moved similarly wide around both the unrewarded and the rewarded object. Bats 1 and 4 moved significantly further around the VO units than the other two bats.

Sound analysis

In a subset of recorded trials in the conditions ACT and TRIG, the sonar emissions were recorded in parallel to the phantom-target echo generation. As sound presentation in the PAS condition was independent of the bats' sonar emissions, no sound analysis was conducted for trials acquired in the PAS condition. The output of both channels of the real-time processor which generated the virtual-object echoes was recorded using a firewire sound card (Phase 24, Terratec Electronic GmbH, Nettetal, Germany) at a sampling frequency of 192 kHz. As the sounds were recorded after the processor, only those sounds which were loud enough to trigger playback threshold are included in the analysis. The recordings were analysed off-line in terms of the number of calls and the average repetition period in a four-second interval preceding the bats' decision in each trial. Results of the sound analysis are shown in Figure 23. Most bats emitted between 10 and 20 calls towards each of the two VO units in each trial (panel A). The average repetition period of the calls, however, appears to be longer for Bat 1 than for the other bats (panel B). A detailed inspection of the recorded files shows that this result is due to a relatively constant call emission sequence of the bat in contrast to the more burst-like emission pattern of the other bats.

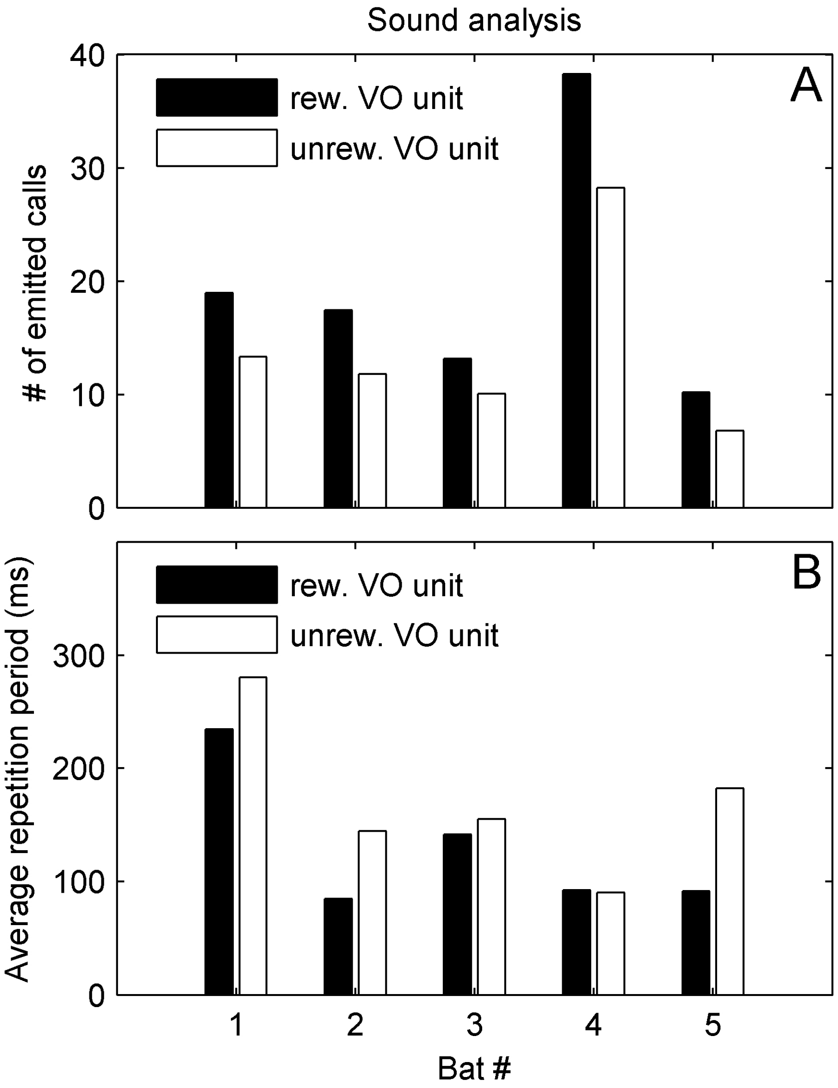


Figure 23: Analysis of recorded sounds from the VO units in the ACT and TRIG conditions
The bars show individual data for the average number of emitted calls (panel A) and the average call repetition period in the last four seconds of an experimental trial (panel B). In these four seconds, Bat 1 emitted a train about 18 echolocation calls a relatively stable period of 300 to 400 ms.

1.2.3 Discussion

The current experiments investigated the behavioural strategy and perceptual sensitivity of echolocating bats inspecting the 3D shape of complex objects. The presentation of holographic echo-acoustic objects was implemented to allow both the systematic variation of echo-acoustic object features and the behavioural strategies of the echolocating bats. Due to their excellent airborne manoeuvring capacity and their dominant recruitment of echolocation to investigate objects, *M. lyra* proved a very well suited animal model to study object-related echolocation strategies. The psychophysical sensitivity to the space-variant echo-acoustic properties of the holographic objects was described in terms of the minimum AM depth required to discriminate a virtual object with a spatial frequency of 0.016 cpd from an angle-invariant virtual object. The threshold spatial AM depth varied between 10 and 23 dB.

Difference between the ACT and TRIG condition on the one hand and the PAS condition on the other hand

In the ACT and TRIG conditions, the bats could only acquire object information when they flew around the objects and echolocated towards the objects. In the PAS condition, however, object information could be acquired by just flying around the objects. Interestingly, threshold spatial AM depths did not differ significantly between these conditions. This finding allows concluding that the call-triggered information when flying around the objects was equally good as the information acquired passively in the PAS condition. This means that the sonar activity of the animals was so high when they performed in the ACT or TRIG condition that it equalled the information provided passively by the synthetic calls presented passively at a repetition rate of 9.85 Hz. The sound analysis reveals average call repetition periods between 75 and 370 ms, corresponding to call repetition rates between 2.7 and 13.3 Hz. In summary, these data indicate that the flow of object information acquired passively in the PAS condition, on the one hand, and the flow of information acquired actively in the ACT and TRIG conditions, on the other hand, was in the same order of magnitude.

Comparison to fluttering-target investigations

The current experiments investigated the time-variant echo information a bat receives when moving around a complex 3D object and ensonifying it from different angles. A complementary, and much more studied case of time-variant echoes is the analysis of echoes generated from fluttering targets like flying insects. In this case, the bat is relatively stationary but the echoes are time-variant due to the time-variant nature of the ensonified object. The amplitude modulations in the echo sequences perceived by a bat ensonifying such a fluttering target are in the range of 15 to 30 dB (Roeder, 1963; Kober and Schnitzler, 1990; Moss and Zagaeski, 1994). Psychophysical experiments with fluttering targets have only been implemented with real objects, specifically rotating propellers, and the animals were trained to detect changes in the rotation speed. Thus, this setup allows varying the modulation frequency but not the modulation depth. A variation of modulation depth could only be achieved by using different-sized propellers, a parameter that has not been systematically investigated. In summary, the AM depths created by insects as fluttering targets are well in the range of spatial amplitude modulations used here, but modulation frequencies in the fluttering-target experiments are typically much higher and sensitivity to modulation depth has not been systematically investigated in fluttering-target experiments.

An electrophysiological study undertaken in 1987 (Reimer, 1987) was able to show that the tuning curves neurons in the inferior colliculus of the bat, *Rhinolophus rouxi*, were modulation frequency dependent, when evoked with sinusoidal amplitude modulated (SAM) stimuli with different modulation depths. In a few neurons a 6 % modulation depth still elicited response synchronization. The lowest psychophysical threshold of 10 dB spatial AM depth we observed in this study would correspond to a modulation depth of 25 %. *Rh. rouxi* actively hunts flying insects and relies on Doppler-effects imprinted on the returning echoes. The prey being small in size will not reflect with great loudness differences. This bat species might therefore need to evaluate not only frequency, but also small amplitude differences imposed on the echoes. The bat *M. lyra* used in this study does not usually actively hunt flying prey. It relies on prey generated rustling noises for prey detection and will then echo-acoustically investigate the object of interest (Neuweiler, 1990). It is therefore not necessarily dependent on having to evaluate small loudness differences of echoes.

Discussion of ensonification-correlated movements

As outlined in the introduction, the echo-acoustic analysis of 3D objects requires movements of the bat around the (stationary) object. The current experimental paradigm, which included a rove both of the overall target strength as well as a rove of the spatial position of the ‘glints’ of the rewarded virtual object, forced the bats to evaluate the correlation of their own movement and the sequence of perceived target strengths. The cognitive analysis of this correlation is mandatory to create an echo-acoustic representation of the ensonified object. The movement analysis, based on both exemplary video analysis and recordings of the temporal sequence of activated IRs, shows that the bats followed quite stereotyped but individually different flight paths around the VO units in the experiments. Bat 4 for example took a relatively long time to decide for a VO unit. This long flight duration is also represented in the high number of visited segments shown in Figure 22 (page 44) for this bat. Bat 2 and 5 were first trained to the PAS condition, where the bats did not need to echolocate and both VO units presented echoes simultaneously. To solve the task, both bats followed a certain strategy: usually flying in between both VO units and then deciding for a VO unit by turning towards the unit. Bat 2 applied the same strategy for the ACT condition. The inferior performance of Bat 2 in the ACT condition appears to be related to this strategy, where the bat only exploited a relatively small range of angles around the VO units. It is likely that this smaller range of exploited angles mediated less reliable information. Specifically, if e.g., an animal exploited each VO unit just from one angle, the target strength it would perceive cannot be used to discriminate the space-variant (rewarded) from the space invariant virtual object because of the roving paradigm. Interestingly, Bat 2 seemed to have slightly changed its flight strategy towards the end of data acquisition for all conditions. Every once in a while it displayed more flight movement around the VO units than at the beginning of training. This first sign of a possible shift in flight strategy is depicted in two examples in the last row of Figure 21 (page 42).

The current experiments show that bats can be trained to evaluate holographic echo-acoustic objects. This technique allows quantification, in psychophysically exact terms, of the bats' sensitivity to the space-variant properties of complex, three dimensional objects. At the same time, this technique allows the investigation of the correlation between the bats movements around an object and movement-correlated sonar activity. Psychophysically, we showed that the bats circling around the holographic object required a minimum spatial modulation depth to 10 to 23 dB to discriminate the space-variant virtual object from the space-invariant object. Interestingly the psychophysical performance of the bats is not always related to the behavioural strategies followed by the individual bats in terms of the range of observation angles exploited and the echo-acoustic call density during this exploitation: Bats 2 and 4 had the worse psychophysical performances in the ACT and TRIG conditions. They represented the two extreme behaviours concerning the exploitation of observation angles around the VO units. Bat 2 exploited the smallest range and used an average call density, which fits to the high threshold the bat needed to still solve the discrimination task, whereas Bat 4 displayed the largest exploitation range and the highest call densities, but still needed high spatial AM depths.

In further experiments with this experimental technique, it appears promising to investigate the bats' sensitivity to movement correlated changes in the echo delay, spectral interference pattern, and temporal characteristics.

1.2.4 Outlook

M. lyra is able to acquire echo-acoustic information about an object through perceptual integration of echoes and is sensitive to temporal changes of reflected echoes.

In a further study employing the same omni-directional playback technique we will be investigating *M. lyra's* sensitivity for movement correlated changes of the echoes' spectral composition. As before the bat will have to discriminate two virtual objects defined by 36 IRs. One is invariant to the bat's position and the other is position-variant. The IRs are generated by using the same filter banks as in the passive listening study. It will be interesting to compare these results with those obtained in the passive listening study; whether the bats' threshold performance degrades or improves, when evaluating the same IRs. In contrast to the passive listening study the bats obtain the echoes through active echolocation and the frequency modulation is correlated with the bats' position in the azimuthal axis. As the modulation rate is dependent on the bats' movement in space, this will not be a variable parameter. Instead we will vary the modulation period defining the position-variant virtual object. For example a modulation cycle of six produces six IRs per modulation period, resulting in six times six IRs; a modulation cycle of one produces 36 IRs per period. Less IRs per period should be less informative, as less IRs define a cycle. A lower cycle rate on the other hand will produce a slower spectral variation with smaller changes in frequency. It will be interesting to investigate, whether the thresholds differ in correlation with the modulation cycle. This study will further help understand the behavioural (flight and echolocation) strategies with which the bat, *M. lyra*, is able to reconstruct the 3D shape of complex objects by perceptual integration of the information acquired through sequences of echoes.

This concludes the first part of this thesis. Part II will lead us to a further topic relevant for echo-acoustic object analysis: the significance of object size for echolocating bats.

PART II: OBJECT SIZE ANALYSIS

2 INTRODUCTION

Not only is the evaluation of the sensory objects necessary, but also a reliable classification. Object recognition is not self-evident, as objects, independent of which modality they are perceived with, occur in different sizes. Size-invariant object recognition has already been shown for many different species, e.g. humans, monkeys, fish and gerbils, using different sensory cues (Larsen and Bundesen, 1978; Logothetis and Sheinberg, 1996; von der Emde, 2004; Sawamura *et al.*, 2005; Schebesch *et al.*, 2007). Single objects with given shapes are perceived as members of given categories regardless of the size.

Humans and monkeys spontaneously form size-invariant representations of visual objects (Larsen and Bundesen, 1978; Fiser and Biederman, 1995; Logothetis and Sheinberg, 1996; Furmanski and Engel, 2000; Sawamura *et al.*, 2005). Visually, objects are encoded via different neuronal pathways in the visual cortex, processing shape, motion and colour (Hubel and Wiesel, 1962; Hubel and Wiesel, 1977; van Essen, 1979; Zeki, 1980; Livingstone and Hubel, 1984). One basic for visual size-invariance is the pattern recognition process, which is achieved by comparing stimulus pattern with memory representations (Larsen and Bundesen, 1978) and evaluating an object's shape (Lloyd-Jones and Luckhurst, 2002b). Bundesen and Larsen (1975) suggest in a study testing visual transformation of size, that by encoding a visual object as a visual image and by transforming this image to the size format of a second object, we test visual objects for matches in shape.

Acoustically, objects can also show size variances. How object size is represented in the auditory domain is not quite understood. Several studies concerning speech processing in humans have showed that our auditory system easily compensates for the effect of speaker size on perceived speech. The same vowel when spoken by a child or adult differs in its spectral content, but is perceived as the same vowel, even when spoken by a speaker with an unnatural size (Ives *et al.*, 2005; Smith and Patterson, 2005; Smith *et al.*, 2005).

This holds for other acoustic stimuli as well. For example, when playing a recording of metal pins dropping onto a table with slower speeds, we automatically perceive falling pins increasing with size, until we have the notion of metal rods crashing onto a surface. The quality of the object, meaning the shape, and the size are processed separately, allowing a size-invariant, normalized object representation.

So acoustically we can compensate for size-induced variations of sound sources, does the same hold for echo-acoustic sounds? As discussed in the first part of this thesis, bats can recognize 3D objects in complete darkness through echolocation, but are they able to echo-acoustically normalize for object size, meaning is an object size-invariantly represented? To approach this question we need to outline what echo-acoustic parameters determine an object's size.

Object surface area and depth change proportionally with object size. With increasing object size the IR becomes louder, due to the increasing reflecting surface. It has been shown that bats are sensitive for loudness differences. They can perceive very fine echo sound differences of 1 to 3 dB produced by different sized triangles (Neuweiler, 2000). Another size dependent parameter is the relationship of single reflections to each other. Figure 24 shows an IR scaled in size from left to right. An object with a large depth extension (panel C) will produce reflections distributed on an elongated time axis. The single surfaces are further apart from each other, creating longer time delays between each reflection.

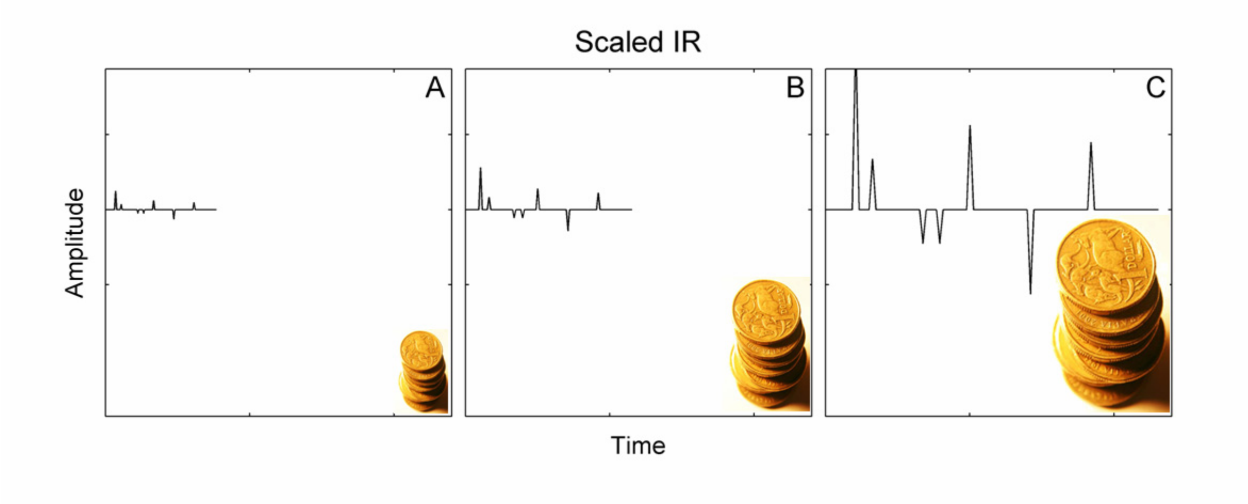


Figure 24: Illustration of a scaled IR
Shown is an IR scaled with the increasing scaling factors. The scaling increases from left to right (panel A to C). Plotted is the amplitude against the time scale. With increasing object size, object surface area and depth change proportionally. With increasing reflecting surface size the stronger the reflection and therefore the louder the IR becomes. The IR is temporally expanded, resulting in an increase in the temporal delay between single reflections. The insets in the panels depict scaled stacks of coins which very theoretically could reflect in such a manner.

This temporal expansion of the IR goes hand in hand with a compression of the spectral interference pattern. This can be seen in Figure 25 illustrating the magnitude spectra of a scaled target made up of two reflections with the same amplitude. As in Figure 24 the scaling increases from left to right. One can clearly see the spectral compression and the increase of magnitude when scaled with a bigger scaling factor. This is due to the increasing time delay between the single reflections. A longer delay results in a spectral ripple with a lower fundamental frequency and consequently more notches. This is due to the fact that the cancellation within a spectrum is dependent on the relationship between the wavelength and the distance between two reflections.

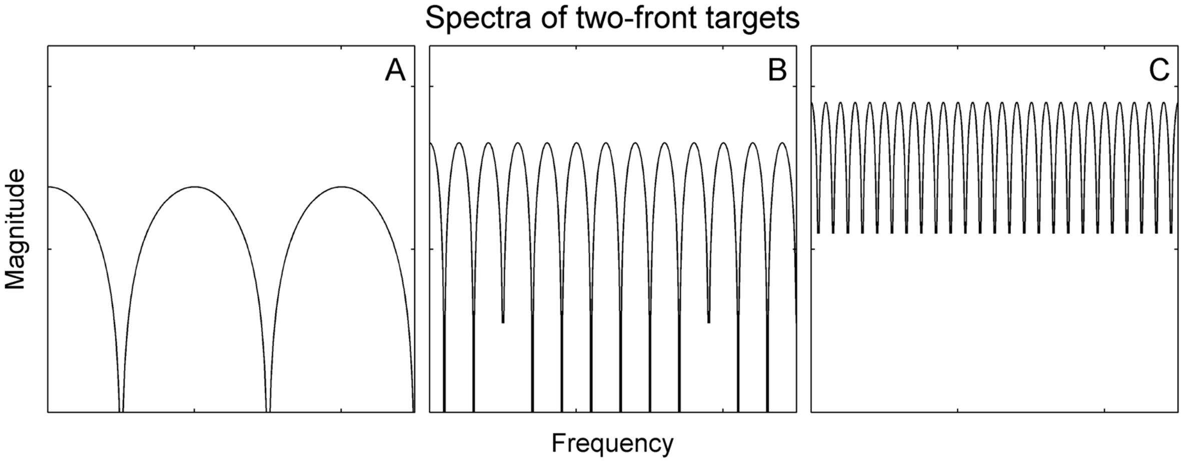
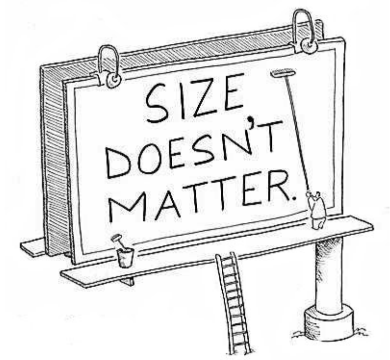


Figure 25: Illustration of the magnitude spectra of two-front targets
Shown are the magnitude spectra of a scaled IR consisting of two reflections. The size increases from left to right (panel A to C). Plotted is the magnitude against the frequency. With increasing object size the spectrum is compressed and increases in amplitude.

Whether an object is encoded temporally or spectrally is dependent on the size of an object. It has been shown that objects smaller than 4 to 6 cm are encoded mainly in the spectral domain and bigger objects in the time domain (Weissenbacher and Wiegrebe, 2003). This is dependent on the resolution of the cochlear filters for the single reflection peaks. Cochlear filter integration times measured for different bat species indicate that for two temporally separated reflection peaks to be encoded temporally they would have to display a delay of at least 200 µs to each other (Wiegrebe and Schmidt, 1996; Wittekindt *et al.*, 2005; Schörnich and Wiegrebe, 2007). Successive peaks too close to each other temporally can not be resolved as single reflections, but are encoded in the spectral domain by analysing the spectral pattern generated by destructive and constructive interference of the reflections.

By evaluating these parameters a bat should be able to determine the size of the ensonified object. With the assumption that a bat has a percept of the size and shape of an echo-acoustic object, can it transpose a size-variance onto the internal representation? This will be the topic in the first of the following studies. The second study, included in the appendix of this thesis, will be dealing with the effects of varying the combination of loudness and duration on the perception of object size in echolocating bats.

2.1 Echo-acoustic object normalization in the bat, *Phyllostomus discolor,* and in humans



Previous studies suggest that bats are able to compensate for object size. The bat *Glossophaga soricina* correctly classified scaled versions of previously learned simple real objects (von Helversen, 2004). The extracted magnitude spectra of the IRs from the objects displayed highly characteristic directional echo patterns in dependence of the ensonification angle; indicating that dynamic object analysis and the thereby generated echo patterns might be significant for an unambiguous object categorization. One critical point is that the bats always received an object pair; allowing the possibility that they compared the objects to solve the task, not by evaluating according to an internal object representation. In a later behavioral and electrophysiological study the bat, *Phyllostomus discolor*, was able to normalize to some extent for the size of previously learned complex virtual objects, represented by ensonification angle independent IRs (Firzlaff *et al*., 2007). This successful classification, despite the static IRs, does not corroborate the hypothesis, that an object's spectral or temporal reflection patterns produced when ensonified from different angles are necessary for the generation of a size-invariant internal representation of an object. The results though were not as unambiguous, suggesting that the perception of distinct echo patterns might facilitate object categorization.

In the following study we want to investigate whether the bat *P. discolor* is able to extract general features of an object and transpose these general features to an object of the same shape, but of different size. We trained the bat *P. discolor* to discriminate two real objects (standard-objects). Unlike the previous study (Firzlaff *et al*., 2007), the bats were able to move around the objects and collect information about the object from different angles. The reflection characteristics of both objects did not vary with the horizontal ensonification angle, but one object differed slightly along the vertical axis. Furthermore in contrast to the study by von

Helversen (2004) we presented only one object at a time and never both simultaneously. We therefore forced the bats to create an internal representation of each object, preventing correct performances based exclusively on comparison. As soon as their discrimination performance was above 80 %, scaled versions of these objects (test-objects) were interspersed, obtaining the bats' spontaneous decision for size-induced variances of previously learned objects.

In a control experiment we tested the bats' evaluation performance for two new objects with the same size and shape as the standard-objects used in the first experiment, but made up of a different material. We expected the bats to easily segregate the shape and not to be influenced by the material.

In an additional experiment we wanted to investigate how human listeners deal with the IRs of the real objects presented to the bats. The IRs of the objects are very short in duration, which impedes the evaluation temporally or spectrally (Robinson and Patterson, 1995a; Robinson and Patterson, 1995b). We generated an IR-train of each IR with short interpulse intervals, producing a spectral pattern consisting of many harmonics. This spectral composition is similar to spectral patterns found in human speech. The spectral shape of vowels in human speech consists of a fundamental frequency and formant frequencies above the fundamental. The formant frequencies produce a vowel specific pattern, meaning their alignment to each other is proportionally always the same, but their distribution along the frequency axis is dependent on gender, age and size of the speaker. When a child grows, the resonators in the vocal tract grow and the formant frequencies of the vowels decrease. Studies concerning size-induced variations of human spoken vowels have already shown that humans are capable of compensating for speaker size on vowel recognition (Ives *et al.*, 2005; Smith and Patterson, 2005; Smith *et al.*, 2005). Processing of acoustic scale and size perception of musical sounds seems to be similar (van Dinther and Patterson, 2006). We asked human subjects to discriminate in a psychoacoustic experiment IR-trains of the standard-objects used in the first experiment and to spontaneously classify interspersed IR-trains of the test-objects. We wanted to thereby acquire the spontaneous evaluation of scaled acoustic objects for humans. The results obtained in the bat and human psychoacoustic experiments were then compared for similarities and differences.

2.1.1 Material and Methods

2.1.1.1 Experiment 1: size-invariant object recognition

Animals

Phyllostomus discolor is a New World sub-tropical omnivorous bat, feeding on fruit, pollen and insects (Nowak, 1994). Figure 26 shows the bat and its typical echolocation call. It belongs to the Phyllostomidae family and can be found on the South-American continent. It emits short (<3 ms) broadband downward modulated multi-harmonic echolocation calls in the frequency range of 40 and 90 kHz.

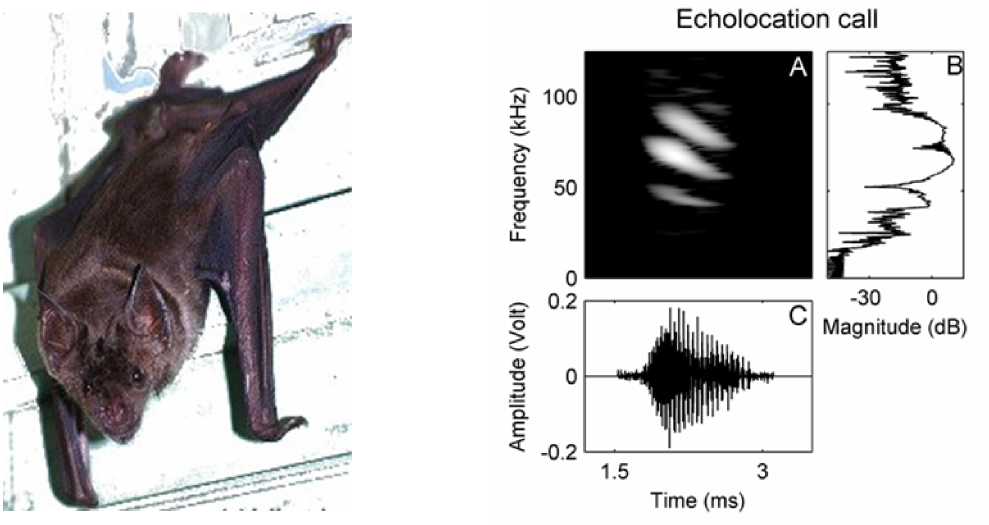


Figure 26: Phyllostomus discolor
The echolocation call is about 1.5 ms long and features multi-harmonic, downward modulated frequency sweeps with three dominant harmonics. Panel A displays the spectrogram of the call. Plotted is the frequency in kHz as a function of the time signal in ms. Panel B shows the magnitude spectrum, plotted is again the frequency in kHz as a function of the magnitude in dB. Panel C shows the time signal of the call; plotted is the amplitude in Volt as a function of the time signal in ms.

Four adult male *P. discolor* took part in the training. They were kept in cages (80 x 40 x 50 cm) with free access to water. During the training periods consisting of five consecutive days the bats were fed with banana pulp as a reward. On days without training they were fed mealworms and a fruit-mix consisting of banana, melon and mango.

Experimental setup

The bats were trained in a 2-AFC setup as used in former experiments with this bat species. All experiments were performed in an echo attenuated chamber (2.1 x 1.8 x 2.1 m) with a wall foam coating. The setup (Fig. 27) consisted of a Y-shaped panel, inversely mounted onto a metal fixture at an angle of 60°. Wire mesh was attached to the sides to ease the bats' action on the setup. The panel was otherwise not enclosed. An enclosable starting box was mounted at the top end and a feeder was attached at the end of each leg of the Y. The angle between each leg was 90°. The experimenter was seated next to the setup, observing the experimental procedure in total darkness via three small LEDs mounted onto a control panel. Three infrared-light barriers, positioned before each feeder and in the starting box, activated dependent on the bat's position on the Y one of the three LEDs. Both feeders could be activated and deactivated manually via the control panel.

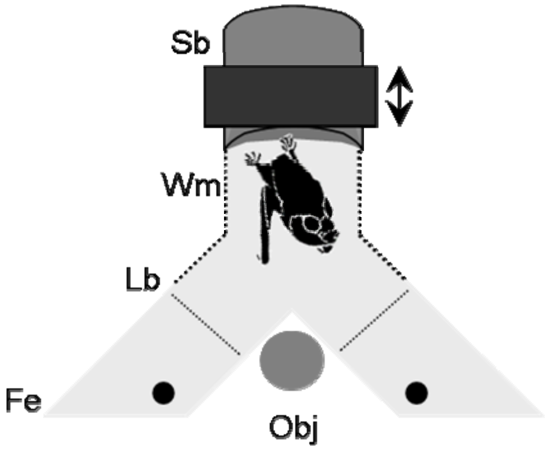


Figure 27: Illustration of the setup for the real targets
The Y-shaped setup is depicted from the front. Indicated are the feeders (Fe), the object's position (Obj), the light barriers (Lb), the surrounding wire mesh (Wm) and the closable starting box (Sb).

Objects

We chose as standard-objects a sphere (Object 1) and an hourglass shaped object (Object 2) both with a diameter of six cm and made up of Styrofoam. Attached to the bottom of each object was a metal rod with a halter on the bottom end for positioning. The test-objects consisted of scaled versions of both objects with scaling factors of: 0.67, 1.33 and 1.67. These were also made up of Styrofoam and had a rod with halter attached. Consequently the diameters of the scaled objects were four, eight and ten cm. This adds up to eight objects for presentation: two standard-objects and six test-objects. Schematics of each object can be seen in Figure 28.

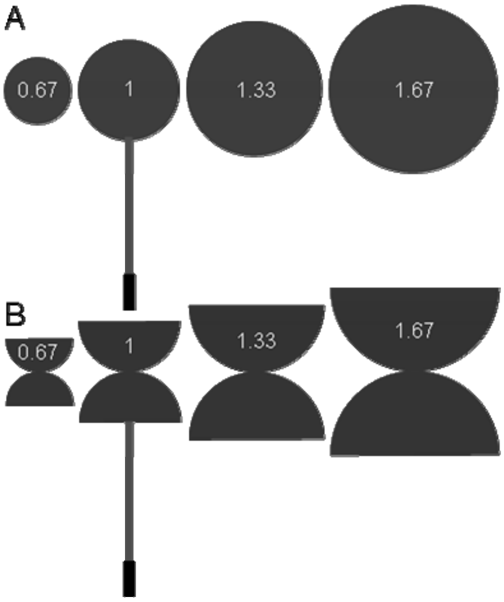


Figure 28: Real objects
Illustrated are the standard-objects with the attached metal rod, indicated with the scaling factor 1. Object 1 is shown in row A and Object 2 in row B. The corresponding scaled versions are indicated with the scaling factors 0.67, 1.33 and 1.67.

Procedure

In a 2-AFC experiment with food reward, four *P. discolor* were trained to discriminate the two standard-objects. Dependent on the presented object the bat had to crawl onto the left (Object 1) or right (Object 2) leg to obtain a food reward. The bats were only rewarded for correct decisions. After each decision the bat had to crawl back into the starting box, which was then closed. The experimenter exchanged the objects manually with the help of a then activated red lighting. An object was always removed, even when presented more then once in a row to avoid the bat using non echo-acoustic cues. Visual cues were excluded during object exchange due to the closed starting box. Test-objects were randomly interspersed with a probability of 25 %, when a bat's discrimination performance for the standard-objects exceeded 80 %. In these test-trials test-objects consisted of scaled versions of either Object 1 or Object 2. Test-objects were always rewarded independent of the bat's decision to assess the bats' spontaneous classification of the test-objects as either Object 1 or Object 2. During data acquisition for the test-objects it was ensured that the performance for the standard-objects was above 75 % correct decisions.

Data acquisition was stopped when at least 30 trials for each test-object had been obtained. The performance was calculated as the decisions for the side of the corresponding standard-object in percent correct as a function of the four scaling factors (0.67, 1, 1.33 and 1.67). Highly significant ($p < 0.01$) and significant ($p < 0.05$) correct were set to 75 and 67 % correct choices, respectively.

2.1.1.2 Experiment 2: control experiment

Three of the animals which had participated in Experiment 1 took part in the following experiment. Housing and feeding procedures were the same as in Experiment 1.
The experimental setup was the same as in Experiment 1.

Objects

The objects had the same shape as the standard-objects used in Experiment 1. As before the diameter was six cm, but the objects were made up of wood (wooden standard-objects). Again for positioning a metal rod with halter was attached.

Procedure

The same 2-AFC procedure with food reward was applied as in Experiment 1. The bats had to crawl to the left when the wooden Object 1 and right when the wooden Object 2 was presented. After each decision the bat had to crawl back into the starting box, which was then closed. The bats were only rewarded for correct decisions. Object exchange was as in Experiment 1. Data acquisition was stopped when at least 30 trials for each wooden standard-object had been obtained. Performance was calculated as the decisions for the side of the corresponding standard-object in percent correct. The significance level was set to 67% correct choices ($p < 0.05$).

2.1.1.3 Experiment 3: size-invariant IR recognition

Subjects

Six human subjects (four female, two male) aged between 22 and 30, participated in the experiment. Four subjects had already participated in psychoacoustic experiments.

Experimental setup

All experiments were performed in an echo attenuated, double walled chamber (1.2 x 1.2 x 2.2 m) built by G+H (G+H Schallschutz GmbH, Germany). The signals were presented over headphones (AKG K240 Studio, Germany). The subjects controlled signal playback via touch screen (WES TS, ELT121C-7SWB-1, Nidderau-Heldenbergen, Germany). The experimental program was written in Matlab 5.3 (Mathworks, Natick, MA).

Echo-acoustic objects

For the presentation of the IRs each object was ensonified and the reflected echo recorded. Each object was placed 20 cm in front of an ultrasonic speaker (Matsushita EAS 10 TH 800D, Osaka, Japan) and ultrasonic ¼ inch microphone (B&K 4135 with 2610 preamplifier and 2807 power supply). The objects were ensonified with an amplified (Rotel RB976 MKII, Germany) 10 s broadband noise sent out by a real-time processor (RX6, sampling rate 260 kHz, Tucker Davis Technologies, Gainesville, FI). The outgoing ensonification signal was computer generated (Matlab 5.3, Mathworks, Natick, MA). The microphone, recording the reflected echo, was connected to the real-time processor, which simultaneously played back the outgoing and recorded the incoming signal. The computer cross-correlated the incoming echo with the outgoing ensonification noise, producing the object specific IR. The IRs were saved for later presentation.

As seen in Figure 29 the time delays and amplitudes of the IRs of the objects are proportional to surface area and thus the radius. The high correlation of the temporal pattern between each standard-object and its scaled versions is visible. Object 1 and its scaled versions display one prominent amplitude peak. Object 2 and its scaled versions display two, but less prominent peaks.

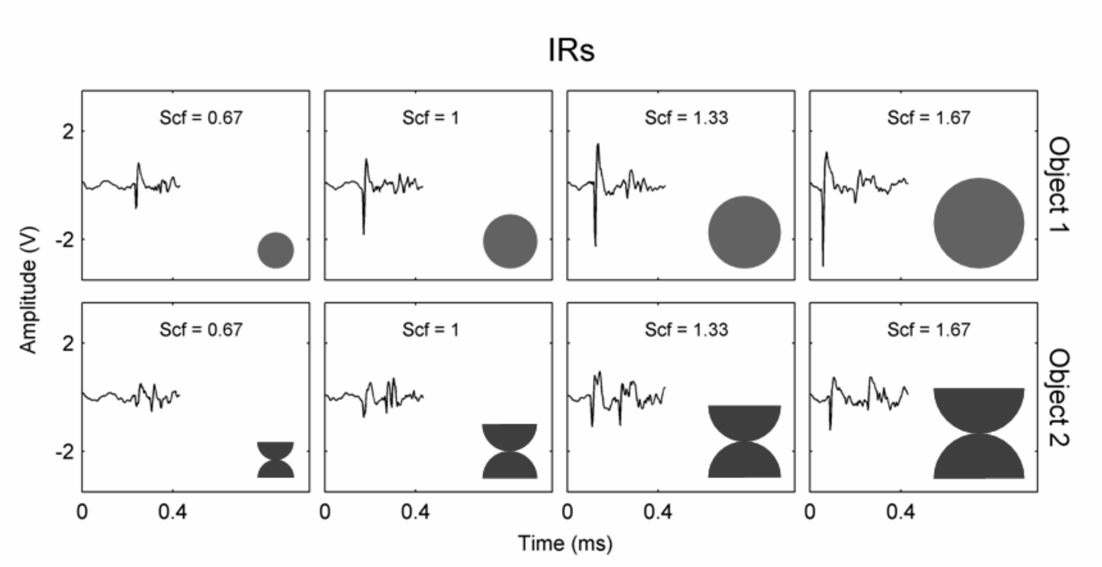


Figure 29: Illustration of the IRs of the real objects
The upper row shows the IRs of the standard and test-objects of Object 1, the bottom those of Object 2. Plotted is the amplitude in volt as a function of the time scale in ms. Object 1 is characterized by one prominent reflection peak, Object 2 displays two main reflections. The insets of the panels illustrate the objects' shape and size relations.

Stimuli

For each IR a signal train with a duration of 0.5 s was generated consisting of an IR with a repetition rate of 200 Hz. This produces the same pitch for all trains equal to the repetition rate of 200 Hz. Each IR was windowed with a raised-cosine window with a 1 ms rise time, 8.9 ms steady state and 1 ms decay time before generating the train. The trains were windowed with a raised-cosine window with a 50 ms rise time, 400 ms steady state and 50 ms decay time. All in all eight IR-trains were generated: one from each standard-object (standard-train) and six from the test-objects (test-train). Figure 30 depicts the trains for the standard objects.

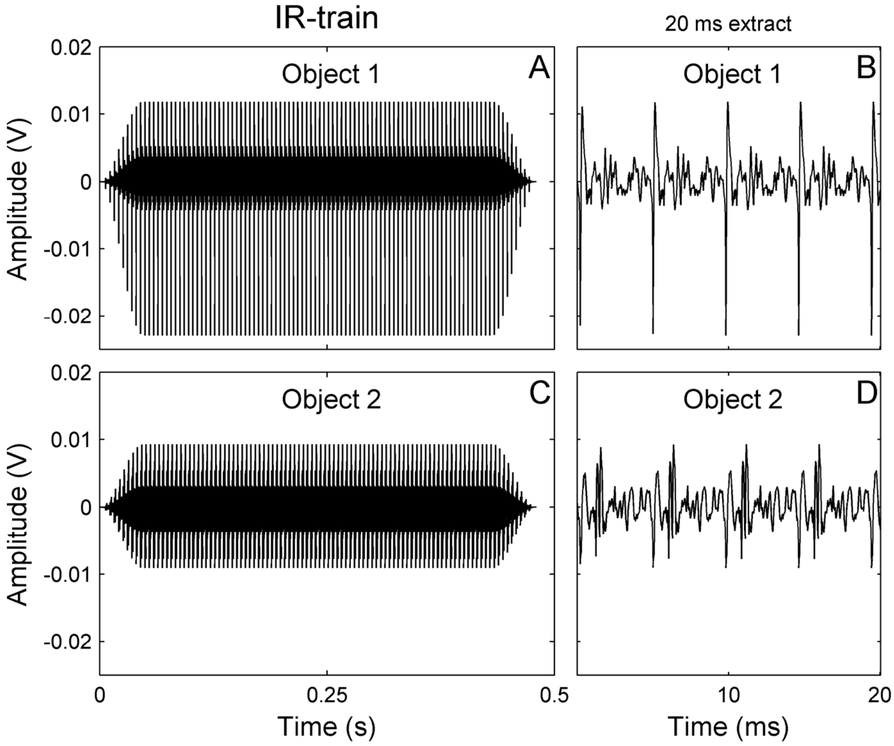


Figure 30: Illustration of the standard IR-trains
Panel A and C depict the IR-trains for Object 1 and Object 2, respectively. Plotted is the amplitude in Volt as a function of the time in s. Panels B and C show a 20 ms extract of each train.

During the presentation the trains were computer generated (Matlab 5.3, Mathworks, Natick, MA), digital-analog converted with a soundcard (DIGI 96/8 PST, RME Synthax Audio AG, Haimhausen, Germany) and played back with a sampling rate of 24 kHz. Due to the chosen sampling rate a train was played back with about a 10x smaller sampling rate than the original IR was recorded with. Connected to the soundcard were the headphones in the chamber. The signals were presented binaurally.

Procedure

In a 2-AFC experiment the spontaneous classification of the test-trains as belonging to either Object 1 or Object 2 were obtained. In a training-program the subjects were allowed to become accustomed to the standard-trains. For each trial a standard-train was presented once. The subject then had to make a decision. The subjects received feedback, whether a correct or wrong decision had been made. When a subject was able to discriminate the standard-trains with more than 80 % correct decisions and at least 30 trials, the test-program was started. The test-trains were then randomly interspersed with a probability of 25 %. In these test-trials subjects always received a positive feedback (correct choice), independent of the decision. This allowed the assessment of the subjects' spontaneous classification of the test-trains as belonging to either Object 1 or Object 2. Data acquisition was stopped when at least 30 trials for each test-train had been obtained. Performance was calculated as the decisions for the corresponding standard-object in percent correct as a function of the four scaling factors (0.67, 1, 1.33 and 1.67). Highly significant ($p < 0.01$) and significant ($p < 0.05$) correct were set to 75 and 67 % correct choices, respectively.

2.1.2 Results

2.1.2.1 Experiment 1: size-invariant object recognition

Four bats were successfully trained to discriminate the sphere, standard Object 1, from the hourglass, standard Object 2. During data acquisition for the test-objects the performance of Bat 4 for both standard-objects dropped slightly under the 75 % level. As this was a criterion for data acquisition for the test-objects the data for Bat 4 is not included. Panel A of Figure 31 depicts the decisions of the remaining three bats for Object 1 and its scaled versions and panel B the decisions of the bats for Object 2 and its scaled versions. All bats were able to significantly correctly classify at least four of the six test-objects as the corresponding standard-object. Highly significant and significant correct decisions are indicated by one or two asterisks above the plots, respectively. None of the bats was able to correctly classify the test-object of Object 1 with a scaling factor of 0.67.

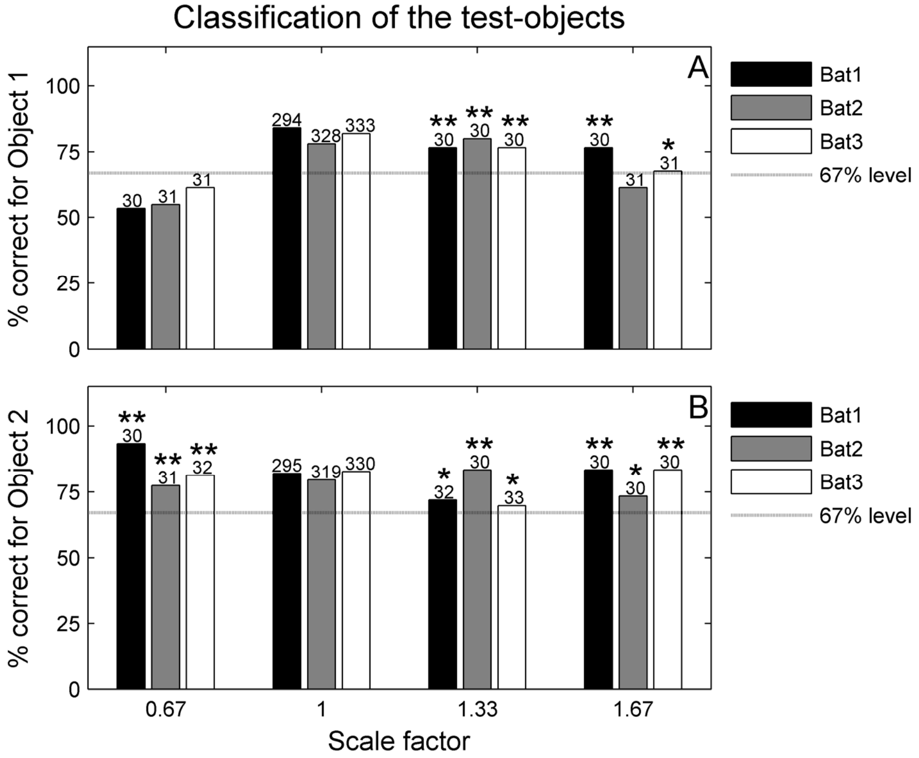


Figure 31: Classification results for three bats for each test-object as its corresponding standard-object Depicted are the decisions of three bats for Object 1 and its scaled versions (panel A) and for Object 2 and its scaled versions (panel B). Plotted are the percent correct decisions as the corresponding standard-object as a function of the scaling factor. The significant level of 67 % is indicated by the dotted line. Highly significant and significant correct classifications are marked by two and one asterisks above the plots, respectively.

2.1.2.2 Model: spectral or temporal pattern recognizer

We developed a model which classified the echoes of the test-objects as corresponding to Object 1 or 2. Echoes were generated by convoluting a standard *P. discolor* echolocation call with the IRs of each object, extracted for Experiment 3. The model compares either the spectral or the temporal pattern of the test-echoes with the standard-echoes. For the temporal pattern recognition the time signal of each test-echo was shifted sample per sample along a standard-echo, to obtain the sample point, where the cross-correlation between both echoes was maximal. For the spectral pattern recognition the same method was applied only using the magnitude spectra instead of the time signal. The time signals were additionally zero-padded at the beginning and end to ensure that both echoes had the same starting and end sample. A test-echo or spectra was then subtracted from a standard-echo or spectra. The result was squared, the mean calculated and the square root extracted. This resulted in a value for each test-echo or spectra which was divided by the sum of the values of this test-echo or spectra for each standard-echo or spectra. This resulted in a percent probability of a test-echo or spectra as belonging to a corresponding standard-echo or spectra. The calculation in the first row shows how a value for a test-echo is extracted. The probabilities of this test-echo as belonging to Object 1 or 2 are shown in the following two rows, respectively.

$$\mathrm{val}_{\text{test-stand1}} = \mathrm{sqrt}(\mathrm{mean}((\mathrm{stand1}-\mathrm{test})\hat{}2))$$

$$p_{\text{test-stand1}} = (\mathrm{val}_{\text{test-stand1}}/(\mathrm{val}_{\text{test-stand1}} + \mathrm{val}_{\text{test-stand2}}))\mathrm{x}100$$

$$p_{\text{test-stand2}} = (\mathrm{val}_{\text{test-stand2}}/(\mathrm{val}_{\text{test-stand1}} + \mathrm{val}_{\text{test-stand2}}))\mathrm{x}100$$

Note that this model does not have a sense of scaling, and therefore acts as a null hypothesis for our behavioural experiment of size-invariant object recognition in echolocating bats. The model failed to significantly correctly classify the test-echoes either temporally or spectrally as the corresponding standard-echo (Fig. 32). Panel A and B of Figure 32 depict the probability that the echoes of the scaled versions of Object 1 are correctly classified as of Object 1 due to a spectral and temporal pattern, respectively. Panel C and D depict the same only for Object 2. In comparison the bats were able to correctly classify at least four of six test-objects.

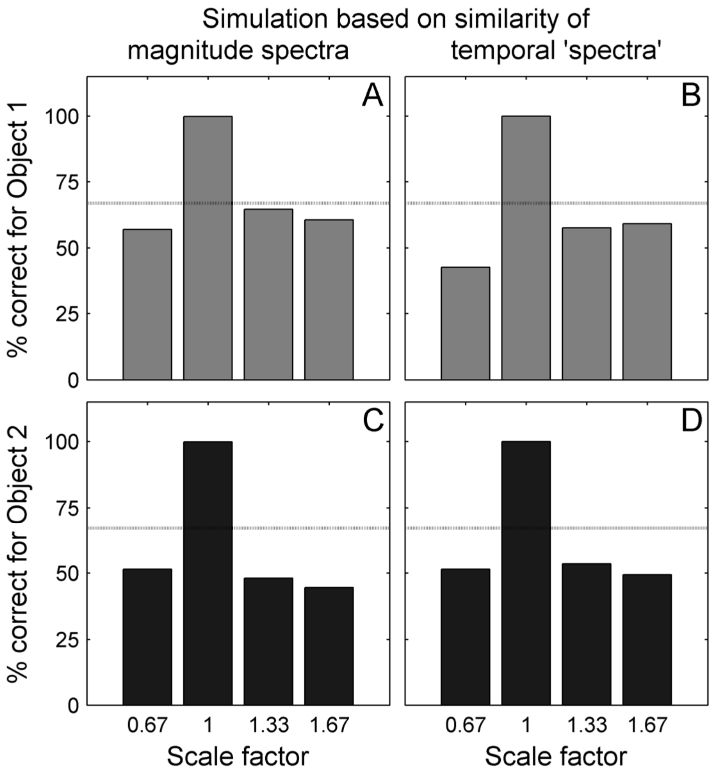


Figure 32: Evaluation of the test-echoes by a spectral or temporal pattern recognizer
Depicted are the evaluations of the model for each test-echo. Plotted are the correct decisions of a test-echo belonging to the corresponding standard-echo as a function of the scaling factor. In the panels A and C the model evaluated based on the similarity of the magnitude spectra, in the panels B and D based on the similarity of temporal spectra. Panels A and B show the results for Object 1 and its scaled versions, panels C and D show the results for Object 2 and its scaled versions.

2.1.2.3 Experiment 2: control

Three of the four bats were tested for the effect of an object's material on object recognition. For this, objects were chosen with the same shape and size of the standard-objects used in Experiment 1, but with different material (wood). Before the data acquisition for the wooden standard-objects was started, data for the Styrofoam standard-objects, used in Experiment 1, were obtained. As in Experiment 1 the bats had to crawl to the left, when Object 1 was presented and to the right, when Object 2 was presented, to ensure a performance above 75 % correct decisions. When at least 30 trials for each Styrofoam standard-object had been acquired, the wooden standard-objects were presented instead. The obtained results are shown in Figure 33. Performance was calculated as percent correct decisions for the wooden standard-objects as corresponding to one of the two standard-objects. Panel A depicts the decisions of all bats for the Styrofoam and wooden Object 1 and panel B the decisions of all bats for the Styrofoam and wooden Object 2. All bats easily recognized the wooden standard-objects as the previously learned standard-objects.

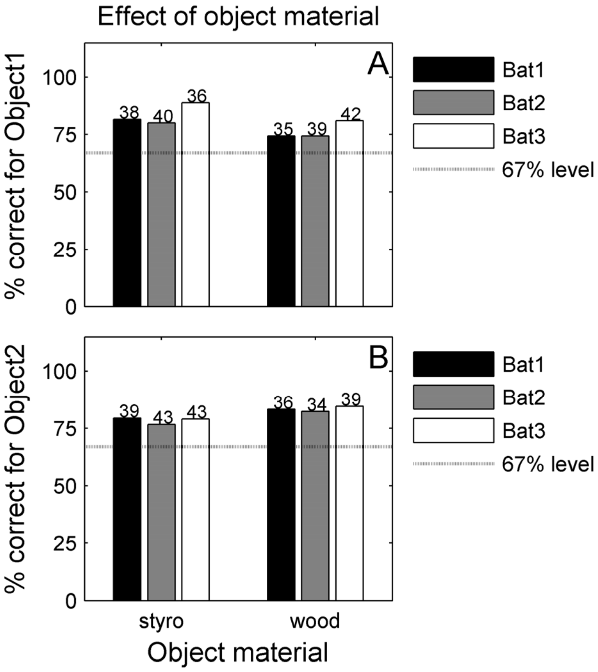


Figure 33: Evaluation results for three bats for the wooden standard-objects as the corresponding Styrofoam standard-object
Depicted are the decisions of all three bats for the Styrofoam and wooden Object 1 (panel A) and for the Styrofoam and wooden Object 2 (panel B). Plotted are the percent correct decisions as the corresponding standard-object as a function of object material. The significant level of 67 % is indicated by the dotted line.

2.1.2.4 Experiment 3: size-invariant IR recognition

Six human subjects were able to discriminate the IR-train of the sphere, standard Object 1, and the hourglass, standard Object 2. The obtained results are shown in Figure 34. Panel A depicts the decisions of all subjects for the standard-trains of Object 1 and the corresponding test-trains. The decisions of all subjects for the standard-trains of Object 2 and the corresponding test-trains are shown in panel B. All subjects were able to correctly classify at least two of the six test-IRs as the corresponding standard-IR. Highly significant and significant correct decisions are indicated by one or two asterisks above the plots, respectively. Subject 2 was able to correctly classify all test-trains as the corresponding standard-train. Subject 1 and 3 correctly classified all test-trains except the smallest scaled version of Object 1. Subject 4, 5 and 6 were only able to correctly classify the two bigger scaled versions of Object 1 and none of the test-trains for

Object 2. Subject 3 and 4 predominantly evaluated the test-IR of Object 1 with a scaling factor of 0.67 as belonging to Object 2.

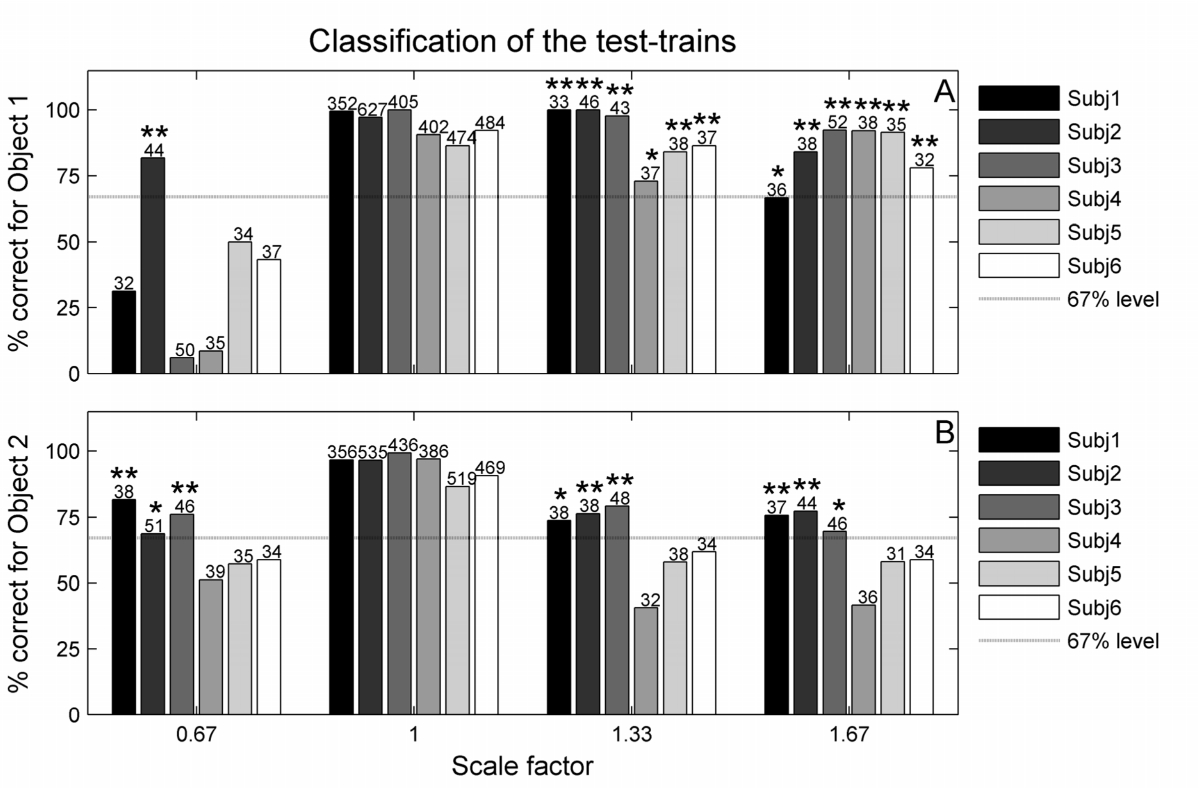


Figure 34: Classification results for six human subjects for each test-train as its corresponding standard-train Depicted are the decisions of all six human listeners for the trains of Object 1 and its scaled versions in panel A and for Object 2 and its scaled versions in panel B. Plotted are the percent correct decisions for the test-trains as the corresponding standard-train as a function of the scaling factor. The significant level of 67 % is indicated by the dotted line. Highly significant and significant correct classifications are marked by two and one asterisks above the plots, respectively.

2.1.3 Discussion

In a psychophysical experiment we tested the ability of the echolocating bat *P. discolor* to normalize for real objects with the same shape, but of different sizes. The results show that the bats spontaneously classified most scaled objects as the corresponding standard-object. Furthermore, the bats easily associated a material-induced variation of each standard-object with the corresponding standard-object.

A simulation based on either spectral or temporal pattern recognition was not able to correctly any of the test-objects as the corresponding standard-object.

In the third psychophysical experiment we tested the ability of human subjects to normalize for IRs of the objects used in the first experiment. Three listeners were able to compensate for most of the size-induced variations of the standard-trains. The remaining three listeners only classified two of the six test-trains correctly. Difficulties with discriminating between the two objects can not have hindered the evaluation of the test-trains, as the performance for the standard-trains is for almost all listeners well above 90 % correct decisions. This exceeds the performance the bats showed for the standard-objects.

In the following we will discuss the obtained results in regard to three points: the analysis of object size in echolocation, the evaluation of scaled sounds in humans and the comparison between the classification results obtained for the bats and the human listeners.

Echo-acoustic object size analysis

When evaluating an object in nature a bat will most likely acquire echoes from different ensonification angles and integrate the information from successive echoes (see Part I). The nectar- and pollen-feeding bat, *G. soricina*, is dependent on the ripeness of the flowers of interest. Bat-pollinated flowers display a certain petal formation when ready for pollination, generating a distinct shape (von Helversen and von Helversen, 1999; von Helversen and von Helversen, 2003; von Helversen *et al*., 2003). The bat should therefore be specialized to evaluate distinct echo patterns reflected by these flowers for object categorization. Later publications corroborate this hypothesis (von Helversen, 2004; Simon *et al*., 2006), by showing that this bat might indeed employ spectral pattern variations obtained from successive echoes from different ensonification angles for object analysis. *G. soricina* not only was able to distinguish two real objects, each characterized by an ensonification angle dependent reflection pattern, but was furthermore able to correctly identify size-induced variations. One critical point is that the results had been obtained in a discrimination paradigm. The bats might have solved the task by comparing the echo patterns of the object pairs and not by normalizing for size-variances of an internal representation of an object. The later following study by Firzlaff et al. (2007) had successfully tested *P. discolor's* echo-acoustic object normalization ability. In contrast to the study by von Helversen only one object, represented by one static IR, was presented in each trial, therefore excluding comparison possibilities. They stated that *P. discolor's* performance might improve when allowed to evaluate echo sequences from different ensonification angles. In our study, the bats only perceived one object per trial and were able to ensonify the objects from different angles. Both objects though reflected invariantly along the horizontal axis. Object 2 varied only along the vertical axis. The results display an improved classification performance for size-induced variations of objects. This corroborates the hypothesis that acquiring echoes from different angles facilitates object recognition, but that these echoes not have to display spectral variations. Firzlaff et al. (2007) proposed neuronal candidates for encoding this size-compensation mechanism in their study. Additionally to the behavioural experiment they conducted electrophysiological experiments concerning size-compensation of echo-acoustic objects. They state that important stages towards a neural correlate of echo-acoustic object recognition independent of size exist in the auditory cortex.

The question arises, whether the bats used spectral or temporal object features for classification. Wiegrebe and Schmidt (Wiegrebe and Schmidt, 1996) showed that the cochlear filter integration time for the bat, *M. lyra,* is about 200 µs. This means for two temporally separated reflection peaks to be encoded temporally they would have to display a delay of at least 200 µs to each other. Schörnich and Wiegrebe (2008) calculated similar integration times for the bat *P. discolor* based on data by Wittekindt et al. (2005). Every feature in the time domain has its equivalent in the frequency domain (spectral pattern). Temporal peaks closer together than 200 µs are encoded in the spectral domain (Weissenbacher and Wiegrebe, 2003), as multiple reflections with short time delays cause through interference reinforcement and cancellation of the reflections, resulting in spectral composition patterns. The two prominent peaks of Object 2 used in our study scaled with a factor of 0.67 are about 60 µs apart from each other. The peaks of the standard have a delay of about 100 µs. This indicates that these IRs must have been encoded in the spectral domain. The peaks of the size-induced version of Object 2 scaled by 1.67 are separated by 160 µs. This is in the transition range where encoding in the spectral to in the temporal domain takes place. All tested bats were able to classify the size-induced versions of Object 2 correctly. If the bats had evaluated the standard of Object 2 and the smaller size-induced version (scaling factor 0.67) spectrally and the version of Object 2 scaled by a factor of 1.67 temporally, this would suggest that the bats can transfer echo-acoustic properties of an object encoded either spectrally or temporally into the other domain.

Evaluation of scaled sounds in humans

It has already been hypothesized that the human auditory system applies a scale transform to all sounds to segregate size information from resonator shape information, and thereby enhance both size perception and sound recognition. This size-compensation ability of our auditory system has been shown for speech in earlier psychoacoustic (Ives *et al*., 2005; Smith and Patterson, 2005; Smith *et al*., 2005). Neuronal candidate for the first stage of processing speaker size might be the auditory thalamus and of the speaker’s fundamental frequency, a further size indicator, the non-primary auditory cortex (von Kriegstein *et al.*, 2006). Size-compensation ability seems to apply for musical instruments. Human listeners can discriminate the scale of musical instrument sounds reliably and recognize the family of an instrumental sound (van Dinther and Patterson, 2006). Other studies have successfully investigated non-speech sounds for the effect of shape, material and size on the perception of the sound, e.g. judgement of length and material of struck bars, geometric features of struck bars, length and width of struck plates and size evaluation of rolling bars (Freed, 1990; Lakatos *et al.*, 1997; Carello *et al.*, 1998; Kunkler-Peck and Turvey, 2000; Houben *et al.*, 2004; Houben *et al.*, 2005). The listeners in our study displayed ambiguous results concerning the size compensation of our IRs. Grassi (2005) conducted an experiment concerning the question whether it is possible to recover the size of an object from sound of an impact. Human listeners had to tell the size of a ball when dropped onto plates of different diameters, but were not informed of these plate differences. In this paradigm most of the reflection differences are produced by the plates, not by the carrier signal (ball). The subjects were influenced by the plate size, judging balls to be larger, when dropped onto larger plates. This indicates that the evaluation of the size of a sound source may be influenced by the expectation a listeners has towards the quality of a sound source. The listeners in our study may have had certain expectations towards the quality of the presented objects which may have impeded correct classifications.

It has already been shown, that the spectral profile of a sound is a key feature for the identification of sound quality (Green, 1996; Griffiths and Warren, 2004; Warren *et al.*, 2005). Other studies concerning human psychoacoustic have shown that humans are sensitive for spectral inference patterns (Sams and Salmelin, 1994; 1995b; Larsby and Arlinger, 1998; Macpherson and Middlebrooks, 2003; Alves-Pinto and Lopez-Poveda, 2005). Houben et al. (Houben *et al.*, 2005) found in their study concerning the contribution of spectral and temporal information to the auditory perception of the size and speed of rolling balls that the judgment of size is dominated by spectral information. We asked the human listeners in our study what cues they used to evaluate the standard-trains. The standard IR-train of Object 2 was always identified as being quieter and higher in timbre in comparison to the standard IR-train of Object 1. The loudness cue though could not have been applied for the reliable classification of the test-trains, as loudness of the IRs depends on the object size, meaning scaling factor: the amplitude of the test-train of Object 1 with a scaling factor of 0.67 was lower than that of the test-train of Object 2 with a scaling factor of 1.67. Instead of relying on loudness, some listeners stated of having perceived ‘darker’ or ‘lighter’ versions of the standard-IRs, indicating that they used the relative spectral structure for the classification of the test-trains. The first three subjects might have employed different spectral structures for classifying test-trains than the latter three, leading to the discrepancy in performance of these two groups. Subject 6 stated that sometimes amplitude was used as a cue for the classification of trains, and Subject 4 and 5 stated varying between different spectral cues for evaluation, which might be the reason for the more inferior classification performance of these three listeners. Note that the subjects did not know when a test-trial was conducted and therefore did not necessarily recognize the presented train as a test-train.

Comparison between bat and human performance

The employment of the spectral profile as a cue for classification of the test-objects might have been the case for bats as well. Other studies concerning bats have shown the importance of spectral profile analysis for object recognition in echolocation (Schmidt, 1988; Schmidt, 1992; Preisler and Schmidt, 1998; Krumbholz and Schmidt, 1999). The first study discussed in this thesis corroborated this. The earlier study by von Helversen (2004) investigating echo-acoustic normalization for object size indicated that the spectral profile may reflect the size of an object. Here studies are quoted which show that birds demonstrate a perceptual constancy for relational spectral structures for the generalization across objects with different sizes; meaning the birds transfer between signals with different timbre (Braaten and Hulse, 1991; Guttinger *et al.*, 2002). Bats using the spectral profile as a cue could generalize for size-induced variations of object independent of the ensonification angle. A further indication for the importance of the spectral profile for animals was shown for identity coding within vocal signatures for lambs and ewes (Searby and Jouventin, 2003). Experiments with carp fish and goldfish demonstrated that fish learn to discriminate signals featuring differences in timbre (Chase, 2001) and that they exhibit timbre-like perceptual dimensions (Fay, 1995).

Figure 35 compares the mean performance for all bats for all test-objects with the mean performance of all human listeners for all test-trains. In general the human listeners performed better than the bats for the standard-objects, but interestingly the classification results obtained in our study only showed slight differences between bats and humans. Both the bats and human listeners poorly evaluated the smaller scaled version of Object 1 (scaling factor 0.67), but otherwise were mostly able to compensate for object size. The human listeners were better in their classification performance than the bats for the two bigger versions of Object 1. The bats on the other hand seemed to be better than the human subjects in classifying all scaled versions of Object 2.

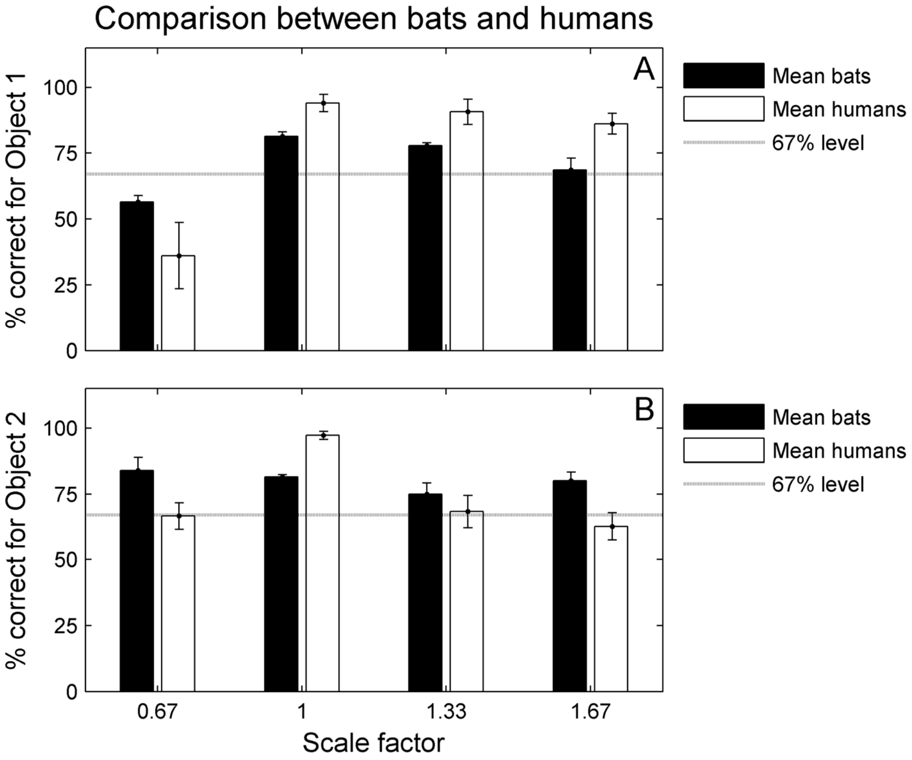


Figure 35: Comparison between bats and human subjects
Plotted are the mean correct decisions as a function of the scaling factor for Object 1 and 2 in panels A and B, respectively. The mean performance for the bats is shown with the black bars, the same for the human listeners with the white bars. The 67 % level is indicated by the dotted line.

We had a closer look at the spectra of the IRs to get an idea of what the bats or humans may have listened to. In Figure 36 and 37 the magnitude spectra in dB for all IRs are plotted as a function of frequency in kHz. Figure 36 shows the IRs in the frequency range presented to the human listeners. The trains were sampled with a 10 times slower rate (24 kHz), consequently the audible range is only up to 12 kHz. Note that none of the information provided by the IRs is thereby missing. Panel A shows the spectra for Object 1 and its scaled versions and panel B the same for Object 2.

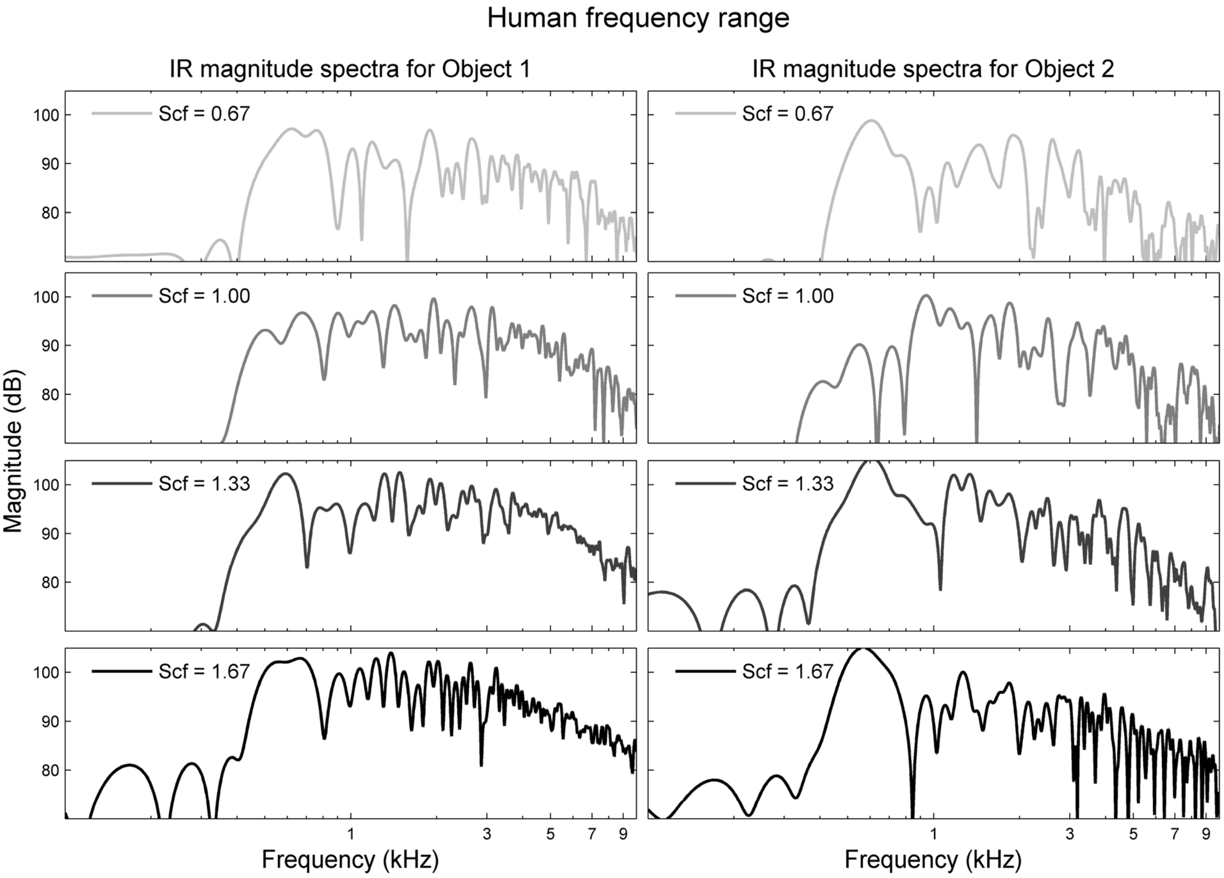


Figure 36: Illustration of the magnitude spectra of the IRs for the human listeners
Depicted are the magnitude spectra of all IRs for a frequency range of 100 Hz to 10 kHz. The magnitude in dB is plotted as a function of frequency in kHz. The left panels depict the spectra for the IRs from Object 1 and its scaled versions. The right panels the same for Object 2. The corresponding scaling factors (Scf) are indicated in the left corner. The scaling factor increases from top to bottom.

Figure 37 depicts the spectra for a frequency range of 20 to 120 kHz, the range relevant for the bats. The panels are otherwise in the same format as in Figure 36.

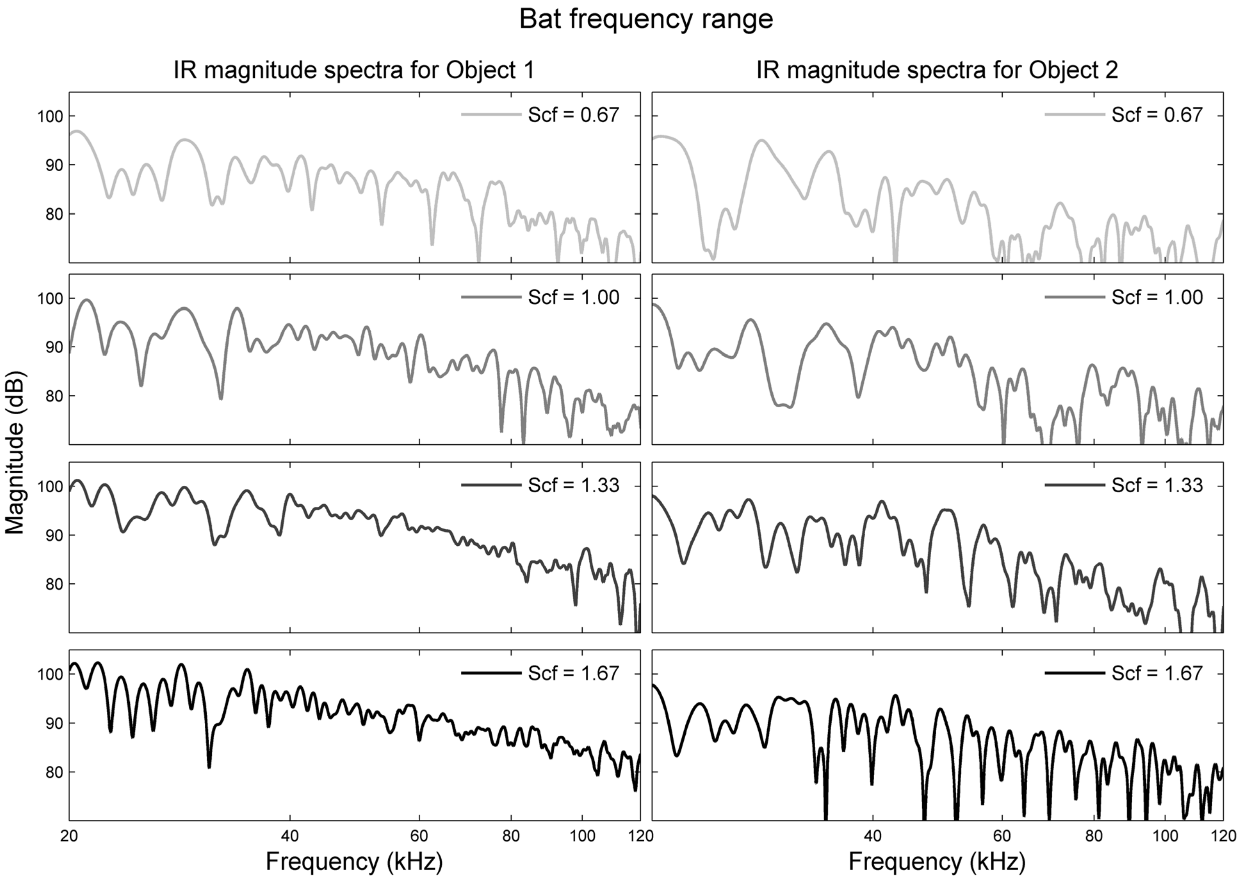


Figure 37: Illustration of the magnitude spectra of the IRs for the bats
Depicted are the magnitude spectra of all IRs for a frequency range of 20 to 120 kHz. The magnitude in dB is plotted as a function of frequency in Hz. The left panels depict the spectra for the IRs from Object 1 and its scaled versions. Panel B the same for Object 2. The corresponding scaling factors (Scf) are indicated in the right corner. The scaling factor increases from top to bottom.

Visible in the spectra of both figures is the increasing amplitude and compression of the spectral ripple in correlation with increasing object size. This is to be expected, as with increasing object size the temporal expansion of the IR goes hand in hand with a compression of the spectral interference pattern. In general the notches produced by Object 2 and its scaled versions are more pronounced. A spectral cue both bats and humans may have attended to. The spectrum of the scaled version of Object 1 stands out in both figures, as it displays more prominent spectral notches, as seen by Object 2 and its scaled versions. This might explain why this test-object was classified so ambiguously by both bats and humans.

One must not forget that the humans did not actively acquire a reflection signal and only perceived one static IR of each object as a train. The bats were allowed to freely move and echolocate on the setup to evaluate a presented object. Especially the reflection difference of Object 2 along the vertical axis may have been exploited, information missing for the human listeners. This might have been an advantage for the bats for the evaluation of the test-objects. By being able to scan the objects from different angles the bats had the possibility to achieve a sequence of echoes and form a 3D acoustic image of the objects. Additionally, the bats might have been able to vary the frequency content of their call, pronouncing or attenuating certain bands, and therefore achieving more focus on certain object characteristics. Earlier studies have shown that bats adjust their calls to the reflection characteristics of prey or background (Kalko and Schnitzler, 1993; Leippert *et al*., 2002; Kingston *et al.*, 2003; Macias *et al.*, 2005; Bartonicka and Rehak, 2005; Holderied *et al.*, 2006; Gillam and McCracken, 2007; Guillen-Servent and Ibanez, 2007).

To summarize, the current study corroborates the hypothesis proposed in earlier studies (von Helversen, 2004; Simon *et al*., 2006; Firzlaff *et al*., 2007) that the bat's auditory system has hard-wired mechanisms to deal with size-induced variations of echo-acoustic objects. However, in contrast to earlier studies the bats evaluated our test-objects according to an internal representation of an object. The results corroborate that information from successive echoes from different ensonification angles probably facilitate object normalization. We did not find that the acquisition of spectral pattern variations of echoes reflected from different angles is necessary for reliable object categorization. The investigation of the temporal and spectral features of the IRs suggests that bats might be capable of transferring an object's characteristic evaluated spectrally onto a size-induced variation temporally and vice versa. The results obtained during the human psychoacoustic experiment indicate that individual human listeners attend to different spectral sound characteristics. We propose that timbre may have been employed as a cue for classification by the human listeners and possibly by the bats.

GENERAL DISCUSSION

This thesis encompasses a series of psychophysical experiments which were designed to investigate fundamental cognitive mechanisms of echo-acoustic object recognition. Previous experiments have shown that bats can discriminate real and virtual objects with high fidelity (Simmons *et al.*, 1974; Schmidt, 1988; Simmons *et al.*, 1990; Schmidt, 1992; von Helversen, 2004; Weissenbacher and Wiegrebe, 2003; Firzlaff *et al.*, 2006). More recent work has shown that bats can also extract and memorize object related echo-acoustic information and use this extracted information to classify unknown real and virtual objects (von Helversen, 2004; Grunwald *et al.*, 2004; Firzlaff *et al.*, 2007). While these studies have provided important insights into the neural basis of echo-acoustic object recognition, important questions have remained open: first, the cognitive mechanisms underlying the integration of echo-acoustic object information obtained from different observation angles has not been investigated in an explicit psychophysical paradigm, and second, the echo-acoustic information mediating object-size information have not been systematically characterised. In an attempt to address these central questions in echo-acoustic object recognition, we measured first, in a formal psychophysical 2-AFC experiment the capability of the bat, *M. lyra*, to follow rapid changes in the spectral composition of synthesized echolocation-call sequences as they would occur when a bat ensonifies a complex object while flying around the object. The data show that *M. lyra* is quite sensitive to changes in the echo spectral composition and retains this sensitivity even when the spectral changes occur at relatively high rates, corresponding to quite fast flight speeds around the object. In this passive-acoustic paradigm, however, the occurrence and repetition rate of echoes was defined by the experimenter. Moreover, the perception of echoes was not correlated with movements of the bat around an object. These echo-acoustically critical issues were addressed in the second set of experiments which involved an important extension of previous virtual-object experiments: We designed an omni-directional phantom-target technique for the generation of holographic echo-acoustic objects. This technique enables us for the first time to acquire perceptual data on the sensitivity of *M. lyra* to changes in target strength of a virtual object with an ensonification-angle dependent target strength. These data showed that, under highly controlled experimental conditions, the bats recruited a flight- and echolocation strategy which allowed them to detect observation-angle dependant changes in target strength down to 10 dB. While this sensitivity is inferior to the amplitude-modulation sensitivity observed passive-acoustically in other small mammals (Henderson *et al.*, 1984; Cooke *et al.*, 2007;

Wiegrebe and Sonnleitner, 2007) this measure encompasses the whole behavioural context including the requirement of flight, its correlation with sonar activity and forward masking of the echo by the sonar emission. These first experimental data also serve to validate the omni-directional phantom-target technique. Future experiments involving this technique are in progress. These include angle dependent variations of spectral and echo-delay information. In summary, the omni-directional phantom-target technique allows for the first time to investigate the correlation of flight and sonar activity under highly controlled echo-acoustic and motion conditions.

In the second part of this thesis, a topic is addressed which has already received some scientific attention: to what extent can echolocating bats segregate echo-acoustic size from echo-acoustic structure information. Unlike in a previous experiment (von Helversen, 2004), a classification paradigm with real objects, not a simple discrimination paradigm was used. This paradigm forces the bats to create an internal representation of the objects' 3D shape and size and compare this internal representation with the echo-acoustic properties of the one object presented in a specific experimental trial. Using this refined paradigm, together with a quite different class of echo-acoustic objects corroborates previous findings that echolocating bats can indeed segregate echo-acoustic size from shape information: The bats spontaneously classified objects of different size but the same shape into the same class. The classification performance can be qualitatively explained under the assumption that the bats evaluated the objects IRs in terms of the presence or absence of a second strong reflection. In terms of spectral analysis, this corresponds to the evaluation of the distinctiveness of a harmonic spectral ripple in the echo spectrum irrespective of the absolute position and size of individual spectral peaks and notches. Those perceptual cues that could have mediated the performance of the bats in the study by von Helversen (2004), i.e., the angle-dependent changes in echo amplitude, can be excluded here because of the shape and orientation of the presented objects. Thus, these data, together with the virtual-object data by Firzlaff et al. (2007) confirm the size invariance of echo-acoustic object classification. These experiments however, do not clarify how size information is actually extracted from echo information. Specifically, object size could be deducted either from echo loudness (target strength) but also from spatial information, explicitly the spatial extent, of the perceived echoes. This differentiation is critical as, while the phantom-target technique can mediate target-strength information, it cannot mediate spatial-extent information. In an additional project in this thesis, this question is addressed in a field study with a high number of untrained echolocating bats.

This project, as it is still in progress and will extend beyond the thesis as it is presented here, is introduced and described in terms of materials and methods in the appendix.

To conclude, the current thesis provides both new experimental techniques and psychophysical data on one of the most fascinating mammalian senses, the bats' capability to extract an internal image from the auditory analysis of self generated sounds which are reflected by surrounding objects.

APPENDIX I

The effect of virtual echo-acoustic objects on free flying bats

Earlier studies have investigated flight patterns in bats and have found that commuting bats display very constant flight paths every night over many years (Eisentraut, 1952; Bateman and Vaughan, 1974; Rieger *et al.*, 1990; Rieger and Adler, 1993; Rieger, 1997; Rahmel and Dense, 1997; Rahmel and Dense, 1998; Keil *et al.*, 2005; Schaub and Schnitzler, 2007a). It is probable that a bat will correct for its flight when suddenly confronted with objects standing in the commuting flight path. The extent and the time of the correction should be correlated with the object's size. As briefly discussed in the first part of the thesis, loudness and duration of an echo are the main parameters determining object size: the bigger the reflecting surface, the louder the echo and the bigger the object's spatial extension, the longer the temporal expansion of the echo. Figure 38 depicts the IR of two surfaces with different distances to each other. The panels A to C illustrate how loudness of and temporal delay between the single reflections increase with an increase of the reflecting surfaces' size and distance to each other.

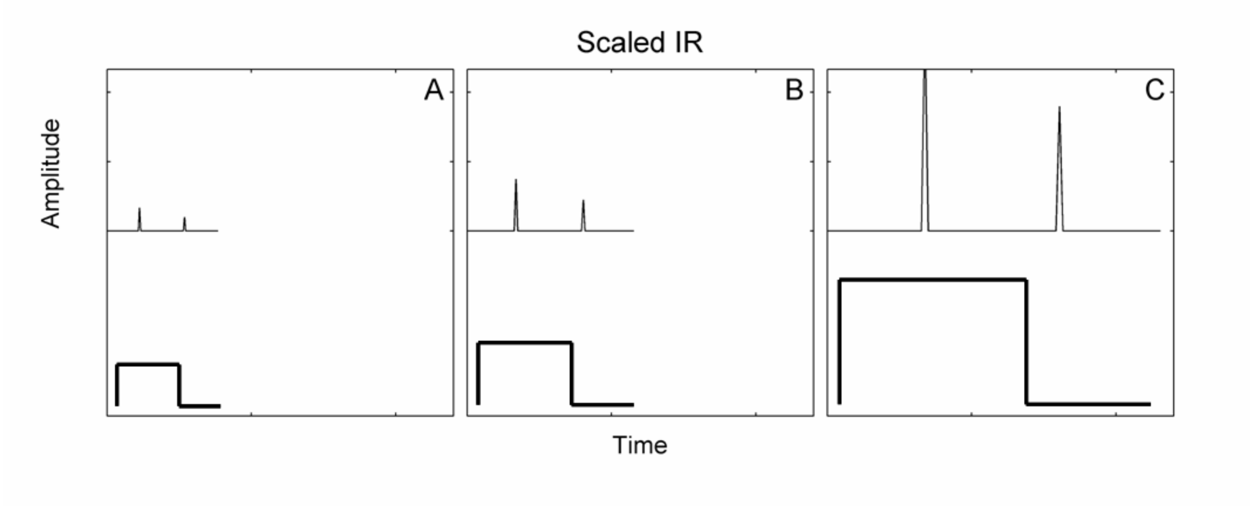


Figure 38: Scaled surface depths
Illustrations of two reflections scaled in size. Plotted is the amplitude as a function of time. The amplitude and duration of the IRs and the time delay between the two reflections increase with increasing scaling factor. The object's size increases from panel A to C. The insets of the panels depict the two exemplary surfaces with different depths.

Consequently, the extent of a bat's flight corrections around an object should reflect the perceived object size.

The exact relationship between loudness and duration for the representation of object size has not been explored before. We therefore wanted to investigate how the size of an object influences flight paths of commuting bats. As we have successfully employed the phantom target technique in the previous studies with laboratory bats, we are interested in how untrained wild bats, familiar only with real objects react to unfamiliar virtual objects. The advantage of such a field study is that it does not require training which takes a large amount of time and limits the number of animals one can test. As a location we chose a cave in Trinidad, already known to us, through an earlier visit and from the internet literature (Kenny, 2007; Riskin, 2007). The Tamana-Cave provides the unique opportunity of testing this phantom target technique on wild bats, as towards nightfall thousands of bats of at least 12 different species fly out of the cave opening for about two hours to forage in the surrounding woods. Due to the cave's surrounding configuration the bats, when exiting the cave through an opening in the ground, fly about 10 meters parallel to the ground until they fan out. This is an ideal requirement for placing the microphone-speaker combination into the animals' flight path. As in the previous phantom target playback experiments, the microphone recorded echolocation calls are manipulated in real-time and played back as virtual echoes over the speaker to the calling animals. The flight paths will be extracted by recording the flying bats with a high speed video camera positioned on the ground aimed upwards and by off-line editing of the recordings with a self-designed calculation software. The flight path characteristics and the high sample number of animals will allow us to obtain a large data size in a short time range.

We created several IRs defined by their loudness and duration for the presentation of numerous size-variant objects with variable spatial (time) extensions. The IRs will be convolved with the bats' echolocation calls and played back as echoes. Note that a virtual object represented by only one loudspeaker can only display a width and height expansion by increasing its amplitude. A representation of a spatial extension along the horizontal or vertical axis is not possible. The variations along the time axis represent size differences along the longitudinal axis away from the bats. This is represented by varying the number of reflecting surfaces. A parameter defining this echo-acoustic property is echo roughness. In a behavioural and electrophysiological study it has been shown that the bat *P. discolor* can distinguish and evaluate echoes according to their roughness (Grunwald *et al*., 2004; Firzlaff *et al*., 2006). Roughness is defined by the echo's envelope fluctuations: higher fluctuations, generated through bigger reflection surfaces, will produce a rougher echo and lower fluctuations, generated through smaller reflection surfaces, a smoother echo. This roughness parameter is given in log10M4. Three signals with same duration (~7300 µs), but with different roughness, 0.5, 1.5 and 2.5 log10M4, are illustrated in the three panels of Figure 39, respectively.

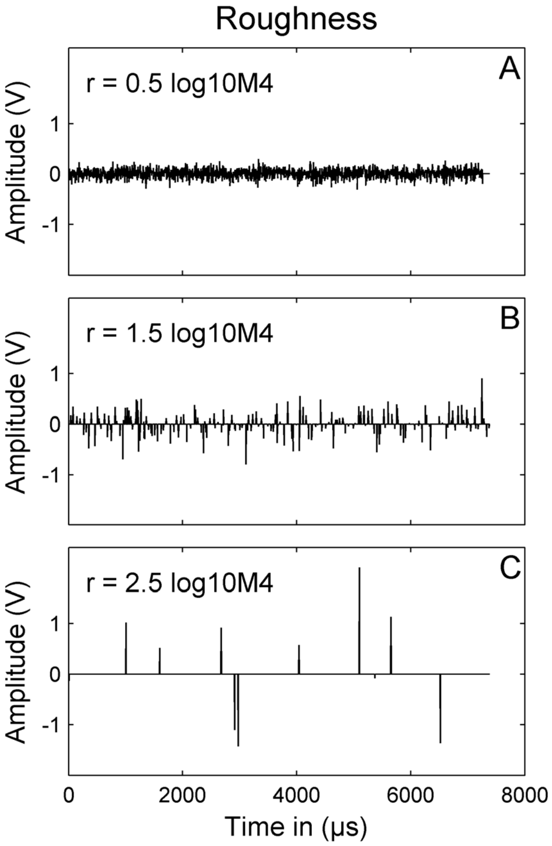


Figure 39: Roughness examples
Depicted are three signals with the same duration (7383 µs), but with different roughness values. In panel A, B and C the amplitude in Volt is plotted as s function of time in µs with the roughness value in log10M4 varying from 0.5 to 1.5 to 2.5, respectively.

An additional parameter which should influence the bats' flight path should be the time of echo presentation. The time when the bats alter their flight path should depend on when an object is detected, meaning the delay between the echolocation call and the returning echo. The smaller the delay, the earlier a bat should correct for its flight path. We will therefore additionally vary the delay of echo presentation. Figure 40 depicts schematically the expected flight corrections in dependence of the virtual object's size and delay.

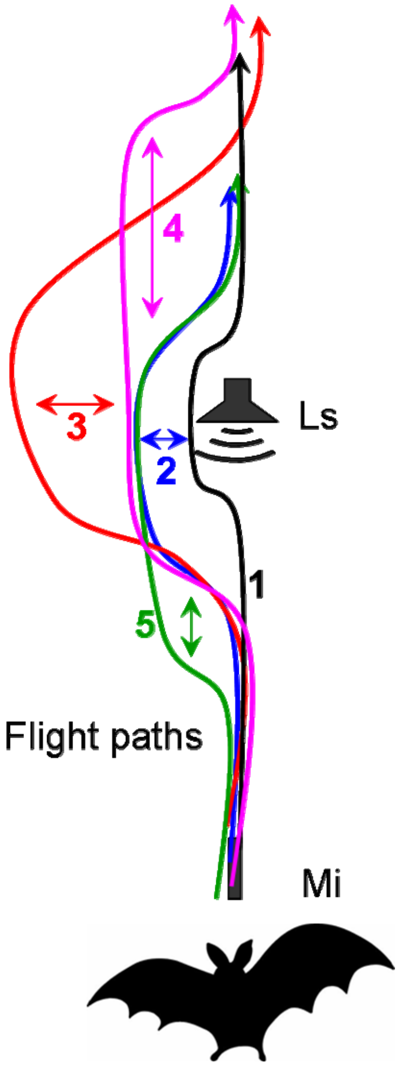


Figure 40: Theoretical flight paths
Schematic of how we expect the flight paths to be influenced. The position of loudspeaker (Ls) and microphone (Mi) are indicated (not true to scale). The flight path when nothing is presented is indicated in black and by the number 1. The bats will perceive the physical echo of the loudspeaker and will correct their flight around it. The blue line (nr 2) illustrates a flight correction when a simple reflector is presented. The virtual object will always be positioned before the loudspeaker. The red line (nr 3) reflects an increased extent of the correction when the reflector's size, meaning the amplitude, is increased. The magenta coloured (nr 4) line illustrates how the flight path correction should extend along the longitudinal axis when the spatial extent of an object, in this case the duration, is increased. The green line (nr 5) depicts how the time of flight correction should be initiated earlier when the time of presentation is decreased.

The goal of this study is to see a change in the flight paths when a virtual target compared to when none is presented, meaning the bats would try to avoid flying into the virtual object. The echoes of the virtual objects should act as objects with variable longitudinal and diagonally extensions into the animals' flight path. We expect avoidance flight manoeuvres around the speakers in relation and correlation to the depth and width dimension of the virtual object. The results will deliver more insight into how the different parameters defining an object's size play together.

Material and Methods

Animals

The experimental animals will consist of the different bat species dwelling in the Tamana-Cave (e.g., *Natalus tumidirostris, Anoura geoffroyi, Carollia perspicillata, Chilonycteris personata, Chilonycteris rubiginos, Glossophaga soricina, Phyllostomus hastatus, Pteronotus dauyi, Myotis nigricans, Desmodus rotundus*). The number of animals will depend on how many bats are active to that time and will trigger recording by flying through the camera's observation area.

Experimental setup

All experiments will be performed in front of the Tamana-Cave opening. The camera's observation field is with a distance of 2 m about 5x3 m. The setup consists of an ultrasonic microphone (CO 100k, Sanken, Germany), ultrasonic speaker (RT-3Pro, Expolinear, Berlin, Germany), a high-speed video camera (Basler A602f 1394 camera, Basler, Ahrensburg, Germany) and infra-red lightings (IR-294, Infrared Illuminator Microlight Co. Ltd, Moscow, Russia). Microphone and speaker are positioned in about 1 m height. The camera is placed on the ground between these, with about 118 cm distance to the microphone and 60 cm to the speaker. The lightings are positioned on the ground along the axis between microphone and speaker. The computer and other technical equipment are kept in a further distance to the setup, to avoid producing disturbing cues.

Stimuli

We can create several virtual echo-acoustic objects, all defined by different parameters of their IR. The variable parameters are loudness (attenuations: 0 to -24 dB, in 3 dB steps), duration (59 to 7383 µs), echo roughness (roughness values: 0.5, 1.5 and 2.5 log10M4) and time delay (0.01 to 10 ms). The roughness-duration combinations are limited due to the physical characteristic of the roughness parameter.

Technical setup

For the presentation of phantom targets the microphone records the emitted calls and sends them amplified (Quadmic, RME, Haimhausen, Germany) to a real-time processor (RX6, sampling rate 220 kHz, Tucker Davis Technologies, Gainesville, FI). The processor high-pass filters all incoming signals and convolves them with an IR, determined by parameters set by the experimenters in a computer program (Matlab 5.3, Mathworks, Natick, MA). The high-pass filter's frequency can be adjusted in the program. This echo is amplified (TDA-7560, STMicroelectronics, Genf, Switzerland) and sent to the ultrasonic speaker. Camera recording is triggered by the incoming signals on the processor. Parameters for recording duration, frame rate and trigger threshold is set by the experimenters in the program. The infra-red lighting is active during the whole session.

Procedure

Control-experiments

To be able to evaluate the flight paths while presenting phantom targets we first will have to extract the flight paths the bats produce in their natural environment during the commuting flight, meaning without positioning microphone and speaker. The camera should not influence the bats as it is positioned on the ground.

As a second step we will have to record the influence our technical setup might have on the bats' flight paths. Microphone and speaker will be placed as in the playback situation, but will not be active. A comparison of these flight paths with those acquired above will show possible avoidance manoeuvres due to the physical echoes of these real objects.

As a third control experiment we will present artificial signals (e.g. constant noise, sequences of impulses) to evaluate, whether artificial sounds uncorrelated to the bats' echolocation activity influence the animals' flight path.

Each of these control experiments should be completed during one night of recording.

Test-experiments

In a first block of test-experiments we want to study the effect several echo-acoustic parameters have on the flight path of bats.

As a first test-experiment we will simulate a simple reflector (i.e. a wall) with variable loudness as a degree for the extent of an object.

We will then simulate a more complex reflector with graded depth dimensions (roughness). Again the attenuation and additionally duration will be varied.

In a second block of test-experiments the influence of the time delay between emitted call and echo, meaning the distance between bat and object, on the bats' flight path will be studied. The bats should not react to the virtual object if the distance exceeds a certain value. The reasons for the animals to ignore this echo could be that it either is beyond a certain time window, in which the bats evaluate reflected echoes, or that it is perceived as too far away for the bats to have to avoid it. As stimuli we will use three IRs from the IRs described in the first block, but vary the time delay between call and echo.

During each session we will be recording some echolocation call sequences emitted by the bats. In a last experiment we will present these recordings filtered in the same manner as before and evaluate how the bats react to echoes uncorrelated to their own emissions.

We are planning to complete all test-experiments in a maximum of 16 nights.

Flight path extraction

The camera recordings will be evaluated every night after an experimental session. With a computer program (Matlab 5.3, Mathworks, Natick, MA) each recorded frame is subtracted from the preceding frame. This produces images only depicting changes from one frame to the next, subtracting unvarying parameters, e.g. recordings of trees. From each subtracted image the local pixel-minima and maxima are calculated. Grouping the extracted peaks into one image produces an image depicting the minimal or maximal variations in time. This will then produce one image depicting flight paths and occurring flight variations over time.

The experiments will be conducted from the beginning of December to about the December 20th. In agreement with DAAD supporting this study, the obtained results will not be elaborated and discussed in this dissertation. We hope to present and discuss the data in a later publication.

REFERENCES

Alves-Pinto, A. and Lopez-Poveda, E. A. (2005). Detection of high-frequency spectral notches as a function of level. *J.Acoust.Soc.Am* **118**, 2458-2469.

Au, W. W. and Martin, D. W. (1989). Insights into dolphin sonar discrimination capabilities from human listening experiments. *J.Acoust.Soc.Am.* **86**, 1662-1670.

Aubauer, R., Au, W. W. L., 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**, 2750-2754.

Audet, D., Krull, D., Marimuthu, G., Sumithran, S. and Singh, J. B. (1991). Foraging Behavior of the Indian False Vampire Bat, Megaderma-Lyra (Chiroptera, Megadermatidae). *Biotropica* **23**, 63-67.

Bartonicka, T. and Rehak, Z. (2005). Variability in echolocation calls of Pipistrellus pygmaeus (Chiroptera : Vespertilionidae) during search flight in different habitats. *Acta Theriologica* **50**, 145-160.

Bartsch, E. and Schmidt, S. (1993). Psychophysical frequency modulation thresholds in a FM-bat, Tadarida brasiliensis. *Hear.Res.* **67**, 128-138.

Bateman, G. C. and Vaughan, T. A. (1974). Nightly Activities of Mormoopid Bats. *J.Mammal.* **55**, 45-65.

Braaten, R. F. and Hulse, S. H. (1991). A Songbird, the European Starling (Sturnus-Vulgaris), Shows Perceptual Constancy for Acoustic Spectral Structure. *J.Comp.Psychol.* **105**, 222-231.

Bulthoff, H. H. and Edelman, S. (1992). Psychophysical support for a two-dimensional view interpolation theory of object recognition. *Proceedings of the National Academy of Sciences of the United States of America* **89**, 60-64.

Bundesen, C. and Larsen, A. (1975). Visual transformation of size. *J.Exp.Psychol.Hum.Percept.Perform.* **1**, 214-220.

Carello, C., Fitzpatrick, P., Flascher, I. and Turvey, M. T. (1998). Inertial eigenvalues, rod density, and rod diameter in length perception by dynamic touch. *Percept.Psychophys.* **60**, 89-100.

Chase, A. R. (2001). Music discriminations by carp (Cyprinus carpio). *Animal Learning & Behavior* **29**, 336-353.

Cooke, J. E., Zhang, H. M. and Kelly, J. B. (2007). Detection of sinusoidal amplitude modulated sounds: Deficits after bilateral lesions of auditory cortex in the rat. *Hear.Res.* **231**, 90-99.

Delong, C. M., Au, W. W., Lemonds, D. W., Harley, H. E. and Roitblat, H. L. (2006). Acoustic features of objects matched by an echolocating bottlenose dolphin. *J.Acoust.Soc.Am.* **119**, 1867-1879.

Edwards, H. H., Schnell, G. D., Dubois, R. L. and Hutchison, V. H. (1992). Natural and Induced Remanent Magnetism in Birds. *Current Biology* **109**, 43-56.

Eisentraut, M. (1952). Beiträge über Jagdroute und Flugbeginn bei Fledermäusen. *Bonner Zoologische Beiträge* Heft 3-4.

Esser, K. H. and Kiefer, R. (1996). Detection of frequency modulation in the FM-bat Phyllostomus discolor. *J.Comp Physiol [A]* **178**, 787-796.

Fay, R. R. (1995). Perception of Spectrally and Temporally Complex Sounds by the Goldfish (Carassius-Auratus). *Hear.Res.* **89**, 146-154.

Firzlaff, U., Schornich, S., Hoffmann, S., Schuller, G. and Wiegrebe, L. (2006). A neural correlate of stochastic echo imaging. *J.Neurosci.* **26**, 785-791.

Firzlaff, U., Schuchmann, M., Grunwald, J. E., Schuller, G. and Wiegrebe, L. (2007). Object-oriented echo perception and cortical representation in echolocating bats. *PLoS.Biol.* **5**, e100.

Fiser, J. and Biederman, I. (1995). Size Invariance in Visual Object Priming of Gray-Scale Images. *Perception* **24**, 741-748.

Fish, J. F., Johnson, C. S. and Ljungblad, D. K. (1976). Sonar Target Discrimination by Instrumented Human Divers. *J.Acoust.Soc.Am* **59**, 602-606.

Freed, D. J. (1990). Auditory correlates of perceived mallet hardness for a set of recorded percussive sound events. *J.Acoust.Soc.Am.* **87**, 311-322.

Furmanski, C. S. and Engel, S. A. (2000). Perceptual learning in object recognition: object specificity and size Invariance. *Vision Res.* **40**, 473-484.

Gillam, E. H. and McCracken, G. F. (2007). Variability in the echolocation of Tadarida brasiliensis: effects of geography and local acoustic environment. *Animal Behaviour* **74**, 277-286.

Graff, C., Kaminski, G., Gresty, M. and Ohlmann, T. (2004). Fish perform spatial pattern recognition and abstraction by exclusive use of active electrolocation. *Current Biology* **14**, 818-823.

Grassi, M. (2005). Do we hear size or sound? Balls dropped on plates. *Perception & Psychophysics* **67**, 274-284.

Green, D. M. (1996). Discrimination changes in spectral shape: Profile analysis. *Acustica* **82**, S31-S36.

Griffiths, T. D. and Warren, J. D. (2004). What is an auditory object? *Nature Reviews Neuroscience* **5**, 887-892.

Grunwald, J.-E. (2004). Echo-acoustic evaluation of real and phantom objects in phyllostomid bats. *PhD thesis, Ludwig-Maximilian-Universität Munich, Germany*.

Grunwald, J. E., Schornich, S. and Wiegrebe, L. (2004). Classification of natural textures in echolocation. *Proc.Natl.Acad.Sci.U.S.A.* **101**, 5670-5674.

Guillen-Servent, A. and Ibanez, C. (2007). Unusual echolocation behavior in a small molossid bat, Molossops temminckii, that forages near background clutter. *Behavioral Ecology and Sociobiology* **61**, 1599-1613.

Guttinger, H. R., Turner, T., Dobmeyer, S. and Nicolai, J. (2002). Melody learning and transposition in the bullfinch (Pyrrhula pyrrhula). *Journal fur Ornithologie* **143**, 303-318.

Hamm, J. P. and McMullen, P. A. (1998). Effects of orientation on the identification of rotated objects depend on the level of identity. *J.Exp.Psychol.Hum.Percept.Perform.* **24**, 413-426.

Harley, H. E., Putman, E. A. and Roitblat, H. L. (2003). Bottlenose dolphins perceive object features through echolocation. *Nature* **424**, 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. *Journal of Experimental Psychology-Animal Behavior Processes* **22**, 164-174.

Helweg, D. A., Au, W. W. L., Roitblat, H. L. and Nachtigall, P. E. (1996a). Acoustic basis for recognition of aspect-dependent three-dimensional targets by an echolocating bottlenose dolphin. *J.Acoust.Soc.Am* **99**, 2409-2420.

Helweg, D. A., Roitblat, H. L., Nachtigall, P. E. and Hautus, M. J. (1996b). Recognition of aspect-dependent three-dimensional objects by an echolocating Atlantic bottlenose dolphin. *Journal of Experimental Psychology-Animal Behavior Processes* **22**, 19-31.

Henderson, D., Salvi, R., Pavek, G. and Hamernik, R. (1984). Amplitude-Modulation Thresholds in Chinchillas with High-Frequency Hearing-Loss. *J.Acoust.Soc.Am* **75**, 1177-1183.

Holderied, M. W., Jones, G. and von Helversen, O. (2006). Flight and echolocation behaviour of whiskered bats commuting along a hedgerow: range-dependent sonar signal design, Doppler tolerance and evidence for 'acoustic focussing'. *J.Exp.Biol.* **209**, 1816-1826.

Houben, M. M. J., Kohlrausch, A. and Hermes, D. J. (2004). Perception of the size and speed of rolling balls by sound. *Speech Communication* **43**, 331-345.

Houben, M. M. J., Kohlrausch, A. and Hermes, D. J. (2005). The contribution of spectral and temporal information to the auditory perception of the size and speed of rolling balls. *Acta Acustica United with Acustica* **91**, 1007-1015.

Hubel, D. H. and Wiesel, T. N. (1962). Receptive Fields, Binocular Interaction and Functional Architecture in Cats Visual Cortex. *Journal of Physiology-London* **160**.

Hubel, D. H. and Wiesel, T. N. (1977). Functional Architecture of Macaque Monkey Visual-Cortex. *Proceedings of the Royal Society of London Series B-Biological Sciences* **198**.

Ives, D. T., Smith, D. R. and Patterson, R. D. (2005). Discrimination of speaker size from syllable phrases. *J.Acoust.Soc.Am.* **118**, 3816-3822.

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

Kalko, E. K. V. and Schnitzler, H. U. (1993). Plasticity in Echolocation Signals of European Pipistrelle Bats in Search Flight - Implications for Habitat Use and Prey Detection. *Behavioral Ecology and Sociobiology* **33**, 415-428.

Kalko, E. K. V., Schnitzler, H. U., Kaipf, I. and Grinnell, A. D. (1998). Echolocation and foraging behavior of the lesser bulldog bat, Noctilio albiventris: preadaptations for piscivory? *Behavioral Ecology and Sociobiology* **42**, 305-319.

Keil, M., Keil, A., Zahn, A. and . (2005). Die Flugwege von Wimpernfledermäusen (*Myotis emarginatus*) in Quartiernähe. *Nyctalus* **10(1)**, 61-66.

Kenny, J. S. (2007). Floor plan, environment, and fauna of Tamana caves. http://www.wow.net/ttfnc/livworld/Papers/tamana/tamana.html (retrieved 11/07).

Kingston, T., Jones, G., Akbar, Z. and Kunz, T. H. (2003). Alternation of echolocation call's in 5 species of aerial-feeding insectivorous bats from Malaysia. *J.Mammal.* **84**, 205-215.

Kober, R. and Schnitzler, H. U. (1990). Information in Sonar Echoes of Fluttering Insects Available for Echolocating Bats. *J.Acoust.Soc.Am* **87**, 882-896.

Korine, C. and Kalko, E. K. V. (2005). Fruit detection and discrimination by small fruit-eating bats (Phyllostomidae): echolocation call design and olfaction. *Behavioral Ecology and Sociobiology* **59**, 12-23.

Krajick, K. (2007). Discoveries in the dark. *National Geographic* **09**, 134-147.

Krumbholz, K. and Schmidt, S. (1999). Perception of complex tones and its analogy to echo spectral analysis in the bat, Megaderma lyra. *J.Acoust.Soc.Am.* **105**, 898-911.

Kunkler-Peck, A. J. and Turvey, M. T. (2000). Hearing shape. *J.Exp.Psychol.Hum.Percept.Perform.* **26**, 279-294.

Lakatos, S., McAdams, S. and Causse, R. (1997). The representation of auditory source characteristics: simple geometric form. *Percept.Psychophys.* **59**, 1180-1190.

Langemann, U. (1991). Frequenzunterscheidung beim Staren (*Sturnus vulgaris*): Eine Schwellenbestimmung mit verschiedenen Reizmustern. *Diploma Thesis, Technical University of Munich, Germany*.

Larsby, B. and Arlinger, S. (1998). A method for evaluating temporal, spectral and combined temporal-spectral resolution of hearing. *Scand.Audiol.* **27**, 3-12.

Larsen, A. and Bundesen, C. (1978). Size scaling in visual pattern recognition. *J.Exp.Psychol.Hum.Percept.Perform.* **4**, 1-20.

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.

Lekagul, B. and McNeely, J. A. (1988). Mammals of Thailand. *Thailand Association for the Conservation of Wildlife, Bangkok.*

Liang, L., Lu, T. and Wang, X. (2002). Neural representations of sinusoidal amplitude and frequency modulations in the primary auditory cortex of awake primates. *J.Neurophysiol.* **87**, 2237-2261.

Livingstone, M. S. and Hubel, D. H. (1984). Anatomy and Physiology of A Color System in the Primate Visual-Cortex. *J.Neurosci.* **4**, 309-356.

Lloyd-Jones, T. J. and Luckhurst, L. (2002b). Outline shape is a mediator of object recognition that is particularly important for living things. *Mem.Cognit.* **30**, 489-498.

Lloyd-Jones, T. J. and Luckhurst, L. (2002a). Effects of plane rotation, task, and complexity on recognition of familiar and chimeric objects. *Mem.Cognit.* **30**, 499-510.

Logothetis, N. K. and Sheinberg, D. L. (1996). Visual object recognition. *Annu.Rev.Neurosci.* **19**, 577-621.

Lu, T., Liang, L. and Wang, X. (2001). Temporal and rate representations of time-varying signals in the auditory cortex of awake primates. *Nat.Neurosci.* **4**, 1131-1138.

Lyzenga, J. and Carlyon, R. P. (1999). Center frequency modulation detection for harmonic complexes resembling vowel formants and its interference by off-frequency maskers. *J.Acoust.Soc.Am.* **105**, 2792-2806.

Lyzenga, J. and Horst, J. W. (1997). Frequency discrimination of stylized synthetic vowels with a single formant. *J.Acoust.Soc.Am.* **102**, 1755-1767.

Macias, S., Mora, E. C., Koch, C. and von Helversen, O. (2005). Echolocation behaviour of Phyllops falcatus (Chiroptera : Phyllostomidae): unusual frequency range of the first harmonic. *Acta Chiropterologica* **7**, 275-283.

Macpherson, E. A. and Middlebrooks, J. C. (2003). Vertical-plane sound localization probed with ripple-spectrum noise. *J.Acoust.Soc.Am.* **114**, 430-445.

Mogdans, J. and Schnitzler, H. U. (1990). Range resolution and the possible use of spectral information in the echolocating bat, Eptesicus fuscus. *J.Acoust.Soc.Am.* **88**, 754-757.

Moore, B. C. and Sek, A. (1995). Effects of carrier frequency, modulation rate, and modulation waveform on the detection of modulation and the discrimination of modulation type (amplitude modulation versus frequency modulation). *J.Acoust.Soc.Am.* **97**, 2468-2478.

Moore, B. C. and Sek, A. (1996). Detection of frequency modulation at low modulation rates: evidence for a mechanism based on phase locking. *J.Acoust.Soc.Am.* **100**, 2320-2331.

Moss, C. F. and Surlykke, A. (2001). Auditory scene analysis by echolocation in bats. *J.Acoust.Soc.Am.* **110**, 2207-2226.

Moss, C. F. and Zagaeski, M. (1994). Acoustic information available to bats using frequency-modulated sounds for the perception of insect prey. *J.Acoust.Soc.Am.* **95**, 2745-2756.

Murray, J. E., Jolicoeur, P., McMullen, P. A. and Ingleton, M. (1993). Orientation-invariant transfer of training in the identification of rotated natural objects. *Mem.Cognit.* **21**, 604-610.

Neuweiler, G. (1990). Auditory adaptations for prey capture in echolocating bats. *Physiol Rev.* **70**, 615-641.

Neuweiler, G. (2000). The Biology of Bats. *Oxford University Press, New York.*

Nowak, R. M. (1994). Walker's Bats of the World. *Johns Hopkins University Press, London*.

Oertel, D. (1999). The role of timing in the brain stem auditory nuclei of vertebrates. *Annual Review of Physiology* **61**, 497-519.

Ohl, F. W., Schulze, H., Scheich, H. and Freeman, W. J. (2000). Spatial representation of frequency-modulated tones in gerbil auditory cortex revealed by epidural electrocorticography. *J.Physiol Paris* **94**, 549-554.

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**, 1-15.

Preisler, A. and Schmidt, S. (1998). Spontaneous classification of complex tones at high and ultrasonic frequencies in the bat, Megaderma lyra. *J.Acoust.Soc.Am.* **103**, 2595-2607.

Rahmel, U. and Dense, C. (1997). Untersuchung zur Fledermausfauna im Hasbruch bei Delmenhorst. *Untersuchung im Auftrag des DLV e.V, Osnabrück.* p. 32.

Rahmel, U. and Dense, C. (1998). Untersuchungen zum räumlichen und zeitlichen Quartier- und Nahrungshabitat der Grossen Bartfledermaus (Myotis brandtii) im Hasbruch. Teilprojekt F +E-Vorhaben 'Untersuchungen zur Erhaltung der Fledermäuse in Wäldern' des Bundesamtes für Naturschutz. *Untersuchung im Auftrag des DLV e.V. Osnabrück.* p. 39.

Reimer, K. (1987). Coding of sinusoidally amplitude modulated acoustic stimuli in the inferior colliculus of the rufous horseshoe bat, Rhinolophus rouxi. *J.Comp Physiol [A]* **161**, 305-313.

Rieger, I. (1997). Flugstrassen von Wasserfledermäusen (*Myotis daubentonii*) finden und dokumentieren. *Nyctalus* **6(4)**, 331-353.

Rieger, I. and Adler, H. (1993). Weitere Beobachtungen an Wasserfledermäusen, *Myotis daubentonii*, auf Flugstassen. *Mitt natf Ges Schaffhausen* **38**, 1-34.

Rieger, I., Walzthöny, D. and Adler, H. (1990). Wasserfledermäuse, *Myotis daubentonii*, benutzen Flugstrassen. *Mitt natf Ges Schaffhausen* **35**, 37-68.

Riskin, D. (2007). Noctilio. http://www.noctilio.com/specieslist.html (retrieved 11/07).

Robinson, K. and Patterson, R. D. (1995b). The Duration Required to Identify the Instrument, the Octave, Or the Pitch Chroma of A Musical Note. *Music Perception* **13**, 1-15.

Robinson, K. and Patterson, R. D. (1995a). The Stimulus-Duration Required to Identify Vowels, Their Octave, and Their Pitch Chroma. *J.Acoust.Soc.Am* **98**, 1858-1865.

Rodda, G. H. and Phillips, J. B. (1992). Navigational Systems Develop Along Similar Lines in Amphibians, Reptiles, and Birds. *Ethology Ecology & Evolution* **4**, 43-51.

Roeder, K. D. (1963). Echoes of Ultrasonic Pulses from Flying Moths. *Biological Bulletin* **124**.

Rose, J. E., Brugge, J. F., Andreson, D. J., and Hind, J. E. (1968). Patterns of activity in single auditory nerve fibers of the squirrel monkey. *Hearing Mechanisms in Vertebrates (eds. de Reuck, A. V. S. and Knight, J.), Churchill, London.*

Sams, M. and Salmelin, R. (1994). Evidence of Sharp Frequency Tuning in the Human Auditory-Cortex. *Hear.Res.* **75**, 67-74.

Sawamura, H., Georgieva, S., Vogels, R., Vanduffel, W. and Orban, G. A. (2005). Using functional magnetic resonance imaging to assess adaptation and size invariance of shape processing by humans and monkeys. *J.Neurosci.* **25**, 4294-4306.

Schaub, A. and Schnitzler, H. U. (2007b). Echolocation behavior of the bat Vespertilio murinus reveals the border between the habitat types "edge" and "open space". *Behavioral Ecology and Sociobiology* **61**, 513-523.

Schaub, A. and Schnitzler, H. U. (2007a). Flight and echolocation behaviour of three vespertilionid bat species while commuting on flyways. *J.Comp Physiol A Neuroethol.Sens.Neural Behav.Physiol.*

Schebesch, G., Lingner, A., Firzlaff, U., Wiegrebe, L. and Grothe, B. (2007). Perception and neural encoding of size-variant human vowels in the Mongolian gerbil (*Meriones unguiculatus*). Pers. Comm.

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**, 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**, 975-988.

Schmidt, S., Türke, B. and Vogler, B. (1983). Behavioural audiogramm from the bat, *Megaderma lyra* (Geoffroy, 1810; Microchiroptera). *Myotis* **21-22**, 62-66.

Schörnich, S. and Wiegrebe, L. (2007). Phase sensitivity in bat sonar revisited. Journal of *Comparative Physiology A-Neuroethology Sensory Neural and Behavioral Physiology* . In Press; DOI 10.1007/s00358-007-0290-2

Schulze, H. and Langner, G. (1997). Periodicity coding in the primary auditory cortex of the Mongolian gerbil (Meriones unguiculatus): two different coding strategies for pitch and rhythm? *J.Comp Physiol [A]* **181**, 651-663.

Schwarz, S. and von der Emde, G. (2000b). Identification and categorization of artificial natural objects in the weakly electric fish Gnathonemus petersii. *Eur.J.Neurosci.* **12**, 202.

Schwarz, S. and von der, E. G. (2000a). Distance discrimination during active electrolocation in the weakly electric fish Gnathonemus petersii. *J.Comp Physiol [A]* **186**, 1185-1197.

Searby, A. and Jouventin, P. (2003). Mother-lamb acoustic recognition in sheep: a frequency coding. *Proceedings of the Royal Society of London Series B-Biological Sciences* **270**, 1765-1771.

Sedlmeier, H. (1992). Tonhöhenwahrnehmung beim falschen Vampir *Megaderma lyra*. *PhD thesis, Ludwig-Maximilian-Universität Munich, Germany.*

Sek, A. and Moore, B. C. (1995). Frequency discrimination as a function of frequency, measured in several ways. *J.Acoust.Soc.Am.* **97**, 2479-2486.

Sek, A. and Moore, B. C. (2000). Detection of quasitrapezoidal frequency and amplitude modulation. *J.Acoust.Soc.Am.* **107**, 1598-1604.

Simmons, J. A., Lavender, W. A., Lavender, B. A., Doroshow, C. A., Kiefer, S. W., Livingston, R., Scallet, A. C. and Crowley, D. E. (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**, 449-470.

Simon, R., Holderied, M. W. and von Helversen, O. (2006). Size discrimination of hollow hemispheres by echolocation in a nectar feeding bat. *J.Exp.Biol.* **209**, 3599-3609.

Smith, D. R. and Patterson, R. D. (2005). The interaction of glottal-pulse rate and vocal-tract length in judgements of speaker size, sex, and age. *J.Acoust.Soc.Am.* **118**, 3177-3186.

Smith, D. R., Patterson, R. D., Turner, R., Kawahara, H. and Irino, T. (2005). The processing and perception of size information in speech sounds. *J.Acoust.Soc.Am.* **117**, 305-318.

Stich, K. P. and Winter, Y. (2006). Lack of generalization of object discrimination between spatial contexts by a bat. *J.Exp.Biol.* **209**, 4802-4808.

Uiblein, F., Durand, J. P., Juberthie, C. and Parzefall, J. (1992). Predation in Caves - the Effects of Prey Immobility and Darkness on the Foraging Behavior of 2 Salamanders, Euproctus-Asper and Proteus-Anguinus. *Behavioural Processes* **28**, 33-40.

van Dinther, R. and Patterson, R. D. (2006). Perception of acoustic scale and size in musical instrument sounds. *J.Acoust.Soc.Am* **120**, 2158-2176.

van Essen, D. C. (1979). Visual Areas of the Mammalian Cerebral-Cortex. *Annual Review of Neuroscience* **2**, 227-263.

von der Emde, G. (2004). Distance and shape: perception of the 3-dimensional world by weakly electric fish. *Journal of Physiology-Paris* **98**, 67-80.

von der Emde, G. (2006). Non-visual environmental imaging and object detection through active electrolocation in weakly electric fish. *Journal of Comparative Physiology A-Neuroethology Sensory Neural and Behavioral Physiology* **192**, 601-612.

von der Emde, G. and Schwarz, S. (2000). Three-dimensional analysis of object properties during active electrolocation in mormyrid weakly electric fishes (Gnathonemus petersii). *Philosophical Transactions of the Royal Society of London Series B-Biological Sciences* **355**, 1143-1146.

von Helversen, D. (2004). Object classification by echolocation in nectar feeding bats: size-independent generalization of shape. *J.Comp Physiol A Neuroethol.Sens.Neural Behav.Physiol* **190**, 515-521.

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

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

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

von Kriegstein, K., Warren, J. D., Ives, D. T., Patterson, R. D. and Griffiths, T. D. (2006). Processing the acoustic effect of size in speech sounds. *Neuroimage* **32**, 368-375.

Warren, J. D., Jennings, A. R. and Griffiths, T. D. (2005). Analysis of the spectral envelope of sounds by the human brain. *Neuroimage* **24**, 1052-1057.

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

Wiegrebe, L. and Schmidt, S. (1996). Temporal integration in the echolocating bat, Megaderma lyra. *Hear.Res.* **102**, 35-42.

Wiegrebe, L. and Sonnleitner, V. (2007). Perceptual interaction between carrier periodicity and amplitude-modulation in the gerbil (Meriones unguiculatus). *J.Comp Physiol A Neuroethol.Sens.Neural Behav.Physiol* **193**, 305-312.

Willems, B. and Wagemans, J. (2001). Matching multicomponent objects from different viewpoints: mental rotation as normalization? *J.Exp.Psychol.Hum.Percept.Perform.* **27**, 1090-1115.

Wittekindt, A., Drexl, M. and Kossl, M. (2005). Cochlear sensitivity in the lesser spear-nosed bat, Phyllostomus discolor. *J.Comp Physiol A Neuroethol.Sens.Neural Behav.Physiol* **191**, 31-36.

Zeki, S. (1980). The Representation of Colors in the Cerebral-Cortex. *Nature* **284**, 412-418.

ACKNOWLEDGEMENTS

We thank the ‘Deutsche Forschungsgemeinschaft’ (DFG Wi 1518/8) for supplying the grant for the both studies in Part I of this thesis. We thank the ‘Volks-Wagen-Stiftung Initiative’ (VW Project 179/780) for supplying the grant for the first study in Part II of this thesis and the ‘Deutscher Akademischer Austausch Dienst’ (DAAD D/06/49876) for supporting the last study in Part II to be conducted in December.

I want to thank PD Dr. Lutz Wiegrebe for giving me the scientific opportunity to work on this thesis. He has enthusiastically guided and supported me throughout the years. I was able to learn a lot from him and gain immense experience in conducting and designing experimental studies. Additional thanks for all the cheese dinners!

Thank you RX6 for being so imaginative and numerous in your identities, working with your yous makes life interesting…

Thanks to Holger Görlitz, Sven Schörnich, Maike Schuchmann, Uwe Firzlaff and Frank Borina for their immense scientific support and helpful advice. You created a warm, friendly and welcome atmosphere in the lab and office and I really enjoyed working with you! Extra thanks go to Holger for the team work, to Sven for his computer know-how and Maike for giving me such an openhearted welcome when I started.

I want to thank all colleagues in the neurobiology group for great discussions and generally a fun time. Working at this department really was constructive and enjoyable.

I thank Silvana Siebert for her medical expertise concerning the bats and the time she always took when help was needed, and Monika Holik and Sabrina Schenk for the great care they take of all our animals and fun talks we had. Additionally thanks to Dieter Leippert for helpful suggests concerning animal keeping. Thanks of course to the guys of our workshop, they always managed to get things done in time!

Many thanks to Leonie Baier for the great working and personal time we had and will have together. Our bikes will meet again…

Special thanks to Andrea Lingner and Miriam Wolf for yummy lunches, beers and bras over t-shirts! Thailand will see us….! Additional thanks to Andrea for help winning the war over Microsoft’s formatting software (a day down the drain).

Additional thanks to the best friend in the world, Victoria Forster. Watch out, a stoat!

Big thanks to my family, without their support this all would not have been possible! Especially to my mother, who is the most courageous and strongest person. I am very proud of her.

Least, but not last, I give a special warm big thank you to Florian Schnabel, my significant other, who endures all my moods, supports me in all I do, picks me up when requested and loves me in spite of everything! (Of course the ‘Thanks’ includes the brat-cat!)

CURRICULUM VITAE

Personal:

Name: Daria Eva Irene Genzel
Address: Neufeldstr. 49
81243 München
Deutschland
Email: daria.genzel@gmx.net
Date of birth: 18.09.1979
Place of birth: Boston, Massachusetts, U.S.A.

Education:

1984–1987 Primary school in Berkeley, USA: Sierra school

1987–1989 Primary school in Munich: Fröttmaningerschule

1989–1998 Secondary school in Munich: Maximiliansgymnasiums

University study:

10/98-12/03	Studies in Biology: Ludwigs-Maximilians-Universität in Munich; Major in Zoology Minor in Neurobiology, Anthropology und Palaeontology
10/02-06/03	Thesis work for the diploma (Master) conducted at the Institute of Medical Psychology in Munich on ‘Untersuchung des geruchlichen Diskriminationsvermögens von Totenkopfaffen (*Saimiri sciureus*) für neun Enantiomerenpaare mit einer Isopropenylgruppe am chiralen C-Atom’
Since 10/04	PhD studies in the Department Biology II / Neurobiology on ‘Dynamic object analysis in echo imaging’ From 10 / 2004 – 09 / 2006 supported by a grant of the Deutsche Forschungsgemeinschaft (DFG) From 10 / 2006 – 11 / 2007 supported by a grant of the Volkswagen Stiftung Initiative (VW Stiftung) In 12 / 2007 supported by a grant of the Deutscher Akademischer Austausch Dienst (DAAD)

Occupations:

- Several temporary student jobs
- Temporary student job in a laboratory in the Department of Palaeontology
- 1998-2004 employed as a temporary salesperson at Fa. Schraudolph
- Since 08/05 employed as an assistant dog trainer at the dog school 'Artgerechte Hundeerziehung Fr. Dr. Kleist'
- Since 02/06 active as a honorary dog trainer in the animal pound ‘Tierheim München Riem’
- Since 05/07 self-employed as an animal trainer

Additonal information:

Internships:
- 2 observational studies at the zoo ‚Münchener Zoo Hellabrunn’ concerning parrots and lions
- Behavioural monitoring of bats while presenting acoustic stimuli
- Internship for animal keeping in the zoo ’Münchener Zoo Hellabrunn’
- Behavioural monitoring of squirrel monkeys with the main focus concerning biological relevant aromatic substances

Advanced training concerning animal training at a training school in Los Angeles, U.S.A

PUBLICATION LIST

Matthias Laska, Daria Genzel and Alexandra Wieser (**2005**): The Number of Functional Olfactory Receptor Genes and the Relative Size of Olfactory Brain Structures Are Poor Predictors of Olfactory Discrimination Performance with Enantiomers
Chem. Senses 30: 171–175

Daria Genzel and Lutz Wiegrebe: Time-variant spectral peak and notch detection in echolocation-call sequences in bats
Journal of Experimental Biology, in press DOI 012823

EHRENWÖRTLICHE ERKLÄRUNG

Diese Arbeit wurde von mir selbstständig und nur unter Verwendung der angegeben Hilfsmittel angefertigt.

München,

Daria Genzel

...and somewhere an evil monkey is lurking....in the closet?