

**Spatial Echo Suppression  
and  
Echo-Acoustic Object Normalization  
in Echolocating Bats**



**Dissertation  
der Fakultät für Biologie  
der Ludwig-Maximilians-Universität München**

**Maike Schuchmann**

**Munich, 22.12.2006**



**Tag der mündlichen Prüfung:**

**Freitag, der 13.04.2007**

**Gutachter:**

- 1. PD Dr. Wiegrebe**
- 2. Prof. Dr. Schuller**
- 3. Prof. Dr. Grupe**
- 4. Prof. Dr. Starck**
- 5. Prof. Dr. Foitzik**
- 6. Prof. Dr. Boshart**

*to my parents*

---

---

## Index

---

---

<b>Zusammenfassung</b>	<b>i</b>
<b>Summary</b>	<b>i</b>
<b><i>Echo suppression in bats</i></b>	<b>1</b>
<b>1.1 Introduction</b>	<b>2</b>
<b>1.2 Experimental procedures</b>	<b>10</b>
1.2.1 Animals	10
1.2.1.1 <i>Megaderma lyra</i>	10
1.2.1.2 <i>Phyllostomus discolor</i>	11
1.2.2 Experimental setup	12
1.2.2.1 <i>Megaderma lyra</i>	12
1.2.2.2 <i>Phyllostomus discolor</i>	15
1.2.3 Experimental design	16
1.2.4 Stimuli	18
1.2.4.1 Active-acoustic paradigm: echo suppression in echolocation	18
1.2.4.2 Passive-acoustic paradigm: echo suppression with external sounds	19
<b>1.3 Results</b>	<b>24</b>
1.3.1 Active-acoustic paradigm: echo suppression in echolocation	24
1.3.2 Passive-acoustic paradigm: echo suppression with external sounds	27
1.3.2.1 Echo suppression with impulses	28
1.3.2.2 Echo suppression with contact calls	29
1.3.2.3 Echo suppression with inverted contact calls	30
1.3.2.4 Echo suppression with the first syllable of the contact call	31
<b>1.4 Discussion</b>	<b>32</b>
<b>1.5 Perspectives</b>	<b>39</b>

<b><i>Object-oriented echo perception</i></b>	<b>42</b>
<b>2.1 Introduction</b>	<b>43</b>
<b>2.2 Experimental procedures</b>	<b>48</b>
2.2.1 Experimental animal	48
2.2.2 Experimental design	49
2.2.3 Impulse responses	49
2.2.4 Experimental setup	52
2.2.5 Training procedure	54
2.2.6 Computation of the significance of the data	54
2.2.7 Simulation of the classification of scaled virtual objects via auditory spectrograms in P discolor	55
<b>2.3 Results</b>	<b>59</b>
<b>2.4 Discussion</b>	<b>61</b>
2.4.1 Comparison with previous studies	61
2.4.2 Physiological correlates for auditory object normalization in echolocating bats	63
2.4.3 Towards a functional model of echo-acoustic object normalization	65
2.4.3.1 The extraction of impulse responses	65
2.4.3.2 Auditory IR normalization	66
2.4.3.2.1 Auditory IR normalization in the time domain	66
2.4.3.2.2 Auditory IR normalization in the frequency domain:	66
2.4.4 Problems and perspectives	68
<b>References</b>	<b>72</b>
<b>Figure List</b>	<b>78</b>
<b>Acknowledgements</b>	<b>80</b>
<b>Curriculum vitae</b>	<b>81</b>
<b>Publication list</b>	<b>82</b>
<b>Ehrenwörtliche Erklärung</b>	<b>83</b>

## Zusammenfassung

Die Verarbeitung von akustischen Reizen ist für alle Lebewesen in vielen Bereichen des Verhaltens einschließlich Orientierung, Jäger-Beute-Interaktionen und sozialer Kommunikation von entscheidender Bedeutung.

Echoortende Fledermäuse haben eine außergewöhnliche Fähigkeit entwickelt, um mit akustischen Reizen umzugehen. Sie sind zu einem großen Teil auf die Bewertung von Echos ihrer Echoortungsrufe angewiesen. Dieses Echoabbildungssystem ermöglicht es ihnen unter anderem, auch kleinste Beute wie Insekten zu detektieren, jagen und zu fangen sowie Hindernissen auszuweichen. Dies geschieht gewöhnlich in völliger Dunkelheit. Auch die Verarbeitung externer, echoortungsunabhängiger Schalle spielt für Fledermäuse eine große Rolle, sei es bei der passiven Beutelokalisation anhand Beute-generierter Geräusche oder bei sozialer Interaktion.

Die vorliegende Arbeit befasst sich mit zwei unterschiedlichen Aspekten der sehr komplexen echo-akustischen Situation, mit der diese außergewöhnlichen Tiere in ihrem Alltag konfrontiert sind.

Im ersten Abschnitt wird die Frage behandelt, wie Fledermäuse mit irreführender räumlicher Information in Echos umgehen.

Akustische Orientierung findet meist in hallenden Umgebungen statt. Die präzise Lokalisation von Schallquellen in natürlichen, hallenden Umgebungen ist eine wesentliche Aufgabe des auditorischen Systems. Verhaltensstudien zeigen, dass das auditorische System für die präzise Lokalisation von Schallquellen nur die räumliche Information des Schalls, welcher zuerst das Ohr erreicht, nutzt. Die räumliche Information in den zeitlich verzögert am Ohr eintreffenden Echos wird unterdrückt („Precedence effect“).

Für Fledermäuse ist der Precedence effect in der Echoverarbeitung sinnvoll, wenn es um die Lokalisation von externen, echoortungsunabhängigen Schallquellen geht. Geht es jedoch um die Verarbeitung der Echos von Echoortungsrufen scheint die Unterdrückung von Echos kontraproduktiv zu sein.

In dieser Studie wurde in einem 2AFC (two- alternative, forced- choice) Paradigma untersucht, ob und in welchem Ausmaß die echoortenden Fledermäuse *Megaderma lyra* und *Phyllostomus discolor* spontan die räumliche Information eines zweiten Echos ihrer

---

## Zusammenfassung

---

Echoortungsrufe unterdrücken. Weiterhin wurde untersucht, wie Fledermäuse mit Echos von unterschiedlichen, echoortungsunabhängigen Schallquellen umgehen.

Die Ergebnisse zeigen, dass *M. lyra* und *P. discolor* generell die räumliche Information eines zweiten Echos ihres Echoortungsrufes nicht unterdrückten. Dies war unabhängig von der zeitlichen Verzögerung zwischen ersten und zweiten Echo. Nur ein *M. lyra*-Individuum zeigte eine signifikante Unterdrückung der räumlichen Information des zweiten Echos. In einer exakten Wiederholung des Experimentes trat diese Echounterdrückung jedoch nicht erneut auf. Weiterhin konnte gezeigt werden, dass Echounterdrückung bei der Fledermaus *M. lyra* auf externe, echoortungsunabhängige Schalle begrenzt ist, die eine semantische Bedeutung für die Fledermaus tragen, wie das der Fall bei einem arttypischen Kontaktlaut ist. Abstrakte Schalle wie ein akustischer Impuls, ein zeitinvertierter Kontaktlaut oder die erste Silbe des Kontaktlautes riefen keine spontane Echounterdrückung hervor.

Diese Daten zeigen, dass Fledermäuse zwar grundsätzlich in der Lage sind, die räumliche Information in Echos zu unterdrücken; dies scheint aber nicht der ‚default‘ Mechanismus in ihrer Echoverarbeitung zu sein. Es ist denkbar, dass der Grund für dieses ungewöhnliche Fehlen des Precedence effects bei Fledermäusen möglicherweise in den sehr kurzen Integrationszeiten der auditorischen Filter im Ultraschallbereich liegt. Für die beobachtete fakultative Echounterdrückung spielt somit der Einfluss von kognitiven Komponenten, die mit dem Precedence effect assoziiert sind, eine wichtige Rolle. Diese Studie unterstreicht den Einfluss höherer auditorischer Verarbeitungsebenen für die Echounterdrückung.

Das Ziel des zweiten Abschnitts dieser Arbeit war es zu untersuchen, wie Fledermäuse die Unterschiede in Echos von unterschiedlich großen Objekten bewerten.

Echoortende Fledermäuse können dreidimensionale Objekte allein durch die Analyse der Echos ihrer Echoortungsrufe identifizieren. Allerdings können sich Objekte mit gleicher Struktur in der Größe unterscheiden. Damit Objekte trotz dieser Größenvarianz ohne Probleme erkannt werden können, muss das auditorische System eine größen-invariante, normalisierte Repräsentation der Objekte erstellen.

Diese Verhaltensstudie beschreibt die Klassifikation von Echos von komplexen virtuellen Objekten unterschiedlicher Größe. In einem Phantom-Ziel-Rückspielexperiment konnte gezeigt werden, dass die Fledermaus *P. discolor* spontan die meisten skalierten Versionen

## **Zusammenfassung**

---

zweier Objekte als das entsprechende Standardobjekt klassifiziert. In einer begleitenden elektrophysiologischen Studie wurden Populationen von kortikalen Neuronen gefunden, die diese psychophysische Leistung in einem größen-invarianten Antwortverhalten widerspiegeln. Es konnte in diesem zweiten Projekt erfolgreich gezeigt werden, dass echoortende Fledermäuse in der Tat ein Konzept von auditorischer Objektnormalisierung haben.

### Summary

The processing of acoustic cues is critical for all animals in a wide range of behaviours including orientation, predator-prey interactions and social communication. The auditory system can process these sound information with amazing precision.

Echolocating bats have developed an extraordinary ability to deal with acoustic cues. Their echo-imaging system has enabled them to detect, pursue and capture tiny prey like insects, to avoid obstacles and to interact with their environment, often in total darkness. Bats heavily rely on the evaluation of echoes for orientation and hunting. The evaluation of external, echolocation-independent sounds also plays an important role for bats, e.g. while localizing prey via prey-generated noise or for social purposes.

The current thesis addresses two different aspects of the very complex echo-acoustic situation these extraordinary animals are confronted with in their daily life.

The first approach of this thesis is concerned with the question how bats deal with misleading spatial information of echoes.

Acoustic orientation most often takes place in echoic environments. Accurate sound localization in natural, echoic environments is a vital task of the auditory system. Many behavioral studies have shown that for accurate sound localization, the auditory system relies only on the spatial information provided by the first wave front and that spatial information of the (delayed) echoes is suppressed ('precedence effect'). For a bat, this approach is also useful when localizing external, echolocation-independent sound sources, but it is in conflict with the processing of the echoes of self-generated sounds in an echolocation context.

In a two-alternative, forced choice paradigm, it is investigated whether and to what extend the echolocating bats *Megaderma lyra* and *Phyllostomus discolor* spontaneously suppress the spatial information of either a second echo of their sonar emission or echoes of different external, echolocation-independent sounds. In general, *M. lyra* and *P. discolor* did not suppress the spatial information of a second echo independent of the delay. Only one *M. lyra* showed significant echo suppression. However, this suppression could not be confirmed in an exact repetition of the experiment. Furthermore, it is shown that in the bat *M. lyra*, spatial echo suppression is restricted to an external sound which carries semantic

---

## Summary

---

meaning for the bat, in this case, a typical contact call. Abstract sounds like an acoustic impulse, a time-inverted contact call, or only the first syllable of the contact call do not induce spontaneous echo suppression.

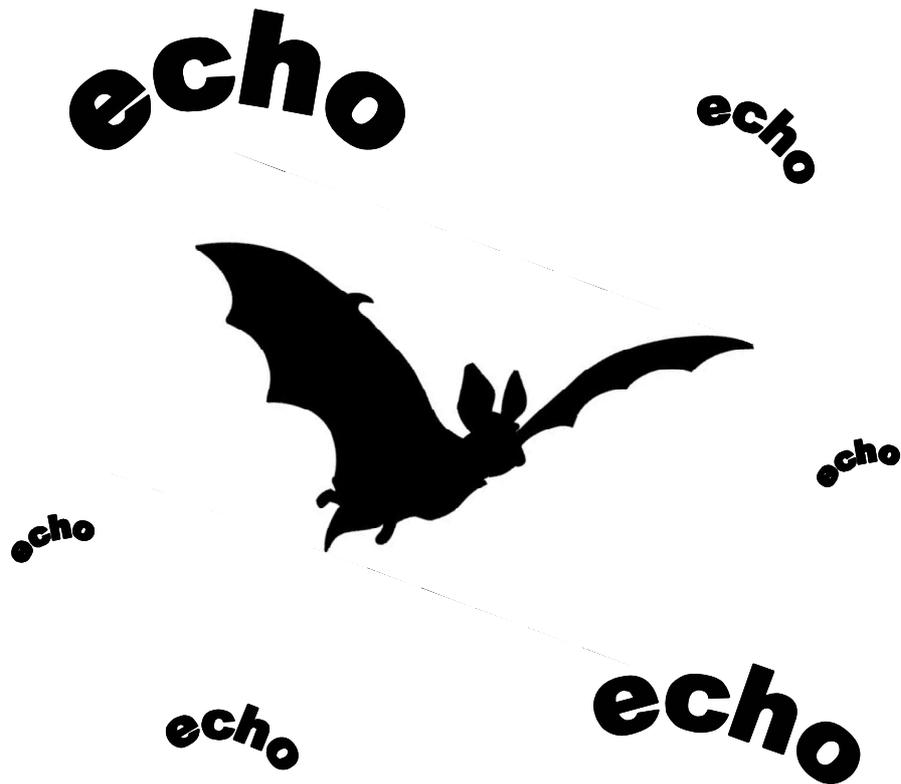
The current data indicate that while bats may be able to suppress the spatial information of echoes, this is not their default mode of auditory processing. The reason for this exceptional absence of spatial echo suppression may lie in the shorter time constants of cochlear processing in the ultrasonic frequency range and the strong influence of cognitive components associated with the precedence effect. This study emphasises the contribution of high-level semantic auditory processing to echo suppression.

The aim of the second approach was to characterize how echolocating *Phyllostomus discolor* deals with size-induced variations in echoes due to different-sized ensonified objects. Echolocating bats can identify three-dimensional objects exclusively through the analysis of acoustic echoes of their ultrasonic emissions. However, objects of the same structure can differ in size and the auditory system must achieve a size-invariant, normalized object representation for reliable object recognition.

This study describes the behavioral classification of echoes of complex virtual objects that vary in object size. In a phantom-target playback experiment, it is shown that the bat *P. discolor* spontaneously classified most scaled versions of objects according to trained standards. This psychophysical performance is reflected in electrophysiological responses of a population of cortical units received from a cooperated study, which showed an object-size invariant response. The current results indicate that echolocating bats have indeed a concept of auditory object normalization.

# Echo suppression in bats

Is there an echo in here?



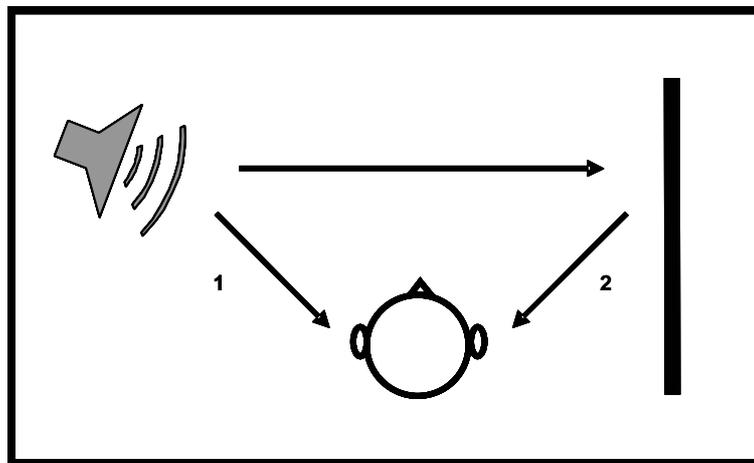
Part of this chapter is already published in The Journal of Experimental Biology (2006) under the title “Spatial echo suppression in the echolocating bats *Megaderma lyra* and *Phyllostomus discolor*” by Maïke Schuchmann, Mathias M. Huebner and Lutz Wiegerebe.

A second publication under the title “Semantic meaning determines spatial echo suppression in bats” by Maïke Schuchmann and Lutz Wiegerebe is in preparation for Current Biology.

As this chapter is a continuation of my Diploma Thesis “Lokalisationsdominanz im Rahmen der Echoortung” (2003), some data out of this thesis is presented again to provide a more detailed insight in the current topic.

### 1.1 Introduction

Humans spend most of their time in reverberant environments. Echoes are a fundamental feature of these acoustic environments. While communicating in e.g. lecture halls, in offices, public transportation and other enclosed rooms listeners are confronted with a mixture of echoes. The speaker's voice- or any other sound- reaches the listeners' ear not only directly (see Figure 1 (1)), but is also reflected by nearby surfaces like walls. These echoes reach the listeners' ear indirectly some milliseconds after the direct sound (see Figure 1 (2)). Each of these echoes carries its own set of localization cues and provides misleading information about the speaker's spatial position. Thus, evaluating confusing spatial information contained in echoes would complicate a correct speaker localization. Nevertheless, despite these existing echoes, correct localization of a speaker's position is easily possible. In the first chapter of this thesis, a mechanism for this accurate sound localization is investigated.



**Figure 1: sound travel path**

In reverberant environments, a sound (here displayed by the loudspeaker symbol) reaches the listeners' ear not only via a direct path (1), but can also be reflected by nearby objects like a wall (2). This echo reaches the listeners' ear shortly after the direct sound and carries misleading spatial information about the position of the sound source. Evaluating the spatial information in the echo would lead to a wrong localization of the sound source nearby the reflecting surface.

As mentioned above, localization of sound sources in a natural environment is complicated by echoes related to the sound source. In order to avoid localization errors the auditory system must resolve which cues come directly from the sound source and has to suppress the misleading spatial information of echoes.

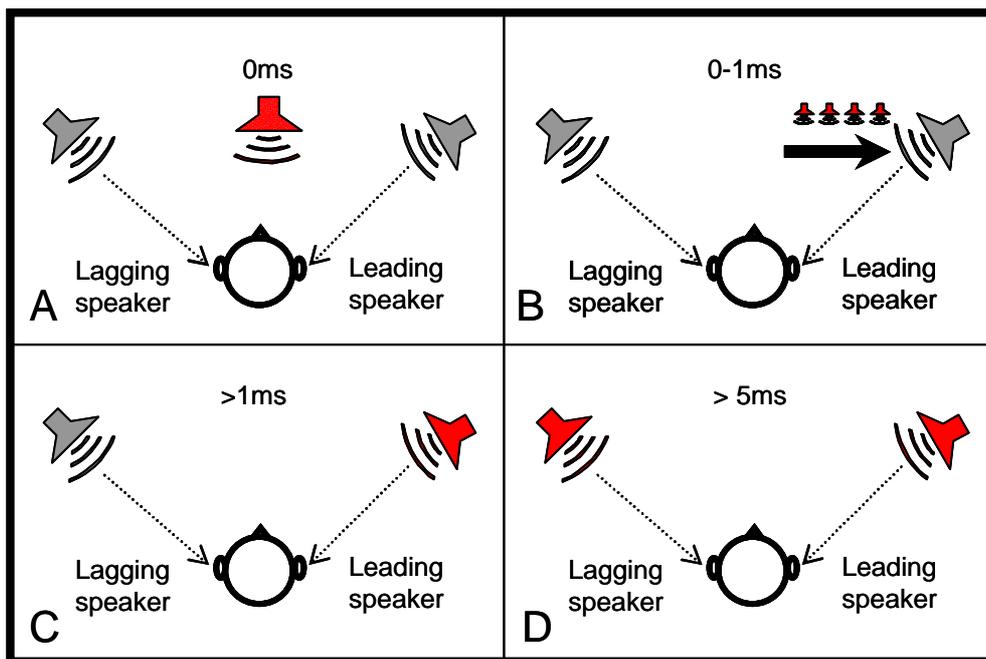
A psychophysical phenomenon relevant to sound localization in reverberant environments is commonly known as the ‘precedence effect’ (Wallach et al., 1949).

Accurate sound localization is possible for most animals because the spatial information cues from the direct and shortest path from sound source to listener carries greater perceptual weight than the spatial information in the echoes. Only the directional information of the sound which reaches the ear first dominates the perceived position of a sound source. The spatial information of echoes is suppressed.

A common behavioural paradigm used to examine the precedence effect is to simulate a direct sound source (‘lead’) and a single echo (‘lag’) with two loudspeakers in an anechoic chamber (see Figure 2). The speakers are placed in the left and right hemi field equidistant to the listener, each presenting identical sounds offset in time such that the lagging sound is delayed relative to the leading sound. These lead-lag trials usually were interspersed between trials, where only one of the two speakers emits a sound. Listeners were asked to indicate where they perceive a sound. The performance of the listeners in the trials with a single sound source serves as baseline performance. This ‘lead-lag paradigm’ allows for the investigation of the mechanisms used by listeners to resolve spatial information for sounds in reverberant environments.

When two sounds were presented simultaneously (lead-lag delay = 0 ms), they are perceptually fused and this fused percept is located midway between the two speakers (see Figure 2A). Two sounds delayed by less than 1 ms are spatially fused and are heard as a single sound located between the leading and lagging sound source position (‘summing localization’ (Blauert, 1997), see Figure 2B). For lead-lag delays between about one and five ms, the sound and its echo remain perceptually fused, but the perceived position of this fused sound image is dominated by the position of the lead (Figure 2C). This percept is referred to as ‘localization dominance’.

When lead and lag are temporally segregated by  $>5$  ms, the ‘echo threshold’ is reached and the two sounds are perceived as independent entities with their own spatial location (Blauert, 1997; Litovsky et al., 1999) (Figure 2D).



**Figure 2: Perceived position of a sound source in a lead-lag paradigm.**

Two speakers are placed in the same distance and angle in front of the listener. The presentation of the sound has an additional delay between the leading and the lagging speaker. In red: perceived position of the sound source. A and B: For lead-lag delays of 0-1 ms, the two sounds from the speakers are perceptually fused and perceived as a single sound source between the leading and lagging speaker ('summing localization'). C: For lead-lag delays > 1 ms, the position of the perceived fused sound source is dominated by the position of the leading speaker ('localization dominance'). D: Reaching the echo threshold (lead-lag delays > 5 ms, stimulus-dependent), the two sounds are perceived as independent entities with their own spatial location.

Echo-threshold values can vary tremendously. Echo threshold depends on both signal duration and complexity. They are much shorter for brief stimuli such as clicks than for longer duration stimuli like noise and ongoing complex stimuli such as running speech (up to 50 ms and more). For example, echo threshold for clicks is reached at lead-lag delays around 2 ms. For speech, echo threshold is around 20-50 ms. The echo thresholds for sinusoids of 0.25, 0.5 and 2 kHz vary, dependent on filtering and duration, between 5 and 100 ms (for review see Blauert (1997)).

Behaviourally, the precedence effect has not only been investigated in humans, but also in several other species.

Tollin and Yin (2003) investigated the precedence effect in cats (*Felis catus domesticus*). They used the search coil technique to measure the eye positions of cats which were trained to saccade to the apparent locations of sounds. Brief broadband stimuli were presented from two locations, with a delay between their onsets; the delayed sound meant to simulate a single reflection. The cats first exhibited summing localization for delays  $< \pm 400 \mu\text{s}$ . They also showed localization dominance for delays from  $400 \mu\text{s}$  to about 10 ms. Finally, the echo threshold was reached for delays  $> 10$  ms. These data reveal that cats experience the precedence effect phenomena similarly to humans.

Cranford (1982) tested cats in a classical lead-lag paradigm with impulses with a frequency spectrum from 500-16000 Hz. The cats were trained to release foot levers corresponding to the side of the sound. The animals showed localization dominance for lead-lag delays of 0.5 to 2 ms.

The albino rat (*Rattus norvegicus*) has also shown behavioural correlates to the precedence effect: Kelly (1974) showed that localization dominance occurred when either click in a pair was delayed or reduced in intensity over a certain range. Hoeffding and Harrison (1979) varied the time, intensity, or both of a lagging click in a pair of clicks and also found that localization dominance occurred in rats at certain interstimulus delays.

Ashmead et al (1986) investigated the development of auditory localization in a litter of dogs. The pups' orienting responses to dog vocalizations presented from loudspeakers 90 degrees to each side were recorded. Sounds were presented in the classical lead-lag configuration as either a single source or a lead-lag presentation. Localization began around 16 days after birth, for single-source sounds. Single-source sounds were localized earlier during development than precedence-effect sounds. Nevertheless, the dogs showed precedence effect.

Babushkina and Poliakov (2001) tested the localization of a sum of acoustic signals by two northern fur seals (*Callorhinus ursinus*) in air depending on sound parameters. This was investigated by using the method of instrumental conditioned reflexes with food reinforcement. It was found that fur seals showed the precedence effect in a similar time range as humans: localization dominance was found for clicks at a delay from around one ms up to six ms; for tonal pulses up to 11 ms.

Echo suppression was also examined in birds. Dent and Dooling (2003a) used behavioural methods to measure the discrimination performance of budgerigars (*Melopsitta undulatus*) of a stimulus presented between + and - 90 degrees azimuth with a delay (left-right), from the same two stimuli presented with the opposite delay (right-left). At short delays, where humans experience summing localization, budgerigars have difficulty discriminating between the two presentation types. With increasing delays, budgerigars showed improved discrimination performance. At even longer delays, where echo thresholds are found in humans, discrimination performance worsens again. These results show that budgerigars exhibit the phases of the precedence effect similar to humans. Similar results were found in two additional bird species: canaries (*Serinus canaria*) and zebra finches (*Taeniopygia guttata*), using a discrimination paradigm (Dent and Dooling, 2004).

Spitzer and Takahashi (2006) tested the accuracy and precision with which the barn owl (*Tyto alba*) turns its head toward sound sources under conditions that evoke the precedence effect in humans. Stimuli consisted of 25-ms noise bursts emitted from two sources, separated horizontally by 40 degrees, and temporally by 3-50 ms. At delays from 3 to 10 ms, head turns were always directed towards the leading source, indicating that the owls perceived localization dominance. Echo threshold was reached at a delay  $\geq 20$  ms. These findings demonstrate that also owls exhibit the precedence effect.

Furthermore, the precedence effect was examined in insects: Wytttenbach and Hoy (1993) demonstrated the precedence effect in the Polynesian field cricket (*Teleogryllus oceanicus*). A standard two-pulse paradigm with a directional escape response to ultrasound was used. The crickets showed summing localization for delays of 0 to 2 ms. Localization dominance was examined for delays of approximately 4 to 75 ms. At delays over approximately 75 ms, the echo threshold was reached. Precedence effect was also found in neotropical katydids (*Neoconocephalus spiza*), when they were tested via two-choice phonotaxis trials. The authors found that female katydids avoided male calls occurring during an interval beginning immediately after the onset of a leading call (Snedden and Greenfield, 1998).

Summarized, all species tested so far experience the precedence effect in nearly the same manner as humans do.

Today, little is known about the actual physiological mechanisms that mediate the precedence effect. There are some candidate structures in the auditory system which play a role for echo suppression.

Several studies showed the role of the inferior colliculus (IC) for echo suppression. It was found that IC neurons are usually sensitive to specific interaural time differences (ITD) and azimuthal locations as it is a prerequisite for the precedence effect, hence the measurements were consistently made at a neuron's "best" azimuth or ITD (for detailed review see Litovsky (1999)).

But also earlier stages of the auditory system contribute to the precedence effect. The IC contains abundant inhibitory synapses, primarily from the lateral superior olive (LSO, (Saintmarie et al., 1989)) and the dorsal nucleus of the lateral lemniscus (DNLL, e.g. Moore and Moore (1987)). Fitzpatrick et al. (1995) have postulated that the DNLL projections may play a prominent role for the observations in context of the precedence effect made in the IC. The DNLL was also described as a key structure for echo suppression by Burger and Pollak (2001). They were able to show that in the Mexican free-tailed bats (*Tadarida brasiliensis mexicana*) the response of DNLL neurons can be suppressed for tens of milliseconds due to GABAergic inhibition. This time range fits well with the times found for the precedence effect. Similar results are reported from the Mongolian Gerbil (*Meriones unguiculatus*) in a work by Pecka et al. (2006).

Furthermore, Hartung and Trahiotis (2001) emphasize the role of the auditory periphery for the precedence effect. The cochlear filters in the auditory periphery are frequency-specific band pass filters. When a sound arrives at these filters, the filters ring for a specific time depending on their specific frequency. The lower the specific frequency of the cochlear filter is the longer is its ringing time. This ringing time is the integration time of the filter. When two sounds excite the filter with a short delay, it leads to two overlapping filter responses which again lead to constructive or destructive interferences of the whole filter response. There's no discrete resolution of the sounds when they arrive at the auditory filters within their integration times. Thus, when an echo arrives at an auditory filter after the direct sound within its integration time, it won't be solved as a single sound event. Therefore, the authors argue that a large amount of psychophysical findings can be explained by the integration times of the auditory filters. The measured ringing times are in good agreement with the time ranges found in precedence studies.

But: although initial stages of echo suppression may occur in the auditory brainstem, single neuron results cannot account for all perceptual phenomena related to the precedence effect.

It is conceivable that behavioural manifestation of the precedence effect is mediated at higher levels in the auditory pathway and is therefore sophisticated. This would be consistent with studies which suggest that the auditory cortex is essential for behavioural precedence effect (e. g. (Mickey and Middlebrooks, 2005; Damaschke et al., 2005)

Note: the precedence effect is not exclusively a hard-wired phenomenon. It can also be seen as an active perceptual process that depends in part on the listeners' expectations. Thus, the perception of echoes appears to be modulated by listeners' expectations about room acoustic information contained in reflected sound (Clifton et al., 2002). Furthermore, echo suppression can break down when the lead-lag arrangement is switched ('Clifton effect' (Clifton, 1987)). It was also shown, that listeners can learn to use the directional cues available in the lagging sound (Saberi and Perrott, 1990). These are strong indicators for a facultative, situation-dependent processing of echoes, presumably executed in higher brain centres.

Although a possible neuronal substrate for echo suppression in bats is found (Burger and Pollak, 2001), yet there is still an investigation of a behavioural manifestation of the precedence effect in bats to be done.

At first sight, the precedence effect is counterproductive to the bats' ability to localize objects via echolocation. It is just the spatial information in echoes which enables the localization of objects in the dark. Bats rely heavily on the evaluation of these echoes for orientation and hunting. Each echo includes information about the properties of the environment, but there are also higher-order echoes: An ensonified object reflects an echolocation call not only directly back to the bat's ears, but also in other directions. There, the echo can be reflected once more by other objects and impinge on the bats' ear indirectly. This results in a higher-order echo which includes misleading spatial information on the object. But it contains information about the environmental properties. The perceptual suppression of the misleading information could be essential for the bats' echo-acoustic orientation and localization abilities, but can also cause a loss of essential environmental information.

Bats produce sounds not only for orientation and prey capture but also for acoustic communication. Thus, in their natural environment, they are not only confronted with higher-order echoes of their echolocation calls. Since bats live in complete darkness, they have developed complex communication calls which enable them to exchange information in the context of their various social interactions. As they live in large colonies, mostly residing in reverberant caves, bats not only have to cope with echoes of their own vocal emissions but also with echoes of communication calls from conspecifics. In this acoustic situation, the suppression of echoes with misleading spatial information of a sender would facilitate the perception of the origin of communication signals from other bats and would ease social interaction. So, while the suppression of spatial information of higher-order echoes of their own echolocation calls may cause a loss of information about the environment, a suppression of misleading spatial information in echoes of social calls might be quite beneficial for a bat. Thus, one can expect a stimulus-dependency in the echo suppression behaviour in bats.

The first part of the current thesis is designed to investigate whether and to what extent the echolocating bats *Megaderma lyra* and *Phyllostomus discolor* exhibit echo suppression behaviour consistent with the precedence effect when processing echoes of their echolocation calls. Furthermore, the echo suppression behaviour of *M. lyra* to an external sound which carries semantic meaning for the bat, in this case, a contact call, is investigated in contrast to abstract sounds like an acoustic impulse, a time-inverted contact call, or the first syllable of the contact call.

## 1.2 Experimental procedures

### 1.2.1 Animals

#### 1.2.1.1 *Megaderma lyra*

*Megaderma lyra* (Geoffroy, 1810), the false vampire bat, is an old world, tropical gleaner bat. It detects its prey both via echolocation (Schmidt et al., 2000) and based on passive listening to prey-generated rustling noises (Neuweiler, 2000). A total of seven adult *M. lyra* took part in the training: four males (Gepetto, Murcielago, Pinocchio and Vladimir) and three females (Aspodia, Avicula and Tikiri). The animals were kept together in two 12 m<sup>2</sup> rooms with free access to water. In these rooms, they were only fed on days without training sessions, i.e., only for two days after 5-day periods of training. For a balanced nutrition, they were fed with mice during the training breaks. During the training period they were only fed with mealworms for reward. The weight of the animals lay between 32 and 38 g. All animals were experienced in performing psycho-acoustic experiments in former studies (e.g. (Weissenbacher and Wiegbe, 2003; Weissenbacher et al., 2002).



**Figure 3:** *Megaderma lyra*

**1.2.1.2**            *Phyllostomus discolor*

*Phyllostomus discolor* (Wagner, 1843), the lesser spear-nosed bat, is a new world, tropical bat. The animals originated from a breeding colony in the Department Biologie II of the Ludwig-Maximilian-University in Munich. It feeds mainly on fruits, pollen and insects (Novak, 1994); the insects mainly being gleaned from the vegetation.



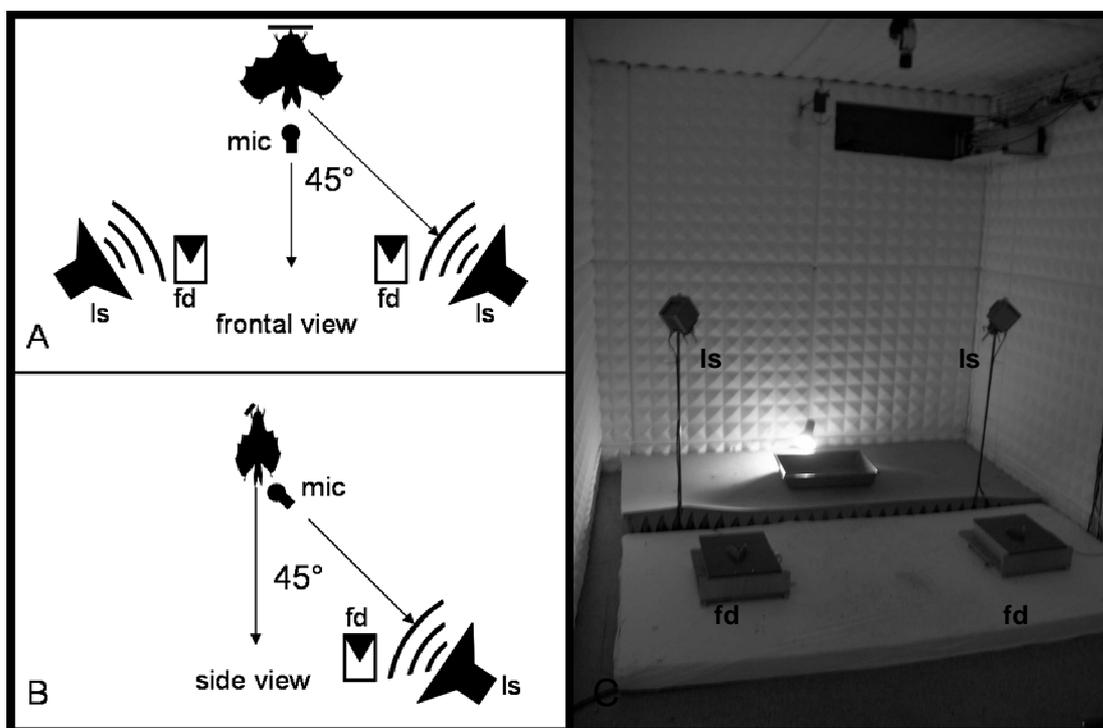
**Figure 4:** The lesser spear-nosed bat (*Phyllostomus discolor*)

Two adult *P. discolor* (both female) took part in the training. The animals were housed together in a box (80cm x 40cm x 50cm) with free access to water. In this box, they were only fed on days without training sessions, i.e. only for two days after a 5-day training period. For a balanced nutrition, the bats were fed with mealworms during the training breaks. During the training period, they were fed with banana pulp (containing of mashed banana, honey and infant milk powder) for reward. The weight of the animals lay between 30 and 36 g. All animals had no experience in performing psychoacoustic experiments when they started joining the current training.

## 1.2.2 Experimental setup

### 1.2.2.1 *Megaderma lyra*

The experiments took place in a weakly illuminated, echo-attenuated chamber (3.5 m x 2.2 m x 2.2 m). As in former studies (Weissenbacher and Wiegrebe, 2003; Wiegrebe and Schmidt, 1996) the setup consisted of a starting perch on one side of the room and two ultrasonic speakers, one in the left and one in the right hemi field. A schematic experimental setup is shown in Figure 5A and B; an image of the experimental chamber is shown in Figure 5C. A block diagram of the setup is depicted in Figure 6.



**Figure 5: Experimental setup for *Megaderma lyra***

The figure shows a cartoon with the frontal (A) and side (B) view of the setup and an image of the experimental chamber (C).

The setup consisted of a starting perch and two ultrasonic speakers (ls), each of those is associated with a feeding dish (fd) next to the speakers. The angle between the speakers in the *M. lyra* setup was 90°. This angle was positioned at 45° elevation (B). A microphone (mic) was placed in front of the bats' head to pick up their ultrasonic emissions.

---

## Experimental Procedures

---

Each speaker was associated with a feeding dish. The distance from the loudspeakers to the bats' head was 1.2 m. The angle between the speakers and the bats' head was 45°. To pick up the bats' sonar emissions, a ¼ inch microphone (Brüel & Kjaer, 4135, Naerum, Denmark) was installed 0.12 m in front of the bats' head. The emissions were amplified (B&K 2636), band-pass filtered (10-99 kHz, Krohn Hite 3342, Brockton, Massachusetts, USA) and digitized by a data acquisition board (Microstar DAP 5200, Bellevue, Washington, USA) at a sampling rate of 250 kHz. On this board, the emissions were digitally delayed and played back over either one channel (lead-only trials) or over two channels (lead-lag trials). These signals were attenuated (custom passive attenuation network), amplified (Rotel RB 976 MK II, Worthing, England) and presented via ultrasonic speakers (Matsushita EAS 10 TH 800D, Osaka, Japan) at an overall attenuation of 30 dB re. emission level. The playback delay of the lead was always 4.5 ms including the acoustic travel times.

The impulses were generated by a pulse generator (Hewlett-Packard 8012B). The contact call, the inverted contact call and the syllable of the contact call were played back via a phase 28 soundcard (TerraTec, Nettetal, Germany) which was controlled by MatLab via the interface software SoundMex. The frequency response of all setup components including speakers was flat within  $\pm$  five dB between five and 100 kHz.

On the other side of the room, opposite to the perch, the experimentator was seated, controlling the experimental procedure and the data storage with a touch screen (WES TS, ELT121C-7SWA-1, Nidderau-Heldenbergen, Germany). The experimental program was written in MatLab 5.3.

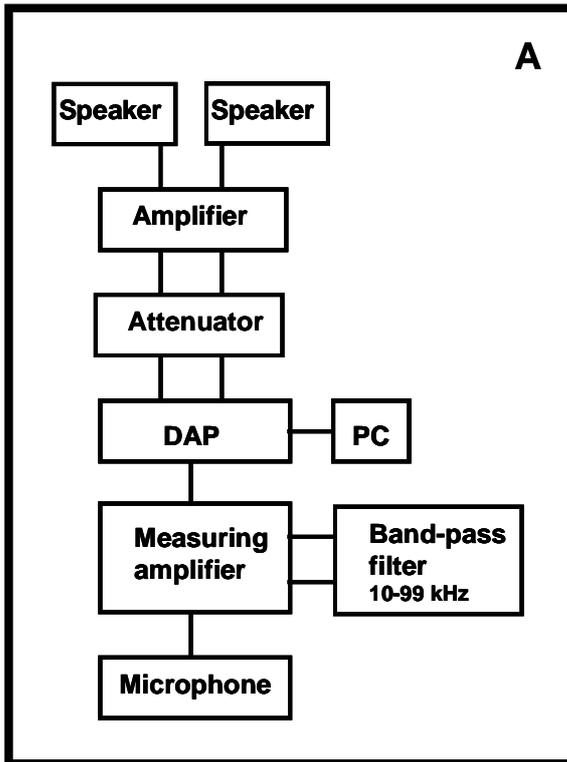
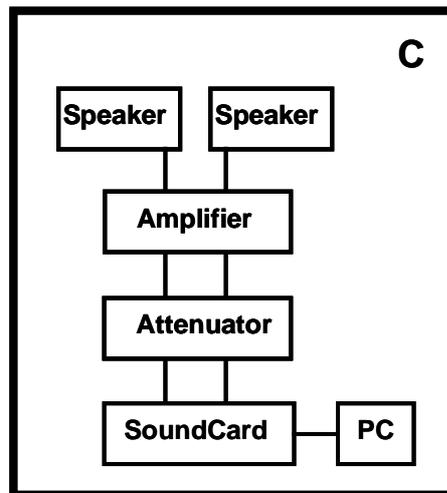
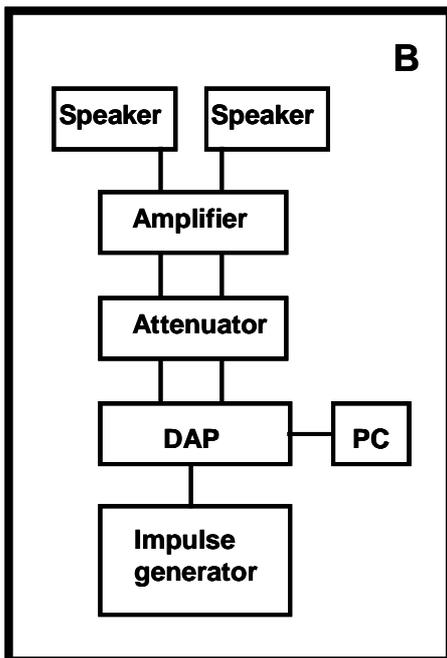


Figure 6: Block diagram of the lead-lag setup

A: Playback setup for the echolocation calls

B: Setup for the impulse presentation

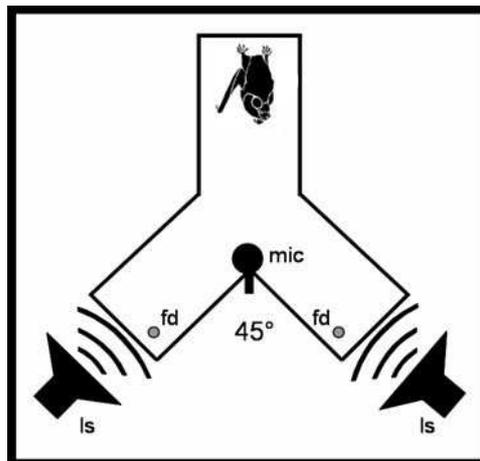
C: Setup for the presentation of the (inverted) contact call and its first syllable via soundcard



### 1.2.2.2 *Phyllostomus discolor*

The experiments took place in a horizontal Y-shaped maze in an echo-attenuated chamber. A starting perch was located at the bottom leg of the Y, and a feeder was mounted at the end of each upper leg. The angle between the legs was 45°. An ultrasonic speaker (Matsushita EAS 10 TH 800D) was placed in a distance of 15 cm in front of each upper leg (see Figure 7). Further, a ¼ inch microphone (B&K 4135) was located in the middle of the maze to pick up the bats' sonar emissions. The stimulation apparatus was identical to that used with *M. lyra* (see Figure 6A).

The experimentator was seated outside the chamber, controlling the experimental procedure via a computer interface and an infrared camera.



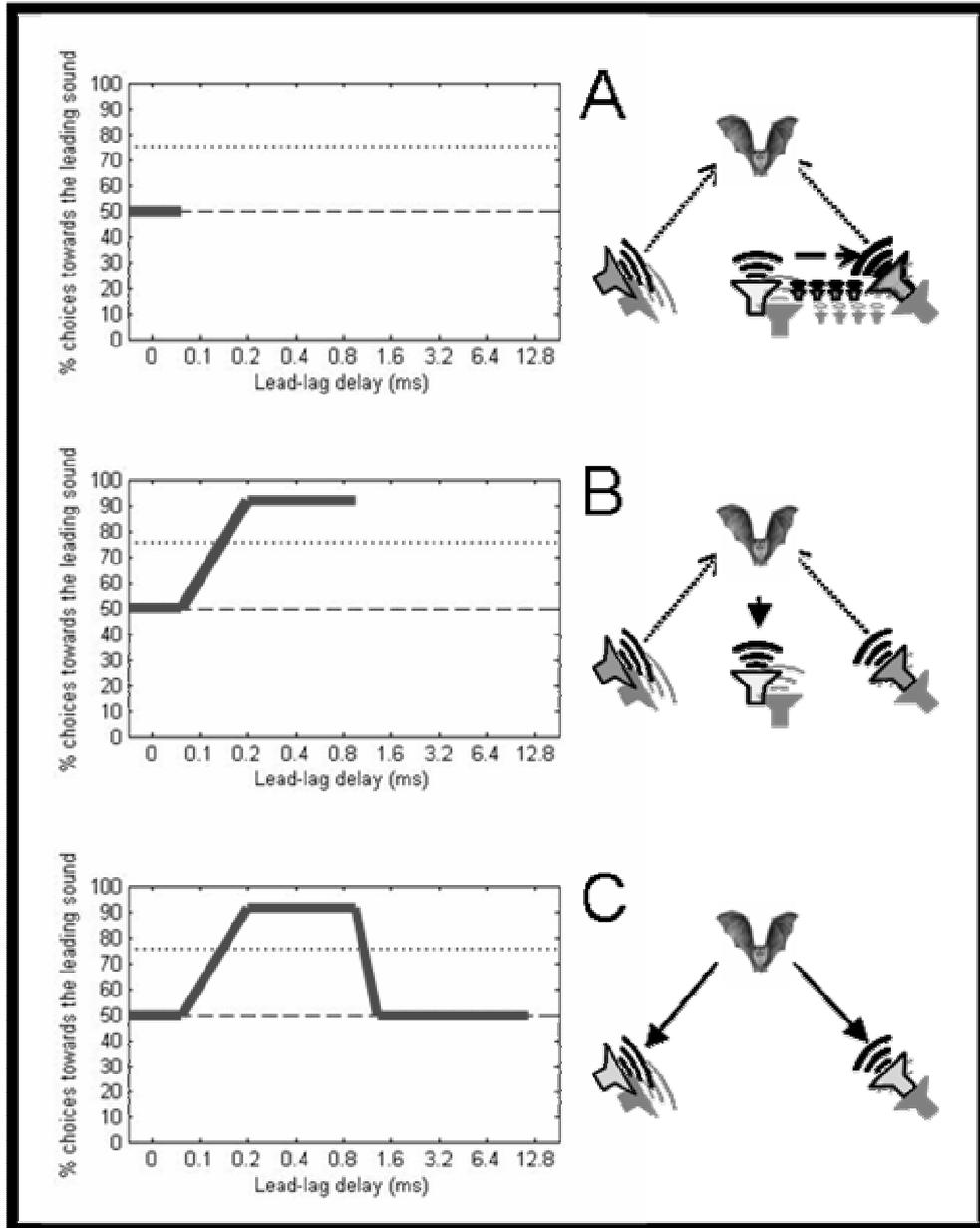
**Figure 7: Experimental setup for *Phyllostomus discolor***

The *P. discolor* setup was a Y-shaped maze (45x30 cm; wire mesh); the angle between the two legs was 45°. The inner width of each leg was 10 cm. The setup consisted of a starting position and two ultrasonic speakers (ls), each of those is associated with a feeding dish (fd) next to the speakers. A microphone (mic) was placed in front of the bats' head to pick up their ultrasonic emissions.

### **1.2.3 Experimental design**

A stimulus and its echo were simulated in a lead-lag paradigm. In this paradigm, two ultrasonic speakers were placed at the same distance and angle in each hemi field to the bats' starting position. The bats were trained in a two-alternative, forced choice experiment. In an initial training period, only one of the two speakers emitted a stimulus to train the bats to move to that speaker emitting the stimulus to get a food reward ('lead-only trials'). The presenting speaker alternated randomly. When the bats were able to solve this task with a stable performance of >85 % correct choices over several days, test trials were randomly interspersed between these lead-only trials with a probability of 25 %, such that two of eight trials were test trials. In these test trials, both speakers emitted a stimulus but the stimulus from one speaker had an additional delay of 0 ms and 0.1 ms up to 12.8 ms in doublings ('lead-lag trials'). The lead-lag delays were presented randomly across lead-lag trials. As in the lead-only trials, the speaker emitting the leading stimulus in the lead-lag trials varied randomly. In these trials the bats were free to choose any speaker to receive a food reward; i.e., the spontaneous performance of the bats was assessed. 30 lead-lag trials were collected for each of the nine lead-lag delays, therefore each bat completed 270 lead-lag trials interspersed between 810 lead-only trials. The performance was calculated as decisions for the side of the leading stimulus in percent as a function of the lead-lag delay.

In the lead-lag trials, the following behaviour would be expected: The bats were trained to move to the side of the speaker where they perceive a stimulus. If the bats experience summing localization (one fused auditory event between the two speakers) with a lead-lag delay of 0 ms, they could not show any side preference, thus the performance level would be around 50 % (Figure 8A). With increasing lead-lag delay, the bats would be expected to prefer the side of the leading speaker. If the bats perceive localization dominance, meaning the leading stimulus dominates the perceptive position of the auditory event, they would significantly (performance > 75 % for 30 trials (Miller, 1996)) move to the side of the leading stimulus (Figure 8B). Reaching the echo threshold, the bats would perceive two separate stimuli in distinct positions and would again choose one of the sides randomly (Figure 8C).



**Figure 8: Expected response behavior of echo-suppressing bats in the lead-lag paradigm**

A: With a lead-lag delay around 0 ms, the bats would perceive a fused sound source in the middle between the two speakers; therefore the bats would respond by chance and a performance around 50% would be reached.

B: With increasing lead-lag delay, the side of the leading speaker would dominate the localization of the fused sound source; therefore the bats would significantly choose the side of the leading speaker (>75% for 30 trials per condition).

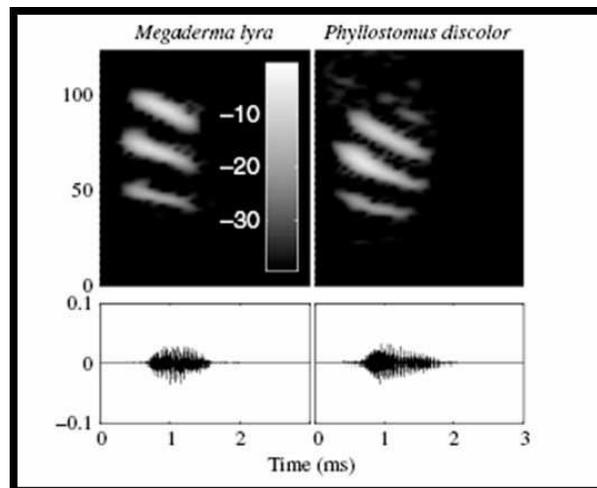
C: Reaching the echo threshold, the bats would perceive two distinct sound sources and would again choose by chance (50% performance).

## 1.2.4 Stimuli

### 1.2.4.1 Active-acoustic paradigm: echo suppression in echolocation

#### 1.2.4.1.1 Echolocation call

Many echolocation calls are short and broadband signals. Possibly, their emission triggers special attention in the bat (Metzner, 1993). Thus, it might be possible that the active emission of echolocation calls activates circuits which enhance the evaluation of returning echoes. *M. lyra* emits brief (0.5-1.5 ms), broadband, multi-harmonic echolocation calls. The fundamental frequency is frequency modulated from about 23 to 19 kHz. The strongest harmonics are the 3rd, 4th or 5th. *P. discolor* emits short (<3ms) broadband downward-modulated multiharmonic echolocation calls in the frequency range between 40 and 90 kHz. Thus, its echolocation calls are also short, multi-harmonic frequency sweeps but slightly longer than *M. lyra* calls and with a stronger dominance of lower harmonics. Exemplary echolocation calls of *M. lyra* and *P. discolor* are shown in Figure 9.



**Figure 9: Echolocation calls**

Sample of a characteristic *Megaderma lyra* echolocation call (left column) and a *Phyllostomus discolor* echolocation call (right column). The upper row shows the spectrograms (the colorbar shows the magnitude in dB); the lower row shows the oscillograms. *M. lyra* emits brief (0.5-1.5 ms), broadband, multi-harmonic echolocation calls. The fundamental frequency is frequency modulated from about 23 to 19 kHz. The strongest harmonics are the 3rd, 4th or 5th. *P. discolor* also emits brief (<3ms), broadband, multi-harmonic echolocation calls covering the frequency range between 40 and 90 kHz. The fundamental frequency is modulated from about 23.5 to 16 kHz.

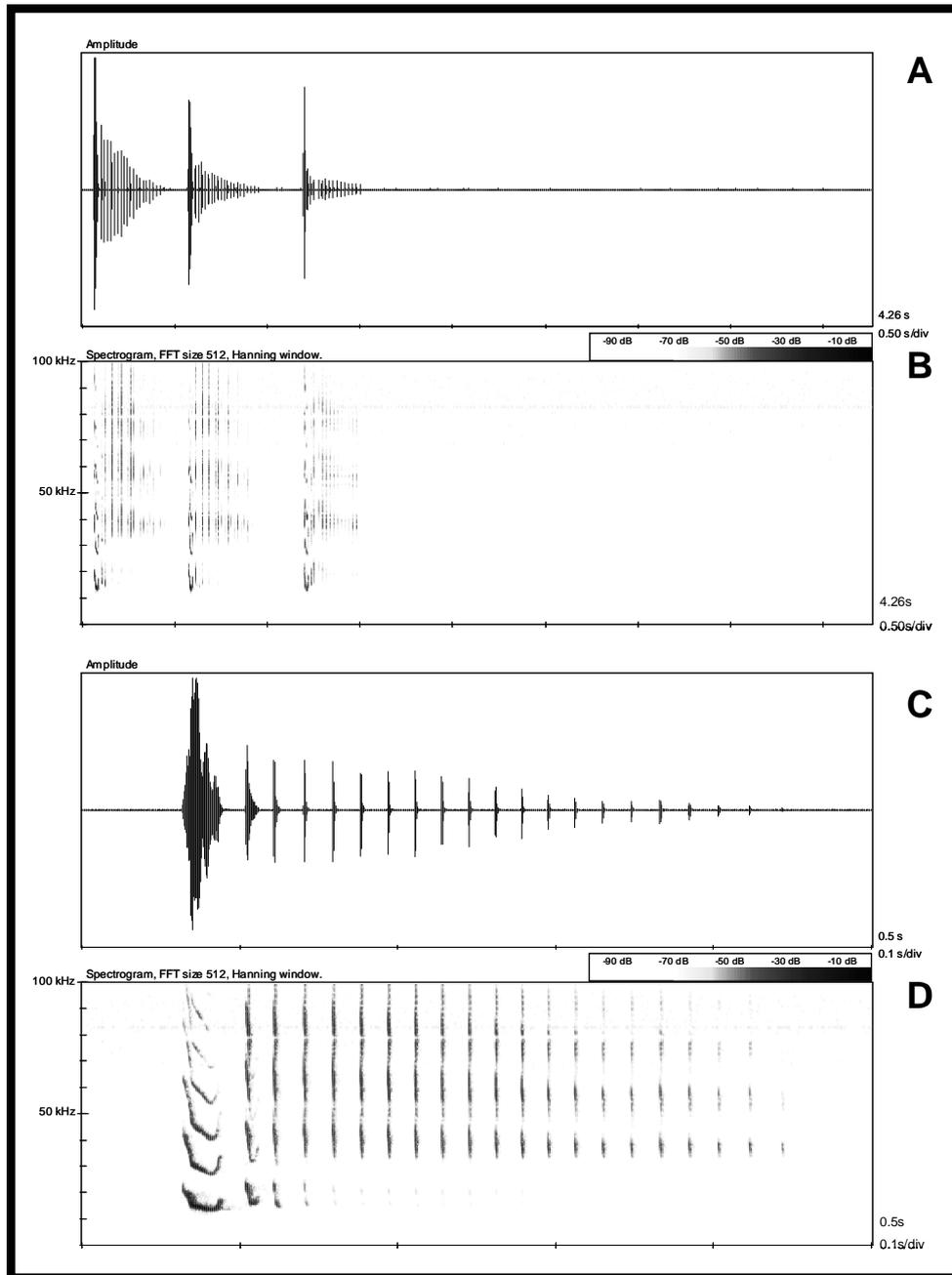
### **1.2.4.2 Passive-acoustic paradigm: echo suppression with external sounds**

#### **1.2.4.2.1 Impulses**

Impulses are often used in echo suppression studies. The current impulse had a duration of 7 $\mu$ s which results in a white spectrum in the frequency range from 5 kHz up to 100 kHz. Thus, the impulses were also short and broadband like an echolocation call, but the presentation was independent from the bats' echolocation activity. The peak- equivalent SPL of the impulses was 94 dB.

#### **1.2.4.2.2 Contact call**

The stimulus was a recording of an adult male *M. lyra* contact call (Figure 10). Corresponding to the nomenclature used for bird songs (Leonardo and Konishi, 1999), a call is composed of motifs and the motifs are comprised of syllables. The current contact call consisted of three motifs. The first syllable of each motif is a frequency-modulated, multiharmonic complex. It had a duration of 15-20 ms; the strongest harmonic was  $f_0$  with a frequency of 15 kHz. The first syllable is followed by a series of multiharmonic syllables, which are similar to echolocation calls. The presentation of the contact call was also independent from the bats' echolocation activity.



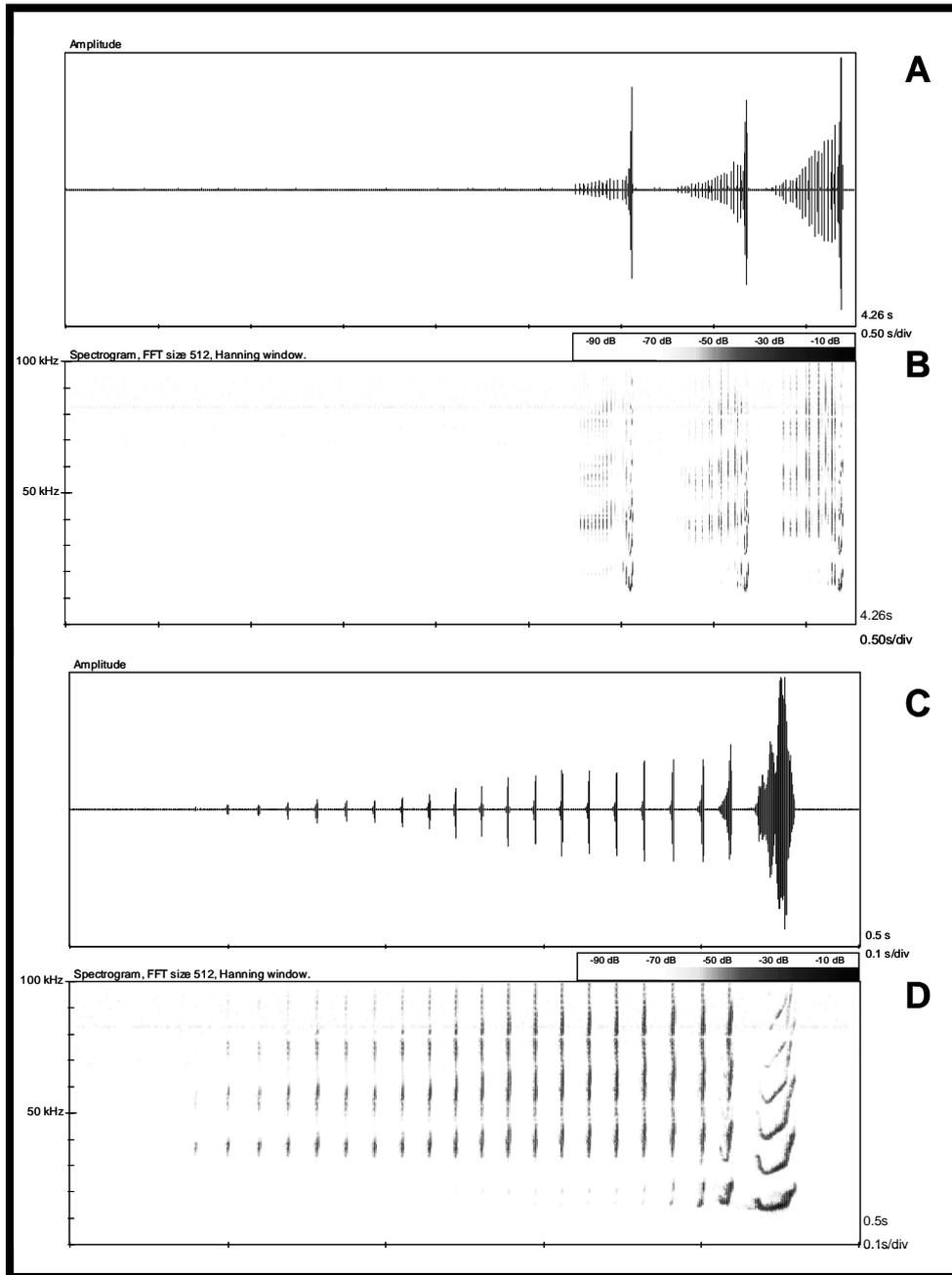
**Figure 10: *M. lyra* contact call**

A and C: Oscillogram; B and D: spectrogram. A and B show the whole contact call (duration: 4.26 s) as it was presented to the bats in a loop until they made their decision. The call consists of three motifs, each is composed of several syllables.

C and D show the first motif of the call for a more detailed impression of the fine-structure. This motif contained a frequency-modulated, multiharmonic complex (=first syllable; duration: 15-20 ms;  $f_0=15$  kHz the strongest harmonic). This complex was followed by a series of multiharmonic syllables, similar to echolocation calls (Figure 9).

### 1.2.4.2.3 Inverted contact call

The stimulus is an temporally inverted contact call (Figure 11). Therefore, it contains the same spectral content as the *M. lyra* contact call (see 1.2.4.2.2), but its temporal features are inverted. Due to this inversion the semantic meaning most likely was lost. This can be compared to a sentence which is spoken backwards.

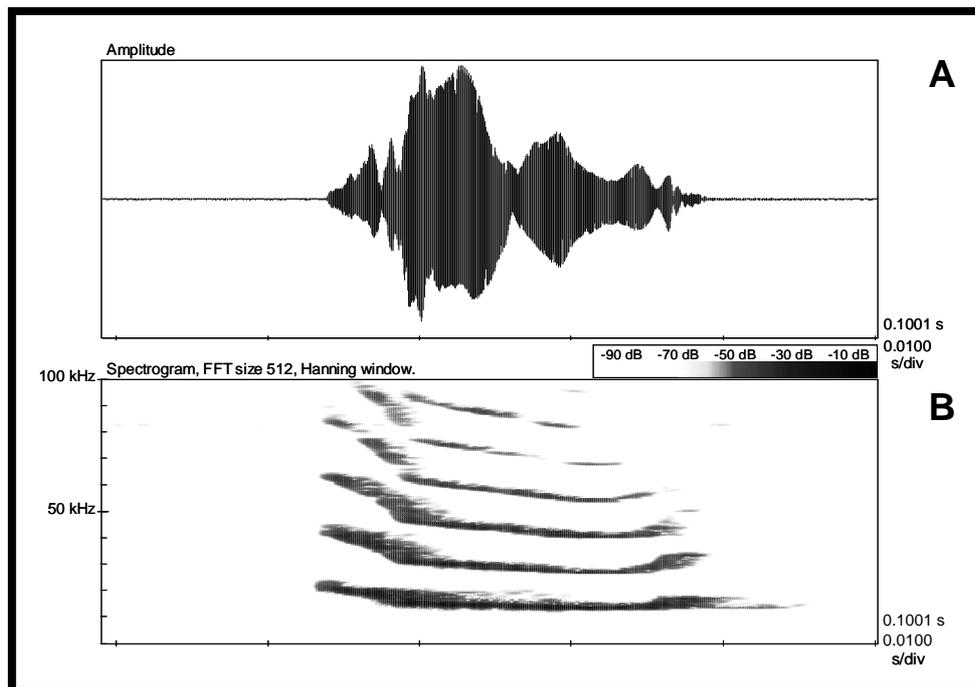


**Figure 11: Inverted *Megaderma lyra* contact call**

A and C: Oscillogram; B and D: Spectrogram. The call contained the same spectral characteristics as the contact call but is temporally inverted. A and B show the whole inverted contact call sequence as it was presented to the bats in a loop until they made their decision. C and D show the inverted first motif of the call for a more detailed impression how the call is constructed.

#### 1.2.4.2.4 First syllable of the contact call

Consisting of a longer and low-frequent element, a possible influence of the onset of the first syllable of the contact call over the social relevance of the call on the echo suppression behavior should be excluded. Therefore, the first syllable of the contact call was presented. This first syllable had a duration of around 17 ms,  $f_0$  was around 15 kHz and was the strongest harmonic (Figure 12).



**Figure 12: First syllable of the *M. lyra* contact call**

A: Oscillogram; B: spectrogram. The first syllable of the contact call had a duration of 15-20 ms;  $f_0= 15$  kHz is the strongest harmonic:

## **1.3 Results**

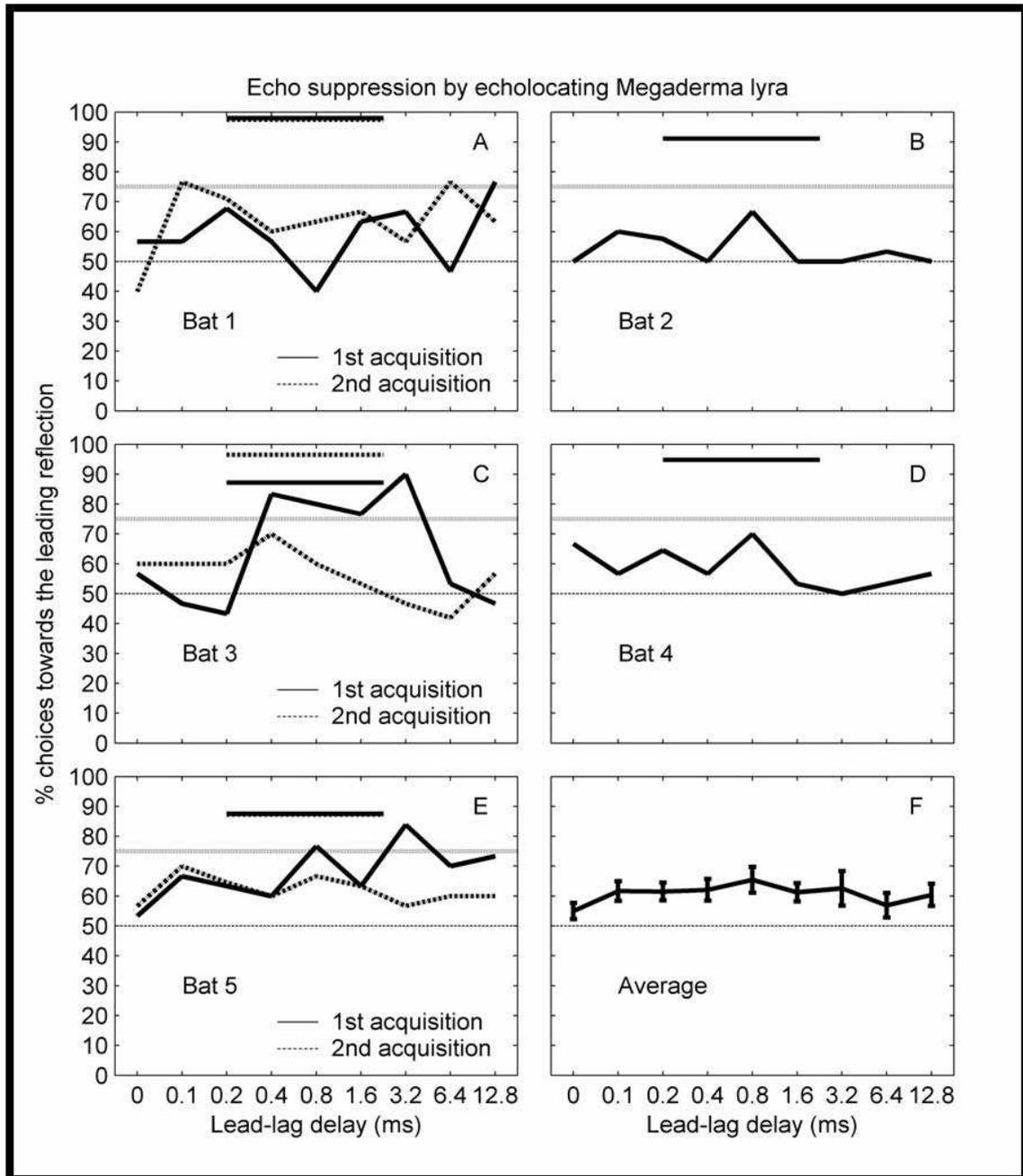
### **1.3.1 Active-acoustic paradigm: echo suppression in echolocation**

In a two-alternative forced-choice task, the bats were successfully trained via lead-only trials to move towards the speaker from which the single reflection of their echolocation call was presented. After this training period, the performance in these lead-only trials was always above 85 % correct. The data reported below were obtained while the bats maintained this good performance for the lead-only trials. The data are based on an overall of about 12,000 trials obtained from 7 animals. In three animals, a second data acquisition was done to make sure that the observed results are valid.

Figure 13 shows the spontaneous decisions of the five *M. lyra* individuals in the interspersed lead-lag trials as a function of the lead-lag delay. Individual data are plotted in panels A to E, average data are shown in panel F. Based on 30 trials per condition, significant performance is reached at about 75 % correct. In general, the bats did not significantly prefer the leading reflection of their sonar emission over the lagging reflection. While the first acquisition of Bat 3 (Panel C) indicates a significant preference for the leading reflection for lead-lag delays between 0.4 and 3.2 ms, a second acquisition could not confirm this finding. The average data show that there is only a weak trend towards a preference for the first reflection.

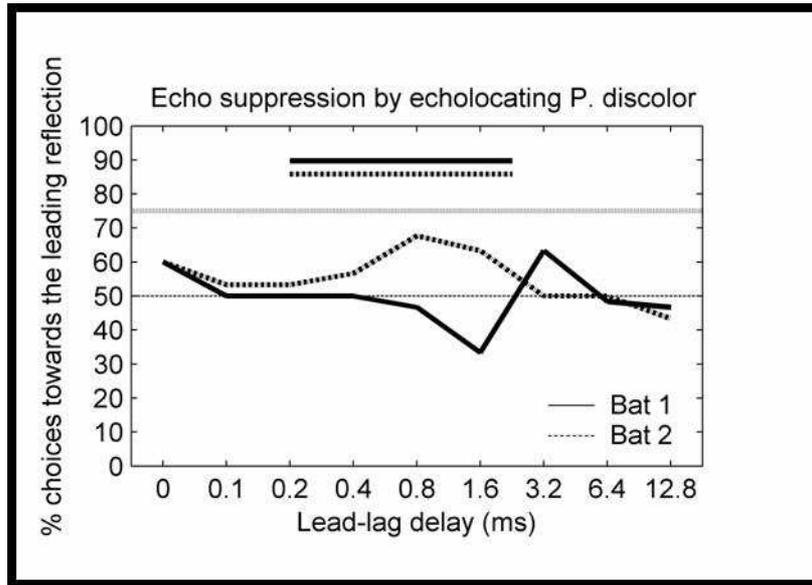
Data for the two *P. discolor* individuals are shown in Figure 14. Again, both individuals do not significantly prefer the first reflection over the second.

Taken together, the results show that both bat species generally do not reveal localization dominance for the echoes of their sonar emissions.



**Figure 11: Echo suppression in the echolocating bat *M. lyra***

The figure shows the percent decisions towards the first of two reflections plotted against the lead-lag delay. Chance performance is represented by the horizontal line at 50 %; significant performance is represented by the fine dotted line at 75 %. The bats' performance in the lead-only trials is shown by the strong horizontal lines. Panels A to E show individual data; Panel F shows the mean performance of all bats; the error bars indicate the standard deviation. The two data plots in Panels A, C, and E show a first and second data acquisition of these bats. Only one of five *M. lyra* individuals spontaneously showed significant preference (>75%) for the first of two reflections (C; first acquisition). Bat 1: Avicula; Bat 2: Gepetto; Bat 3: Murcielago; Bat 4: Pinocchio and Bat 5: Vladimyr.



**Figure 12: Echo suppression in the echolocating bat *P. discolor***

The figure shows the decision towards the first of two reflections plotted against the lead-lag delay. Data are shown in the same format as in Figure 11. The data plot represent data from the two *P. discolor* individuals. None of these showed significant spontaneous preference for the first of two reflections.

### **1.3.2 Passive-acoustic paradigm: echo suppression with external sounds**

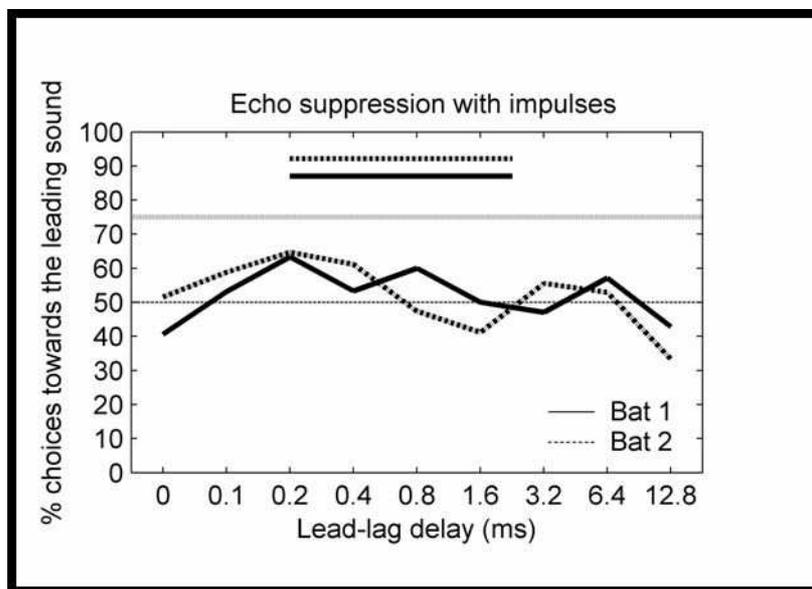
In a two-alternative forced-choice (2-AFC) task, stimulus-dependent echo suppression in passive-listening individuals of *M. lyra* was examined.

In lead-only trials, the bats were successfully trained to move towards the speaker from which a single stimulus was presented. The stimuli used to investigate stimulus-dependent echo suppression in bats were impulses, a *M. lyra* contact call, an inverted contact call and the first syllable of the contact call. After the training period, the performance in lead-only trials was always above 85 % correct. The data reported below were obtained while the bats maintained this good performance for the lead-only trials. Figure 13-16 show the spontaneous decisions for the leading stimulus of two to three *M. lyra* individuals in the interspersed lead-lag trials as a function of the lead-lag delay. Based on 30 trials per delay condition, significant performance is reached at about 75 % correct. Due to the 2-AFC-paradigm, chance level is at 50% correct.

### 1.3.2.1 Echo suppression with impulses

The echo suppression behaviour of two *M. lyra* individuals (Gepetto and Tikiri) was tested when they were confronted with leading and lagging impulses. Figure 13 shows the spontaneous decisions of the two *M. lyra* individuals in the interspersed lead-lag trials as a function of the lead-lag delay. Based on 30 trials per condition, significant performance is reached at about 75 % correct.

Both individuals did not significantly prefer the leading impulse over the lagging one.



**Figure 13: Suppression of impulse echoes in *M. lyra***

The figure shows the percent decisions towards the first of two impulses plotted against the lead-lag delay of two *M. lyra* individuals. Chance performance is represented by the horizontal line at 50 %; significant performance is represented by the fine dotted line at 75 %. The performance in the lead-only trials of the two tested bats is shown by the strong horizontal lines. None of the bats showed significant spontaneous preference for the leading impulse. Bat 1: Gepetto; Bat 2: Tikiri.

1.3.2.2 Echo suppression with contact calls

The echo suppression behaviour of three *M. lyra* individuals (Gepetto and Aspodia) was tested when they were confronted with leading and lagging contact calls. Figure 14 shows the spontaneous decisions of three *M. lyra* individuals in the interspersed lead-lag trials as a function of the lead-lag delay. Individual data is plotted in panel A-C, average data is plotted in panel D. Based on 30 trials per condition significant performance is reached at about 75 % correct. All three individuals preferred the leading contact call over the lagging one when there was lead-lag delay of 0.2 to 0.4 ms (Figure 14, A-C). This trend is also consistent when the average of the data is calculated (Figure 14 D).

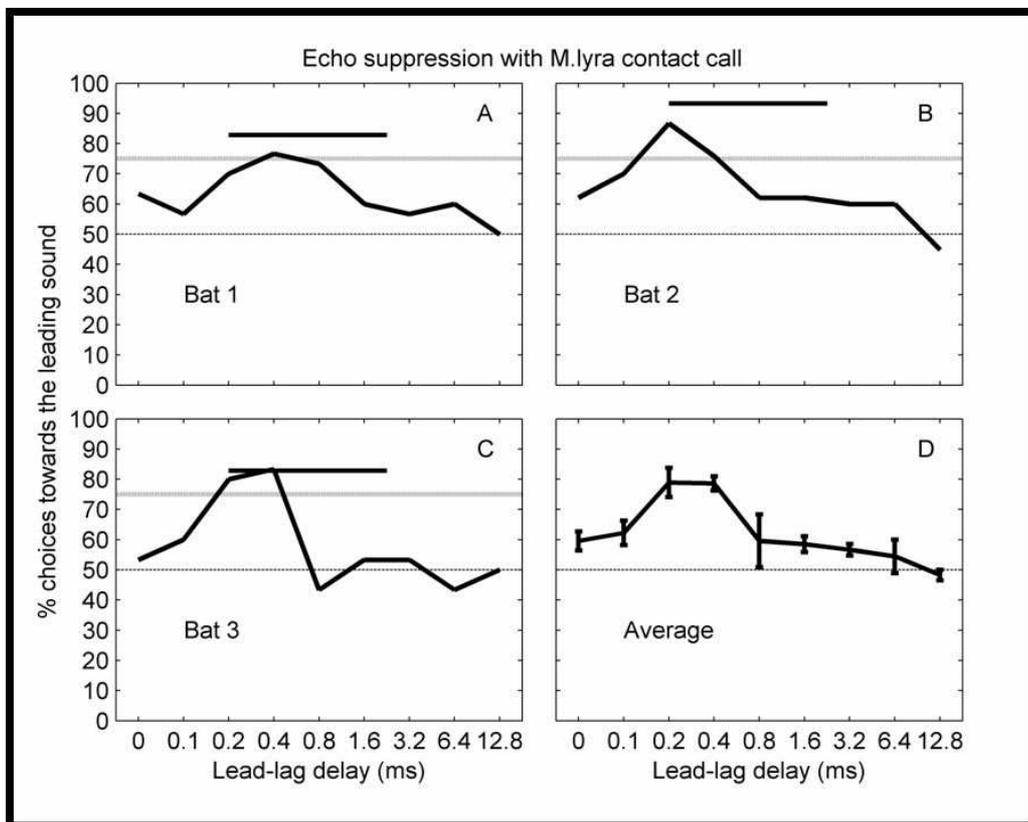


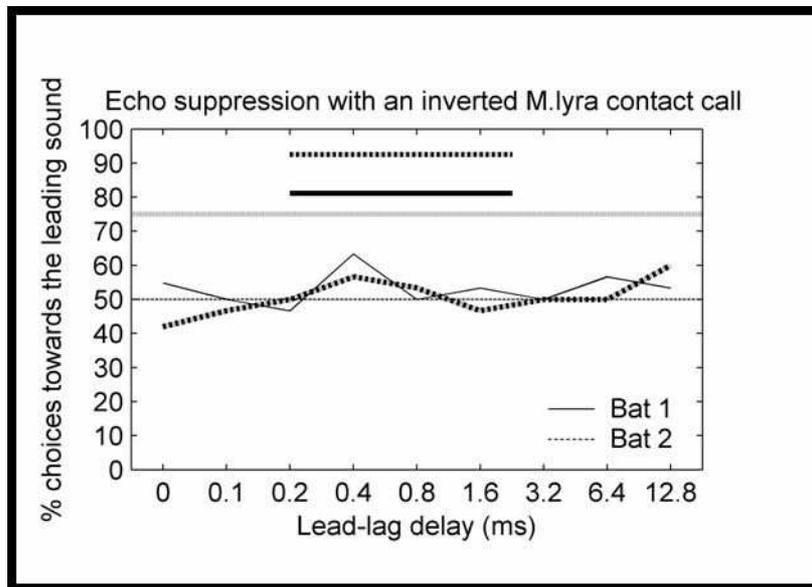
Figure 14: Suppression of contact call echoes in *M. lyra*

The figure shows the percent decisions towards the first of two contact calls plotted against the lead-lag delay of three *M. lyra* individuals. Data are shown in the same format as in Figure 13. The bats' performance in the lead-only trials is shown by the strong horizontal lines. In panel A-C the performance of bat 1 to 3 is plotted; average is plotted in D. All animals showed significant preference for the leading contact call when there was a lead-lag delay of 0.2 to 0.4 ms. This trend is also consistent in the averaged data. Bat 1: Gepetto; Bat 2: Aspodia.

**1.3.2.3 Echo suppression with inverted contact calls**

Two individuals (Gepetto and Aspodia) were confronted with leading and lagging inverted contact calls. Figure 15 shows the bats' spontaneous decisions in the interspersed lead-lag trials as a function of the lead-lag delay. Again, based on 30 trials per condition, significant performance is reached at about 75 % correct.

As observed in the impulse condition (Figure 13), both individuals did not show significant preference for the leading inverted contact call.



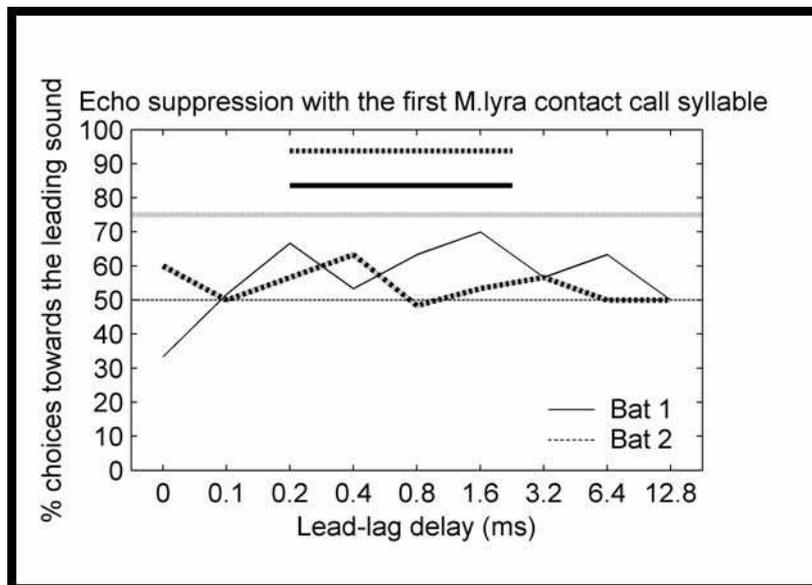
**Figure 15: Suppression of echoes of an inverted contact call in *M. lyra***

The figure shows the percent decisions towards the first of two inverted contact calls plotted against the lead-lag delay of two *M. lyra* individuals. Chance performance is represented by the horizontal line at 50 %; significant performance is represented by the fine dotted line at 75 %. The performance in the lead-only trials of the two tested bats is shown by the strong horizontal lines. None of the bats showed significant spontaneous preference for the leading inverted contact call. Bat 1: Gepetto; Bat 2:Aspodia.

### 1.3.2.4 Echo suppression with the first syllable of the contact call

To exclude an onset influence of the first syllable of the contact call on echo suppression, it was presented to two animals (Gepetto and Aspodia) in the lead-lag paradigm. Figure 16 shows the spontaneous decisions of the bats in the interspersed lead-lag trials as a function of the lead-lag delay. Based on 30 trials per condition, significant performance is reached at about 75 % correct.

Again, both individuals did not show significant preference for the leading stimulus. There is only a tendency in favour for the leading stimulus.



**Figure 16: Suppression of echoes of the first syllable of a contact call in *M. lyra***

The figure shows the percent decisions towards the first of two first syllables plotted against the lead-lag delay of two *M. lyra* individuals. Chance performance is represented by the horizontal line at 50 %; significant performance is represented by the fine dotted line at 75 %. The performance in the lead-only trials of the two tested bats is shown by the strong horizontal lines. None of the bats showed significant spontaneous preference for the leading first syllable. Bat 1: Gepetto; Bat 2:Aspodia.

## 1.4 Discussion

Only one of five tested individuals of *M. lyra* and none of *P. discolor* spontaneously showed localization dominance in the active-acoustic echolocation paradigm. This one bat spontaneously showed localization dominance in a first data acquisition period but failed to do so in a repetition of the experiment. Note that in both these acquisition periods, the animal's performance in the lead-only trial was above 90 percent correct at all times. Thus, the data from both acquisition periods are valid. In the passive-acoustic paradigm, the bats showed quite different results, depending on the stimulus. Echo suppression behaviour could only be observed when the bats were confronted with a typical contact call. Echo suppression occurred in a lead-lag delay range from 0.2 to 0.4 ms. Echoes of impulses, an inverted contact call or the first syllable of a contact call were not suppressed by the bats. Again, in all conditions the bats' performance in the lead-only trials was above 85% correct at all times; thus the data are valid.

The current results are quite different from those of all other species tested yet (see introduction). Summing localization similar to humans has been found in cats (Populin and Yin, 1998). Tested on clicks in a lead-lag paradigm, cats showed localization dominance for delays between 0.5 and two ms (Cranford, 1982). Fur seals show the precedence effect in a similar time range as humans (for clicks from around one ms up to six ms; for tonal pulses up to 11 ms; (Babushkina and Poliakov, 2001)). Other studies found similar values for rats (Kelly, 1974) and crickets (Wytenbach and Hoy, 1993). It appears that these animals experience localization dominance at nearly the same delays as human listeners do. Birds, for example budgerigars and owls, reveal localization dominance in a manner similar to humans (Dent and Dooling, 2004; Dent and Dooling, 2003a; Dent and Dooling, 2003b; Keller and Takahashi, 1996a; Keller and Takahashi, 1996b).

Why do most of the echolocating bats tested in the active-acoustic experiments show no echo suppression when confronted with two reflections of their own echolocation calls or artificial external sounds? And why is echo suppression found, when the bats were confronted with an echolocation independent typical contact call?

The current results might be related to the auditory periphery of echolocating bats: Hartung and Trahiotis (2001) showed that the integration times of the peripheral auditory filters may play an important role for echo suppression. They showed that the ringing of auditory filters in response to the lead-lag stimulation generates changes in the central auditory representation of interaural time differences which can quantitatively predict a large variety of psychophysical findings related to the precedence effect.

In the frequency region which dominates the precedence effect in human listeners (around 750 Hz, (Tollin and Henning, 1999)), the ringing times of auditory filters are in the range of several milliseconds. This is in good agreement with the time range of localization dominance. For ultrasonic hearing bats, the ringing times are very short because the auditory filters are very broad. Weissenbacher et al. (2002) and Wiegrebe and Schmidt (1996) showed that the ringing time of *M. lyra* auditory filters does not exceed about 200  $\mu$ s. Behaviourally, in this time range, both sounds contribute to the perceived sound source position (fusion). The time range for localization dominance starts at a lead-lag delay around 0.63 to 1 ms. Thus, if auditory-filter ringing were the main physiological basis of localization dominance, it cannot be expected that localization dominance is seen in bats while they are echolocating. Note that both the sonar emissions and consequently the reflections do not provide significant acoustic energy below about 40 kHz (cf. Figure 7).

In contrast to peripheral mechanisms, high level cognitive processes might also influence the precedence effect. A study in humans from Damaschke et al. (2005) showed that there is a relation between the mismatch negativity component of cortical auditory evoked potentials for double click pairs and psychophysical data of the precedence effect. In contrast, auditory brainstem responses to double click-pairs reflect stimulus features rather than the perceptive qualities of the stimulus. This indicates that the precedence effect seems to be a result of cognitive processes on higher stages of the auditory pathway rather than being purely based on mechanism located in the auditory periphery (see Hartung and Trahiotis, 2001) or in the brainstem (see Pollak et al. 2002). Recent work on echo processing and the precedence effect have been done for example by Clifton et.al (1994), Freyman et al. (1998; 1991) and McCall et al. (1998). These studies suggest that echo

processing may depend on the listeners' prior experience listening to the source and its echoes in the acoustic environment. Blauert and Col (1989) reported that if lead and lag locations are switched, the listeners' echo threshold stabilized when the switch occurred regularly, but, if switching was done irregularly, echo threshold broke down and did not recover. They recognized this as evidence for a cognitive role in the precedence effect. Also Rakerd and Hartmann (1985) proposed that localization of sound is a dynamic, interactive process that has inputs from higher cognitive levels.

Accurate sound localization is a major requirement for a functional auditory system. Thus, it is essential to suppress the misleading spatial information of echoes. However, if there is a risk that the system might suppress the sound from a source rather than from an echo, then the system stops suppressing echoes until it can redetermine which sounds are from the source and which sounds are echoes ('Clifton effect'). These observations suggest that some parts of the auditory processing involve the auditory system's ability to gain experience of its acoustic environment in order to process sound effectively (Yost and Guzman, 1996; Litovsky et al., 1997; Freyman et al., 1991; Freyman et al., 1997; Blauert, 1997). The results of these experiments indicate that for certain stimulus configurations, the spatial perception of echoes is only suppressed when their occurrence is consistent with the listeners' expectations about the sound source and the room acoustics.

In line with these psychophysical experiments, physiological studies confirm the assumption that higher cortical areas are involved in sound localization in acoustically complex environments: It was shown that unilateral ablation of the auditory cortex of cats eliminated the lead dominance in localization on the side of the lesion (Cranford et al., 1971). This demonstrates the importance of the auditory cortex for the precedence effect.

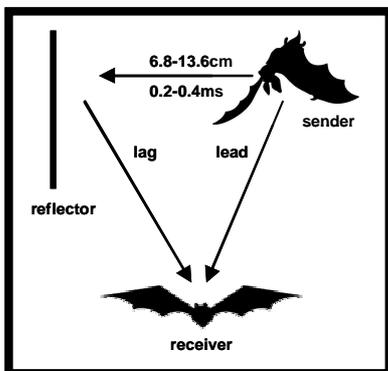
In the echolocation paradigm, the bats were confronted with reflections of their own ultrasonic emissions. Bats orient through the auditory analysis of reflections of their own vocal emission. Conceivably, bats need all the information contained in these reflections both for an auditory assessment of space and for localizing objects in this space. For this reason they don't suppress the spatial information contained in these reflections. It cannot be excluded that bats are able to suppress the spatial information of their own ultrasonic emission (as also shown in Figure 11C); but it may be dependent on the acoustic situation the bats are confronted with. They may be able to recruit localization dominance when it

is beneficial for them. Such a facultative adoption of localization dominance may underlie the different patterns of results obtained from *M. lyra* individual 3 (cf. Figure 11C). While in the first acquisition, the animal chose to adopt localization dominance, in the replication of the experiment, about six months later; it chose to switch its strategy.

In the passive-acoustic paradigm, the bats were confronted with three artificial stimuli such as impulses, an inverted contact call and the first syllable of the contact call and one typical external sound, namely a contact call. Naturally, bats have to cope with a lot of external echoes as mostly resided in reverberant environments like caves. Many of these echoes are echoes of communication calls from conspecifics. Simulating this natural situation, the tested bats were confronted with echoes of one of these communication calls—a typical contact call. In this condition, all tested bats showed echo suppression for the echoes of the contact call (Figure 14) in a lead-lag range of 0.2 to 0.4 ms. But—when the inverted contact call was presented with a preserved spectral content but temporal inverted, and thus with a loss of the semantic meaning, none of the bats showed echo suppression (Figure 15).

The delay range in which the echo suppression for the contact call was observed seems very short compared to those found in other animals.

Considering the speed of sound (340 m/s), these delays are corresponding to a distance of 6.8 cm to 13.6 cm. Thus, the reflecting surface has to be near to the sound source— in this case the calling conspecific to receive this delay range in which echo suppression operates. This is illustrated in Figure 17.



**Figure 17: Distance between sender and reflector for the observed delay range**

To get a lead-lag delay of 0.2 to 0.4 ms, the sender of the call has to be only 6.8 up to 13.6 cm away from the reflecting surface.

Considering the natural habitat of bats, which are mostly resided in cave niches, in hollow trees or are hanging in the vegetation while calling it is likely that a sender of a contact call is quite near a reflecting surface.

But is the observed delay range reasonable in the context of echo suppression? An important feature of a bat's external anatomy is the bat's ears. In contrast to human ears, the ears of *M. lyra* are not placed left and right of the head but at its top, more or less oriented forwards. Additionally, these ears are relatively big, not at least due to the species' well developed ability for passive listening. Furthermore, the bats' auditory system works in the ultrasonic range, where sounds are very directed. The relation between these big, oriented ears and the frequency range in which the bats' auditory system works provides an enormous directionality of the bats' outer ears. Thus, echoes from farther surfaces are automatically ignored due to this directionality. Within the directive hearing field, the suppression of echoes with misleading spatial information might provide a keen resolution of the position of sound sources. Therefore, echo suppression in higher delay ranges might be unnecessary in the ultrasonic range as hearing is more directed.

Echo suppression was only investigated when the bats were confronted with a typical contact call. In birds, the contact call is well investigated. For example, in the budgerigar sound repertoire, it is the most frequently used vocalization (Brittan-Powell et al., 1997). Here, the contact call is used when the flock or mated pair are in flight or about to fly, when individuals are separated from their flock, and when mates greet one another after separation (Farabaugh et al., 1994; Brockway, 1964; Wyndham, 1980).

Among *M. lyra*, the contact call also serves different social purposes. It is used for mother-pup recognition (Goymann et al., 1999), but also sent out by males to attract females. The first syllable of the contact call is thought to attract attention as it is less directional due to its lower frequency and more tonal characteristic. The following multiharmonic calls are thought to serve as a directional cue both for sender and receiver of the contact call (Dieter Leippert, personal communication). Thus, the presented contains an important meaning for the bat. It enables the bats to find its pups or to attract mating partners. Suppression of misleading information contained in echoes of this call is useful for the bats as it eases social interaction.

When presenting the first syllable, a tendency in favour of the leading stimulus was observed. In the social context, this syllable is used to call attention to the sender. Therefore, it might be useful to suppress spatial information in echoes of this first syllable. But nevertheless, confronted with echoes of the first syllable of the contact call, none of the tested bats revealed echo suppression. As it also represents the onset of the contact call, the contribution of the sound onset to the observed echo suppression can be excluded. Interestingly, neither the first syllable alone nor echolocation calls, which resemble the syllables which follow the first syllable, evoked echo suppression. Reassembled as contact call these sounds cause echo suppression.

It seems as if the complete call which contains a semantic meaning is necessary to initiate echo suppression. Thus, the semantic meaning of a sound causes echo suppression in bats. Until now, it is well known that different types of stimuli cause differences in echo suppression (for review see (Blauert, 1997; Litovsky et al., 1999)), but has never been reported that only the semantic meaning of a sound causes echo suppression.

This emphasizes the role of higher-level echo processing. The brain dynamically adjusts to the surrounding auditory environment. Mechanisms in the auditory system exist for sophisticated analysis of the environment. It can be assumed that echo suppression is dependent on in which acoustic situation the listener is confronted with echoes and of course is dependent on what acoustic stimulus is reflected. This is affirmed by findings in human studies pointing out that echoes are only suppressed when they are consistent with the listeners' expectations about the sound source and the room acoustics.

The bats not only have to process echoes with misleading spatial information, they first have to recognize the type of sound and evaluate its meaning. Confronted with meaningful sounds like tupal calls, echo suppression would facilitate the perception of the origin of communication signals from other bats and would ease social interaction. Furthermore, it can also reduce search costs associated with localizing the sender of the call. Therefore, the suppression of meaningful sounds is socially and energetically reasonable.

Note: It was observed that the precedence effect is not only important for simply localizing a sound source. It consequently also has influence on social interactions. Snedden and Greenfield (1998) described the favouring of leading mating calls in katydids (*N.spiza*). They stated that natural selection should favour the preference of leading mating calls in order to reduce search costs for females localizing a mate.

---

---

## Discussion

---

---

Furthermore, this selective mechanism causes the males to avoid placing their mating calls shortly after those of their competitors and to call at faster rates, a signal character often favoured by females. Therefore, female precedence effect will result in sexual selection pressure on males, and hence influence their signal interaction and resulting chorus structure. These findings support the theory that echo suppression of meaningful sounds may be reasonable and useful in behavioural contexts.

In summary, the current data show that echolocating bats did not reveal spontaneous localization dominance when they were trained to lateralize reflections of their ultrasonic sonar emissions. This finding can be explained by different peripheral auditory processing in the far ultrasonic frequency range and a possible strong influence of cognitive processes on the precedence effect. A singular contrary data set suggests that echolocating bats may be able to recruit localization dominance facultatively in an echo-acoustic situation in which localization dominance is beneficial for them. Furthermore, the semantic meaning of a stimulus seems to be important for echo suppression. This, again, underlines the important role of higher-level cognitive processing for echo suppression and therefore for accurate sound localization in a reverberant environment.

## 1.5 Perspectives

In the current thesis, the influence of semantic meaning of sounds on spatial echo suppression is investigated. In numerous studies it was shown, that the time ranges associated with the precedence effect and therefore the upper limit of the validity of the precedence effect –the echo threshold- varies with stimulus type (see introduction). But the detailed influence of frequency on echo suppression is unclear. Especially the comparison of the influence of frequencies in the ultrasonic range compared to those in the audio range on echo suppression has been so far ignored. This comparison, however, could give helpful hints to map out echo suppression mechanisms.

Meanwhile, a study, which I prepared and supervised, is being dealt with which tries to clarify how different frequencies affect echo suppression in the Mongolian Gerbil, *Meriones unguiculatus*.

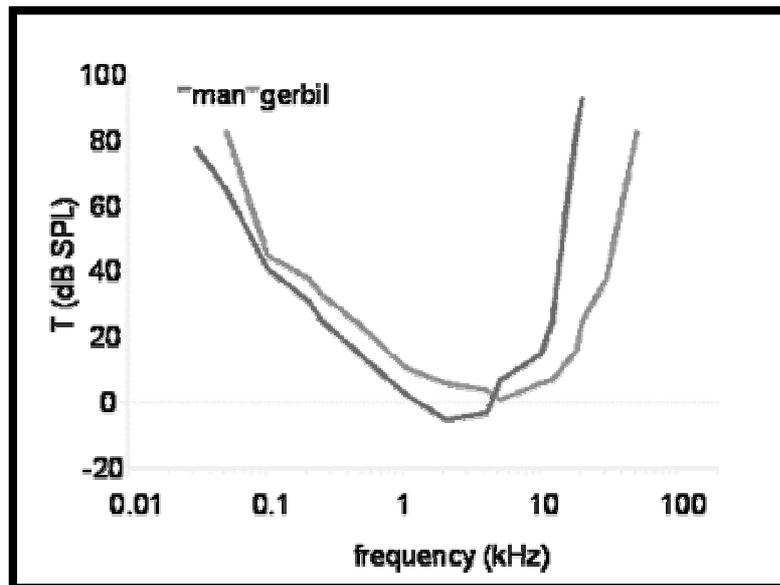
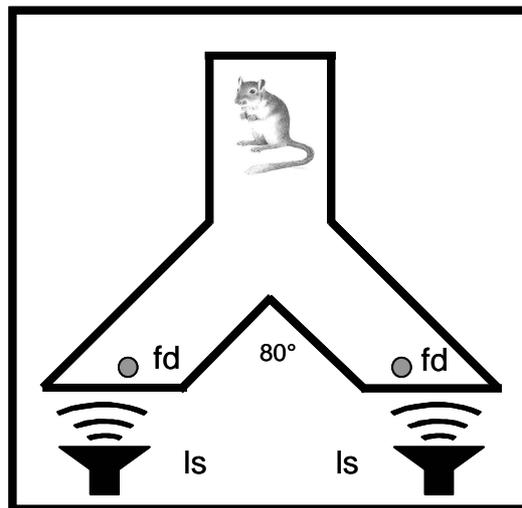


Figure 18: Audiogram of man and Mongolian gerbil

Mongolian gerbils can easily hear sounds from 100Hz up to 40 kHz. Figure 18 shows a comparison of a human audiogram with the audiogram of Mongolian gerbils. Note, that in the audio range, gerbils hear approximately well as humans (Ryan, 1976). Compared with other small rodents this well developed low frequency hearing is very unusual (Heffner et

al., 2001). Additionally, in contrast to humans, gerbils have well developed hearing abilities in the lower ultrasonic range. Due to this very broad hearing ability these animals serve as a perfect animal model for auditory observations in the audio as well in the low-frequency ultrasonic range.

Echo suppression in the Mongolian Gerbil is being examined with the same 2AFC, lead-lag paradigm as designed for this study. Figure 19 shows a depiction of the experimental setup. Short signals with a well-defined narrow frequency content, namely pip trains, serve as stimuli. Each pip train consists of four pips, each with a duration of 5 ms, and an inter-pip-interval of 100 ms, resulting in a total stimulus-length of 420 ms. The centre frequencies of the pips are ranging from 750 Hz up to 24 kHz.



**Figure 19: Experimental setup for the investigation of frequency-dependent echo suppression in the Mongolian Gerbil**

The setup is a Y-shaped maze (36x12x10 cm; wire mesh); the angle between the two legs was 80°. The setup consisted of a starting platform and two speakers (ls), each of those is associated with a feeding dish (fd) next to the speakers.

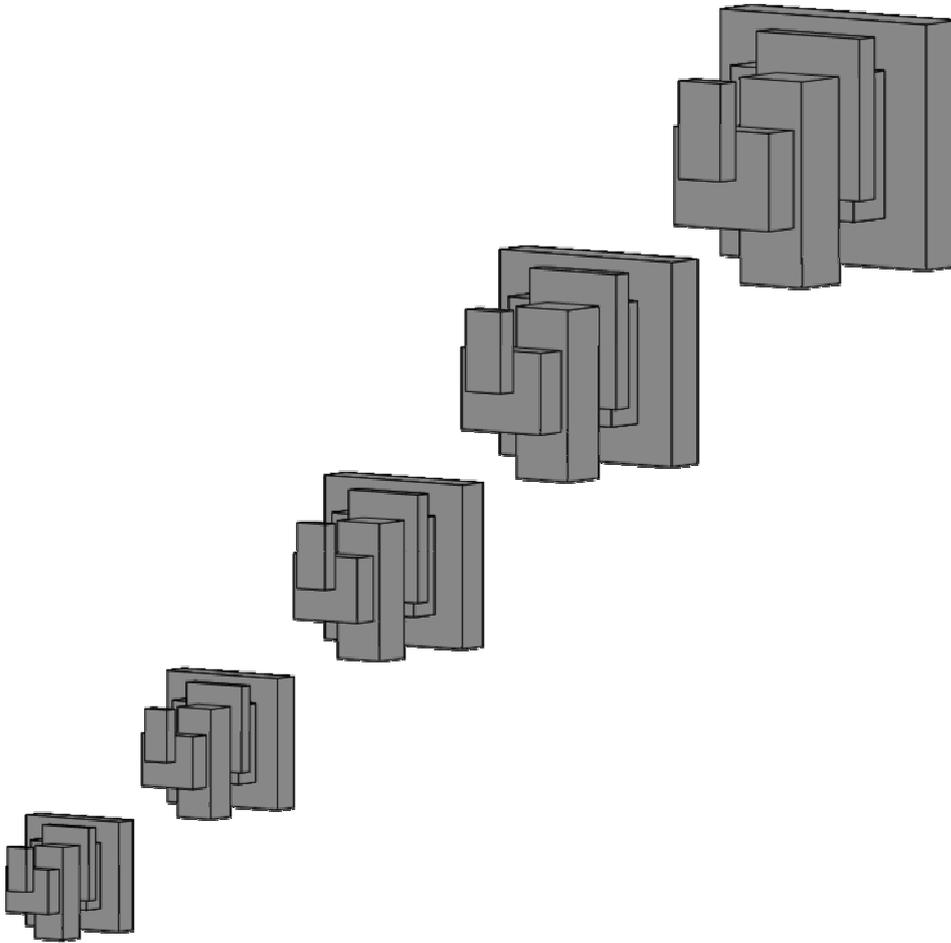
In the current 2AFC experiment with food reward, the pip trains are used to train at least four gerbils to move to the side where they perceive a pip train (lead-only trials). As soon as they have learned this localization task, test trials with lead and lag are randomly interspersed (25% probability). In these lead-lag trials, the gerbils always get a food reward; i.e. the spontaneous performance of the gerbils is assessed.

Four animals have learned this task so far. Miriam Wolff is acquiring this data for her diploma thesis. So far, the tested gerbils show echo suppression for 1500 Hz pip trains and for 6000 Hz pip trains. The lead-lag delays for the observed echo suppression are between 0.1 ms and 3.2 ms for 6000 Hz and 6.4 ms for 1500 Hz. Thus, echo threshold seems to be frequency - dependent. Further data obtained for pip trains with other centre frequencies might support the theory that the higher the frequency is the lower is the echo threshold.

This would support the influence of the auditory periphery on echo suppression as proposed by Hartung and Trahiotis (2001). Auditory integration times shorten with higher centre frequencies of the auditory filters. As these integration times play a role in echo suppression, it is reasonable that the observed time range for echo suppression also shortens with increasing frequency. It is conceivable that in higher frequency ranges this hard - wired echo suppression disappears. This fits with the observed missing hard-wired echo suppression in bats. Therefore, the next step in the current diploma thesis is to test pip- trains with 24 kHz. Due to the possible comparison between echo suppression in the audio and the lower ultrasonic range, the current study might give helpful hints for the analysis of mechanisms involved in echo suppression.

# Object-oriented echo perception

Size matters...



This chapter is in review for PLOS Biology under the title “Object-oriented Echo Perception and Cortical Representation in Echolocating Bats” by Uwe Firzlaff\*, Maike Schuchmann\*, Jan-Eric Grunwald, Gerd Schuller and Lutz Wiegrebe (\*= equally contributed).

## **2.1 Introduction**

The formation of perceptual objects from physical stimuli is an essential task of any sensory system. The recognition of these objects is indispensable for surviving. Vision is a dominant, but not the only sensory modality used for object recognition. Also in the auditory domain, object recognition plays an important role as hearing provides essential information about the environment. The auditory system is well able to analyze complex auditory scenes and thereby forming auditory objects from acoustic events that perceptually belong together (Bregman, 1990), each with its own pitch, timbre, location and loudness. Mammals have an extraordinary ability to identify auditory objects like e.g. speech or other kinds of communication calls. Without this ability, social interaction and orientation in natural scenes would be a trying task.

Reliable object recognition in both the visual and auditory domain is complicated by the variability of naturally occurring objects, for example in object size. In the visual system, the effect of object size on object recognition and underlying neural substrates has been investigated in detail (for review see Logothetis and Sheinberg (1996)). Neurophysiologically, e.g. Ito et al. (1995) found out that neurons in the inferior temporal lobe can exhibit object-size invariant responses. Also a large amount of behavioral studies were conducted concerning the influence of object size of object recognition in the visual domain. For example, Peissig et al. (2006) investigated the pigeon's (*Columba livia*) ability to generalize object discrimination performance to smaller and larger versions of trained objects. Overall, the experiments suggest that the pigeon encodes size as a feature of objects and that the representation of size is most likely logarithmic. Furmanski et al. (2000) measured the transfer of perceptual learning in object recognition. They found that learning transferred across changes in image size and their results indicate that a large amount of learning took place in object-specific mechanisms that are insensitive to image size. Thus, object recognition can easily be performed without an interfering effect of object size information in the visual domain.

The perception of size information in the auditory domain has only recently been addressed (Houben et al., 2004; Smith et al., 2005; Carello et al., 1998).

However, objects can vary substantially from one occurrence to the next. Despite this variations reliable object recognition is possible. This is possible due to a transformation, which allows the comparison of the perceived object with a generalized internal representation of known objects. The most important transformation is normalization. Normalization is required to compensate for the effect of object scale on the analysis of object structure. Without object normalization any perception would be unique and categorization would be impossible.

This becomes obvious from human speech: The same vowel pronounced by an adult and a child differs in its spectral content. However, it is perceived as the same and one can easily identify the speaker group. Smith et al. (2005) showed that the ability to judge relative size of speakers as well as recognition of vowels works well beyond the normally occurring range of speaker size. For human speech as well as communication sounds of other mammals, vocal tract length correlates with body size (Fitch, 1997) whereas the shape of the vocal tract determines the spectral envelope of the sound (e.g. vowel type).

Functionally, it is assumed that a mathematical transform which is immune to the effects of size dilation, is performed by the auditory system to segregate size information from shape information. Via this so-called two-dimensional Mellin transform (Bertrand et al., 1996) a size invariant, normalized version of the sound is achieved.

Object normalization in the auditory domain is especially important for echolocating bats which identify three-dimensional objects exclusively through the analysis of the echoes of their ultrasonic emissions. Parameters like echo duration, echo amplitude and spectral content of the echoes enable them not only to determine the position of the ensonified object in the dark, but also its structure and size.

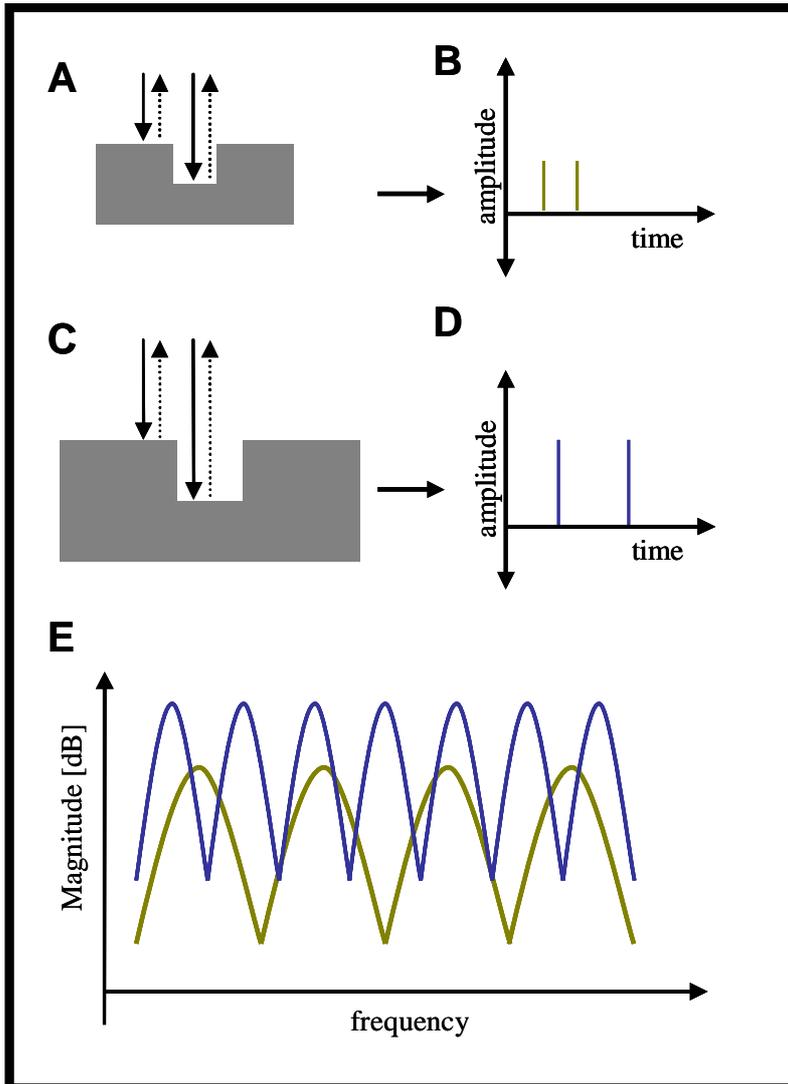
It was shown that bats use echolocation for object identification to find fruit (Kalko and Condon, 1998; Thies et al., 1998) and flowers (von Helversen and von Helversen, 2003). Schmidt et al. (2000) showed that the gleaning bat *Megaderma lyra* uses echolocation to identify and catch moving prey of different size from the ground. The bats also discriminated edible prey from dummies of the same size.

In all the above mentioned cases object normalization is crucial for compensating inevitable variations in object size. In fact, it has been shown in behavioral experiments that bats are able to discriminate simple shapes independent of their size (von Helversen, 2004).

Acoustically, an object is defined by its impulse response (IR). The IR is the sum of the reflections when an object is ensonified with a Dirac impulse, i.e., an impulse of theoretically infinite shortness and amplitude. Unlike a visual (two-dimensional) image of an object, the IR is a one-dimensional image. It defines the object mainly along the depth dimension but provides comparably less information about the contours along the dimensions height and width. The reverse is the case in vision. An impulse covers the whole range of frequencies with same amplitudes. Therefore, the IR represents the acoustic image with optimal information content. Magnitude spectra of IRs show object- specific interference patterns. This spectral composition is due to the surface composition of the ensonified object and is caused by constructive or destructive interference of the reflections.

With increasing object size there is a proportional change in both object surface area and object depth. As the strength of the reflection depends on the size of the reflecting surface the IR becomes louder with increasing object size. In addition, the IR becomes temporally expanded because of object depth, and thus the temporal delay between single reflections from the reflecting surfaces of a complex structured object, increases with increasing object size (Figure 20B and D). This expansion corresponds to a compression of the spectral interference pattern (Figure 20E). Taken together, much of the information about the three-dimensional structure of an object is represented in its IR, and much of the variability of the IR relates to the size of the object.

An echolocating bat perceives the IR convolved with its echolocation call. Thus, the acoustic image of an object is imprinted on the echolocation call. Although echolocation calls are usually short and cover a broad band of frequencies, they cannot be compared with a dirac impulse. Due to spectral gaps in the usually harmonic calls, bats do not receive the complete information of an ensonified object as it would be contained in the reflection of an impulse. Nevertheless, echolocating bats are able to discriminate and classify these convolved signals both in the auditory frequency- and time domain (Weissenbacher and Wiegrebe, 2003; Grunwald et al., 2004).



**Figure 20: IR and spectral compression**  
 A and C: schematic cartoon of different-sized two-front targets, ensonified by a Dirac-impulse. Solid arrows: ensonification; dotted arrows: reflections. B and D: schematic cartoon of recorded impulse responses from different-sized two-front targets. The bigger the target is, the more the two reflections are temporally separated and the louder they are. E: schematic cartoon of spectral compression patterns for different-sized objects

The aim of this part of this thesis was to search for a size-invariant echo-acoustic object representation in a psychophysical approach. In a two-alternative, forced-choice phantom-object experiment, fruit-eating bats (*Phyllostomus discolor*) were trained to discriminate echoes of their sonar emissions. These echoes consisted of the emission convolved with the IR of one of two objects. Each IR consisted of 12 randomly spaced reflections of different amplitude (see Figure 22A). Once they had learned this task, test trials were randomly interspersed where a scaled version of one of the standard objects was presented (Figure 22A) and the bats' spontaneous classification of these scaled objects was assessed. The bats' psychophysical performance is compared to the performance of a spectro-

temporal pattern recognizer being fed with auditory representations of the echoes, as they are perceived by the bats.

## 2.2 Experimental procedures

### 2.2.1 Experimental animal

The experimental animal used in this study was the new world bat *Phyllostomus discolor* (family: Phyllostomidae). As in the former study, the animals originated from a breeding colony in the Department Biologie II of the Ludwig-Maximilian-University in Munich. *P. discolor* emits short (<3ms) broadband downward-modulated multiharmonic echolocation calls in the frequency range between 40 and 90 kHz (see Figure 7). It feeds mainly on fruits, pollen and insects (Novak, 1994), the insects mainly being gleaned from the vegetation.



**Figure 21:** *Phyllostomus discolor*

This picture was taken while the bat received additional banana pulp

Five adult *P. discolor* (one male, four females) took part in the training. Three of the females had a subcutaneous injected transponder (No. 000 631 D 247; 000 632 8 A39; 000 631 E AB0). The animals were housed in boxes (80 cm x 40 cm x 50 cm) with free access to water, separated for sexes. In these boxes, they were only fed on days without training sessions, i.e. only for two days after a 5-day training period. During the training period, they were fed with banana pulp for reward. On the days without training, they had access to mealworms (larvae of *Tenebrio molitor*) *ad libitum*. All animals had no experience in performing psychoacoustic experiments when they started joining the current training.

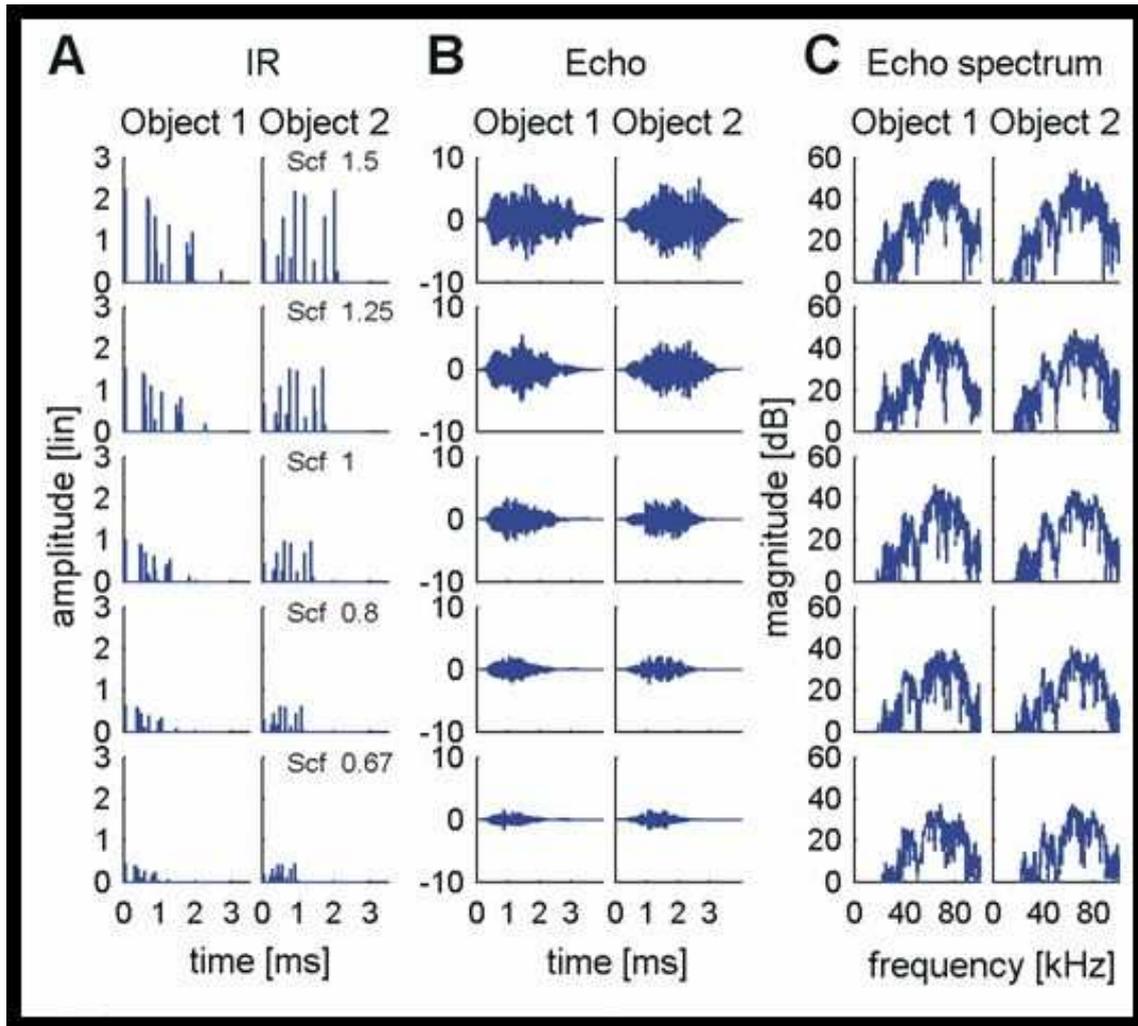
### **2.2.2 Experimental design**

This psychophysical experiment was implemented as a virtual-object playback experiment. The bats were required to evaluate echoes of their own echolocation calls. These echoes were generated by convolving the calls with impulse responses (IRs) of a virtual object in real time. Thus, unlike in classical psychoacoustic experiments, the bats did not hear sound unless they emitted echolocation calls.

### **2.2.3 Impulse responses**

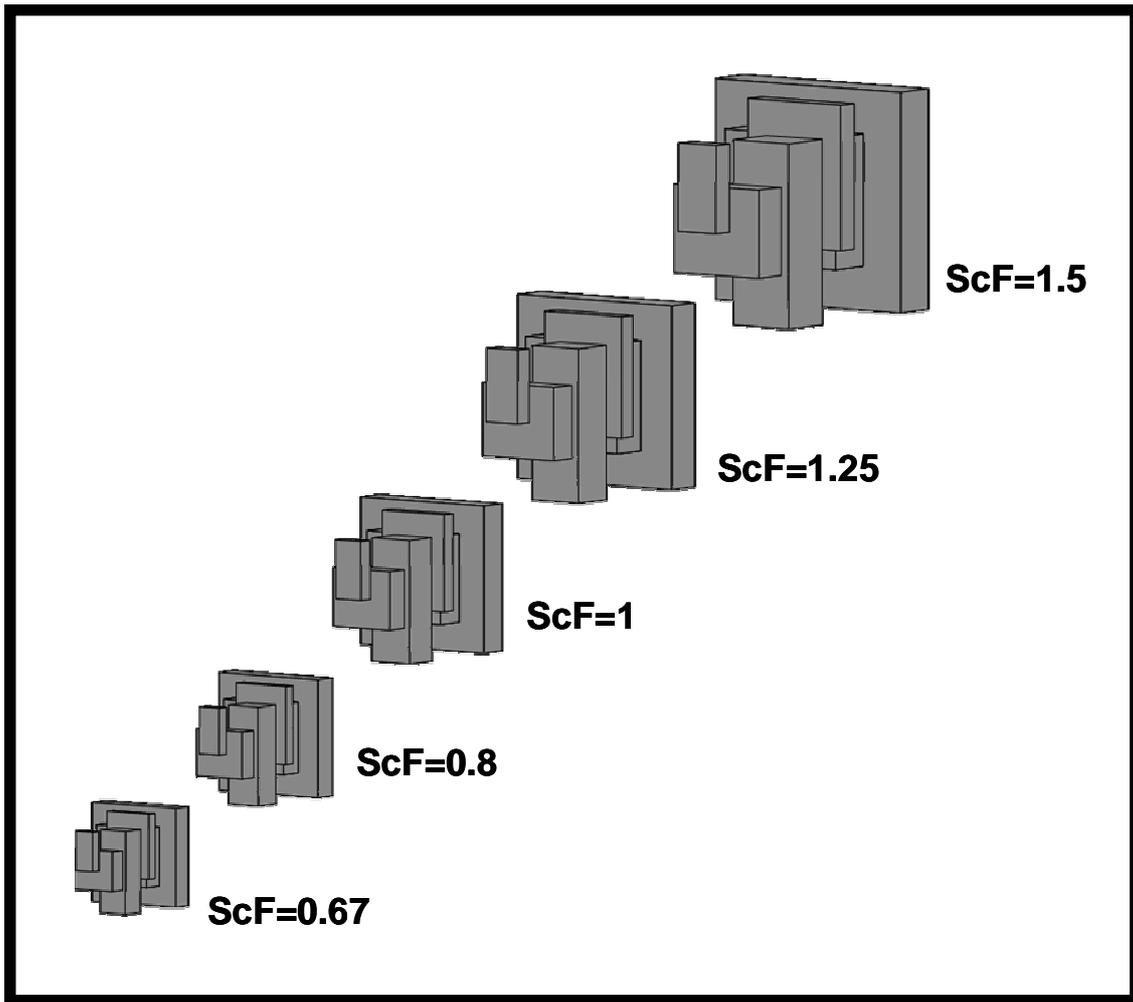
100 IRs with 12 single reflections each of random relative level were generated. Each of these IRs was 750 samples long. This corresponds to a length of 1.86 ms. The size of these virtual objects would be around 64 cm. Then, those two IRs with the largest mean-squared difference in the time domain were determined. These two IRs were taken as standard IRs of virtual object 1 and 2.

Both IRs then were scaled with the following scaling factors (Scf): 0.67, 0.8, 1.25 and 1.5 to cover a reasonable range of naturally occurring differences in size. The time delay of each reflection was scaled with Scf. The amplitude of each reflection was scaled with  $Scf^2$  because the level of a reflection is proportional to surface area and the latter is proportional to the square of the radius. In consequence, the IR produced with a Scf of 0.67 was 15 dB fainter than the same IR scaled with Scf of 1.5. The scaled IRs of both objects are shown in Figure 22A. Note, that all IRs had frequency-independent, white magnitude spectra. This is the case because the 12 reflections with random delays do not produce a systematic spectral ripple. Figure 22B shows the waveforms of the echo after convolution of an echolocation call with the IR. In Figure 22C, the magnitude spectra of the echoes of object 1 and object 2 are plotted. Figure 23 shows a simplified illustration how the corresponding objects to the IRs may look like: they can be imagined as a staple of plates of different size and height. The size of the scaled objects ranged between around 42.5 cm and 95.7 cm.



**Figure 22: Stimuli used for the psychophysical experiments**

(A) Impulse responses, (B) waveforms of the echo after convolution of an echolocation call with the IR and (C) magnitude spectra of the echoes of object 1 and object 2 (left and right column, respectively). In the psychophysical experiments, the bats were trained to discriminate echoes of the standard objects shown the 3rd row of A (scaling factor, Scf, 1). Once the bats had learned this task, presentations of scaled objects (scaling factors 0.67, 0.8, 1.25 and 1.5) were interspersed and the spontaneous classification of these scaled objects was assessed (Firzlaff et al., 2006b).

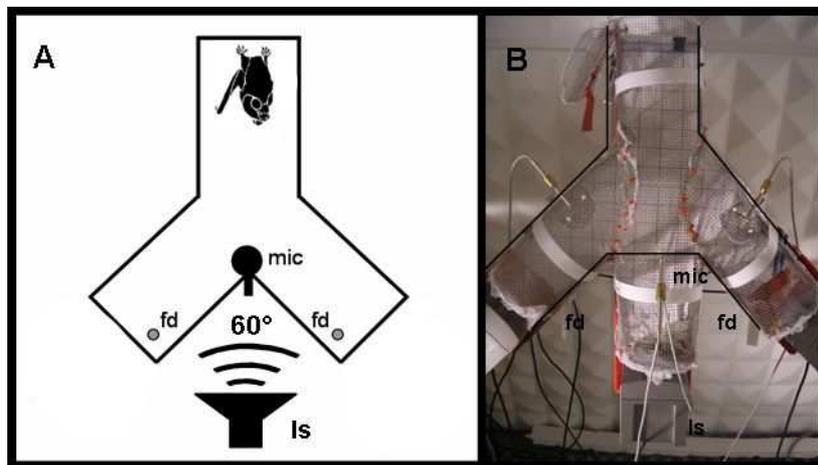


**Figure 23: Simplified illustration how scaled versions of one of the presented virtual objects could look like**

The objects can be thought as a staple of plates of different size and height, scaled with the corresponding scaling factor (ScF).

### 2.2.4 Experimental setup

The bats were trained in a two-alternative, forced-choice playback setup as used in former experiments with *P. discolor* (Grunwald et al., 2004; Schuchmann et al., 2006; Firzlaff et al., 2006a). It consists of a Y-shaped maze, inversely mounted on the wall of an echo-attenuated chamber at an angle of 45 degrees. A starting perch was located at the top end, and a feeder was mounted at the end of each leg. The angle between the legs was 60°. A ¼ inch microphone (Microtech Gefell MK 301, Gefell, Germany) was located in the middle of the maze to pick up the bats' sonar emissions. Figure 24 shows the setup, in Figure 25, the block diagram of the setup is illustrated.



**Figure 24: Experimental setup for *Phyllostomus discolor***

The figure shows a cartoon (A) with the frontal view of the setup and an image of the experimental chamber (B). The crawling setup was a Y-shaped maze (45x30 cm; wire mesh); the angle between the two legs was 60°. The inner width of each leg was 10 cm. The setup consisted of a starting position and an ultrasonic speaker (Is). Feeding dishes were positioned at the end of the left and right leg. A microphone (mic) was placed in front of the bats' head to pick up their ultrasonic emissions.

The ultrasonic emissions were amplified (model 2160; Bruel&Kjaer, Naerum, Denmark) and then digitized by a data-acquisition board (data acquisition processor 5200a; Microstar, Bellevue, WA) at a sampling rate of 250 kHz. On the processor, a software trigger was implemented to look for input values larger than about a tenth of possible input range. When triggering occurred, 500 samples (50 before the trigger event, 450 after the event) were processed. The data-acquisition board convolved this input with the desired IR by

---

---

## Experimental Procedures

zero-padding both the recorded call and the IR to 2048 samples, and multiplying the complex spectra of the recorded call and the IR. The resulting artificial echo was then again amplified (model 6110; Harman/Kardon; Château du Loir, France) and played back via an ultrasonic speaker (model EAS10 TH800D; Matsushita, Osaka, Japan) which was placed in a distance of 20 cm from the starting position in the middle between the two legs. The gain of the whole system depended on the presented IR and ranged between -6 dB and -21 dB. The digital delay of the system was 6 ms.

The experimentator was seated outside the chamber, controlling the experimental procedure via a computer interface and an infrared camera. Data acquisition occurred via Matlab 6.5 (Mathworks, Natick, MA).

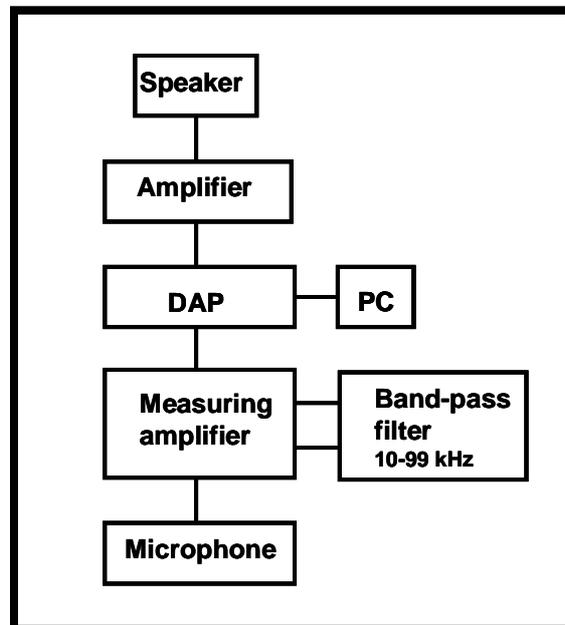


Figure 25: Block diagram for the impulse response playback set-up

### **2.2.5 Training procedure**

In a two-alternative, forced-choice experiment with food reward, five individuals of *P. discolor* were trained to discriminate the two standard IRs, representing two virtual objects. Dependent on the presented object, the bat had to crawl in leg one (object 1) or two (object 2) to obtain a food reward. Three out of five trained animals were able to solve this task. When the bats' performance in these standard trials exceeded 80% correct, test trials were randomly interspersed with a probability of 25%. In these test trials, scaled versions of object 1 or 2 were presented. Test trials were always rewarded, independent of the bats' decision. Thus, in the test trials, the bats' spontaneous classification of the scaled objects as either object 1 or object 2 was assessed. Note: Test trials were only interspersed on those days when the animal first absolved five to eight standard trials correctly, thus showed a stable performance on the training day. This was to guaranty a valid data set. As the animals performed between 10 and 40 trials per day, depending on the individual bat and the motivation of the animal, data acquisition stopped when the slowest animal had performed at least 30 trials for each scaled object. Training of the faster animals was continued to gather more data. The spontaneous performance for each scaled object was numerically tested for significance by simulating 10,000 repetitions of the 2-AFC experiment with the given number of trials and a random performance (see 2.2.6). Significance was set at  $p < 0.05$ .

### **2.2.6 Computation of the significance of the data**

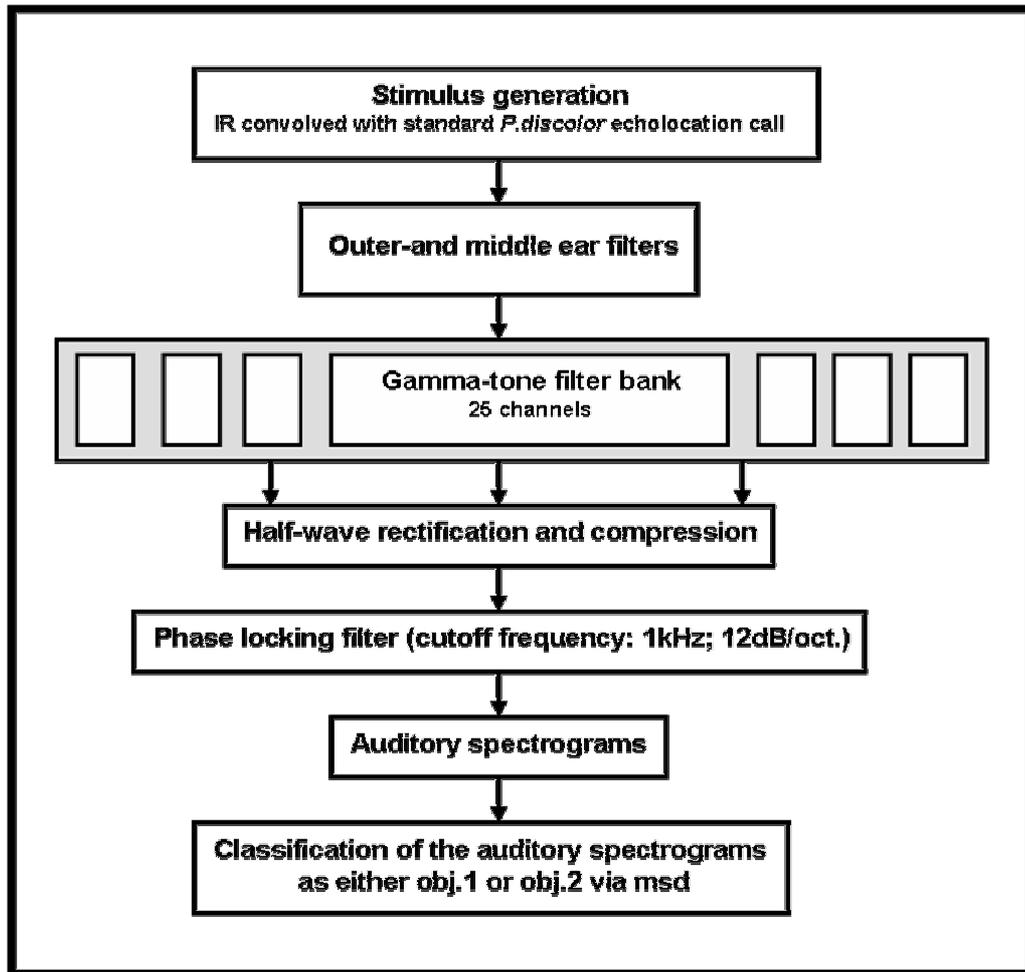
Unlike in the first experiment of this current thesis, where a total 270 lead-lag trials was obtained for each animal, variable numbers of trials were obtained for each animal in this study. This was due to the fact that some animals have been faster in solving the task and thus performed more trials in the same time as others did. To assure a correct comparison of these variable numbers of data, the significance of each data set was computed and then set in relation.

The spontaneous performance for each scaled object was numerically tested for significance by simulating 10,000 repetitions of the 2-AFC experiment with the given number of trials and a random performance.

To calculate if a given performance for a given scaled object was significant for a given number of trials, the p-values for the performance, depending on the number of trials were calculated. A p-value<0.05 hereby was significant, a p-value<0.01 was highly significant. To obtain the p-values, we first simulated a 2-AFC-experiment for 10.000 times and checked how often one alternative was chosen. That way we got the theoretical statistical distribution for each decision possibility. Via the resulting Gaussian distribution we could calculate the p-value of each result, based on the percent correct performance, the number of alternatives and the number of trials used to obtain the percent correct of the result. If the resulting p-value was less than 0.05, the according data bar was marked with one star. If the resulting p-value was less than 0.01; the according data bar was marked with two stars.

### **2.2.7 Simulation of the classification of scaled virtual objects via auditory spectrograms in *P. discolor***

A bat does not hear the IR of an object itself, but the IR convolved with its echolocation call. Moreover, these convolved calls are modified in the auditory periphery of the bats. To simulate the classification of the scaled objects used in this study, we have to take account of these modifications. We simulated the auditory periphery up to the stage of the auditory nerve. First, we convolved the IR of each object with a standard echolocation call. Then, the convolved signal was send through outer- and middle ear filters which mimicked the absolute thresholds of *P. discolor* as described by Esser and Daucher (1996). To simulate the inner ear characteristics, we applied a gamma-tone filter bank, consisting of 25 channels with centre frequencies equally spaced on a log frequency axis between 20 and 110 kHz. The transfer function of the gammatone filters is designed to mimic the shape of the distortion-product, otoacoustic-emission tuning curves of *P. discolor* as measured by Wittekindt et al. (2005). After half-wave rectification and exponential compression, we applied a phase-locking filter (cutoff frequency: 1 kHz; 12 dB/octave). The resulting signal is then resampled at 20 kHz. A schematic illustration of the simulation is shown in Figure 26.

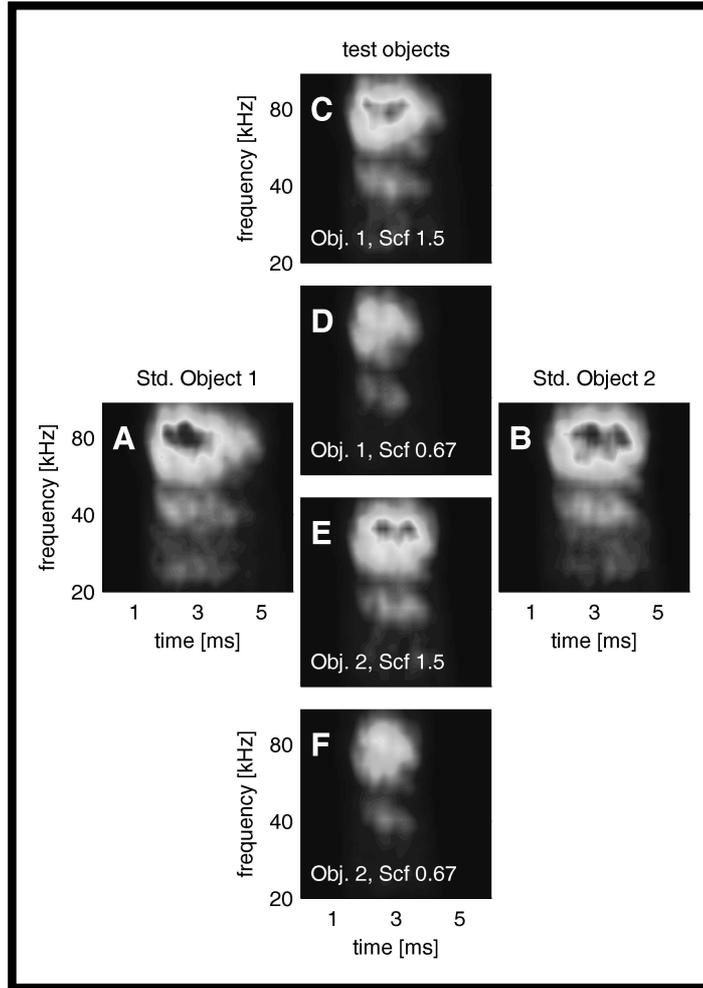


**Figure 26: Schematic illustration of the simulation of the classification of scaled virtual objects via auditory spectrograms in *P. discolor***

Input signal is the object echo, consisting of a standard *P. discolor* call convolved with the IR of the object. Then, the transfer functions of the outer and middle ears of *P. discolor* are applied. A 25-channel gamma tone filter bank simulates the filter functions of the inner ear. Via half-wave rectification, compression and applying a phase-locking filter, the auditory spectrogram is received. Then, the mean square differences are calculated to classify the auditory spectrograms as either object 1 or object 2.

Thus, we received the spectro-temporal excitation patterns- the auditory spectrograms-generated by echoes as they are perceived by the bats in the experimental setup.

The auditory spectrograms generated with the two standard objects are shown in Figure 27A and 27B. Examples of auditory spectrograms for four of the eight scaled versions of the standard objects are shown in Figure 27C-F.



**Figure 27: Auditory spectrograms generated by echoes as they are perceived by the bats.**

Examples of auditory spectrograms produced by echoes generated with the IRs of both standard and scaled objects. The spectrograms incorporate the limits of spectral and temporal resolution of the auditory periphery of *P. discolor*. These spectrograms serve as inputs to the spectro-temporal pattern recognizer which compares the spectrograms generated with different scaled objects (middle column) with those generated with the two standard objects. Simulation results are shown in Figure 28D (Firzlaff et al., 2006b).

Then the mean square differences between the auditory spectrograms generated with scaled objects and those generated with standard objects were calculated. Based on these mean square differences, the simulation classified the auditory spectrograms of the scaled objects as either similar to the spectrograms of standard object 1 or standard object 2. Thus, the simulation works as a spectro-temporal pattern recognizer. Note that this simulation has no concept of scaling and thus, it serves as a null hypothesis for a

---

---

## **Experimental Procedures**

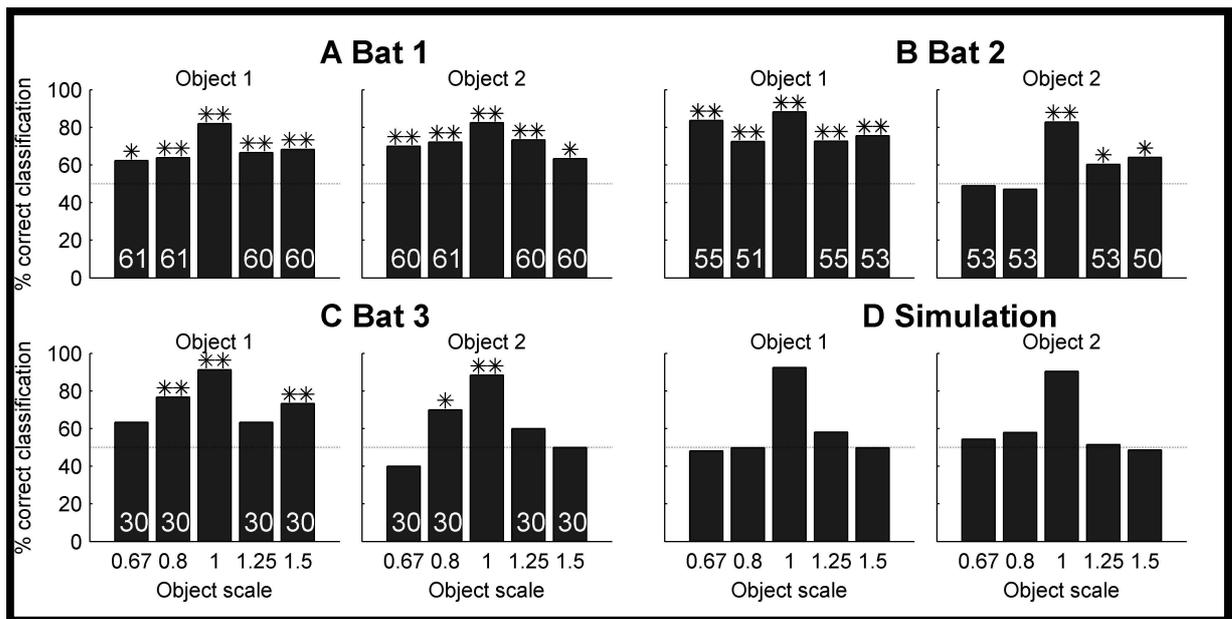
behavioral test of echo-acoustic object normalization. The results of the simulation are shown in Figure 28D. Correct classification is plotted against the corresponding ScFs.

## 2.3 Results

Three of the five trained individuals of *P. discolor* were able to solve the task properly. Thus, behavioral results are based on a total of 4500 trials obtained from three bats. In Figure 28, the classification performances of the animals and the results of the spectral pattern recognizer are shown. Correct classification as the corresponding standard object is plotted against scaling factors for both objects for each animal and the recognizer. The bars corresponding to scaling factor 1 represent the standard trial performance. Bat 1 showed a stable standard trial performance above 80% correct for both standard objects. It also spontaneously classified all (8/8) scaled objects significantly correct (Figure 28A). Bat 2 also showed a very good standard trial performance above 80% correct for both standard objects. The standard trial performance for object 1 was even better than for object 2. It spontaneously classified all four scaled objects of object 1 and two of the scaled versions of object 2 correctly; object 1 scaled with a scaling factor of 0.67 was classified even better than the standard object 1 (Figure 28B). But it failed in classifying both smaller scaled versions of object 2 correctly. Also bat 3 had a very high standard trial performance above 80% for standard object 2 and for object 1, the standard trial performance was even above 90% correct. In contrast to bat 1 and bat 2, this individual classified only two scaled versions of object 1 correctly and one scaled version of object 2. Although it failed to achieve significant performance in five of eight cases it showed a similar trend in evaluating the scaled objects (Figure 28C). The scaling factors the animals failed to classify varied unsystematically. There is no trend in systematically preferring a certain object or scaling factor. Total data acquisition from all animals took around two years.

The performance of a spectro-temporal pattern recognizer is shown in Figure 28D. This pattern recognizer (described in 2.2.7) worked on a representation of the object echoes as it is generated by the bat's auditory periphery (Figure 27). The mean-square differences between the auditory spectrograms of the two standard objects (Figure 27A and B) and those of the scaled objects (examples in Figure 27C-F) were calculated. Based on the mean square differences, the recognizer can reliably classify the peripheral representation of the standard objects as themselves. But it failed in reliably classifying the peripheral representation of the scaled objects as one of the trained standard objects. Thus, even the

bat that showed the weakest normalization behavior (Bat 3) performed considerably better than the spectro-temporal pattern recognizer. Note that such a model could successfully predict responses in other echo-acoustic playback experiments (Weissenbacher and Wiegrebe, 2003; Grunwald et al., 2004). These simulation results show that a dedicated neural mechanism beyond the auditory periphery is required to explain the bats' compensation for size-induced echo variations.



**Figure 28: Psychophysical classification performance and simulation results**

Spontaneous classification of scaled virtual objects by three bats (A, B and C) and a simulation (D) based on a spectro-temporal pattern recognizer. Significantly correct classification of scaled objects is marked by a single star ( $p < 0.05$ ) or two stars ( $p < 0.01$ ). The number of test trials for each condition is superimposed on each bar. While the spectro-temporal pattern recognizer cannot classify the scaled objects correctly, the bats' performance is significant in the majority of test conditions (Firzlaff et al., 2006b).

## **2.4 Discussion**

In a psychophysical approach, this study tested the ability of the echolocating bat *Phyllostomus discolor* to normalize for size-induced variations of virtual echo-acoustic objects. The results showed that the bats spontaneously classified most scaled virtual objects similar to their unscaled selves. Data acquisition only occurred when the bats' standard trial performance was above 80% correct at all time, thus the data are valid. A simulation of the psychophysical paradigm based on spectro-temporal pattern recognition indicates that a complex central auditory circuitry is required to explain the bats' performance.

### **2.4.1 Comparison with previous studies**

In the current study, echolocating *P. discolor* had to extract and to evaluate IRs of different sized virtual objects. In a study by Weissenbacher and Wiegrebe (2003), it was successfully shown that the echolocating bat *M. lyra* was able to evaluate IRs of different virtual objects. The bat was trained to classify unknown virtual objects according to learned reference objects of different temporal and spectral composition. This study showed that the bats developed an accurate internal representation of the objects' IRs. Furthermore, they stated that IRs of small objects (< 4-6 cm) are coded along the tonotopic frequency axis. Considering the work from Weissenbacher, it was to be expected that the bat *P. discolor* would be able to deal with virtual objects. Furthermore, this species was already successfully used in playback experiments (Firzlaff et al., 2006a; Grunwald et al., 2004). This presumption was confirmed: the bats were able to discriminate the standard objects properly and also classified most of the scaled objects correctly.

The current study used virtual acoustic objects, defined by their IRs. As mentioned in the introduction, the IR is a one-dimensional image of an object with less information about the objects shape.

In nature, a bat can evaluate an object by integrating the echo information of the object from different ensonification angles. In fact, the echo spectra of bat pollinated flowers changed significantly with the angle of sound incident whereas echoes from single position often were similar (von Helversen and von Helversen, 2003). Von Helversen (2004) showed correct size-independent discrimination of real targets performed by the

---

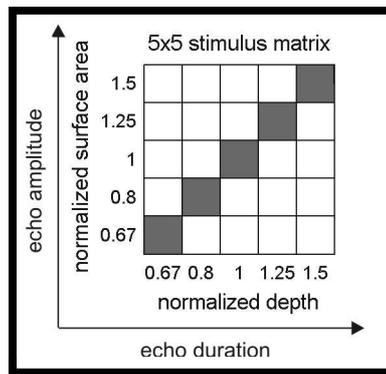
---

## Discussion

echolocating bat *Glossophaga soricina*. Von Helversen proposed that correct object classification is guided by the changes in the spectral pattern of the returning echoes at different ensonification angles. Object recognition could then depend on serial integration of acoustic signals. The ability to integrate over a sequence of signal has been demonstrated in songbirds and mammals (e.g. (Hulse and Cynx, 1986; Ralston and Herman, 1995). In contrast, the bats in the current study were confronted with the IRs of virtual, static objects. Thus, the bats receive only a one-dimensional echo-acoustic image of these virtual objects. Nevertheless, the bats were able to classify scaled versions of these IRs correctly, although the information from different ensonification angles was missing. Hypothetically, the bats' echo-acoustic object normalization ability will improve further when they can evaluate sequences of echoes from different ensonification angles.

### 2.4.2 Physiological correlates for auditory object normalization in echolocating bats

In an accompanying electrophysiological study done by Uwe Firzlaff (Firzlaff et al., 2006b), a neurophysiological correlate of the current perceptual accomplishment of auditory object normalization in echolocating bats was found. This study uses echoes as they are perceived by the bats in the behavioral experiments to test the ability of units in the bats' auditory cortex to encode a normalized auditory representation of an object. The IRs of the two standard objects (scaling factor 1) used in the current behavioral study were scaled both in terms of the delay and amplitude of the reflections (ScF: 0.67, 0.8, 1, 1.25, 1.5). This results in two 5x5 matrices (see Figure 29).



**Figure 29: 5x5 stimulus matrix for the electrophysiological experiments**

Echo amplitude, which encodes object surface areas, is plotted against echo duration, which encodes object depth. The corresponding scaling factors are noted. The red squares indicate those IRs were echo duration /object depth and echo amplitude/object surfaces area co- vary in a meaningful manner as they did in the behavioral experiments. From Firzlaff et al. (2006b)

This results in a set of 50 stimuli, 25 for each object. The diagonal of this 5x5 matrix contains the scaled IRs as presented behaviorally. Here, echo duration, which encodes object depth and echo amplitude, encoding the object surface area, were scaled with the same ScF and therefore co-varied in a meaningful manner as they did in the psychophysical experiment. The IRs were then convolved with a standard *P. discolor* echolocation call. Examples for such echoes of the virtual scaled objects are shown in Figure 22 C. These echoes were presented to anaesthetized *P. discolor*. The responses from cortical units to these echoes were recorded. Several units responded preferentially either

---

---

## Discussion

---

---

to a particular object depth or to a combination of object depths, but were largely independent of object-surface area. Also units that responded strongest either to particular object-surface area/echo amplitude or to a combination of surface areas, but were largely independent of object depth/echo duration were found. Most interestingly, a small number of cortical units which showed an object-size invariant response was found. These units responded preferentially to echoes from objects along the diagonal axis of the 5x5 matrix where echo duration and echo amplitude co-varied meaningfully. This means that the response was not simply dependent on echo amplitude or duration but on a meaningful combination of both. Thus, these neurons reflected normalized object features in their response rates. This population may serve as a substrate for the psychophysically observed perceptual compensation of size-induced object variations. The results suggest that the activity of a neural subpopulation in the auditory cortex reflects a normalized auditory object feature representation.

### **2.4.3 Towards a functional model of echo-acoustic object normalization**

An object is acoustically defined by its IR, and the IR differs with object size. Behavioral evidence exists (Weissenbacher and Wiegrebe, 2003; Simmons et al., 1990) that bats have an accurate internal representation of an objects' IR. Consequently, one way to achieve a normalized representation of an object in echolocation may involve IR extraction from the perceived echo and subsequent processing to compensate for size-induced IR variations.

#### **2.4.3.1 The extraction of impulse responses**

The extraction of an IR out of the echo requires a detailed comparison of the bats' sonar emission and the returning echo. Behavioural evidence for IR extraction comes from (Simmons et al., 1990) and Weissenbacher and Wiegrebe (2003). These results indicate that the bats developed an accurate internal representation of the objects' IRs. Mathematically, the extraction can be implemented in the time domain as a cross-correlation of the emission and the echo (Simmons et al., 1990) or in the frequency domain as a multiplication of the emission- and echo spectra. As both the emission and the echo are transformed into an auditory time-frequency representation by the bat's auditory periphery, the physiological mechanism of IR extraction could be a hybrid of these two mechanisms. Specifically, reflections in the IR which are separated by more than the integration times of the bat's auditory filters are preserved in the auditory time domain. Reflections that are separated less produce spectral interference and their information is transformed into the auditory frequency domain. Thus, dependent on the IR duration, the IR is represented either exclusively along the auditory frequency axis or along both the auditory frequency and time axis. The outcome of this is that there have to exist mechanisms to perform normalization in either the frequency or time domain. Physiologically, a bank of neurons which are spectrally narrow tuned and also tuned to a specific echo delay by means of a delayed coincidence detection could generate such a detailed auditory representation of an object's echo-acoustic image. Nevertheless, the physiological mechanism underlying IR extraction is still unclear.

### **2.4.3.2 Auditory IR normalization**

Dependent on the IR duration, the IR is represented either exclusively along the auditory frequency axis or along both the auditory frequency and time axis. The outcome of this is that there have to exist mechanisms to perform normalization in either the frequency or time domain.

#### **2.4.3.2.1 Auditory IR normalization in the time domain**

When an IR is scaled in time, the gap between the single reflections is either dilated or compressed. Therefore, a mechanism is required which compensates for these size-induced temporal variations in the auditory representation of the IR. There is a mathematical transform, which is immune to the effects of size dilation: the Mellin transform (Bertrand et al., 1996). It maps IRs that differs in temporal scale onto a single distribution and encodes the size information separately as a scalar constant. Consequently, the Mellin transform can be used to segregate the shape information of an IR from its size information (Irino and Patterson, 2002).

#### **2.4.3.2.2 Auditory IR normalization in the frequency domain:**

Reflections in the IR that interact within the integration time of the bat's auditory filters produce a spectral interference pattern along the basilar membrane. With varying object size, frequency peaks and notches of this pattern are simply shifted along a logarithmic frequency axis; the spacing of the peaks and notches between each other remains the same. A spectral pattern recognizer, which is tolerant for shifts in the absolute position of a spectral interference pattern, could thus serve for auditory object normalization in the auditory frequency domain. In this context, the tonotopic axis, which for most bats is indeed a logarithmic frequency axis, is well suited for frequency-domain object normalization.

---

---

## Discussion

---

---

The current behavioral and the additional physiological data from Dr. Firzlaff support the existence of an object-oriented perception and auditory representation of echoes in echolocating bats. The functional model outlined above will serve as a roadmap to identify the neural building blocks in the auditory brainstem and midbrain which underlie this highly abstracted cortical representation. Specifically, one can hypothesize that the medial superior olive, which is surprisingly well developed in bats (Grothe et al., 2001; Grothe, 2000), may serve as a neural substrate for a spectro-temporal auditory representation of an object's IR. A possible neural circuit which could compensate for size-induced variations in the temporal representation of the IR may be implemented with mechanisms as described by Hopfield (2004) involving the interaction between the temporal structure of the auditory IR representation with subthreshold neural oscillations. Along the tonotopic axis, the auditory representation of the IR may be subjected to a similar analysis as it has been suggested for the analysis of human speech sounds by the medial geniculate body (von Kriegstein et al., 2006).

The current study supports the hypothesis that the auditory system has dedicated mechanisms to deal with the compensation of size-induced variations of acoustic sources. The fact that the size-invariant neural responses were obtained from individuals which had not been exposed to the stimuli behaviorally (Firzlaff et al.), indicates that this mechanism is not experience dependent but hard-wired.

#### **2.4.4 Problems and perspectives**

The current work on echo-acoustic object normalization shows that the echolocating bat *P. discolor* is indeed able to discriminate and classify the echoes of different sized virtual objects and that this ability is due to dedicated, hard-wired mechanisms in the auditory system. This study contributes to the research in echo-acoustic object normalization. It might be very interesting to further track this issue.

This chapter discusses the problems of the current study and introduces possible approaches for the continuation of the investigation of echo-acoustic object normalization in echolocating bats.

Although the bats were able to perform the task properly, all in all training of the animals took quite some time, approximately two years. The animals needed several months to learn to discriminate the two complex virtual standard objects with a stable and high correct performance. When they had learned this ambitious task, it again took several months to obtain the data for the scaled objects, as they generally only participate for an untypical small number of trials per day: most of the data acquisition period, the bats performed around 10 to 20, at most 40 trials per day. On several days, the bats additionally did not show a stable performance in the standard trials, displaying a very low motivation. From other studies (e.g. Firzlaff et al., 2006a; Grunwald et al., 2004; Schuchmann et al., 2006), personal communications with other scientists working with *P. discolor* and from own experience, it is known that *P. discolor* is easily able to perform more than 50 trials each day without a loss of their performance quality. This is normally the case, when the given task is easily solvable for the animal and is highly motivated to perform more trials than necessary to satisfy its hunger. If the task becomes almost too difficult for an animal, it will only perform as much trials as necessary to just satisfy its hunger. Therefore, considering the training behaviour of the bats in this study, the current paradigm might have been very complicated for the animals and given them difficulties in solving the task. A further important fact one has to keep in mind is the fact that they had to create a precise cognitive representation of the two very complex virtual objects, had to associate these objects with a corresponding side to receive a food reward and also had to compare scaled versions of the virtual objects with the remembered templates. This might have

overstrained their abilities. Presumably, it would ease the task to present both objects at the same time from two loudspeakers to allow a direct comparison of the objects. This would be a more simple discrimination task of two objects, with one object offering reward and the other not. When achieving this task, scaled versions of the object of interest can be interspersed. The performance of the bat would give information about the bats' normalization abilities. A negative aspect of this paradigm would be that mainly the bats' discrimination abilities would be assessed.

The IRs of the current virtual objects contained nearly no spectral information for the bats, which in turn plays an important role in auditory processing. Auditory information can be resolved either in the frequency domain, in the time domain or-more common- in both domains. The current work showed the influence of time on echo-acoustic object normalization. Following work could deal with the role of spectral cues on echo-acoustic object normalization. Auditory objects which only differ spectrally would be needed. They could be obtained by generating a virtual two-front target, resulting in an impulse response consisting of two distinct reflections. Depending on the temporal separation of the two reflections, this two-front target would lead to known spectral interference patterns which easily could be scaled. These scaled IRs of virtual two-front targets could be used in the same playback paradigm as used in the current study. Thus, it would be possible to examine echo-acoustic object normalization in the spectral domain with well-defined stimuli.

A further problem in the current study might have been the static object representation.

In a natural echolocation situation, bats are not always confronted with the same "view" of the virtual object. An IR defines an object mainly in the depth dimension; it is a one-dimensional acoustic image. It provides only sparse information about the object's shape. Naturally, a bat ensonifies objects of interest from different angles. They string the received echo information together and thus receive more information about the object by integrating the echo information. Therefore, presenting only a one-dimensional acoustic image of an object causes a loss of information about the objects' characteristics. As already mentioned in 2.4.1, the bats' echo-acoustic object normalization ability might

improve further if they are allowed to evaluate sequences of echoes of an object obtained from different ensonification angles.

This hypothesis leads to two different paradigms: the presentation of virtual objects which change depending on the ensonification angle or to present real targets, which can be evaluated by the bat from different positions.

The first paradigm requires an enormous technical complexity. The technique necessary to realize this paradigm is momentarily being developed by Daria Genzel for her PhD-thesis. Here, a loudspeaker presents different IRs convolved with the bats' echolocation calls; which IR depends on the spatial position of the bat. Therefore, the bats' position in space has to be tracked and a processor has to convolve in real time the corresponding IR with a recorded echolocation call. IRs recorded from real targets of different size could serve as stimuli. The same paradigm as used in the current study could be used. Note: for this paradigm the bats have to be excellent flyers, as the animals have to move around the speaker to receive the different IRs and to evaluate them.

The second paradigm would certainly be less technically complex but more natural. Simple shaped real targets would be presented to echolocating bats. Training would be quite similar as in the current thesis: the bats would first have to learn to discriminate two different objects. When achieving this task, different sized versions of the trained standard objects would be presented and the bats' spontaneous classification of the scaled objects would be assessed. Note: the bats need the possibility to ensonify the objects from different angles. Thus again, excellent flyers would be useful as experimental animals, but also crawling bats like *P. discolor*. At the moment, Daria Genzel is working on this project. Here, *P. discolor* is trained to stay on a platform without a wire mesh, to ensonify a presented object and to associate an object with a corresponding feeder.

The size of an object determines whether a bat can extract mainly spectral or temporal (or both) information from the IR. Narrowly spaced reflecting surfaces of smaller objects cause mainly spectral interference patterns in echoes. The bigger an object gets, the greater the temporal separation between the reflections. Thus, a bat can use the temporal information in the echo as the temporal separation of the reflections get longer than the temporal integration times of the cochlear (around 200 $\mu$ s). Small objects up to around 4-6 cm (Weissenbacher and Wiegrebe, 2003) mainly provide spectral information for the bats.

The objects in the current study mainly provide temporal cues for the bats as their size ranged from around 42.5 cm and 95.7 cm. Considering the long training period needed the provide temporal information in echoes might not have been enough for the bats to be able to evaluate the presented virtual object with a constant performance. But nevertheless, they showed echo-acoustic object normalization, although mainly temporal cues were available for them.

Von Helversen (2004) proposed that correct object classification in *G. soricina* is guided by changes in the spectral pattern of the echoes. Additionally, it was shown by Weissenbacher (2003) that *M. lyra* is able to evaluate IRs which provide mostly spectral information.

As a consequence, it would be interesting to compare the influence of spectral and temporal cues on echo-acoustic object normalization. Bats could be confronted with virtual objects with sizes ranging from around 4 cm or less up to 45 cm or more, scaled in reasonable steps. These IRs would offer the bats temporal and spectral cues for object analysis: the smallest objects would mainly provide spectral information, middle-sized object would provide both spectral and temporal cues and the biggest objects would mainly provide temporal cues. The classification performance for the different sized objects would shed light on the weighting of temporal and spectral processing for echo-acoustic object normalization.

## References

1. **Ashmead DH, Clifton RK, Reese EP** (1986) Development of auditory localization in dogs: single source and precedence effect sounds. *Dev Psychobiol* 19: 91-103.
2. **Babushkina ES, Poliakov MA** (2001) Lokalizatsiia summy akusticheskikh signalov severnym morskim kotikom v vozdushnoi srede. [Localization of a sum of acoustic signals in air by the northern fur seal]. *Biofizika* 46: 557-562.
3. **Bertrand J, Bertrand P, Ovarlez J-P** (1996) In: *The Transforms and Application Handbook* (Poularikas AD, ed), Florida: CRC Press (IEEE Press).
4. **Blauert J** (1997) *Spatial Hearing: The Psychophysics of Human Sound Localization*. Cambridge MA: MIT Press.
5. **Blauert J, Canevet G, Voinier T** (1989) The Precedence Effect - No Evidence for An Active Release Process Found. *J Acoust Soc Am* 85: 2581-2586.
6. **Bregman AS** (1990) *Auditory scene analysis*. Cambridge, MA: MIT.
7. **Brittan-Powell EF, Dooling RJ, Farabaugh SM** (1997) Vocal development in budgerigars (*Melopsittacus undulatus*): contact calls. *J Comp Psychol* 111: 226-241.
8. **Brockway BF** (1964) Ethological studies of the budgerigar (*Melopsittacus undulatus*) : non-reproductive behaviour. *Behaviour* 22: 193-222.
9. **Burger RM, Pollak GD** (2001) Reversible inactivation of the dorsal nucleus of the lateral lemniscus reveals its role in the processing of multiple sound sources in the inferior colliculus of bats. *J Neurosci* 21: 4830-4843.
10. **Carello C, Anderson KL, Kunkler-Peck AJ** (1998) Perception of object length by sound. *Psychological Science* 9: 211-214.
11. **Clifton RK** (1987) Breakdown of Echo Suppression in the Precedence Effect. *J Acoust Soc Am* 82: 1834-1835.
12. **Clifton RK, Freyman RL, Litovsky RY, McCall D** (1994) Listeners Expectations About Echoes Can Raise Or Lower Echo Threshold. *J Acoust Soc Am* 95: 1525-1533.
13. **Clifton RK, Freyman RL, Meo J** (2002) What the precedence effect tells us about room acoustics. *Perception & Psychophysics* 64: 180-188.
14. **Cranford J, Ravizza R, Diamond IT, Whitfield IC** (1971) Unilateral Ablation of Auditory Cortex in Cat Impairs Complex Sound Localization. *Science* 172: 286-&.

---

---

## References

---

---

15. **Cranford JL** (1982) Localization of Paired Sound Sources in Cats - Effects of Variable Arrival Times. *J Acoust Soc Am* 72: 1309-1311.
16. **Damaschke J, Riedel H, Kollmeier B** (2005) Neural correlates of the precedence effect in auditory evoked potentials. *Hear Res* 205: 157-171.
17. **Dent ML, Dooling RJ** (2003a) Investigations of the precedence effect in budgerigars: Effects of stimulus type, intensity, duration, and location. *J Acoust Soc Am* 113: 2146-2158.
18. **Dent ML, Dooling RJ** (2003b) Investigations of the precedence effect in budgerigars: The perceived location of auditory images. *J Acoust Soc Am* 113: 2159-2169.
19. **Dent ML, Dooling RJ** (2004) The precedence effect in three species of birds (Melopsittacus undulatus, Serinus canaria, and Taeniopygia guttata). *J Comp Psychol* 118: 325-331.
20. **Esser KH, Daucher A** (1996) Hearing in the FM-bat Phyllostomus discolor: a behavioral audiogram. *J Comp Physiol [A]* 178: 779-785.
21. **Farabaugh SM, Linzenbold A, Dooling RJ** (1994) Vocal plasticity in budgerigars (Melopsittacus undulatus): evidence for social factors in the learning of contact calls. *J Comp Psychol* 108: 81-92.
22. **Firzlaff U, Schornich S, Hoffmann S, Schuller G, Wiegrebe L** (2006a) A neural correlate of stochastic echo imaging. *J Neurosci* 26: 785-791.
23. **Firzlaff U, Schuchmann M, Grunwald JE, Schuller G, Wiegrebe L** (2006b) Object-oriented Echo Perception and Cortical Representation in Echolocating Bats. *in Prep.*
24. **Fitch WT** (1997) Vocal tract length and formant frequency dispersion correlate with body size in rhesus macaques. *J Acoust Soc Am* 102: 1213-1222.
25. **Fitzpatrick DC, Kuwada S, Batra R, Trahiotis C** (1995) Neural responses to simple simulated echoes in the auditory brain stem of the unanesthetized rabbit. *J Neurophysiol* 74: 2469-2486.
26. **Freyman RL, Clifton RK, Litovsky RY** (1991) Dynamic Processes in the Precedence Effect. *J Acoust Soc Am* 90: 874-884.
27. **Freyman RL, McCall DD, Clifton RK** (1998) Intensity discrimination for precedence effect stimuli. *J Acoust Soc Am* 103: 2031-2041.
28. **Freyman RL, Zurek PM, Balakrishnan U, Chiang YC** (1997) Onset dominance in lateralization. *J Acoust Soc Am* 101: 1649-1659.
29. **Furmanski CS, Engel SA** (2000) Perceptual learning in object recognition: object specificity and size Invariance. *Vision Res* 40: 473-484.

---

---

## References

---

---

30. **Goymann W, Leippert D, Hofer H** (1999) Parturition, parental behaviour, and pup development in Indian false vampire bats, *Megaderma lyra*. *Z Säugetierkunde* 64: 321-331.
31. **Grothe B** (2000) The evolution of temporal processing in the medial superior olive, an auditory brainstem structure. *Prog Neurobiol* 61: 581-610.
32. **Grothe B, Covey E, Casseday JH** (2001) Medial superior olive of the big brown bat: neuronal responses to pure tones, amplitude modulations, and pulse trains. *J Neurophysiol* 86: 2219-2230.
33. **Grunwald JE, Schornich S, Wiegrebe L** (2004) Classification of natural textures in echolocation. *Proc Natl Acad Sci U S A* 101: 5670-5674.
34. **Hartung K, Trahiotis C** (2001) Peripheral auditory processing and investigations of the "precedence effect" which utilize successive transient stimuli. *J Acoust Soc Am* 110: 1505-1513.
35. **Heffner RS, Koay G, Heffner HE** (2001) Audiograms of five species of rodents: implications for the evolution of hearing and the perception of pitch. *Hear Res* 157: 138-152.
36. **Hoeffding V, Harrison JM** (1979) Auditory discrimination: role of time and intensity in the precedence effect. *J Exp Anal Behav* 32: 157-166.
37. **Hopfield JJ** (2004) Encoding for computation: recognizing brief dynamical patterns by exploiting effects of weak rhythms on action-potential timing. *Proc Natl Acad Sci U S A* 101: 6255-6260.
38. **Houben MMJ, Kohlrausch A, Hermes DJ** (2004) Perception of the size and speed of rolling balls by sound. *Speech Communication* 43: 331-345.
39. **Hulse SH, Cynx J** (1986) Interval and Contour in Serial Pitch Perception by a Passerine Bird, the European Starling (*Sturnus Vulgaris*). *J Comp Psychol* 100: 215-228.
40. **Irino T, Patterson RD** (2002) Segregating information about the size and shape of the vocal tract using a time-domain auditory model: The stabilised wavelet-Mellin transform. *Speech Communication* 36: 181-203.
41. **Ito M, Tamura H, Fujita I, Tanaka K** (1995) Size and Position Invariance of Neuronal Responses in Monkey Inferotemporal Cortex. *J Neurophysiol* 73: 218-226.
42. **Kalko EKV, Condon MA** (1998) Echolocation, olfaction and fruit display: how bats find fruit of flagelliferous cucurbits. *functional ecology* 12: 364-372.
43. **Keller CH, Takahashi TT** (1996a) Binaural cross-correlation predicts the responses of neurons in the owl's auditory space map under conditions simulating summing localization. *J Neurosci* 16: 4300-4309.

---

---

## References

---

---

44. **Keller CH, Takahashi TT** (1996b) Responses to simulated echoes by neurons in the barn owl's auditory space map. *J Comp Physiol [A]* 178: 499-512.
45. **Kelly JB** (1974) Localization of Paired Sound Sources in Rat - Small Time Differences. *J Acoust Soc Am* 55: 1277-1284.
46. **Leonardo A, Konishi M** (1999) Decrystallization of adult birdsong by perturbation of auditory feedback. *Nature* 399: 466-470.
47. **Litovsky RY, Colburn HS, Yost WA, Guzman SJ** (1999) The precedence effect. *J Acoust Soc Am* 106: 1633-1654.
48. **Litovsky RY, Rakerd B, Yin TCT, Hartmann WM** (1997) Psychophysical and physiological evidence for a precedence effect in the median sagittal plane. *J Neurophysiol* 77: 2223-2226.
49. **Logothetis NK, Sheinberg DL** (1996) Visual object recognition. *Annu Rev Neurosci* 19: 577-621.
50. **Mccall DD, Freyman RL, Clifton RK** (1998) Sudden changes in spectrum of an echo cause a breakdown of the precedence effect. *Percept Psychophys* 60: 593-601.
51. **Metzner W** (1993) An Audio Vocal Interface in Echolocating Horseshoe Bats. *Journal of Neuroscience* 13: 1899-1915.
52. **Mickey BJ, Middlebrooks JC** (2005) Sensitivity of auditory cortical neurons to the locations of leading and lagging sounds. *J Neurophysiol* 94: 979-989.
53. **Miller J** (1996) The sampling distribution of d'. *Percept Psychophys* 58: 65-72.
54. **Moore JK, Moore RY** (1987) Glutamic acid decarboxylase-like immunoreactivity in brainstem auditory nuclei of the rat. *J Comp Neurol* 260: 157-174.
55. **Neuweiler G** (2000) *The Biology of Bats*. Oxford University Press.
56. **Novak RM** (1994) *Walker's Bats of the world*. John Hopkins University Press.
57. **Pecka M, Zahn TP, Saunier-Rebori B, Siveke I, Felmy F, Wiegrebe L, Klug A, Pollak G, Grothe B** (2006) Inhibiting the inhibition: a neuronal network for sound localization in reverberant environments. *in Prep*.
58. **Peissig JJ, Kirkpatrick K, Young ME, Wasserman EE, Biederman I** (2006) Effects of varying stimulus size on object recognition in pigeons. *J Exp Psychol Anim Behav Process* 32: 419-430.
59. **Populin LC, Yin TCT** (1998) Behavioral studies of sound localization in the cat. *J Neurosci* 18: 2147-2160.
60. **Rakerd B, Hartmann WM** (1985) Localization of Sound in Rooms .2. the Effects of A Single Reflecting Surface. *J Acoust Soc Am* 78: 524-533.

---

---

## References

---

---

61. **Ralston JV, Herman LM** (1995) Perception and Generalization of Frequency Contours by A Bottle-Nosed-Dolphin (*Tursiops-Truncatus*). *J Comp Psychol* 109: 268-277.
62. **Ryan A** (1976) Hearing sensitivity of the mongolian gerbil, *Meriones unguiculatis*. *J Acoust Soc Am* 59: 1222-1226.
63. **Saberi K, Perrott DR** (1990) Lateralization Thresholds Obtained Under Conditions in Which the Precedence Effect Is Assumed to Operate. *J Acoust Soc Am* 87: 1732-1737.
64. **Saintmarie RL, Ostapoff EM, Morest DK, Wenthold RJ** (1989) Glycine-Immunoreactive Projection of the Cat Lateral Superior Olive - Possible Role in Midbrain Ear Dominance. *J Comp Neurol* 279: 382-396.
65. **Schmidt S, Hanke S, 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.
66. **Schuchmann M, Hubner M, Wiegrebe L** (2006) The absence of spatial echo suppression in the echolocating bats *Megaderma lyra* and *Phyllostomus discolor*. *J Exp Biol* 209: 152-157.
67. **Simmons JA, Moss CF, 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.
68. **Smith DR, Patterson RD, Turner R, Kawahara H, Irino T** (2005) The processing and perception of size information in speech sounds. *J Acoust Soc Am* 117: 305-318.
69. **Snedden WA, Greenfield MD** (1998) Females prefer leading males: relative call timing and sexual selection in katydid choruses. *Anim Behav* 56: 1091-1098.
70. **Spitzer MW, Takahashi TT** (2006) Sound localization by barn owls in a simulated echoic environment. *J Neurophysiol* 95: 3571-3584.
71. **Thies W, Kalko EKV, Schnitzler HU** (1998) The roles of echolocation and olfaction in two Neotropical fruit-eating bats, *Carollia perspicillata* and *C-castanea*, feeding on Piper. *Behavioral Ecology and Sociobiology* 42: 397-409.
72. **Tollin DJ, Henning GB** (1999) Some aspects of the lateralization of echoed sound in man. II. The role of the stimulus spectrum. *J Acoust Soc Am* 105: 838-849.
73. **Tollin DJ, Yin TC** (2003) Psychophysical investigation of an auditory spatial illusion in cats: the precedence effect. *J Neurophysiol* 90: 2149-2162.
74. **von Helversen D** (2004) Object classification by echolocation in nectar feeding bats: size-independent generalization of shape. *J Comp Physiol [A]* 190: 515-521.

---

---

## References

75. **von Helversen D, von Helversen O** (2003) Object recognition by echolocation: a nectar-feeding bat exploiting the flowers of a rain forest vine. *J Comp Physiol [A]* 189: 327-336.
76. **von Kriegstein K, Warren JD, Ives DT, Patterson RD, Griffiths TD** (2006) Processing the acoustic effect of size in speech sounds. *Neuroimage* 32: 368-375.
77. **Wallach H, Newman EB, Rosenzweig MR** (1949) The Precedence Effect in Sound Localization. *Am J Psychol* 62: 315-336.
78. **Weissenbacher P, Wiegrebe L** (2003) Classification of virtual objects in the echolocating bat, *Megaderma lyra*. *Behav Neurosci* 117: 833-839.
79. **Weissenbacher P, Wiegrebe L, Kossl M** (2002) The effect of preceding sonar emission on temporal integration in the bat, *Megaderma lyra*. *J Comp Physiol [A]* 188: 147-155.
80. **Wiegrebe L, Schmidt S** (1996) Temporal integration in the echolocating bat, *Megaderma lyra*. *Hearing Research* 102: 35-42.
81. **Wittekindt A, Drexl M, Kossl M** (2005) Cochlear sensitivity in the lesser spear-nosed bat, *Phyllostomus discolor*. *J Comp Physiol [A]* 191: 31-36.
82. **Wyndham E** (1980) Diurnal Cycle, Behavior and Social-Organization of the Budgerigar *Melopsittacus-Undulatus*. *Emu* 80: 25-33.
83. **Wytttenbach RA, Hoy RR** (1993) Demonstration of the Precedence Effect in An Insect. *Current Directions in Psychological Science* 94: 777-784.
84. **Yost WA, Guzman SJ** (1996) Auditory processing of sound sources: Is there an echo in here? *Current Directions in Psychological Science* 5: 125-131.

## Figure List

Figure 1: sound travel path .....	2
Figure 2: Perceived position of a sound source in a lead-lag paradigm.....	4
Figure 5: Experimental setup for <i>Megaderma lyra</i> .....	12
Figure 7: Experimental setup for <i>Phyllostomus discolor</i> .....	15
Figure 8: Expected response behavior of echo-suppressing bats in the lead-lag paradigm .....	17
Figure 9: Echolocation calls .....	18
Figure 10: <i>M. lyra</i> contact call.....	20
Figure 11: Inverted <i>Megaderma lyra</i> contact call .....	22
Figure 12: First syllable of the <i>M. lyra</i> contact call .....	23
Figure 11: Echo suppression in the echolocating bat <i>M. lyra</i> .....	25
Figure 12: Echo suppression in the echolocating bat <i>P. discolor</i> .....	26
Figure 13: Suppression of impulse echoes in <i>M. lyra</i> .....	28
Figure 14: Suppression of contact call echoes in <i>M. lyra</i> .....	29
Figure 15: Suppression of echoes of an inverted contact call in <i>M. lyra</i> .....	30
Figure 16: Suppression of echoes of the first syllable of a contact call in <i>M. lyra</i> .....	31
Figure 17: Distance between sender and reflector for the observed delay range.....	35
Figure 18: Audiogram of man and Mongolian gerbil.....	39
Figure 19: Experimental setup for the investigation of frequency-dependent echo suppression in the Mongolian Gerbil .....	40
Figure 22: Stimuli used for the psychophysical experiments .....	50
Figure 23: Simplified illustration how scaled versions of one of the presented virtual objects could look like .....	51
Figure 24: Experimental setup for <i>Phyllostomus discolor</i> .....	52
Figure 25: Block diagram for the impulse response playback set-up.....	53
Figure 26: Schematic illustration of the simulation of the classification of scaled virtual objects via auditory spectrograms in <i>P discolor</i> .....	56
Figure 27: Auditory spectrograms generated by echoes as they are perceived by the bats.....	57
Figure 28: Psychophysical classification performance and simulation results.....	60

**Figure 29: 5x5 stimulus matrix for the electrophysiological experiments .....63**

### Acknowledgements

I want to thank PD Dr. Lutz Wiegrebe for giving me the opportunity to do this thesis. Since the last five years he gave me persistent enthusiastic guidance and support. I am grateful for all the scientific opportunities the work with him offered. Thank you for giving me the chance to join the DFG Graduiertenkolleg 267 ‘Sensorische Interaktionen in biologischen und technischen Systemen’ and to work on the VW-project ‘auditory object normalization’.

Many thanks go to Daria Genzel and Sven Schörnich for their enormous scientific and personal support and motivation during the last years. Thank you for your friendship and a great time in the lab and in the office. I really enjoyed working with you!

I want to thank Uwe Firzlaff for all the constructive discussions. Thanks for the support and all helpful advices.

Many thanks also to Jan Grunwald for his preparatory work for the VW-project and for all helpful suggestions.

Thanks to Miriam Wolff, who is enthusiastically working on a continuation of the precedence effect study and became a good friend.

I want to thank Holger Görlitz, Frank Borina, Leonie Baier and Andrea Lingner for many constructive discussions, motivation, good advices and support. Thank you for creating an exceptionally warm working atmosphere and for your friendship!

Thanks to Dieter Leippert for his support concerning the animals and for helpful comments.

Furthermore, I want to thank all members of the neurobiology group for a great time. I enjoyed working at the department very much over the last couple of years.

Special thanks go to Rebekka von Leoprechting. Thanks for your friendship, your support in good and bad times and the power you gave me.

At least, I want to thank my parents Hiltrud and Werner for their constant motivation support in all areas. Many thanks that you make this work possible.

## **Curriculum vitae**

### **Personal data**

Name: Maike Schuchmann  
Address: Amalienstrasse 51  
80799 München  
Date and Place of Birth: 05.11.1977, Gross- Gerau  
E-Mail: m-schuchmann@gmx.de

### **Education:**

Since 09/ 2003 PhD- Student, Ludwig-Maximilian- University Munich, Department Biologie II / Neurobiology  
“Localization dominance in bats” and “Auditory object normalization in echolocation”  
From 01/2004-12/2004: supported by a scholarship of the DFG Graduiertenkolleg GK 267 “Sensomotorische Interaktionen in biologischen und technischen Systemen“  
  
Since 01/2005 supported by a grant from the Volkswagen Stiftung Initiative

Since 1998 Ludwig- Maximilians- University Munich; Department of Biology  
Major in Neurobiology,  
Minor in Cell biology, Anthropology and Pharmacology and Toxicology  
  
Thesis work for Diplom (Master) conducted at the Institute of Neurobiology on “ Lokalisationsdominanz im Rahmen der Echoortung”

1997 Johannes-Gutenberg University Mainz; Department of Biology

1988-1997 Secondary school: Justus Liebig Gymnasium, Darmstadt

1984-1988 Primary school: Carl Ulrich Schule, Darmstadt

**Publication list**

Maïke Schuchmann., Matthias M. Hübner and Lutz Wiegrebe (2006):

*The absence of spatial echo suppression in the echolocating bats Megaderma lyra and Phyllostomus discolor*

J. Exp. Biol. 2006 Jan; 209(Pt 1): 152-7

Uwe Firzlaff, Maïke Schuchmann, Jan-Eric Grunwald, Gerd Schuller and Lutz Wiegrebe

*Object-oriented Echo Perception and Cortical Representation in Echolocating Bats*

Submitted to PLoS Biology

Maïke Schuchmann and Lutz Wiegrebe

*Semantic meaning determines spatial echo suppression in bats*

In prep.

## **Ehrenwörtliche Erklärung**

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

München,

Maike Schuchmann