Morphological and behavioral mechanisms underlying sound localization in barn owls

Von der Fakultät für Mathematik, Informatik und Naturwissenschaften der RWTH Aachen
University zur Erlangung des akademischen Grades eines Doktors der Naturwissenschaften
genehmigte Dissertation

vorgelegt von
Diplom-Biologe
Lutz Kettler
aus Siegburg

Berichter:
Univ.-Prof. Dr. rer. nat. Hermann Wagner
Prof Dr. rer. nat. Jörg Mey

Tag der mündlichen Prüfung: 26. Mai. 2015

Diese Dissertation ist auf den Internetseiten der Universitätsbibliothek online verfügbar.
Content

LIST OF FIGURES AND TABLES..................................................................................................................III

ABBREVIATIONS..............................................................................................................................................IV

1 SUMMARY ................................................................................................................................................1

2 ZUSAMMENFASSUNG .................................................................................................................................3

3 GENERAL INTRODUCTION......................................................................................................................5

3.1 Investigating sound localization in barn owls .....................................................................................5

3.2 Neural processing of binaural cues in barn owls ...............................................................................8

3.3 Aim and structure of the thesis .........................................................................................................9

4 INFLUENCE OF DOUBLE STIMULATION ON SOUND-LOCALIZATION BEHAVIOR.................................11

4.1 Introduction ..........................................................................................................................................11

4.2 Methods ...............................................................................................................................................12

4.2.1 Animal handling...............................................................................................................................12

4.2.2 Experimental procedure ...............................................................................................................12

4.2.3 Double stimulus paradigm ............................................................................................................13

4.2.4 Data analysis ...................................................................................................................................14

4.3 Results ................................................................................................................................................15

4.3.1 General characteristics of head turns .........................................................................................15

4.3.2 Influence of an adapter stimulus on the number of rewarded trials ..........................................17

4.3.3 Influence of an adapter stimulus on response latency.................................................................18

4.3.4 Influence of an adapter stimulus on localization accuracy and precision ..................................20

4.3.5 Combined effect on response latency and localization accuracy ...............................................24

4.4 Discussion ..........................................................................................................................................25

4.4.1 Head turns in the single-stimulus paradigm .................................................................................25

4.4.2 Influence of an adapter stimulus on sound localization .............................................................25

4.4.3 Relation between behavioral and neuronal adaptation ..............................................................26

4.4.4 Relation to attention and forward masking ..................................................................................27

5 SPECTRO-TEMPORAL INTEGRATION IN SOUND LOCALIZATION AS REVEALED BY FREQUENCY-MODULATED TONES ..................................................................................................................29

5.1 Introduction ..........................................................................................................................................29

5.2 Methods ...............................................................................................................................................30

5.2.1 Animal handling...............................................................................................................................30

5.2.2 Stimulation .......................................................................................................................................31

5.2.3 Psychophysical procedure ............................................................................................................32
5.2.4 Data analysis ........................................................................................................33
5.2.5 Cross-correlation model .......................................................................................33
5.3 Results ......................................................................................................................36
  5.3.1 Characteristics of head turns ..............................................................................37
  5.3.2 Influence of bandwidth on phantom-source localizations ..................................41
  5.3.3 Influence of stimulus duration on phantom-source localizations .......................42
  5.3.4 Estimation of the duration of the temporal window of frequency integration ..........44
5.4 Discussion ..............................................................................................................47
  5.4.1 Use of phantom-source localization as a tool to study spectral and temporal integration ..48
  5.4.2 Modelling time windows ..................................................................................48
  5.4.3 Binaural sluggishness .......................................................................................49
6 EARDRUM DIRECTIONALITY INDUCED BY INTERAURAL TRANSMISSION
6.1 Introduction ...........................................................................................................51
6.2 Methods ................................................................................................................54
  6.2.1 Directionality measurements ...........................................................................54
  6.2.2 Transmission measurements ............................................................................55
  6.2.3 Data analysis ....................................................................................................56
6.3 Results ..................................................................................................................57
  6.3.1 The external auditory meatus, the eardrum and columella, and the interaural canal of the barn owl .57
  6.3.2 The sensitivity of the barn owl ear ....................................................................59
  6.3.3 Coherence as a measure of reliability ...............................................................60
  6.3.4 Directionality ...................................................................................................62
  6.3.6 Transmission gain ..............................................................................................64
6.4 Discussion .............................................................................................................65
  6.4.1 Methodological considerations .......................................................................65
  6.4.2 Relation to earlier interaural transmission data ...............................................66
  6.4.3 Relation to sound localization .........................................................................67
  6.4.4 Mechanisms underlying interaural coupling ....................................................67
7 GENERAL DISCUSSION ..........................................................................................69
  7.1 Mechanisms underlying optimization of sound-source localization .....................69
  7.2 Suggestions for auditory processing underlying sound-localization behavior ........71
  7.3 Conclusion and outlook .......................................................................................72
8 REFERENCES ...........................................................................................................74
9 ACKNOWLEDGEMENTS .......................................................................................84
10 CURRICULUM VITAE ............................................................................................85
List of Figures and Tables

Figure 4.1: Scheme of the behavioral setup and the stimulus paradigm. 13
Figure 4.2: Traces of head turning responses of owl Hu. 16
Figure 4.3: Percentage of rewards after the first saccade. 17
Figure 4.4: Head turning latencies in owl Hu. 19
Figure 4.5: Angles of the fixation after the first saccade. 20
Figure 4.6: Linear fits of fixation angles of 4 owls. 21
Figure 4.7: Accuracy of head turns as a function of interstimulus interval. 22
Figure 4.8: Precision of head turns as a function of interstimulus interval. 23
Figure 4.9: Combined effect of an adapter stimulus on latency and accuracy. 24
Figure 5.1: Flow chart of the cross-correlation model. 36
Figure 5.2: Head turning properties and response latency. 38
Figure 5.3: Traces of head turns and distributions of fixation angles of owl Bu. 40
Figure 5.4: Percentage of correct responses as a function of stimulus bandwidth. 42
Figure 5.5: Percentage of correct responses as a function of bandwidth and duration. 43
Figure 5.6: Response of a cross-correlation model with a running temporal window of integration of frequency channels. 45
Figure 5.7: Model performance and time constants. 46
Figure 6.1: Mechanisms of low frequency sound receivers. 53
Figure 6.2: Experimental setup for directionality measurements. 55
Figure 6.3: Eardrum and interaural canal. 58
Figure 6.4: Eardrum transfer function. 60
Figure 6.5: Coherence function. 61
Figure 6.6: Directionality of the eardrum. 63
Figure 6.7: Interaural transmission and directionality simulations. 65

Table 5.1: Model time constants. 47
## Abbreviations

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>AAr</td>
<td>acoustic arcopallium</td>
</tr>
<tr>
<td>CF</td>
<td>center frequency</td>
</tr>
<tr>
<td>CM</td>
<td>cochlear microphonics</td>
</tr>
<tr>
<td>FM</td>
<td>frequency modulated</td>
</tr>
<tr>
<td>HRTF</td>
<td>head-related transfer function</td>
</tr>
<tr>
<td>IC</td>
<td>inferior colliculus</td>
</tr>
<tr>
<td>ICC</td>
<td>central nucleus of the inferior colliculus</td>
</tr>
<tr>
<td>ICCc</td>
<td>core of the central nucleus of the inferior colliculus</td>
</tr>
<tr>
<td>ICCls</td>
<td>lateral shell of the central nucleus of the inferior colliculus</td>
</tr>
<tr>
<td>ICX</td>
<td>external nucleus of the inferior colliculus</td>
</tr>
<tr>
<td>ILD</td>
<td>interaural level difference</td>
</tr>
<tr>
<td>ISI</td>
<td>interstimulus interval</td>
</tr>
<tr>
<td>ITD</td>
<td>interaural time difference</td>
</tr>
<tr>
<td>MSE</td>
<td>mean squared error</td>
</tr>
<tr>
<td>NA</td>
<td>nucleus angularis</td>
</tr>
<tr>
<td>NL</td>
<td>nucleus laminaris</td>
</tr>
<tr>
<td>NM</td>
<td>nucleus magnocellularis</td>
</tr>
<tr>
<td>OT</td>
<td>optic tectum</td>
</tr>
<tr>
<td>SPS</td>
<td>side-peak suppression</td>
</tr>
<tr>
<td>VAS</td>
<td>virtual acoustic stimulus</td>
</tr>
</tbody>
</table>
1 Summary

In this thesis, properties of the acoustic system and sound-localization behavior of barn owls were investigated. The influence of adaptation on sound localization behavior was examined in the first experiment. Payne (1971) was able to observe that barn owls wait for at least a second sound before they approach their prey. This situation was mimicked in a behavioral experiment to investigate how a preceding stimulus accuracy of the response to the second stimulus. It is known that the response of neurons to a second stimulus is decreased compared to the response to the first stimulus. This phenomenon is called response adaptation. This means that the detection threshold of the second stimulus may be elevated stimulus and, therefore, response-adaptation influences localization accuracy of the owl. Response adaptation was examined with a double stimulus paradigm. The owl had to locate a broadband noise token, which was preceded by another broadband noise token. I found out that the accuracy and precision with which the barn owls localized the sound source, decreased with double stimulation compared to the condition with only a single stimulus. By varying the interval between the end of the first and onset of the second stimulus I was able to show that the adaptive or masking effect of the first stimulus expires after a few hundred milliseconds. The results suggest that waiting for the second stimulus actually caused costs in terms of decreasing accuracy.

In the second study, the head-turning behavior was used to compare responses to frequency-modulated and stimuli with stationary stimulus content. Barn owls detect time differences in the arrival of sound at both ears and can thus determine the azimuth of a sound source. When stimulated with narrow-band stationary stimuli, however, barn owls locate so-called phantom sources, i.e. they turn their head to a position that does not correspond to the actual sound source. The position of the phantom source can be predicted by the period of the center frequency and a known factor that converts the time differences in an angle. The percentage of phantom localization was determined as a function of stimulus bandwidth. Phantom sound sources are not localized at high stimulus bandwidths. Integration of frequency information in the auditory pathway of the barn owl leads to a reduction of phantom-source locations. Frequency-modulated tones offered the opportunity to present the same frequency content as with stationary noise, but within a certain time interval. This allows determination of the duration of the time window in which the frequency information is integrated. The behavioral data could be well explained with a model that simulates two important processes in the auditory pathway of the barn owl: 1) binaural interaction 2) integration of frequency information. The time constants of the time windows had a duration between 2 and 17 ms for both model steps and did not depend on the stimulus duration.
Summary

In the third series of experiments I investigated whether the tympanic membrane of the barn owl functions as pressure receiver or as pressure-gradient receiver. In a pressure receiver, as it occurs in mammals, both middle ears are not acoustically coupled. This means that no sound is transmitted through an intracranial, interaural canal. Especially in small lizards, but also birds, however, there are cavities that couple both middle ears. Sound does not only reach the eardrum from outside, but also through the interaural canal. The incoming signals are phase shifted. The phase shift depends on stimulus location. In the case of lossless sound transmission through the interaural canal certain sound directions lead to complete extinction of the eardrum vibration. Consequently, the reduction of the eardrum vibration also depends on the degree of sound attenuation through the canal. To measure ear coupling the eardrum vibration was measured with a laser Doppler vibrometer. Eardrum vibration was measured as a function of stimulus frequency and azimuth. In addition, the actual attenuation of acoustic signals by the interaural canal was measured. The tympanic membrane was directional up to 3 kHz. That is, the eardrum vibration amplitude varied by more than 3 dB in 360° of stimulation angles. These data can be explained by attenuation of sound through the interaural canal. For frequencies higher than 3 kHz attenuation was too high to produce significant directionality.
2 Zusammenfassung


2 Zusammenfassung


3 General introduction

One of the most challenging problems in understanding organisms is how the brain generates behavior and how behavior is related to morphology. The discipline of “Neuroethology” investigates neural mechanism underlying behavior by studying behavior under well-defined and controlled conditions and draws inferences from that behavior about its neuronal basis. Neuroethological research is often performed with model organisms which are specialists in a certain ecological niche, typically connected with excellent sensory performance. This is accompanied by enlarged and more easily accessible brain areas that process sensory information. In addition, insight on evolutionary adaptations to specialized behavior is obtained from neuroethological research by comparing behavioral and sensory performance in different species. Such adaptations can also be found in the morphology of brain and body structures linked to specific sensory and behavioral tasks. The barn owl (Tyto furcata pratincola) is a model organism for correlating behavioral patterns with morphological and physiological data. The barn owl exhibits several behavioral, morphological and neuronal adaptations to sound perception and localization that have been subject to neuroethological research in the past 40 years. Three such adaptations in barn owls shall be described in the present thesis: I) Influence of neural adaptation on sound-localization behavior. II) Spectro-temporal integration in the auditory system. III) Eardrum directionality induced by transmission of sound through the interaural canal.

3.1 Investigating sound localization in barn owls

In contrast to diurnal birds of prey of the order Falconiformes, which predominantly hunt by sight, barn owls hunt in twilight or at night and, thus, cannot solely rely on their sense of vision. From this it follows that the sense of hearing is well adapted to hunting in weak lighting conditions (Orlowski et al. 2012). This includes adaptation of morphological structures like the prominent facial ruff or adaptations of neural processing of sounds concomitant with enlarged brain areas responsible for this task.

Payne and Drury (1958) were the first to show that barn owls are able to detect and localize prey by exclusively using the sense of hearing. The studies of Payne and Drury (1958) were followed up by similar investigations by Payne (1962, 1971). The experiments of Payne (1962, 1971) were performed in controlled conditions outside the natural habitat of the barn owl and excluded the senses of vision, chemo-sensation, and even magnetic and electric senses as potential main source for the excellent prey-capturing capabilities of barn owls. The work of Payne (1971) as well as Konishi (1973) paved the way for the barn owl becoming a model for sound detection and localization. A lot
of behavioral and physiological data have been collected since then, creating detailed insight to the processes underlying sound detection and localization.

The free-flight experiments performed in the early 70ies by Konishi demonstrated an exceptionally large frequency hearing range compared to other birds (200 Hz to 12 kHz, compare Konishi, 1973 but also Dyson et al. 1988). In addition, the hearing threshold was considerably lower than, for example, in humans. The detection threshold of barn owls lies between -10 and -20 dB SPL from 2 to 8 kHz (Konishi 1973). But what are the reasons for such a low detection threshold? Payne (1962) and Konishi (1973) speculated that the facial disc may function as paraboloid antenna and, thus, may amplify the sound prior to reaching the eardrum. Coles and Guppy (1988) later found out that, indeed, the facial disc serves as two exponential horns (left and right hemisphere of the facial disc) that amplify the sound by about 20 dB before reaching the eardrums.

The morphological adaptations are also essential for precise sound localization in barn owls in that they alter physical cues that are indicators for the sound-source position. The main cues for sound localization are interaural time difference (ITD) and interaural level difference (ILD). The ITD is the difference of the arrival time of sound waves at spatially separated ears. The ILD arises from attenuation of sounds traveling around the head at high frequencies. Typically both parameters vary with azimuth, if both ears are symmetrically aligned on a horizontal axis, like in humans. The situation is different in barn owls, since the ears are asymmetrically arranged (Payne, 1971), which generates ILDs also along the vertical axis. The facial ruff, consisting of acoustically transparent auricle feathers and reflective coronal feathers, additionally alters sound location cues. The range of ITDs is increased compared to the ITD range expected from the head size (von Campenhausen and Wagner 2006). The range of ITDs also reaches its maximum beyond 90° azimuth due to the influence of the facial ruff (von Campenhausen and Wagner 2006; Hausmann et al. 2010). In addition, the ruff eliminates front-back confusions which would typically be expected with a spheroid head (Hausmann et al. 2009).

Sound localization in barn owls was first tested in free-flight conditions. Konishi (1973) found out that the best performance was achieved with tonal frequencies between 5 and 9 kHz. Singheiser et al. (2010) and Hausmann et al. (2008) made similar observation also taking into account flight corrections toward concurrent and changing sound sources. Hence, barn owls are obviously capable of fast adaptations of their hunting strategy. However, free-flight experiments were not the only behavioral experiments performed with barn owls. Payne (1971) observed that the barn owl saccadically turns its head toward a sound source before actually striking the potential prey after hearing a second sound. These saccadic head turns in response to hearing a sound created the possibility of in depth investigation of sound-localization accuracy in barn owls. Using the so called
search coil technique, that utilizes the currents induced in a metal coil that changes its orientation within a magnetic field, Knudsen et al. (1979) and Knudsen and Konishi (1979) could measure the accuracy of sound localization in barn owls provided the fact that barn owls are nearly incapable to move their eyeballs. Thus, the head turning amplitude measured with a search coil mounted on the owl’s head delivered a precise measure of sound-localization performance in barn owls. A major step, thereafter, was the utilization of so called virtual acoustical stimuli (VAS). VAS are created by filtering the stimulus with the bird’s head related transfer function (HRTF). Therefore, VAS are presented via earphones. They allow specific manipulation of sound location cues. Poganiatz et al. (2001) and Poganiatz and Wagner (2001) used such manipulated VAS and concluded that indeed ILD is an important factor for elevational sound localization while ITD exclusively determines the azimuthal sound-source location. Barn owls localize horizontal sound sources even if the stimuli that are presented via earphones contain only ITDs as cues for the sound-source location (Wagner 1991; Saberi et al. 1998a). That means, that the signals at the left and right ears are fully correlated but either the left or right signal is leading by a sub-millisecond time interval. In such a case the barn owl responds with a similar saccadic head turn toward a sound source as if the stimulus would be presented from a free-field location corresponding to the respective ITD. Unaffected of the stimulus being presented via earphones or free-field loudspeakers the owls typically tend to undershoot the actual sound-source location, that is the owls turn their head not far enough (Poganiatz et al. 2001; Hausmann et al. 2009). The undershooting is greater the more peripheral the sound source is located. This effect could be well explained by a Bayesian estimator that has a tendency to frontal sound sources (Fischer and Peña 2011). However, utilizing pupillary dilation the azimuthal accuracy of barn owls was determined to be about 3° in the frontal space (Bala et al. 2003).

Parallel to advances in behavioral experimentation with barn owls, the neural mechanisms underlying sound-localization behavior in barn owls have been discovered. Knudsen et al. (1977) and Knudsen and Konishi (1978) discovered space-specific neurons, i.e. neurons in the auditory pathway that have restricted spatial receptive fields. These cells represent a specific location in auditory space. It is striking that the elevational and azimuthal components of these receptive fields are due to the responsiveness to both ITD (varying in horizontal plane) and ILD (mainly varying in vertical plane). The space-specific neurons create a two-dimensional map of the auditory space underlying sound-source localization behavior. In the next paragraph I shall expand on the neuronal computation of the binaural parameters.
3.2 Neural processing of binaural cues in barn owls

Sound signals indicate the location of danger, prey, or may serve as communication signals. Localization of objects and events by using the sense of hearing functions different as with other senses. In the senses of vision or touch, for example, maps at the level of the sensory epithelium predict the locus of the respective stimulus. By contrast, the location of sound stimuli must be computed from different physical parameters that are inherent in the sound waves incoming at each ear. Such parameters are the sound level, the temporal fine structure, or the phase of the sound. These parameters contain either monaural information about the sound source or are compared between the two ears which leads to processing of binaural parameters like the interaural time or phase difference and the interaural level difference (for review see Blauert 1997).

Axon collaterals of the auditory nerve separately carrying level and phase information project to the cochlear nuclei nucleus magnocellularis (NM) and nucleus angularis (NA). At these stations level and time information is separated by morphological specializations (Moiseff and Konishi 1983; Takahashi et al. 1984; Carr and Boudreau 1993) before converging in the auditory midbrain again to create the map of the auditory space (Olsen et al. 1989; Fischer et al. 2009). The first neurons that are sensitive to binaural, phase-sensitive input are located in the nucleus laminaris (NL). Nucleus magnocellularis axons project to both the ipsilateral and contralateral NLs carrying precise information about the temporal structure of the signal. Displaying the temporal structure is achieved by locking action potentials to the phase of the sound signal in NM neurons (Köppl 1997a). A specialization of barn owls is the ability of the auditory nerve and NM neurons to lock to the phase of signals with frequencies up to 9 kHz (Köppl 1997a). The consequence of high-frequency phase locking is that barn owls can use ITD in frequency ranges that also elicit ILDs and, thus, precisely localize sound sources in almost its entire hearing range, especially azimuthal sound sources with high frequencies. The actual detection of ITD is supposed to be consistent with a model of coincidence detection introduced by (Jeffress 1948). The model consists of delay lines and coincidence detectors. The coincidence detector responds maximally when input from both ears arrive simultaneously. More than one coincidence detector must exist to detect a wide range of ITDs. The inputs to the coincidence detectors, therefore, form delay lines which compensate for different external delays (ITD). The nucleus laminaris, together with its inputs from nucleus magnocellularis, remarkably resembles the structures proposed by Jeffress (1948) for the detection of ITD (Carr and Konishi 1990). The somas of the NL neurons function as coincidence detectors while the axons of the NM neurons act as delay lines. The activity of NL neurons in barn owls varies as a function of ITD, and can be described well by a cross-correlation of ipsi- and contralateral inputs. Other reptiles possess similar layout and functionality of NL (emu: MacLeod et al. 2006; alligator: Carr et al. 2009). Another possible source
of sensitivity to ITD is to generate directionality of eardrum vibrations induced by the transmission of sounds through an interaural. Eardrum directionality can in turn lead to ITD sensitivity in nuclei that only receive monaural input. This mechanism creates another possible strategy of coding the auditory as for example found in lizards (Christensen-Dalsgaard et al. 2011), but it is not known whether such a mechanism is established in barn owls.

Since NL neurons are narrowly tuned to frequency, the response to ITDs varies periodically which leads to an ambiguous representation of ITD. The representation of horizontal space remains ambiguous in the downstream central nucleus of the inferior colliculus (ICCc). The ambiguity is solved by the projections of several narrowband tuned neurons in the lateral shell of the inferior colliculus (ICCls) to the external nucleus of the inferior colliculus (ICX). ICX neurons are tuned to a broad frequency range, but typically not to frequencies below 3 kHz. The tuning to ITD of ICX neurons is, therefore, characterized by a maximum response at a specific frequency independent ITD and suppressed side-peaks. The underlying process is called side-peak suppression (SPS). ILD and ITD pathways already started to converge in the ICCls (Mazer 1998), and create a map of the auditory space in the ICX. Lesioning experiments (Wagner 1993) revealed that indeed the ICX is crucial for sound-localization behavior. Deficits in the localization of sound sources were produced by destroying parts of the space map. The space-map in the ICX is congruent with the ITD and ILD varying along different axes (von Campenhausen and Wagner 2006) and conclusions from behavioral experiments with virtual acoustic stimuli (Poganiatz et al. 2001; Poganiatz and Wagner 2001)

3.3 Aim and structure of the thesis

In the following chapters 4-6 I shall describe three major experiments tackling sound source localization behavior in barn owls. In chapter 4, I will describe the influence of double stimulation on sound-source localization in barn owls. This work has already been published (Kettler and Wagner 2014). Payne (1971) observed that barn owls do not immediately approach a source after they hear a sound, but wait for a second sound before they strike. This represents a gain in striking behavior by avoiding responses to random incidents. However, the first stimulus is also expected to change the threshold of perception for the subsequent second sound, thus possibly introducing some costs. I mimicked this situation in a behavioral double-stimulus paradigm utilizing saccadic head-turns of owls. The first stimulus served as adapter, was presented in frontal space, and did not elicit a head turn. The second stimulus, emitted from a peripheral source, elicited the head turn. The time interval between both stimuli was varied. Data obtained with double-stimulation were compared with data collected with a single stimulus from the same positions as the second stimulus in the double-stimulus
3.3 Aim and structure of the thesis

paradigm. Sound-localization performance was quantified by the response latency, accuracy, and precision of the head turns.

In chapter 5, I will describe how I used frequency-modulated sounds to study temporal integration in sound localization in the barn owl. Two barn owls were trained to respond to a sound in which the location was specified by the interaural time difference with a head turn towards the perceived sound direction. In such a situation, barn owls turn their head towards the perceived sound source, if broadband noise is used as stimulus. If the stimulus is narrowband, so-called phantom-source localizations may occur, indicating that the barn owl perceives two images. I hypothesized that frequency-modulated sounds might lead to similar behavior as stationary sounds, if a temporal window is taken into account. I further hypothesized that the size of the temporal window may be estimated from the percentage of phantom-source localizations when the bandwidth and duration of the frequency-modulated sounds were varied.

In chapter 6, I investigated eardrum directionality in anesthetized barn owls with the help of laser vibrometry. Both middle ears of barn owls are coupled by a cranial cavity. I show that sound is transmitted through this so called interaural canal inducing directivity in the eardrum vibration amplitude for frequencies up to 3 kHz. In animals with almost un-attenuated sound transmission through the interaural canal, for example lizards, eardrum directionality induces ITD sensitivity at early monaural stages of the auditory pathway (Christensen-Dalsgaard et al. 2011). In such a case, directionality creates yet another strategy for ITD coding in contrast to the place code found in the midbrain of barn owls. Since barn owls use mainly frequencies above 3 kHz for sound localization, where the eardrum is not directional, internal ear coupling does not influence precise sound localization.
4 Influence of double stimulation on sound-localization behavior

4.1 Introduction

Barn owls mainly rely on auditory cues for predation (Wagner et al. 2012). Auditory perception of the barn owl is characterized by high sensitivity to faint sounds like the rustling of a mouse (Payne 1971; Konishi 1973; Dyson et al. 1998). After hearing a sound, a barn owl typically waits for a second sound before it flies toward its potential prey (Payne 1971). This increases the chances of listening to a real source and not a random incident. However, this first stimulus may also change the threshold for perception for the subsequent second sound. Either the bird may become more sensitive or less sensitive. The former case might include mechanisms of arousal and attention (Posner 1980; Knudsen 2007; Marín et al. 2012), while latter case is known as masking or adaptation. From a formal point of view, the first sound may be functioning as a (constant) reference in the situation described above, while the second sound might be viewed as a test or probe sound that may vary in several respects compared to the first sound. Since the findings reported below demonstrate adaptation, I shall only describe masking phenomena in the following.

Psychophysical forward masking is a widely investigated phenomenon, which is measurable as the increase of the detection threshold of a probe stimulus, if it is preceded by another stimulus (Zwislocki et al. 1959; Deatherage and Evans 1969; Bee and Klump 2004). The neural mechanism that likely accounts for forward masking is response adaptation (Nelson et al. 2009). Response adaptation is defined as the reduction of the response of a neuron to a second stimulus compared to the response to a first stimulus without a change in the stimulus. Response adaptation is modulated by spatial, temporal or loudness parameters. For example, in the temporal domain, adaptation decreases as the interval between the first and the second sound is increased (Gutfreund and Knudsen 2006; Singheiser et al. 2012; Wang and Peña 2013).

In the present study I examined the impact of a preceding stimulus on the localization performance in response to a second stimulus. This stimulus regime will be termed double-stimulus paradigm (Singheiser et al. 2012). Different double-stimulus paradigms exist: 1) the first and the second stimulus may be varied independently; 2) the first stimulus may be varied while the second stimulus is held constant; 3) the first stimulus is held constant while the second stimulus is varied. The third paradigm will be used for the work presented here, because it best reflects the behavioral situation in the barn owl as described above. More specifically, I tested how the first stimulus influences response latency, accuracy (nearness of head direction to the direction of the sound source), and precision (variance of localizations) of the natural head turning toward a sound source. I shall report that a first
stimulus decreases the accuracy of head turns and increases response latency. A variation of the silent interval between the two stimuli demonstrated that the adaptive effect persisted for several hundreds of milliseconds.

4.2 Methods

4.2.1 Animal handling

Four adult North American barn owls (*Tyto furcata pratincola*) were used in this study. Owl Bu and Hu were female and owl Q and Ql male, respectively. The owls were treated and cared for in accordance with the guidelines of the “Landespräsidium für Natur, Umwelt und Verbraucherschutz Nordrhein-Westfalen, Recklinghausen, Germany”. Prior to the experiments a metal bar was attached to the skull with dental cement. Surgery was performed under ketamine/diazepam-anesthesia (for more information see Vonderschen and Wagner 2012). During experiments the sensor of the head-tracking system was mounted to the metal bar.

Before the beginning of the experiments the owls were trained by operant conditioning to fixate a position in front of them at 0° azimuth and elevation. A red light-emitting diode (LED) was mounted on a post to mark that position and the LED was switched on, if the owl fixated the position with a deviation of less than 6° in azimuth and 15° in elevation. This position will be called zeroing window in the following. During the training of orienting toward the zeroing window the owls were rewarded after 2 to 5 seconds of stable fixation. During experiments the owls started an experimental trial by fixating the zeroing window. Note, that he LED was removed during the experimental sessions. Training and experimental sessions took about one hour per owl and day.

4.2.2 Experimental procedure

The tests were carried out in an anechoic and sound-attenuating chamber (A403, Industrial Acoustics Company GmbH, Niederkrüchten, Germany). Sound reflections were absorbed by acoustic foam on the inner walls and ceiling. The owl was placed on a perch in the center of the chamber with the feet loosely tethered to the perch to ensure that the bird could not fly away. An infra-red camera positioned directly above the owl allowed monitoring of the owl’s condition and behavior from outside. The camera also allowed rough guesses of head-turning angles and localization performance during an experimental session which was important for the rewarding scheme. A food dispenser was installed in front of the owl. Small pieces of meat were offered as food reward. During experiments, a head-tracking-system was attached to the metal bar on the bird’s head. It should be noted that the
Influence of double stimulation on sound-localization behavior

The head-tracking device had to be changed during the experiments. Experiments with owl Bu were performed only with an Ascension miniBIRD (Ascension Technology Corporation, Burlington, Vermont, USA). The miniBIRD allowed tracking of the head orientation at 144 Hz update rate. In the other three owls measurements were performed with a miniBIRD as well as an Ascension trakSTAR system. The trakSTAR system performed with an update rate of 420 Hz. Orientation accuracy is 0.5° RMS and resolution is 0.1° in both systems. Seven loudspeakers (Visaton VRS 8) were positioned in front of the owl at 0° elevation and a distance of 60 cm from the owl’s head. The azimuthal position of the speakers was -60° (left) to 60° (right) in 20° steps (Figure 4.1). The owls were stimulated with broadband noise with 100 ms duration and flanked by 5 ms cosine ramps. The stimuli were synthesized digitally with custom written C++ and RpvdsEx-programs (48828 Hz sampling rate), and converted into analog signals with a RX8 Multi-I/O processor (Tucker Davis Technology). The signal was then power-amplified (ECLER MPA 4-80) and eventually delivered to one of the seven loudspeakers. Background noise in the chamber was 27 dB SPL. The stimuli were presented at 45 ± 2.0 dB SPL, which means that the stimuli were presented at a signal-to-noise ratio of about 18 dB, well above the threshold of barn owls (Takahashi and Keller 1992; Wagner et al. 1994).

4.2.3 Double stimulus paradigm

To test whether an acoustic stimulus has an influence on localization of a subsequent stimulus I used a double stimulus paradigm. The double stimulus consisted of a preceding stimulus or adapter from
0° azimuth followed by a second stimulus or probe from one of the six peripheral speakers. The silent interval between the offset of the adapter stimulus and the onset of the probe stimulus was varied between 0 and 3200ms on a pseudo-log2-scale (Figure 4.1). The sound pressure level of adapter and probe was equal. The owl was also stimulated with single stimuli from one the six peripheral loudspeakers. Single-stimulus trials were pseudo-randomly interleaved with double-stimulus trials at a rate of 25% in owls Hu, Q, and Ql and at a rate of 75% in owl Bu.

During the experiment the owls started a trial by fixating the zeroing window (as explained in animal handling). After keeping fixation for 0.5 to 1.5 seconds either the adapter in double stimulus conditions was emitted from the 0° loudspeaker or the single stimulus was emitted from one of the peripheral speakers. Note, that the adapter stimulus and the owl’s orientation were the same at that point in time. Thus, no response to the adapter stimulus with double stimulation was expected. Following the adapter stimulus the probe was emitted from one of the peripheral loudspeakers after a pseudo-randomly chosen ISI. The owl’s natural behavior was a fast head turn or saccade which ended in a fixation of the sound source. The owl was rewarded with a small piece of chick meat from the food dispenser if the deviation between fixation angle and sound-source position was smaller than 7° azimuth and 15° elevation and if the fixation lasted at least 250 ms. The spatial window which led to reward is called target window in the following. An experimental trial ended after 3.5 seconds or after the owl was rewarded. The next trial started again when the owl looked into the zeroing window, but starting a new trial was not possible until 3 seconds elapsed after the end of the last trial.

4.2.4 Data analysis
I analyzed head orientation in respect to fixation angles and response latency with custom written MATLAB-programs. Response latency is the delay time between onset of the single stimulus or probe stimulus in double-stimulation and the beginning of the head turning response. The beginning of the head turn was defined as the point in time when the head orientation angle exceeded the orientation at stimulus onset by more than 5°. I quantified fixation angles by calculating the turning velocity after stimulus onset: the velocity had to fall below 20°/s for at least 80 ms. The fixation angle was then determined as the mean orientation angle within the first 80 ms after the velocity became < 20°/s. Head turning accuracy and precision were calculated from fixation angles. Note, that the owls sometimes performed correction saccades, if they did not hit the target window after the first saccade. In these cases only the amplitude of the first head turn was used to measure accuracy and precision, because the first head-turn reflects best the localization performance of the owls. Furthermore, trials were classified either as valid or invalid trials. Invalid trials included trials at which a) the owl shook
its head instead of turning its head toward the sound source (no fixation angle could be calculated), or 
b) the owl did not turn its head at all (head orientation did not exceed 5°), or c) the owls reaction time 
was too long so that I could assume that it was not alert (response latency > 1000 ms), or d) the 
reaction was not correlated with the stimulus (latency < 50 ms), or e) the owl made a head turn after 
the first stimulus in the double-stimulus conditions but before the second stimulus, which means that 
the head orientation head to stay within the zeroing window (+/- 6° azimuth) after the adapter 
stimulus until the probe stimulus was played. Only valid trials were included in the analysis. The 
statistical analyses were performed with the MATLAB statistics toolbox.

4.3 Results

This study is based on 16,674 trials from four adult barn owls of which 13,651 (81.9%) were valid 
trials (see Methods - chapter 4.2.4). Invalid trials mostly resulted from missing the criterion of 
keeping the gaze in the zeroing window before the onset of the probe stimulus (16%). A minority of 
invalid trials resulted from missing a latency criterion (3%). In a minority of trials (<1%) the owl 
missed both criteria which explains a sum >100%. Further analysis is based on valid trials only (owl 
Bu: 1,659; owl Hu: 3,904; owl Q: 3,643; owl Ql: 4,445).

4.3.1 General characteristics of head turns

The natural head-turning response toward the source of a sound signal served as measure for this 
study. After a stimulus was presented, a barn owl performs a short-latency, fast saccadic head turn 
toward the sound source (Figure 4.2a). In this study, I started a trial only if the owl oriented its head 
toward 0° azimuth and 0° elevation within a zeroing window (dotted rectangle on left side in Figure 
4.2a) and for a certain time (see Methods - chapter 4.2.3). Similarly, the owl had to reach a target 
window and maintain fixation within this window for more than 0.25 s to obtain a reward (dotted 
rectangle on right side of Figure 4.2a). If the owl did not reach a position within the target window 
with the first saccade, it often performed correction saccades toward the sound source, which then 
could lead to a food reward. The trace shown in Figure 4.2a represents such a case. The owl first 
performed a short-latency, large-amplitude head turn that failed to reach the target window. The head 
stayed still for some time (up to 1500 ms), before the owl made a second, small-amplitude saccade 
and reached the target window. The bird was then rewarded with a small food pellet. Note that the 
stimulus ceased after 100 ms. The response latency in this case was 110 ms. That means that the 
head-turn started after stimulus offset. Such a head-turn is called open loop here. About 97% of the 
trials resulted in open-loop head turns. In the case in Figure 4.2a a double-stimulus trial is shown.
4.3 Results

The probe stimulus was presented at 0 ms and lasted 100 ms. The probe stimulus was preceded by the adapter stimulus with an ISI of 400 ms between the offset of the adapter stimulus and the onset of the probe stimulus. Note that the adapter stimulus was presented after the owl had oriented its gaze for more than 500 ms in the zeroing window. The owl did not show a change in head orientation after the adapter stimulus was emitted, but reacted only after the probe stimulus had been presented. The owls performed between 30 and 150 both valid and invalid trials per daily experimental session.

Figure 4.2: Traces of head turning responses of owl Hu.

Traces of head turning responses of owl Hu are shown. a Single trace showing the head turning response toward a probe stimulus at -40° azimuth (left hemisphere). The interstimulus interval (ISI) was 400 ms. The owl responded with a latency of 110 ms. The left dotted rectangle illustrates the zeroing window around 0° azimuth which the owl had to fixate to start a trial and to stay within until the probe sound is emitted so that the trial counted as a valid trial. The right dotted rectangle displays the rewarding window. The owl was rewarded if it fixated an azimuthal position within the rewarding window. The owl fixated at 29.7° after the first saccade (horizontal arrow) but was rewarded not until it performed a correction saccade toward the sound source (vertical arrow). b 28 valid head turns collected during one daily experimental session. Trials with single stimulation are shown in red and trials with double stimulation in black. Solid lines show head turns which led to reward, while dotted line show non-rewarded trials. The traces are aligned in a fashion that the onset of the probe or reference sound, respectively, falls on 0 ms response time, independent of the duration of the ISI.

The probe stimulus was presented at 0 ms and lasted 100 ms. The probe stimulus was preceded by the adapter stimulus with an ISI of 400 ms between the offset of the adapter stimulus and the onset of the probe stimulus. Note that the adapter stimulus was presented after the owl had oriented its gaze for more than 500 ms in the zeroing window. The owl did not show a change in head orientation after the adapter stimulus was emitted, but reacted only after the probe stimulus had been presented. The owls performed between 30 and 150 both valid and invalid trials per daily experimental session.

Figure 4.2b shows a collection of 28 valid trials from a daily session with owl Hu. The trials are aligned in a way that for the single-stimulus trials the onset of the single stimulus is set to 0 ms, while for the double-stimulus trials the time axis was adjusted so that the onset of the probe stimulus was set to 0 ms. In all cases the owl reacted with a short-latency saccade. The target window was reached in most cases (solid lines). Note that all reactions of the bird that failed to reach the target window were double-stimulus trials (dashed black lines). The number of saccades needed to find the target window was also significantly increased in all owls (Wilcoxon-rank-sum-test, p < 0.001). The mean number of saccades with single stimuli was 1.73, while it was 2.15 in double-stimulus conditions.

These observations provided a first indication that the adapter stimulus had an influence on the head turning in response to the probe stimulus. In the following I quantify this influence by comparing
Influence of double stimulation on sound-localization behavior

4.3.2 Influence of an adapter stimulus on the number of rewarded trials

The number of rewarded trials provided a first estimate of the localization performance. A trial was rewarded, if the owl reached the target window either after the first saccade or after a subsequent correction saccade. The percentage of rewarded trials was smaller if the stimulus angle was greater, which indicated that the localization accuracy decreased in the peripheral auditory space. As indicated already in Figure 4.2b the number of rewarded trials was smaller in double-stimulus conditions than in single-stimulus conditions.

Figure 4.3 shows the influence of a preceding stimulus on the amount of rewards gained after the first saccade in response to a second stimulus. Rewards that were gained after a correction saccade are excluded in Figure 4.3. In each owl a double stimulus with an ISI of 0 ms led to a significant decrease of the amount of rewarded trials compared to single stimulus trials (chi-square test, p < 0.001). In contrast an ISI of 3200 ms had no significant effect on the percentage of rewarded trials. The decreased percentage of rewards in double stimulus conditions with short ISIs suggested an adaptive effect of the adapter stimulus on the accuracy of the localization of a probe stimulus. The effect disappears after several hundreds of milliseconds.

Figure 4.3: Percentage of rewards after the first saccade.

The percentage of rewarded trials is shown for single stimulation and double stimulation. Color indicates the owl as shown in the box. The asterisks show the uniform results of separate \( \chi^2 \)-tests indicating a highly significant (p < 0.001) difference for the data collected with 0 ms ISI from the percentage of rewards in single stimulus conditions. The percentage of rewarded trials with an ISI of 3200 ms was not significantly different from the percentage with single stimuli. Each data point includes at least 120 trials.
4.3 Results

4.3.3 Influence of an adapter stimulus on response latency

Response latency was defined as the delay time between stimulus onset and the onset of the saccadic head-turning response. Median response latencies in single-stimulus conditions were typically between 120 and 145 ms in owls Bu, Hu, and Q for stimulation from the different peripheral loudspeakers, while owl Ql was reacting slightly slower with median head-turning latencies depending on stimulus position and scattering between 170 to 220 ms in response to single stimuli. Response latencies were longer if the owl was stimulated with a double stimulus. The ISI in the double-stimulus condition also had an influence on head-turning latency. Figure 4.4 illustrates typical examples of response latency as a function of ISI as measured in owl Hu. The increase of response latency in the double-stimulus conditions is clearly seen. A three-way analysis of variance was performed and the latency parameter was ranked (Sheirer-Ray-Hare-test) to quantify the influence of ISI, speaker position, and inter-individual differences on response latencies. The statistical analysis revealed, that response latency differed significantly between owls ($\chi^2 = 2535.1, p < 0.001$). Besides, the position of the sound source also had a significant effect on response latency ($\chi^2 = 97.6, p < 0.001$). Usually, the owls showed a bias of latencies in head turns to either the right side or the left side. A possible explanation for this bias is that the owls knew the position of the door to the sound proof chamber through which the experimenter entered the chamber.

Finally, and most important for this study, the interstimulus interval had an effect on response latency as well ($\chi^2 = 491.4, p < 0.001$). However, the effect disappeared for large ISIs. Response latency in the double-stimulus paradigm was not significantly different from response latency with single stimuli for an ISI of 3200 ms in owls Bu, Hu, and Q (Bonferroni-post-hoc-test). In owl Ql response latency was significantly shorter with an ISI of 3200 ms than in single stimuli in owl Ql (Bonferroni-post-hoc-test, $p < 0.05$). In summary, response latency is affected by an adapter stimulus such that a preceding stimulus generally increases the latency in response to a second stimulus, if the ISI is short. However, the influence of an adapter stimulus disappears after a few hundreds of milliseconds.
The distribution of response latencies is plotted as boxplots as a function of ISI. The plots show response latencies of owl Hu for each of the six speaker positions. The boxes show the median (red lines), the upper quartile, and lower quartile. The whiskers show the 1.5-fold interquartile range plus the upper quartile or the lower quartile minus the 1.5-fold interquartile range. Black dots show outliers larger or smaller as the whisker range, respectively. The horizontal dashed grey lines indicate the median latency with single stimulation.
4.3.4 Influence of an adapter stimulus on localization accuracy and precision

Faint sounds elicit swift head-turning responses toward the sound source (Figure 4.2). I defined fixation angles as the angular position of the head after the first head-turning saccade. To summarize the head-turning behavior I analyzed both accuracy and precision of head turns. Head turning accuracy is the nearness of the head direction to the sound-source direction, which was determined as the average difference between the fixation angle and the angle of the sound source. In contrast, head turning precision is the variability of fixation angles in response to a specific stimulus and position. Figure 4.5 shows data from owl Hu, and illustrates the median fixation angle, the upper and the lower quartile in boxplots. In addition the slope of the regression line, the y-intercept and the regression coefficient are shown. Note that the interquartile range is a measure of precision, while the median fixation angles are a measure of accuracy. Instead of calculating accuracy for every stimulus position, I used the slope of the regression line between the fixation angles and the position of the stimuli to characterize and quantify accuracy of head-turnings (Figure 4.5, Figure 4.6). Fixation angles depended on the stimulus position in a highly linear manner (r > 0.89 in each of the 40 correlations shown in Figure 4.6).

Figure 4.5: Angles of the fixation after the first saccade.

Data is shown for owl Hu. a Fixation angles in response to single stimuli. Boxes indicate median values with upper and lower quartiles. The interquartile range may serve as a measure of variability and thus precision of head turns. A linear regression line was fitted through the median fixation angles at different stimulus positions. The slope of the regression line as in indicated in the formula serves as measure for accuracy. b Same plot as in a) but with double stimulation and an interstimulus interval of 200 ms. Numbers above the boxplots represent the number of trials collected.
For the data from owl Hu shown in Figure 4.5a the slope of the regression line was 0.878, and the y-intercept 0.48 degrees. While the y-intercept was close to the ideal value of 0 degrees, the slope was lower by 12.2%. This means that the owl undershot the target. Undershooting was typical for all owls. Since a linear regression line fit the data well, the undershooting also meant that peripheral stimuli were localized with less accuracy than central stimuli. Both the y-intercept close to 0 degree and the linear dependence of head-turning angles on stimulus azimuth also meant that the head-turning amplitude of the owls was symmetrical with respect to azimuth. In other words, a stimulus from the left yielded head turns of equal absolute amplitude as a stimulus from a symmetrical position on the right hemisphere. Figure 4.5b shows that in the double-stimulus condition the symmetrical turning behavior remained as indicated by the high linear regression coefficient \( r = 0.98 \) and the small y-intercept. By contrast, the undershooting increased as demonstrated by the lower slope (0.78 compared to 0.88). This again was typically of all owls and short ISIs.

**Figure 4.6: Linear fits of fixation angles of 4 owls.**

Circles indicate median fixation angles. Different colors represent the different interstimulus intervals as in Figure 4.1. The slopes of the regression lines are indicated in the particular plots for each owl.
4.3 Results

Figure 4.6 shows a summary of all data. It becomes obvious that the slope was smallest for short ISIs and increased for longer ISIs. A test for a significant influence of ISI on the slope with a non-parametric analysis of covariance (ANCOVA) revealed a significant dependence of the slope on the ISI in each owl (owl Bu: $\chi^2 = 1.97$; owl Hu: $\chi^2 = 1.10$; owl Q: $\chi^2 = 9.64$; owl Ql: $\chi^2 = 1.17$, $p < 0.001$ in each owl). Multiple-comparison tests using Scheffé’s method revealed that the slope at an ISI of 3200 ms was not significantly different from the slope in the single-stimulus conditions in any owl. Thus, an adapter stimulus reduced localization accuracy, but the influence of an adapter decreased gradually and was absent for an ISI of 3200 ms.

The data was fitted with a single exponential function to obtain an estimate of the duration of the influence of an adapter stimulus (Figure 4.7). The time constants obtained from the four owls had the same order of magnitude in owls Bu (466 ms), Hu (252 ms) and Ql (305 ms), and the fit explained.

Figure 4.7: Accuracy of head turns as a function of interstimulus interval. An exponential function was fitted to derive a time constant (τ) of recovery from the masking effect of the adapter. The dashed line indicates the accuracy in single stimulus conditions. (rmse = root-mean-squared error).
4 Influence of double stimulation on sound-localization behavior

97%, 88%, and 83% of the variability, respectively. By contrast, the time constant for owl Q was much longer (2170 ms). However, also the coefficient of determination of the exponential fit was smallest in this owl. The long time constant in owl Q was mainly due to the low number of correct responses for ISIs of 400 ms, 800 ms, and 1600 ms, while the performance at 3200 ms and at the short ISIs was comparable to those in the other birds. The reason for this decreased performance for ISIs of medium duration in owl Q is unclear.

Not only head-turning accuracy, but also head turning precision was affected by double stimulation (Figure 4.8). A Brown-Forsythe test for equality of variances revealed that the interstimulus interval influenced the variance of fixation angles and therefore the precision of head turns. Figure 4.8 shows the interquartile ranges of fixation angles (as a measure for precision) plotted against the interstimulus interval. Data from all loudspeakers were merged for this plot. The precision was best in owl Hu and worst in owl Ql. Although the scatter is considerable, regression analysis demonstrated that in each owl, the precision increases significantly from the shortest ISI to the largest ISI (Figure 4.8). While in owls Bu and Hu the interquartile ranges for all ISIs were larger for the double stimuli than for the single stimulus, in owls Q and Ql the interquartile ranges for 3200 ms were smaller for the double stimuli than for the single stimulus.

![Figure 4.8: Precision of head turns as a function of interstimulus interval.](image)

The precision of head turns was measured as the variability of fixation angles (i.e. shown as the inter-quartile range, iqr). Data from all 4 owls is shown. The solid lines show logarithmic fits with their correlation coefficients denoted in the box in the upper right. Note that precision increases with ISI.
4.3.5 Combined effect on response latency and localization accuracy

I have so far described a significant influence of an adapter stimulus both on response latencies and localization accuracy. The effect disappeared on a time scale of a few hundreds of milliseconds in three of the four birds tested. Figure 4.9 illustrates the combined effect on response latency and on fixation angles evoked by double stimuli with different ISIs. Data from all owls and speaker positions was pooled. Response latency in double-stimulus conditions was divided by response latency in single-stimulus condition to obtain a normalization factor characterizing the influence of the adapter on the response to the probe. An increase in response latency would lead to a factor greater than 1. Absolute fixation angles in the double-stimulus conditions were also divided by absolute fixation angles in the single-stimulus condition to obtain an equivalent normalization factor. If the relative variables are below 1, the owl reacts faster and turns less in the double-stimulus conditions compared with the single-stimulus conditions, and vice versa. The combined effect of increasing response latencies and decreasing head turning accuracy may be seen by clustering of dots in the lower right quadrant. These data suggest that in double stimulus trials, the birds were slower in responding and tended to fall short of the target.

![Figure 4.9: Combined effect of an adapter stimulus on latency and accuracy.](image)

Data is merged from each owl and loudspeaker. A colored dot represents the response with a specific sound source and owl. Altogether 216 data points are shown. On the x-axes the median latency with a specific ISI in double stimulation is divided by the median latency in single stimulation at the same owl and speaker position. On the y-axis the median fixation angle (°) in double stimulation is divided by the median fixation angle in single stimulation with the same owl and speaker position. The colors indicate the ISI as in Figure 4.1. The dotted lines represent the performance in single stimulation. Note that the warmer colors, corresponding to longer ISIs, tend to scatter around the cross that represents the values obtained single stimulation.
4.4 Discussion

This study investigated the influence of a given adapter stimulus on the localization performance in response to a second stimulus. I found out that the owl’s ability to localize sound sources declines in such a double-stimulus paradigm. Specifically, the accuracy and precision of head turns toward the sound source decreased with the introduction of a second stimulus, while response latency increased. In the following, I shall first discuss the behavioral performance of barn owls in the single-stimulus condition and then discuss the influence of an adapter on the sound-localization behavior. I tackle three points: precision and accuracy, relation to neural adaptation, and relation to attention and forward.

4.4.1 Head turns in the single-stimulus paradigm

The natural response of barn owls when hearing faint sounds is a swift head saccade toward the sound source (Knudsen and Konishi 1979; Poganiatz et al. 2001; Poganiatz and Wagner 2001). The response latencies of the owls in the single-stimulus condition were similar to those observed in other studies (Wagner 1993; Hausmann et al. 2009). After the saccade the owl usually fixated the sound source but often also performed correction saccades after a few tens to hundreds of microseconds (Wagner 1993). I analyzed the angles of the fixation after the first saccade in single stimulus conditions as well as with double stimuli as a measure of head-turning accuracy. The fixation angles linearly depended on the sound source. But the owls tended to undershoot the sound-source position. The undershooting was larger for peripheral sound sources, which led to slopes of the best fit smaller than 1 within the range of ~0.5 close to 1. The slopes calculated in this study are comparable to those found in previous studies that dealt with azimuthal head turning behavior in barn owls (Knudsen et al. 1979; Wagner 1993; Poganiatz et al. 2001; Hausmann et al. 2009) and could be explained by a Bayesian estimator with a bias to central positions (Fischer and Peña 2011). Similar undershooting was observed in experiments with ferrets (Nodal et al. 2008) and cats (May and Huang 1996; Tollin et al. 2005; Gai et al. 2013).

4.4.2 Influence of an adapter stimulus on sound localization

The performance of the owls in sound-source localization decreased when a probe stimulus was preceded by another sound stimulus. The owls responded slower (increased response latency) and with less accuracy and precision. I discuss the mechanisms underlying changes in accuracy separately from the mechanisms underlying changes in precision, because these two phenomena have been shown to vary independently (Tollin et al. 2005; Gai et al. 2013).
4.4 Discussion

Precision is associated with internal uncertainty (Tollin et al. 2005), i.e. the signal to noise ratio or response variability in nuclei which represent auditory space. As long as spontaneous activity is not affected a decreased signal to noise ratio would lead to decreased precision. The stimuli (single, probe, and adapter) presented in this study were all equalized in sound pressure level and well above the hearing threshold of the barn owl so that intensity differences cannot not be responsible for the decreased precision I observed in the double-stimulation experiments. However, response-adaptation paradigms typically cause a reduction in neuronal responses (Ingham and McAlpine 2004; Gutfreund and Knudsen 2006). In other words, neuronal noise becomes larger compared with the signal-driven responses. This may elevate detection threshold (Nelson et al. 2009) and would, thus, decrease signal-to-noise ratio.

The undershooting, serving as an indicator of accuracy, observed in the double-stimulus paradigm may be in part also caused by an increased response threshold. This would add to the inbuilt centrality bias (Fischer and Peña 2011). Yet unpublished data (see ARO abstract 2014 by Ferger and Wagner) demonstrates that in neurons of the external nucleus of the inferior colliculus (ICX) an adapter stimulus at a given ITD leads to a similar adaptation for all ITDs that cover the entire physiological ITD range. This suggests that adaptation is broadly tuned in space, which would lead to increased behavioral response threshold.

Furthermore, Winkowski and Knudsen (2006) showed that the response of direction sensitive neurons in the optic tectum may be enhanced by stimulation of arcopallial gaze fields in the owl’s forebrain which encode for the same location, whereas the response to other locations is suppressed. Gaze control circuits may be involved during processing of the double stimuli, decreasing the representation of peripheral sound sources, whereas the center of the gaze field may be aroused after presentation of the adapter stimulus.

4.4.3 Relation between behavioral and neuronal adaptation

Neuronal adaptation means a reduction of the driven activity when the stimulus stays constant. Neural data obtained from barn owls with the double-stimulus paradigm are available from the central nucleus of the inferior colliculus (ICC) (Singheiser et al. 2012). These data demonstrated both a dependence of adaptation on the level of the second stimulus and on the ISI. The dependence of neural adaptation on ISI could best be fitted with a double exponential, with one short time constant (1.25 ms) and one long time constant (800 ms). I did not see a dependence of the slopes on ISI that suggested applying double-exponential fits (Figure 4.7). The time constant found in this study lies in the middle between the slow and the fast time constant found by Singheiser et al. (2012) with the
order of magnitude matching that of the long time constant in the ICC (Singheiser et al. 2012). The ICC is an intermediate nucleus in the auditory pathway with neurons that do not well reflect sound-localization properties of barn owls, because these neurons exhibit phase-ambiguity in response to stimulation with interaural time difference (Wagner et al. 2002). Wang and Peña (2013), using a similar paradigm, reported an influence of double-stimulation in the external nucleus of the inferior colliculus (ICX) that receives input from ICC. Neurons in this nucleus exhibit side-peak suppression (Takahashi and Konishi 1986). The average recovery from adaptation was well fitted with a single exponential with a time constant of 98 ms (Wang and Peña 2013). This cannot explain the time constants I observed. One reason may be the different paradigm that Wang and Peña (2013) used. In conclusion, more data are necessary to clarify the relation between neural and behavioral adaptation.

Time constants of recovery from neural adaptation found in other animals have a similar duration (approx. 100 ms, rat superior olive, Finlayson and Adam 1997; 270 ms, rat IC, Finlayson 1999; 225 ms, guinea pig IC, Ingham and McAlpine 2004). This suggests that adaptive circuits are working on a similar time basis in different species.

Response adaptation also influenced segregation of auditory streams in songbirds (Bee and Klump 2004; Bee and Klump 2005; Itatani and Klump 2009; Bee et al. 2011). Bee and Klump (2004) reported that forward masking, equivalent to response adaptation, contributes to perceptual segregation of acoustic events. Itatani and Klump (2009) included 800 ms silent intervals between triplets of sinusoidally modulated tones. Field L2-neurons were still adapted after this period.

### 4.4.4 Relation to attention and forward masking

The influence of attention on my results could not be confirmed, but also not entirely excluded. It is well known that sound localization in barn owls is influenced by attention (Johnen et al. 2001). If a visual cue contains information about the hemisphere in which a succeeding stimulus will appear (valid cue), response latency is shorter than if the visual points to the opposite hemisphere (invalid cue). The mean response latency for invalid cues was increased by about 40 ms. This value is comparable to the amount of increase in response latency found in my experiments for short ISIs (median response latency was increased between 15 and 60 ms). The similarity of the influence on latency in both the cueing paradigm and the adaptation paradigm might indicate an attentional effect in the adaptation paradigm but could also occur by coincidence and not by the same physiological mechanism.

Forward masking describes the phenomenon of increased detection thresholds of a sound that follows another sound (Deatherage and Evans 1969; Hafter and Carrier 1970; Kollmeier and Gilkey 1990).
Gai et al. (2013), working with cats, observed a masking effect with 25 ms probe duration but not with 50 ms. Gai et al. (2013) interpreted their data as being consistent with binaural sluggishness (Kollmeier and Gilkey 1990; Culling and Summerfield 1998). The time constants determined from my data are much longer, and outside the ranges of time constants determined for binaural sluggishness. One reason for the difference between my data and those of Gai et al. (2013) may be that the latter authors only used ISIs of 0 and 25 ms, while I used ISIs up to 3200 ms. I showed that the recovery from the masking effect of the adapter stimulus underlies temporal release which is also known for forward masking (Smiarowski and Carhart 1975).

Payne (1971) reported that the barn owl usually strikes its prey only after hearing at least a second sound in dark environment. Since the owls are generally low on energy reserves, the cost for unsuccessful strikes is very high and must be avoided. One strategy might be to wait for a second sound to get a confirmation of the location. My results suggest that waiting for a second sound imposes costs on prey capture. However, I like to mention that the paradigm used here was only a first attempt to mimic such a behavioral strategy. In a next step, this paradigm can be refined to, for example, allow a turn after the first sound, and then present the second sound at a central position, and thus reflect the natural situation even further.
5 Spectro-temporal integration in sound localization as revealed by frequency-modulated tones

5.1 Introduction

Time and frequency carry important sensory information. These two stimulus parameters have been intensively studied as separate entities (e.g. frequency tuning in audition: Vonderschen and Wagner 2009; motion vision: Borst et al. 2010). A combination of the two is used to investigate spectro-temporal receptive fields (Keller and Takahashi 2000; Qiu et al. 2003) or sensitivity to frequency-modulated tones in the auditory system (Rees and Möller 1983; Brown and Harrison 2009). One aspect of such studies is sensory resolution, i.e. how well spatial positions in space may be discriminated (vision: Westheimer 2009, audition: Bala et al. 2003). Another aspect is sensory integration, i.e. how long are responses integrated in sensory filters. Sound localization utilizing interaural time difference (ITD) is a special case and theme of this work. Differences in ITDs as low as 10 µs can be discriminated (see Blauert 1997 for a review; Bala et al. 2003). They are used for spatial hearing, because the ITD depends on the spatial position of a sound source. Temporal resolution in the auditory system may be described by several time scales. On the short end of the time scales, gaps as short as 2 ms may be detected in monaural stimuli processing (Plomp 1964; Snell 1994), while on the long end of the time scales, temporal integration may allow to store information for several seconds. A variety of time constants has been reported for temporal resolution of binaural hearing (Wagner 1991; Bernstein et al. 2001; Siveke et al. 2008; Shackleton and Palmer 2010). It is generally agreed, however, that the binaural system is more sluggish than monaural processing (Viemeister 1979; Moore et al. 1988; Snell 1994).

Sound localization is an important behavior that is mainly based on binaural processing. Sound localization is typically studied with stimuli having stationary frequency content. In such a situation, barn owls turn reliably towards the sound source, if the stimulus is broadband. If the bandwidth of the stimulus is less than 3 kHz, barn owls may also turn away from the sound source; a phenomenon known as phantom-source localization (Saberi et al. 1999). Similar behavior has also been observed in humans for narrowband stimuli (Stern et al. 1988; Trahiotis and Stern 1989; Shackleton et al. 1992; Buell et al. 1994). One interesting, unanswered question is whether the complete frequency information must be available all the time while the computation is going on, or whether spectral information may be temporally integrated. This question may be tested with frequency modulated (FM) signals covering a similar frequency band as the broadband noises that yield unambiguous sound localization. Frequency modulation is a common feature of natural sounds. Frequency modulation is assumed to be a major cue of communication, sound perception, and object
identification across species (Sales and Smith 1978; Merzenich et al. 1995; Doupe and Kuhl 1999; Brudzynski 2005; Amador and Margoliash 2013). Many species evolved preferences on detecting a preferred sweep direction on cellular level (Nelken and Versnel 2000; Brown and Harrison 2009) and on behavioral level, for example in bats (Razak and Fuzessery 2010). Besides, vocalizations of barn owls also contain a wide spectrum of frequency modulations (Bühler and Eppe 1980). However, little is known about how dynamic spectral components influence sound localization in barn owls.

In this study I used FM-sounds presented to the two ears with an interaural time difference (ITD) to study temporal processing in the binaural system. ITD is a major cue for sound localization in many species (for a review see Blauert 1997; Konishi 2003; Ashida and Carr 2011). Barn owls use interaural time differences for localizing sound sources in the horizontal plane (Moisef and Konishi 1981a; Wagner et al. 2012). ITDs change systematically with azimuth in this animal (Keller et al. 1998; Brainard and Knudsen 1998; Poganiatz et al. 2001; von Campenhausen and Wagner 2006; Hausmann et al. 2009). When the period, the inverse of the frequency, of a stimulus is smaller than twice the physiological ITD range, ambiguities with respect to the location of a sound may occur for narrowband stimuli (Saberi et al. 1998a) and these ambiguities are the reason for phantom-source localizations mentioned above. This situation exists in barn owls for frequencies larger than 2 kHz (von Campenhausen and Wagner 2006). The barn owl uses the ITD of the carrier in frequency ranges where such ambiguities occur, while humans and most other animals do not employ carrier ITDs at such high frequencies (Blauert 1997). I reasoned that phantom localization may be used as a tool to study not only to study spectral, but also to study temporal processing.

I report that frequency modulated tones (chirps) with different chirp rates (i.e. the rate of change of the instantaneous frequency) lead to phantom localizations as long as the frequency modulation was too slow to represent enough spectral information in the binaural time window. I substantiate my conclusions by a model containing a running cross-correlation and spectral integration.

5.2 Methods

5.2.1 Animal handling

Behavioral experiments were conducted with two adult American barn owls (Tyto furcata pratincola). The birds were born in the breeding colony of the Institute of Biology II at RWTH Aachen University and raised by hand. The animals were treated and cared for in accordance with the guidelines of animal experimentation and with permission of the Landespräsidium für Natur, Umwelt- und Verbraucherschutz Nordrhein-Westfalen, Recklinghausen, Germany. The tests were carried out in an anechoic and sound-proof chamber (A403, Industrial Acoustics Company GmbH).
The owls were sitting on a perch in the center of the chamber during experiments with their feet tied to the perch by a leather tape to prevent the owl from leaving its central position in the chamber. The general behavior of the owl was monitored from outside with an infra-red camera placed inside the chamber. During the experimental period the owls were kept at about 90% of their free feeding weight to ensure motivation.

### 5.2.2 Stimulation

Stimuli were presented via custom-built earphones. In-ear-speakers (Philips SHE2550) were attached to a movable frame so that the earphone speakers could be placed right in front of the ear canals. The earphone device was mounted on a small aluminum bar which was implanted on the owl’s cranial bone under anesthesia prior to the experiments (For detailed information of surgery see Vonderschen and Wagner (2012). The receiver of the head-tracking device (MiniBird, Ascension Technology Corporation, Burlington, Vermont, USA) was also attached to the headphone frame. The head tracker allowed to record head movements in the azimuthal position with 140 Hz sampling rate.

The stimuli were synthesized digitally (C++, RpvsEX, 48828 Hz sampling rate), converted into analog stereo signals (Tucker Davis Technology RX8 Multi I/O Processor), power amplified (Denon AVR-1905), and delivered to the earphones. The stimulus sound pressure level was 30 dB SPL which is considerably above the hearing threshold of the barn owl. The interaural level difference was set to zero dB. Each stimulus was presented with 5 ms cosine rise/fall-ramps.

Noises of different bandwidths were presented to figure out the dependence between signal bandwidth and localization ambiguity. The center frequency was 5000 Hz. This center frequency has a period of 200 μs which led to phantom sound sources which were misaligned from the real source by approx. 80°, because the conversion factor for ITD to azimuth in the barn owl is 0.4 deg/μs (von Campenhausen and Wagner 2006). The bandwidth was varied between 1000 Hz and 5000 Hz in steps of 1000 Hz. In addition, a 5000 Hz pure tone was used as a stimulus. Stimulus duration was 100 ms.

The occurrence of phantom sound sources was also investigated with linearly frequency modulated tones (chirp). In order to compare results from stationary stimuli with frequency-modulated stimuli the chirp stimuli had the same time-averaged spectral content as the stationary noise stimuli. In the following, the frequency traversed at half of the stimulus duration is denoted as center frequency (corresponding to noise center frequency). Both downward sweeping (from high to low frequency) and upward sweeping (from low to high frequency) chirps were tested. Chirps were generated with Matlab-function *chirp* from:

\[ x(t) = \sin\left[2\pi(f_0 + \frac{f}{2}t)t\right] \]  

(1)
The chirp rate is defined as $k = \frac{f_e - f_0}{d_s}$ with $d_s$ being the chirp duration; $f_0$ is the starting frequency at $t = 0$, and $f_e$ is the frequency at the stimulus end. The instantaneous frequency $f(t) = f_0 + tk$ is the frequency at a given point in time during the stimulus presentation. The chirp bandwidth is given by the term $(f_e - f_0)$.

I also varied the duration of the chirp stimuli to investigate the influence of the chirp rate or the duration of the chirp. The center frequency was again 5000 kHz and the bandwidth was either 2 kHz or 4 kHz. The stimulus duration was varied between 12.5 ms and 100 ms (12.5 ms, 25 ms, 37.4 ms, 50 ms and 100 ms). Accordingly, the chirp rate varied between 10 Hz/ms and 320 Hz/ms including every chirp stimulus used in the experiments. The data was plotted as a function of duration and bandwidth, since chirp rate co-varies with both duration and bandwidth.

### 5.2.3 Psychophysical procedure

An operant-conditioning paradigm that utilized the natural head-turning movements of the barn owls was used (Wagner 1991). The training started with broadband stimuli. In this phase the owl first learned to eat small pieces of meat from a feeder. Then the bird learned to orient to a frontal target. When it did this with a short reaction time and a certain precision, it was rewarded with a small piece of meat from the feeder. In the last step the animals learned to first react to the frontal target and then to a second sound originating from different positions. This phase lasted until the owls had learned the task and localized the second target in more than 80% of the trials. During the following experimental sessions broadband noise was pseudo-randomly interleaved into the test stimuli. At least 30% of the overall stimuli were broadband noise to ensure the owl to be motivated over a complete session.

More precisely, during the test phase, the owls had to orient toward a frontal corridor ($0° +/- 7°$ azimuth, $0° +/- 20°$ elevation) to start an experimental trial. An LED at $0°$ azimuth and elevation lit up if the owl hit the corridor. After 1 to 2.5 seconds the LED was switched off and the signal was emitted via earphones. Stimuli were presented with different ITDs ranging from -150 $\mu$s to +150 $\mu$s in steps of 50 $\mu$s, though 0 $\mu$s was omitted. These ITDs approximately correspond to 20, 40, and 60 degrees in azimuth. In the following, a negative sign corresponds to sound played first at the left ear followed by the right ear and hence to sound sources in the left hemisphere. The typical response of the owl was a swift head turn, also called a saccade, toward the sound-source location specified by the ITD. The animal was rewarded with little pieces of chick meat if it hit a corridor of +/- 7° azimuth and +/- 20° elevation around the target and fixated the target location for 250 milliseconds. In experimental sessions the owl was also rewarded if it hit a window misaligned 80° from the true...
sound source which corresponds to the position of phantom sound sources of narrowband signals with 5 kHz center frequency. If the reward-rate dropped below 70% the experimental session was stopped and broadband noise was used again to complete the session as training session. This way, 50 to 250 trials could be carried out per day.

5.2.4 Data analysis
The orientation of the owl’s head was recorded continuously during each session. In each trial, the positional data were stored on a computer with the help of a custom-written C++-Program. The orientation data consisted of the time track, the azimuthal and elevational head positions at a specific point in time, and stimulation time. From these data the head-turning amplitude and the head-turning latency were obtained. The head-turning amplitude was determined by the difference between head orientation at stimulus onset and orientation at the end of the first saccade after stimulus onset. Head turning latency is the delay time between stimulus onset and beginning of the head turning response. The beginning of the head turning response was defined as the first point in time after stimulus onset when the head position differed more than 4° from the azimuthal head orientation at stimulus onset (also compare chapter 4.2). Trials with latencies above 1000 ms were excluded from the analyses, because such high latencies indicate that the owl was inattentive. Wagner (1993) also found out that responses faster than 50 ms were not stimulus driven in the barn owl; therefore trials with latencies below 50 ms were also excluded from the analyses. The remaining trials were valid for analysis.

5.2.5 Cross-correlation model
Behavioral data on temporal integration have been typically modelled with the assumption that a temporal window that may have one or more time constant is acting in a running fashion (Moore et al. 1988; Wagner 1991; Akeroyd and Summerfield 1999). Such a model needs correct responses, chirp rate, and bandwidth as input parameters. Since the percentage of correct responses was by and large independent of chirp rate, these three parameters do, however, not constrain the model enough to derive a time constant. I found a way to circumvent this problem by using the relation between bandwidth, chirp rate and side-peak suppression in broadband neurons. The argument goes as follows: The responses of the broadband neurons in the midbrain nuclei ICX and OT underlie sound-localization behavior (Knudsen et al., 1993; Wagner 1993). The neurons specify azimuthal sound locations by the dependence of their driven rate on ITD (Takahashi and Konishi 1986). The responses vary in a periodic fashion with ITD, and show one main response peak flanked by smaller side peaks (Takahashi and Konishi 1986; Wagner et al. 2007). The reduction of the height of the side peaks
5.2 Methods

compared to the height of the main peak is called side-peak suppression. The height of the side peaks depends on stimulus bandwidth (Mazer 1998; Saberi et al. 1999; Goeckel et al. 2014) and influences phantom-source localization (Saberi et al. 1998a; Saberi et al. 1998b; Saberi et al. 1999). All of this holds for stationary noise, and is independent of the duration of the time window for this stimulus type. The duration of the time window starts to play a role, however, when chirps are used as stimuli, because then the spectro-temporal information in the broadband neurons may be limited by chirp rate. For example, side-peak suppression would decrease, if the chirp rate was so low that less than 3 kHz bandwidth is stored within the time window. This dependence of side peak suppression on chirp rate can be used to derive the duration and shape of the time window. Furthermore, ICX neurons receive narrowband inputs from the central nucleus of the inferior colliculus and integrate this input across frequency (Wagner et al. 1987). The responses of these broadband neurons that exhibit side-peak suppression have been modelled with a running-cross correlation having a presumed, fixed time constant of 10 ms (Saberi et al. 1998b). I extended the model of Saberi et al. (1998b, 1999) by introducing a variation of the time constant and by splitting the neural computation into two steps to derive the duration of two separate time windows. The first step consisted of the temporal processes in the narrowband neurons and was modelled by a running cross-correlation having a time constant that was termed $t_{\text{corr}}$ (Colburn and Durlach, 1978; Carr and Konishi 1990). The second step reflected the side-peak suppression occurring in the responses of ICX neurons after integration across frequency channels (Takahashi and Konishi 1986; Wagner et al. 1987). The extension, compared to the model of Saberi et al. (1998b), is that here these responses were convolved by a second temporal window having a time constant $t_p$. The time constant $t_p$ influenced side-peak suppression (SPS).

Figure 5.1 shows a flow chart of the model used in this study.

Specifically, in the first step a gammatone filter bank (Holdsworth et al. 1988) was fed with acoustic waveforms that were also used as stimuli in the behavioral experiments. The filter bank consisted of 25 filters with linearly spaced center frequencies from 2.5 kHz to 9.5 kHz. This frequency range corresponds to the range of best frequencies in the ICX of barn owls (Vonderschen and Wagner 2009). The filter bandwidths were derived from acoustic nerve recordings ($Q_{40\text{dB}} = 0.036(CF)^{0.504}$; Köppl 1997b). The filter output was half-wave rectified and further processed with a running cross-correlation (Eq. 2).

$$R(T, \tau) = \int_{T}^{P_{fl}} q(f)X_r(f, t)X_l(f, t - \tau) e^{-T_{\text{corr}} \tau} dt df$$  (2)
with \(X_r\) and \(X_l\) being the waveforms of the right and left inputs. The left input is delayed by an ITD \(\tau\). The two integrals represent the running cross-correlation (time constant \(t_{\text{corr}}\)) and the integration across frequency channels. Since a straightness weighting of frequency as proposed by Stern et al (Stern et al. 1988) has its neural implementation in the midbrain of the barn owl (Wagner et al. 1987; Mazer 1998; Saberi et al. 1998b) the frequency channels are weighted by a factor \(q(f)\). The factor \(q(f)\) is derived from behavioral data from the barn owl (compare Saberi et al. 1998b). Additionally, the cross-correlation output across frequency channels was convoluted with a double sided decay function (TW), representing the smoothing time window running over the stimulus (Eq. 3).

\[
Z_\tau(T) = \int_{-\infty}^{\infty} R(T, t)TW(T - t)dt
\]  

(3)

The time window TW had the shape of a rounded exponential with dynamic range limitation (ROEX\((t_p, w)\); for details see Moore et al. 1988). The time constant \(t_p\) in the ROEX\((t_p, w)\) is the second time constant used in my modeling. The parameter \(w\) is a weighting factor and is responsible for the dynamic range limitation. The convolution was computed as a discrete convolution using the MATLAB function \texttt{conv} for any ITD \(\tau\).

Saberi et al. (1998b) demonstrated that side-peak suppression in the optic tectum, a nucleus downstream to the ICX, correlated well with phantom-source localizations. Therefore, I calculated the mean side-peak suppression of the model output over the course of the stimulus. The side-peak suppression (SPS) was defined as:

\[
SPS = \frac{MP - SP}{MP}
\]  

(4)

where MP is the model response at the main peak and SP is the response at the side peak closest to the main peak. The modeled side-peak suppression was then related to the percentage of correct responses with both noise and chirp stimulation in 2 steps. In step 1, I calculated the linear relation between side-peak suppression and the number of correct responses up to 3 kHz bandwidth measured in my behavioral study when noise was used as a stimulus. Using this correlation and assuming that side-peak suppression and the correct number of responses were similarly related when chirps were used as stimuli as with noise stimulation, I predicted the side-peak suppression from the behavioral responses with chirps in step 2. In detail, the mean squared error (MSE) of slopes with different stimulus duration and with specific time window parameters was determined (Eq. 5).

\[
MSE_{t_p, t_{\text{corr}}, w} = \frac{1}{n} \sum_{i=1}^{n} (C_i - N)^2
\]  

(5)
where \( d \) is an integer between 1 and 5 representing one of the five durations used in the behavioral experiments with chirps; \( n \) is the number of stimulus durations used for this analysis. \( C \) is the slope of the linear correlation between SPS and CR (percentage of correct responses) with chirps and \( N \) the slope with noise stimuli. Note that the model does not include any parameters which contribute to a strong phasic onset response which can be found in ICX neurons. Since side-peak suppression is nearly absent in the onset response of ICX neurons (Wagner 1990) to broadband stimuli I assume that it has no influence on phantom-source localizations and was omitted in the modelling of side-peak suppression.

**Figure 5.1: Flow chart of the cross-correlation model.**
Frequency components of left and right input signal are extracted with a gammatone filter bank. The narrowband filter bank output is rectified and further processed with a running cross-correlation. Each frequency channel is weighted with the straightness weighting factor \( q(f) \). The frequency channels were then integrated. The cross-correlation output is convoluted with a temporal window.

### 5.3 Results

To study spectro-temporal integration in sound localization of barn owls I used 2 owls and collected 30,019 valid trials (see chapter 5.2.4). Of these, 19,507 trials were conducted with noise or tonal stimuli (owl Bu: 12,545; owl Pa: 6,962), while 10,512 trials with chirp stimuli (owl Bu: 6,671; owl Pa: 3,041). Each bandwidth/duration pair consisted of at least 20 trials. In the following I first introduce the characteristics of the head-turns of barn owls, before I turn to the influence of stimulus bandwidth and stimulus duration on localization performance. Finally, I present a model that allows the derivation of two binaural time windows that are effective during localization.
5 Spectro-temporal integration in sound localization as revealed by frequency-modulated tones

5.3.1 Characteristics of head turns

After hearing a faint sound barn owls typically perform a saccadic head turn toward the sound source. The owls also turn their head if a stimulus is presented via earphones and only contains an ITD as directional cue. The amplitude of the head turn elicited by earphone stimulation corresponds to the amplitude of a head turn towards a free-field sound source having the same ITD as the stimulus presented via earphones (Wagner 1991; Poganiatz et al. 2001). I utilized these head turns to quantify the localization of phantom sound sources by measuring the direction, amplitude, and response latency of these turns. Figure 5.2 shows the track of such a head turn of owl Bu in response to a broadband stimulus presented with an ITD of -150 µs via earphones. An ITD of -150 µs corresponds to -60° in azimuth, i.e. the stimulus sound came from the left hemisphere. The owl was trained to fixate a zero position at 0° azimuth and 0° elevation prior to stimulus presentation. This section of the trial is not shown in Figure 5.2a. The stimulus elicited a head saccade. The owl fixated a position of -52° azimuth (red section in Figure 5.2a) after completing the head turn. The stimulus duration was 100 ms. The owl reacted with a latency of 129 ms. Since the response latency was longer than the stimulus duration this turn is called an open-loop head turn. 90.3% of the head turns after stimulation with 100 ms broadband noise stimuli in owl Bu and 94.4% in owl Pa were open-loop turns. Closed-loop head turns (latency < stimulus duration) were only observed with 100 ms stimulus duration. Broadband noise elicited the largest percentage of closed-loop turns. The percentage of open-loop turns after stimulation with either chirps, tones or narrowband noise was higher than when broadband noise was used for stimulation as may be seen from the latency distributions (Figure 5.2b owl Bu; Figure 5.2c owl Pa). Broadband noise (black bars) elicited significantly faster responses than chirps (magenta, merged over each chirp bandwidth with 100 ms duration) in both owls (Mann-Whitney u-test, p<0.001).
While the owls turned toward the sound source as specified by the stimulus ITD in almost 100% of the trials when broadband noise was used for stimulation, an increase in phantom-source localizations was observed with other stimulus types. The position of phantom sound sources of stationary narrowband noise could be estimated from the center frequency of the stimulus. I assume that the phantom source of chirps also correlates with the center frequency of the chirp stimulus. The center frequency and the stimulus ITDs were chosen such that the estimated phantom sound source was always located in the opposite hemisphere of the real sound source. In this way, I could easily distinguish between real source localizations, which are referred to as correct responses (CR) in the following, and phantom-source localization by analyzing the direction of head turns.

**Figure 5.2: Head turning properties and response latency.**

- **a** Trace of a single head turn of owl Bu in response to a broadband noise stimulus with an ITD of -150 µs (sound source from the left). The owl responded with a saccade toward the sound source after a latency of 129 ms after stimulus onset. The saccade ended with a fixation of the sound source (red highlight). The fixation angle was 52°.
- **b, c** histograms showing the distribution of head turning latency with both owls (b: owl Bu, c: owl Pa). The black histograms show the latencies in response to broadband noise with 100 ms duration. The magenta histograms show the latencies in response to upward sweeping chirps with 100 ms duration. Latencies to chirps of any bandwidth are merged.

While the owls turned toward the sound source as specified by the stimulus ITD in almost 100% of the trials when broadband noise was used for stimulation, an increase in phantom-source localizations was observed with other stimulus types. The position of phantom sound sources of stationary narrowband noise could be estimated from the center frequency of the stimulus. I assume that the phantom source of chirps also correlates with the center frequency of the chirp stimulus. The center frequency and the stimulus ITDs were chosen such that the estimated phantom sound source was always located in the opposite hemisphere of the real sound source. In this way, I could easily distinguish between real source localizations, which are referred to as correct responses (CR) in the following, and phantom-source localization by analyzing the direction of head turns. **Figure 5.3** shows collections of head turns obtained with different stimulus types. Stationary noise with a bandwidth of 4 kHz elicited head turns toward the real source in all but one case (**Figure 5.3a**, n = 143) which can be clearly seen by the clustering of fixations (red sections of the tracks) at positive azimuthal head-orientation angles. On the other hand pure tone stimulation led to localization of phantom sound sources (**Figure 5.3b**, n = 90). The fixations are not only clustered near the real sound source but also close to the estimated phantom sound source at -40°. The third example (**Figure 5.3c**) shows head orientation tracks with an upward sweeping chirp stimulus with a bandwidth of 4 kHz. The chirp stimulus (n = 132) also elicited head turns toward phantom sound source, but to a lesser extent as the pure tone. I used the distribution of fixation angles to quantify the percentage of correct response (CR), i.e. responses toward the real sound source. **Figure 5.3d-e**
shows histograms of the distributions of fixation angles obtained from the tracking data shown in the upper row in Figure 5.3 (ITD = 100 µs, black solid lines). The histograms with grey dashed lines show the distributions of fixation angles with the stimulus in the opposite hemisphere (ITD = -100 µs). Fixation angles with noise stimulation with 4 kHz bandwidth are single peaked, and distributed around +40° for an ITD of 100 µs and around -40° for an ITD of -100 µs, respectively (Figure 5.3d). Tonal stimuli elicited responses toward the real as well as toward the phantom sound source (Figure 5.3e). These distributions were clearly bimodal with a notch at 0°. The owl turned its head more frequently toward the left sound source - either phantom or real source - suggesting a bias in head-turning behavior. The double peaked distribution is also seen with chirp stimulation (Figure 5.3f) and narrowband noise (not shown). The notch occurred always at 0°. This allowed me to discriminate correct responses and phantom-source localization by the turning direction to either the hemisphere of the real sound source or the phantom sound source. In the pure tone example shown in Figure 5.3e with an ITD of 100 µs CR was 32% and 67% with -100 µs. The average CR is 50%. Since 50% phantom localizations are expected with pure tone stimulation, the averaging of data from left-side stimulation and right-side stimulation averaged out the bias in head-turning behavior.
5.3 Results

Figure 5.3: Traces of head turns and distributions of fixation angles of owl Bu.
The upper panels (a-c) show the entire traces of valid head turns of owl Bu to narrowband noise (a, bandwidth 4000 Hz, n = 143), pure tone (b, 5000 Hz, n = 90), and upward sweeping chirp (c, bandwidth 4000 Hz, n = 132). The stimulus duration was 100 ms in these examples. The ITD was 100 μs. The resulting real and phantom sound sources are shown with dashed lines. The fixations are shown as red highlights. Fixation angles are the average head orientation during the fixations. The lower panels (d-f) show the distributions of fixation angles shown in a-c (ITD = 100 μs) with solid lines and the distributions of fixation angles with -100 μs with gray dashed lines. Positive ITDs denote sound sources from the right hemisphere while negative ITDs denote sources from the left hemisphere. Fixations in the hemisphere opposite to the real sound source are phantom-source localizations, while fixations in the hemisphere of the real sound source are correct responses.
5.3.2 Influence of bandwidth on phantom-source localizations
As was shown in the last section, narrowband sounds produce an ambiguity concerning the real position of the sound source, while this is not the case when broadband noise is used as stimulus (Figure 5.3). To quantitatively investigate the influence of stimulus bandwidth on the percentage of correct responses and phantom-source localizations, I varied the bandwidth of both noise and chirp stimuli. Chirps were specifically utilized to answer the question whether the frequency content of the stimulus needs to be available instantaneously or is integrated over time. Figure 5.4 shows the dependence of CR on the stimulus bandwidth in both noise and chirp stimuli. The two owls tested showed a similar performance. For example, with tonal stimulation, both owls exhibited about 50% phantom localizations and about 50% real-source localizations (“tone” in Figure 5.4). Moreover, the percentage of correct responses increased linearly for both stimulus types when the bandwidth was increased. With noise stimuli saturation started for bandwidths above 3 kHz. Owl Pa reached 95% correct responses already for a bandwidth of 3 kHz, while owl Bu reached this criterion for bandwidths broader than 3 kHz. The low error rate observed in both owls with broadband-noise stimulation suggested that the owls could localize signals with a center frequency of 5 kHz and a bandwidth broader than 3 kHz unambiguously.

The increase in the number of correct responses with bandwidth was shallower when chirps were used as stimuli than when noise was the stimulus (Figure 5.4c,d). While the number of correct responses in stimulation with upward sweeping chirps increased linearly from 1 kHz to 5 kHz bandwidth, the 95% criterion was never reached. The highest percentages of correct responses occurred for a bandwidth of 5 kHz in both owls and were 83% (owl Bu) and 92% (owl Pa), respectively. The increase of the number of correct responses, and thus the elimination of ambiguity, happened for both upward and downward sweeping chirps, and was not statistically different for these the two sweep directions in both owls (all p-values in the 10 pairwise $\chi^2$-tests were larger than 0.05). These results suggested that spectral integration also takes place when chirps are used as stimuli, but the temporal extent of this integration remained unclear. This aspect was, therefore, examined next.
5.3 Results

Figure 5.4: Percentage of correct responses as a function of stimulus bandwidth.
CROSSES denote the percentage of correct responses (CR) with one of the six stimulus ITDs used in this study (-150 μs, -100 μs, -50 μs, 50 μs, 100 μs, 150 μs). The red crosses show the examples from Figure 5.3 (+/-100 μs). The solid circles show the average across stimulus ITDs. The left panels (a, b) show data from stimulation with noise with 100 ms duration. Data from owl Bu is presented in a. Data from owl Pa is presented in b. The right panels (c, d) show CR with upward sweeping chirp stimuli (owl Bu in the top panel (c) and of owl Pa in the lower panel (d)).

5.3.3 Influence of stimulus duration on phantom-source localizations
If the owl can integrate spectral information over time, as suggested by the results presented in the last paragraph, the instantaneous change of frequency in the chirp stimuli is converted into the contents of a spectro-temporal filter that underlies the computation of sound-source location. Three parameters will influence the content of this filter: bandwidth, duration and chirp rate of the stimulus. Since chirp rate covaries with bandwidth and duration, there are only two free parameters. In other words, the duration and shape of the time window may be derived by varying two stimulus parameters. I varied duration in addition to bandwidth, which resulted in testing of chirp rates ranging between 10 Hz/ms and 320 Hz/s (Figure 5.5). The mainly horizontally oriented color gradients in the surface plots of the number of correct response in the duration-bandwidth plane demonstrate a strong dependence of the percentage of correct responses on bandwidth with both noise and chirp...
stimulation (Figure 5.5). I analyzed these data quantitatively with contingency tables comparing each data point with all other data points in a pairwise fashion. The analysis revealed that the number of correct responses was significantly reduced from 100 ms duration to 12.5 ms duration only with bandwidth of 4 or 5 kHz in both owls, but not for the other combinations ($\chi^2$-test, $p < 0.05$). Thus, stimulus duration had some influence, but stimulus bandwidth was more important for the localization performance than the stimulus duration. The integration time window was found in a further step with a model that related these data to processing in the brain.

![Figure 5.5: Percentage of correct responses as a function of bandwidth and duration.](image)

The percentage of correct responses (CR) is illustrated as surface plots against duration and bandwidth. The surface is derived from a natural neighbor interpolation (using the Matlab function Triscatterinterp) of the actual data points whose bandwidths and durations used in this study are shown by black dots. The left panels (a,b) show data with noise stimulation. The right panels (c,d) show data with stimulation with upward sweeping chirps. The upper row (a,c) shows data from owl Bu. The lower row (b,d) shows data from owl Pa. The black dots show the actual bandwidths and durations used in this study.
5.3 Results

5.3.4 Estimation of the duration of the temporal window of frequency integration

The performance of the model was tested by fitting the side-peak suppression as a function of bandwidth produced by the model to side-peak suppression measured in neurons in the optic tectum of barn owls recorded by Saberi et al. (1999) (Figure 5.6a). The model fitted the physiological data well. The specific parameter values \( (t_p, w, t_{	ext{scorr}}) \) did not constrain the model very much, because all correlation coefficients between model and physiological data were larger than 0.83. This confirmed the implicit assumption mentioned already in Methods (chapter 5.2.5) that side-peak suppression is independent of the duration of the time window for stationary broadband noise. Saberi et al. (1999) used band-pass filtered noise with a normal distributed frequency domain and a bandwidth of twice the standard deviation, while rectangular filtered noise was used in the present study (Figure 5.6b, black line). Additionally, I modeled side-peak suppression with upward sweeping chirps (Figure 5.6b, green line). The percentage of correct responses in the behavioral experiments increased linearly with modeled side-peak suppression up to 3 kHz bandwidth. By this means I could determine the side-peak suppression resulting in a specific behavioral response. Figure 5.6c shows the correlation between side-peak suppression and behavioral response with noise and chirps in owl Bu. The slope of the linear fit of this correlation with noise stimuli did not depend on the model time constants and window shapes (Bu: \( N = 1.2 \) Pa: \( N = 1.3 \), see Methods - chapter 5.2.5). The deviation of the slope of the linear fit of the correlation with chirps from the slope with noise was used to determine the best parameters of the time windows of cross-correlation and frequency integration. The time window parameters yielding the lowest MSE (see Methods - chapter 5.2.5) and thus estimating the time windows shapes best are shown in Table 5.1.
Spectro-temporal integration in sound localization as revealed by frequency-modulated tones

In a first attempt side-peak suppression was modeled without a time window of frequency integration. However, MSE was drastically increased if no time window of frequency modulation was included compared to a model that includes such a time window. The parameters yielding the best model predictions are shown in Figure 5.7. Increasing \( t_{\text{corr}} \) above about 100 ms did not further improve model predictions of the behavioral performance with a single time window (Figure 5.7a). The introduction of a time window of frequency integration (Figure 5.7b-d) improved the model prediction by a factor of 13 in owl Bu and by 200 in owl Pa. The values of \( t_{\text{corr}} \) yielding the lowest MSE in a model with two time windows were much smaller than in a model with a single window (Figure 5.7b; owl Bu: 2 ms, owl Pa: 9 ms). Most important, the time constant \( t_p \) of the time window of frequency integration was in the same range as \( t_{\text{corr}} \) in a double time window model (Figure 5.7c; owl Bu: 6 ms, owl Pa: 17 ms) and not in the range of binaural sluggishness (approx. 55 ms, Kollmeier and Gilkey 1990). The dynamic range limitation \( w \) was < 0.05 in both owls. Horizontal bars in Figure 5.7c-d illustrate the range of the respective parameter yielding less than 110% of the minimum MSE. This range is a measure of the influence of changes of the model parameters on MSE. Variation of the model parameters within a narrow range did scarcely affect the model predictions (MSE). However, note that increasing the time constant \( t_p \) up to values connected with binaural sluggishness MSE was drastically increased (Figure 5.7c). Figure 5.7e illustrates the shapes of the time windows of frequency integration of both owls derived from the parameters \( t_p \) and \( w \) yielding the lowest MSE.

Figure 5.6: Response of a cross-correlation model with a running temporal window of integration of frequency channels.

The black line in a shows the average modeled side-peak suppression of 30 stationary band-pass noise tokens as a function of stimulus bandwidth (normal distributed frequency domain; bandwidth = 2*standard deviation, 200 ms duration). Grey lines in a show electrophysiological data from the 11 optic tectum neurons measured by Saberi et al. (1999). b illustrates the modeled side-peak suppression with bandpass filtered noise used in the behavioral experiments of the present study (black line; rectangularly filtered band-limited noise, 100 ms duration) and upward sweeping chirps (green lines; 100 ms duration). The solid green line shows data with both a time window of frequency integration \( (t_p = 6 \text{ ms}, w = 0.04) \) and cross-correlation \( (t_{\text{corr}} = 10 \text{ ms}) \). c shows the linear correlation between side-peak suppression and behavioral response of owl Bu (percentage of correct responses, CR) with noise up to 3 kHz bandwidth (black line) and chirps (green lines) with the same time constants as in b.

In a first attempt side-peak suppression was modeled without a time window of frequency integration. However, MSE was drastically increased if no time window of frequency modulation was included compared to a model that includes such a time window. The parameters yielding the best model predictions are shown in Figure 5.7. Increasing \( t_{\text{corr}} \) above about 100 ms did not further improve model predictions of the behavioral performance with a single time window (Figure 5.7a). The introduction of a time window of frequency integration (Figure 5.7b-d) improved the model prediction by a factor of 13 in owl Bu and by 200 in owl Pa. The values of \( t_{\text{corr}} \) yielding the lowest MSE in a model with two time windows were much smaller than in a model with a single window (Figure 5.7b; owl Bu: 2 ms, owl Pa: 9 ms). Most important, the time constant \( t_p \) of the time window of frequency integration was in the same range as \( t_{\text{corr}} \) in a double time window model (Figure 5.7c; owl Bu: 6 ms, owl Pa: 17 ms) and not in the range of binaural sluggishness (approx. 55 ms, Kollmeier and Gilkey 1990). The dynamic range limitation \( w \) was < 0.05 in both owls. Horizontal bars in Figure 5.7c-d illustrate the range of the respective parameter yielding less than 110% of the minimum MSE. This range is a measure of the influence of changes of the model parameters on MSE. Variation of the model parameters within a narrow range did scarcely affect the model predictions (MSE). However, note that increasing the time constant \( t_p \) up to values connected with binaural sluggishness MSE was drastically increased (Figure 5.7c). Figure 5.7e illustrates the shapes of the time windows of frequency integration of both owls derived from the parameters \( t_p \) and \( w \) yielding the lowest MSE.
5.3 Results

Figure 5.7: Model performance and time constants.
The top four plots (a-d) show the mean squared error as a measure of quality of model prediction. Model data is shown for both owls (red: Pa; blue: Bu) a MSE as a function of $t_{xcorr}$ with a model without a temporal window of frequency integration. b-d MSE as a function of model parameters in a model including a time window of frequency modulation. The black segments show the range of parameter values that produced less than 110% of the minimum MSE. e Shape of the temporal window calculated from $t_p$ and $w$ yielding the lowest MSE. The grey area shows the range of window shapes derived from $t_p$ and $w$ yielding less than 110% of the lowest MSE.
Table 5.1: Model time constants.
The table shows time constants resulting from different model conditions. The conditions comprise time constant with simulations with only one duration tested (duration only), simulations with any duration but 12.5 ms (w/o 12.5 ms) or 100 ms (w/o 100 ms), and a simulation including every stimulus duration that was also used in the behavioral experiments (every duration).

<table>
<thead>
<tr>
<th>condition</th>
<th>Pa</th>
<th></th>
<th>Bu</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>t_p (ms)</td>
<td>w</td>
<td>t_corr (ms)</td>
<td>t_p (ms)</td>
</tr>
<tr>
<td>12.5 ms only</td>
<td>22</td>
<td>0.11</td>
<td>7</td>
<td>2</td>
</tr>
<tr>
<td>25 ms only</td>
<td>11</td>
<td>0.06</td>
<td>23</td>
<td>1</td>
</tr>
<tr>
<td>37.4 ms only</td>
<td>18</td>
<td>0</td>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td>50 ms only</td>
<td>17</td>
<td>0.13</td>
<td>1</td>
<td>5</td>
</tr>
<tr>
<td>100 ms only</td>
<td>18</td>
<td>0.03</td>
<td>4</td>
<td>4</td>
</tr>
<tr>
<td>w/o 12.5 ms</td>
<td>11</td>
<td>0.01</td>
<td>23</td>
<td>6</td>
</tr>
<tr>
<td>w/o 100 ms</td>
<td>17</td>
<td>0.01</td>
<td>9</td>
<td>8</td>
</tr>
<tr>
<td>Every duration</td>
<td>17</td>
<td>0.02</td>
<td>9</td>
<td>6</td>
</tr>
</tbody>
</table>

I was concerned that the time constants I derived in the model might be related to the lowest chirp rate used in the stimuli, and that even lower chirp rates may have resulted in longer time constants. To test for this possibility, I analyzed the data without the data sets collected with 100 ms stimulus duration that contained the data with the lowest chirp rate. This analysis also suggested two time constants. The duration of these time constants was similar to the durations of the time constants determined with the complete data set (Table 5.1). I also tested the time constants with any individual stimulus duration. The results are also shown in Table 5.1. The time constants did not depend on the stimulus duration.

5.4 Discussion

In this study behavioral responses to frequency-modulated and stationary sounds were compared to examine temporal integration in the sound-localization system of barn owls. I found that mainly the bandwidth and to a much lesser degree the duration of a frequency-modulated sound influenced the localization of phantom targets. The data could well be fit with a running cross-correlation model including a time window of across-frequency integration. In the following I discuss my findings with respect to the occurrence of phantom sources in barn-owl sound localization and the influence of frequency modulation on the behavioral performance, and binaural sluggishness.
5.4 Discussion

5.4.1 Use of phantom-source localization as a tool to study spectral and temporal integration

Phantom-source localization has been reported in humans and in barn owls for stationary stimuli (Trahiotis and Stern 1989; Buell et al. 1994; Saberi et al. 1998a). Phantom-source localization occurred in barn owls, if the stimulus had a bandwidth of less than 3 kHz (Saberi et al. 1999). My data are consistent with these data. These authors also reasoned that during phantom-source location the owl perceived two images one of which is selected. The location of the images depends on the stimulus ITD and the period of the center frequency of the signal. Following the interpretation of Saberi et al. (1999), the decrease of phantom-source localizations with increasing stimulus bandwidth might be interpreted as the image of the phantom source getting weaker compared to the image of the real source. As a consequence the owl turns its head more often toward the real source. The weakening of the image of the phantom source with an increase in stimulus bandwidth correlates well with an increase in side-peak suppression with tuning bandwidth in neurons of the ICCls, the ICX and the optic tectum in the same stimulus situation (Mazer 1998; Saberi et al. 1999; Göckel et al. 2014). The present data showed that frequency-modulated tones increased the probability of responding to the real sound source compared to pure tone stimulation. I interpret this observation as demonstration of the existence of a time window for frequency integration. The logic behind this conclusion is the following: In the chirp stimuli used here and dependent on the chirp rate, more or less spectral information is available within a given time window. If the spectral content is too small within such a time window, phantom localizations should occur, while, if the spectral content was sufficiently broad within the time window, turns should always be in the direction of the real source.

5.4.2 Modelling time windows

To estimate the size of a time window I used a model with two time constants. Moore et al. (1988) and Wagner (1991) found that one time window with one time constant was much less effective than one time window with two time constants. This is in contrast to Akeroyd and Summerfield (1999) who used a model with a single time constant to explain the detectability of varying interaural correlation. I found here that a time window with a single time constant was not sufficient to fit both the effect of bandwidth and of duration with chirp stimuli.

In my modeling attempts I went a step further and did not only try to provide a phenomenological model of the behavioral data, but also to relate my observations to neural responses in broadband neurons in ICX and the optic tectum of the barn owl. It was known that the neurons in these nuclei underlie sound-localization behavior (Wagner 1993; Knudsen et al. 1993). The most important
property of these neurons in the context of phantom localization is their ability to suppress side peaks in dependence of stimulus bandwidth (Takahashi and Konishi 1986; Mazer 1998; Saberi et al. 1999). The computation of ITD consists of several steps (for a review see Vonderschen and Wagner 2014), in which the first major step is the detection of ITD in narrowband channels (Carr and Konishi 1990). These narrowband neurons are not able to suppress side peaks (Wagner et al. 2002). Side-peak suppression results when many of the narrowband neurons that are tuned to different frequencies converge on one neuron in ICX, thus also generating broad frequency tuning (Wagner et al. 1987, 2007). The two time constants in my model reflected the two major processing steps. The first time constant, implemented by a running cross-correlation, was assumed to be effective in the narrowband inputs to the ICX neurons. Running cross-correlation is a well-accepted model for binaural interaction (Sayers and Cherry 1957; Colburn and Durlach 1978). Responses of midbrain neurons to stationary stimuli were well modeled with a running cross-correlation having a time constant of 5 ms (Fischer et al. 2009). Given the important frequency range (2.5-9.5 kHz, see also Vonderschen and Wagner 2009), the spectral information in a 5 ms time window is more or less constant during the whole stimulus duration (100 ms), and sufficient to generate side-peak suppression (Wagner 1992) if stationary broadband is used for stimulation. This effect is captured by the first time constant of my model. The duration of the first time constant was, however, too short to generate enough side-peak suppression in FM stimuli. The second, longer time constant guaranteed that enough spectral information was available within the time window to allow equivalent side-peak suppression with FM stimuli as observed for stationary broadband noise so that a similar number of real-source localizations resulted with both stimulus types. Therefore, I suggest here that the second time constant becomes effective at the output of the broadband neurons, while the first time constant is effective at the input to these neurons. I like to add a note of caution here: while it is clear that two time constants were effective in my experiments, the duration of the second time constant remained somewhat unclear. It was clearly longer than the first time constant, but the absolute value varied between owls (as did the first time constant). I was not so much worried about this variation, but more about the possibility that slower chirp rates than used here may indicate longer time windows. Therefore, I reanalyzed the data excluding the lowest chirp rates. This did not substantially change the time constants.

5.4.3 Binaural sluggishness

Time windows acting as a smoothing window running over the stimulus were often used to describe behavioral responses to stimuli with time varying components (Moore et al. 1988; Kollmeier and
5.4 Discussion

Gilkey 1990; Wagner 1991; Green 1995; Bernstein et al. 2001). Time windows in the range of 12 ms (Akeroyd and Bernstein 2001) to 170 ms (Granatham and Wightman 1979) have been reported. The time window I derived here was in the range of the shorter time constants reported earlier. It seems that the experimental design to some extent influences temporal resolution (Shackleton and Palmer 2010). Siveke et al. (2008) suspected that it was the fine judgments that make the binaural system sluggish, and that time-varying modulations may help with fast binaural detection. My stimulus contained time-varying modulations, and reflected, therefore, in one sense the Phasewarp stimulus used by Siveke et al. (2008). On the other hand, the owls had to precisely localize the perceived source. The time windows I determined were closer to those observed by Siveke et al. (2008) than to those observed by Grantham and Wightman (1978, 1979) and by Kollmeier and Gilkey (1990). This suggests that it is not the fine judgments alone that make the binaural system sluggish. The data by Pollack (1978) and Wagner (1991) also argue against an effect of continuous variations, because in these studies the binaural images changed abruptly. Further studies are necessary to find out what makes the binaural system so sluggish in some stimulus designs.
6 Eardrum directionality induced by interaural transmission

6.1 Introduction

Bird skulls have several air cavities. These cavities probably evolved to decrease the weight of the skull. For hearing they create the possibility to transmit sound internally from the left to the right ears and vice versa, creating internal ear coupling. The continuous anatomical pathway from one middle ear to the other is called interaural canal. In other words, not only the sound from the ear that directly receives the sound, the ipsilateral ear, arrives at the eardrum, but also the sound that first reaches the contralateral ear. The difference of the sound pressure and phase from the left and right ears contains frequency-dependent directional information. The receiver is called a pressure-difference or pressure-gradient receiver to discriminate it from a receiver in which sound only from one side impinges on the eardrum. Such a receiver is called a pressure receiver. Pressure-gradient receivers are most effective in low frequency ranges where the sound is not diffracted and attenuated by the head. Pressure-gradient receivers alter the amplitude of the sound acting on the eardrum depending on the phase difference at the eardrum so that maximally a doubling in amplitude occurs. Figure 6.1 explains the principle of pressure-gradient receivers in comparison to low-frequency pressure receivers. Indeed, small birds, reptiles, anurans, and insects possess pressure-gradient receivers (review: Coles et al. 1980; Rosowski and Saunders 1980; Calford and Piddington 1988; Klump 2000; Larsen 2004; Larsen et al. 2006; Michelsen and Larsen 2008).

Payne (1971) had already mentioned that barn owls have very large cranial air chambers, but had related this observation to the pneumaticity of the skull, and was perplexed about its possible function in sound transmission. Although the relation between interaural transmission and cavity size is complex (Christensen-Dalsgaard and Manley 2005; Christensen-Dalsgaard and Manley 2008; Vossen et al. 2010), it is tempting to assume that at least at low frequencies enough sound might be transmitted through the interaural canal of the barn owl so that the ears may function as pressure-gradient receivers. However, Moiseff and Konishi (1981b) claimed that the interaural canal is not involved in sound localization in the barn owl. These authors had investigated interaural transmission in a frequency range from 2-9 kHz in this bird using both probe-microphone measurements in the middle ear and thresholds of neurons in the cochlear nucleus. These authors found that attenuation was about 10 dB at 2 kHz, between 10 and 20 dB from 2-4 kHz, and above 20 dB for frequencies above 4 kHz. By contrast, Konishi (1973) had found that sound-localization precision was best between 5 and 9 kHz and was degraded by a factor of 3 at 3 kHz. Knudsen and Konishi (1979) had seen an even bigger loss of precision for frequencies down to 1 kHz, if both azimuthal and elevational
localization were considered. Such frequency dependence was also reflected in localization accuracy as well as in pre-take-off latency in free-flight experiments (Singheiser et al. 2010). Since attenuation was high in the frequency range where localization was most precise, Moiseff and Konishi (1981b) came to the conclusion mentioned above. By contrast, Coles and Guppy (1988) concluded from their cochlear-microphonic measurements that a pressure gradient may act at the tympanic membrane. Later measurements of head-related transfer functions (HRTF, Hausmann et al. 2010) did not reveal signs of a pressure gradient. Specifically, the range of interaural time differences (ITD) did not increase significantly in the tested low-frequency range (500 Hz to 1 kHz) compared to the high frequency range (5-9 kHz) as would be predicted, if the ear would function as a pressure-gradient receiver (Vossen et al. 2010). However, the data of Hausmann et al. (2010) have to be viewed with care, because the eardrum is an acoustic radiator. One would not expect to see the effects of the eardrum movements in HRTF measurements that are taken several millimeters away from the eardrum. However, in a close relative of the barn owl, the grass owl (Tyto longimembris), the ITD range is increased for frequencies below 1 kHz, most likely due to interaural coupling (Calford and Piddington 1988). Thus, up to now, the role of interaural transmission in barn-owl hearing has remained unclear. Therefore, I collected new data, concentrating mainly on low frequencies (< 3 kHz) with laser vibrometry that allowed to directly observe eardrum motion. I report here that the barn-owl ear is directional for low frequencies, but not in the frequency range that is important for sound localization.
Figure 6.1: Mechanisms of low frequency sound receivers.
At low frequencies the sound is not attenuated or diffracted by the head. Only the intrinsic properties of the eardrum and the presence or absence of an interaural canal influence the velocity of eardrum vibration at low frequencies. The upper row sketches the basic anatomy of pressure and pressure-gradient receivers, respectively. Sound incoming from the external side is indicated by a red sine. Sound traversing the interaural canal and thus arriving at the internal side of the eardrum is indicated by a green sine. In a pressure receiver the sound is only arriving at the external side of the eardrum, while in a pressure-gradient receiver both sides of the eardrum are affected by incoming sound. However, the sound may be attenuated while traveling through the interaural canal and thus the sound pressure on both sides of the eardrum may not be equal. Both internal and external sounds are also shifted in phase due different traveling times. As a consequence of phase shift and attenuation, the eardrum-vibration velocity depends on the angle of the incoming sound as shown by the polar plots of eardrum velocity in the middle row. The diagram shows the directionality for 6 dB attenuation through the interaural canal and an arbitrary phase shift. Eardrum velocity does not express such a directionality pattern in pressure receivers. The lower row shows the same directionality in dB as the middle row but normalized to the minimal directionality.
6.2 Methods

Data from five captive-bred adult barn owls (*Tyto furcata pratincola*) were obtained for this study. The owls were belonging to the colony at RWTH Aachen University in Aachen, Germany. The owls were equipped with a metal plate that was fixed to the skull under anesthesia at an earlier time. This treatment was approved by the *Landespräsidium für Natur, Umwelt und Verbraucherschutz Nordrhein-Westfalen* (Recklinghausen, Germany). The metal plate allowed to fix the skull to a head holder and thus to keep the head in a fixed position in the setup during experiments. CITES permits were issued by the *Umweltamt der Städteregion Aachen*, to be able to conduct the experiments reported here at the University of Southern Denmark in Odense. The owls were transported to Denmark, and returned to Aachen after the experiments. Experiments in Odense were carried out with permission from the Danish Animal Experimentation Inspectorate (2012-DY-2934-00021/BES).

For recordings in the experiments reported here and earlier, the owls were initially tranquilized with Diazepam (Valium, 1 mg/kg/h) and subsequently kept under anesthesia (Ketamine hydrochloride 4–20 mg/kg/h) for several hours. Atropine sulfate (0.05 mg/kg) was injected to prevent salivation. Prior to experiments glass beads (50 µm in diameter, alumino-silicate microspheres, Omya AG, Oftringen, Switzerland) were carefully placed on the eardrum at the tip of the extra-columella (*umbro*) to enhance reflection of laser light.

6.2.1 Directionality measurements

The anaesthetized bird was fixed in a brass-rod holder in a natural perched position and transported to the anechoic room of the Department of Biology, SDU. Here it was placed in the center of a circle defined by 12 JBL 1-G loudspeakers located at 30° intervals in a horizontal plane at the height of the owl’s head at a distance of 1 m from the center in an approximated free sound field ([Figure 6.2](#)). In this way, stimulation occurred within the equatorial azimuthal plane. The beam of a Polytec (Waldbonn, Germany) laser Doppler vibrometer OV-505 located at a distance of 0.9 m from the center was focused on the glass beads placed on the eardrum to measure vibration velocity. The sensitivity of the laser sensor was 1-2 mm/s/V. To ensure a constant and identical stimulus amplitude from the 12 loudspeakers in the frequency interval of interest, a Brue & Kjaer ½-inch pressure microphone was first placed at the center of the setup at the position normally occupied by the owl’s head and successively recorded the transfer function (*frequency response*) of each speaker as its response to a sweep signal with a constant amplitude. Sweep stimuli (40 ms duration, 400–14,000 Hz, 16-64 sweeps, levels of 65–70 dB SPL) presented to the owl were filtered with frequency responses of the individual loudspeakers (by dividing the spectrum of the sweep by the transfer function of the
Eardrum directionality induced by interaural transmission

Speaker). The duration of the sweeps was kept short to avoid large eardrum deflections caused by a vigorous stapedius reflex, which was elicited by stimuli at these levels (compare stapedius reflex in chicken: Counter and Borg 1979). To monitor the stimulus sound pressure the 1-mm diameter tip of a B&K 4182 probe microphone was positioned at the lower rim of the meatus of the ear. Both sound-stimulation and microphone and laser vibrometer recordings were controlled by Tucker-Davis Technology System 2. The laser and microphone outputs were digitized using a TDT AD-converter (AD2, 44.1-kHz sample rate, 8,192 samples) and processed using customized software (DragonQuest). After calibrating the setup and positioning the anesthetized owl, directionality was measured of first one eardrum as 12 transfer functions in response to successive sound sweep stimulation from each of the 12 speakers. To test the influence of interaural sound transmission on eardrum vibrations the contralateral ear was reversibly blocked (Gold Velvet II; All American Mold Laboratories, Oklahoma City, OK).

6.2.2 Transmission measurements

I used the method developed and first described by Michelsen et al. (1994) to measure interaural sound transmission. Here the vibration velocity of the eardrum is monitored with laser Doppler vibrometry while a local sound source stimulates first the ipsilateral ear and then the contralateral ear. The transmission (gain and phase lag) is then determined as the eardrum vibration response (transfer function) to contralateral stimulation divided by the response to ipsilateral stimulation. A Beyer DT-48A headphone in a tube coupler was used as a local sound source with the opening placed approx.
6.2 Methods

5 mm from the eardrum and monitored the sound sweeps with the probe microphone. Since the barn-owl head including the facial ruff is rather large (8-10 cm), I did not attempt to make an acoustic shield between the ears since geometric spreading alone attenuated the sound to the contralateral ear by about 20 dB.

6.2.3 Data analysis

Directionality. Eardrum vibration (velocity) was visualized by displaying the eardrum vibration transfer functions in response to free-field stimulation. First, the frequency spectra from all 12 directions were obtained and plotted. The amplitude spectra will be used to infer the sensitivity of the eardrum in the frequency domain. Second, the responses were combined and interpolated to cylinder surface plots introduced by Christensen-Dalsgaard and Manley (2005, 2008). The cylinder surface plots were interpolated contour plots of amplitude with direction on the abscissa (-180°-0° to the left and 0°-180° to the right) and frequency as the ordinate. Third, polar plots were obtained for single frequencies to obtain the gain (in dB) as a function of azimuth. The polar plots will be used to quantitatively determine directionality (see below). The directionality of the ear was determined in a frequency range from 0.4 to 10 kHz roughly covering the hearing range of barn owls (Dyson et al. 1998). The directionality at a given frequency was the difference in dB between the minimal transfer function magnitude and the maximal transfer function magnitude. Directionality values were plotted as a function of frequency. Coherence between input (as specified by the probe signal) and output (the laser signal) was used to judge the reliability of the signal. The coherence function (the cross-spectral density of input and output divided by the auto-spectral density of the input multiplied by the auto-spectral density of the output) was used as a measure for reliability of the data. The coherence function can take values between zero (no relation between input and output) and 1 (perfect relation). Typically a value of larger than 0.8 was used as a criterion for sufficient reliability in comparison to previous studies (Michelsen et al. 1994).

Transmission. The transmission of sound arriving at the internal surface of the ipsilateral ear was determined as the eardrum vibration transfer function to contralateral stimulation divided by the transfer function to ipsilateral stimulation (Michelsen et al. 1994; see also Christensen-Dalsgaard and Manley 2008). The transmission gain was read from the resulting amplitude spectrum as the gain in dB relative to that of external ipsilateral stimulation. The transmission phase spectrum resulting from the division contained the unwrapped interaural phase gains. The transmission phase spectrum showed a linear dependence on frequency around 3 kHz in any owl. The transmission delay was
calculated from the linear section of the transmission phase spectrum by dividing the slope of the linear fit by $2\pi$.

Additionally, directionality from the transmission gain of one owl was simulated by summation of a sweep signal and a second sweep signal, which was attenuated, taking into account the transmission gain as a function of frequency. The signals were temporally shifted according to the transmission delay, which was calculated from the transmission phase. These calculations have been done for each stimulus position that was also used in the measurements and thereby the difference between the minimum and maximum gain at a specific frequency was computed.

6.3 Results

For the present study, directionality data of one owl and transmission gains of five American barn owls (*Tyto furcata pratincola*) were analyzed. First, I briefly mention key features of the morphology of the external ear and the eardrum. Then I turn to the directionality and sound-transmission measurements.

6.3.1 The external auditory meatus, the eardrum and columella, and the interaural canal of the barn owl

The external auditory meatus in the barn owl has a length of about 15 mm from the ear opening to the eardrum (Keller et al. 1998). The ear canal is not straight (Keller et al. 1998), but curved in a complex way that differs in the left and right ears. The asymmetry is due to skin folds (Norberg 1968, for a description). Despite these asymmetries the eardrums of both ears could be observed from outside, if the very small applicable spatial angle was found after holding back ear flaps and feathers around the meatus with tape (Figure 6.3). The eardrum of the barn owl is a large triangular shaped structure (Figure 6.3a, covered with white beads). The attachment point of the tip of the extra-columella to the tympanum (*umbro*) can be identified as radiusued structure at the caudal edge of the eardrum (Figure 6.3b, black arrow). Computer tomographic scans of the head (Figure 6.3c, provided by K.L. Willis, University of Maryland) revealed that the interaural canal connecting the two eardrums is a rostral pointing, V-shaped cavity. The V-shape increases the effective distance between the two eardrums from about 2.5 cm (also compare Coles and Guppy 1988) to approximately 3.5 cm.
The left (a) and right (b) eardrums of one owl. The ear canals are about 15 mm deep (Keller et al., 1998). The triangular shaped structure at the bottom of the ear canal is the eardrum. The laser was pointed at the umbo, where the tip of the extracollumella attaches to the eardrum (b, black arrow). White glass beads were placed at the umbo at the tip of the eardrum to enhance the laser reflections (a). c Interaural canal visualized by computer-tomographic scans (with permission of K. L. Willis, University of Maryland).

**Figure 6.3: Eardrum and interaural canal.**
6.3.2 The sensitivity of the barn owl ear

The directionality of the barn owl ear in the horizontal plane was determined by presenting broadband sounds from 12 loudspeakers, one by one, positioned in the horizontal plane at the height of the owl’s ears around the owl. The data were plotted for qualitative overview as cylinder surface plots interpolating between the 12 azimuthal speaker positions and the frequencies as contained in the sweep stimuli. Vertical cuts through the cylinder surface plots yield frequency responses at specific stimulus positions. Horizontal cuts yield polar plots, from which directionality data can be derived.

In the sample data set shown in Figure 6.4a,b one may see that the gain or sensitivity was low for frequencies up to 3 kHz but showed a steep rise of about 20 dB above that range. Polar plots indicate that the ear is directional toward frontal and ipsilateral positions (Figure 6.4c). Directionality may be read from polar plots in which the minimum transfer function magnitude was set to 0 dB (Figure 6.4c) as the transfer function gain at a specific frequency. At both 1.5 kHz and 2.5 kHz, the directionality was above 3 dB, which has been previously used as suitable criterion for significant (Christensen-Dalsgaard and Manley 2005). Directionality was almost gone at 5 kHz. The behavior of the eardrum was qualitatively visualized in the cylinder surface plots and was quantified by two measures: first, the reliability of the recordings was checked by computing the coherence function (Figure 6.5); second, directionality was computed from the polar plots (Figure 6.6).
6.3 Results

Figure 6.4: Eardrum transfer function.

a Typical eardrum transfer function for a stimulus angle of 30° ipsilateral to the stimulated ear. b Cylinder surface plot showing the eardrum vibration amplitude. Colors show the eardrum vibration amplitude. Black arrow shows the transfer function for 30° (compare a). Horizontal dashed lines show transfer function at each frequency shown as polar plots in c. Positive stimulus angles denote stimulation from the ipsilateral side and negative angles denote contralateral stimulation.

c Polar plots of eardrum transfer functions at specific frequencies. Distance between circles is 2 dB.

6.3.3 Coherence as a measure of reliability

Since the laser recordings typically contain some noise, and since it was sometimes difficult to obtain a good signal, coherence functions were used to check the reliability of the data (compare Methods – chapter 6.2.3). The coherence functions from the case shown in Figure 6.4 are plotted in Figure 6.5. Figure 6.5a displays two coherence functions together with the corresponding amplitude spectra (the laser signals as dotted lines). One response was obtained from stimulation with an ipsilateral speaker (orange lines), the other for stimulation with a contralateral speaker (purple lines). The gains of both responses increase from low frequencies up to about 9 kHz. By contrast, one coherence function shows a high reliability (coherence >0.8) for frequencies between 0.7 and 8.8 kHz (purple lines), while in the other coherence function the high reliability was restricted to frequencies between .8 kHz
Eardrum directionality induced by interaural transmission

and 6.5 kHz (ipsi, orange lines). This suggested that both in the low and the high frequency ranges, the data were less reliable than in the middle frequency range. To obtain a measure for the responses obtained by stimulation from each of the 12 speakers, the coherence function for the response obtained from each speaker was first calculated separately (grey lines in Figure 6.5b). These coherence functions were then averaged for the mean coherence function (black line in Figure 6.5b). The data confirmed what was seen in the examples. Coherence was below 0.8 for frequencies up to 1.5 kHz (Figure 6.5b). Reliability (as read from the mean curve) was high between 1 and 8 kHz and then decreased again while the coherence with contralateral stimulation was below 0.8 above 6.5 kHz. Therefore, I concluded that the signal that was recorded with the laser was reliable for frequencies between 1 and 6.5 kHz for the case shown in Figure 6.4 while the data for frequencies outside this range need to be judged with care.

The coherence function was calculated for all recordings and served as criterion for inclusion or omission of data in the following analysis. Only those cases were included in the further analyses, in which the mean coherence of all speakers was above 0.8. I also excluded measurements in which single-speaker coherence did not reach 0.8. In this way 8 individual data sets (transfer functions of each of the 12 stimulus locations) from 1 owl and 1 ear remained for the further analysis. One of those cases was with the contralateral ear plugged and one other after removal of the plug. The lower boundary values for the data sets after unplugging of the contralateral ear was 1.7 kHz and the upper boundary was 6.4 kHz. For the plugged condition the lower boundary was 2.4 kHz and the upper boundary 6.3 kHz.
6.3 Results

6.3.4 Directionality

According to the criteria outlined above, directionality was analyzed as a function of frequency. The length of the vectors in the polar plots (Figure 6.4c) represented the difference between the maximum and minimum gain for all 12 directions, because the minimum gain was set to zero in these plots. Thus, the value of maximum gain corresponded to the maximum difference of the gains for all 12 directions, the directionality at that frequency. Directionality depended on frequency (Figure 6.6).

Figure 6.6a shows the directionality in unblocked conditions before plugging the contralateral ear canal. Data of 6 consecutive measurements was averaged in Figure 6.6a. The black section of the directionality function plot shows the frequency range with coherence > 0.8. The directionality was higher than 3 dB up to a frequency of 3.5 kHz. From 3.5-6.2 kHz, the directionality was below 3 dB. Directionality increased again for frequency above 6.2 kHz. The steep ramp at 3 kHz shown in the frequency responses in Figure 6.4 is accompanied by peak directionality at 3 kHz. These data indicated that directionality may be low in the range where sound-localization precision is high. Blocking the contralateral ear causes the directionality in the low frequency range to drop below 3 dB (Figure 6.6b). The right panels of Figure 6.6 illustrate the difference of directionality between the three conditions: before blocking, contralateral ear canal blocked, and after removal of the block. The effect of blocking is a strong evidence for the interaural canal to induce directionality in the low frequency range. After removal of the block the ear is again directional (> 3 dB) up to 3 kHz. The coherence drops below 0.8 at 6.4 kHz in these measurements. There is again an increase of directionality seen at those frequencies. But note that directionality at frequencies with a coherence < 0.8 should be judged with care.
Eardrum directionality induced by interaural transmission

Figure 6.6: Directionality of the eardrum.

a-c Directionality as a function of stimulus frequency. Directionality is derived from eardrum transfer functions as the difference between the maximum and minimum transfer function magnitude at a specific frequency. a Directionality without blocking the contralateral ear. The plot shows the average of 7 individual measurements of one ear with high coherence. b Directionality with a blocked contralateral ear. Only single measurement with high coherence was available. c Directionality after removal of the contralateral plug. Only single measurement with high coherence was available for the post-blocking condition. The solid black lines in a-c indicate the range with coherence > 0.8. Coherence is below that value for data indicated with grey lines. d-f Difference of directionality between blocked and unblocked conditions. The respective cases are indicated in the insets. The solid black lines in d-f indicate the range with coherence > 0.8 for both conditions considered in the respective plot.
6.3 Results

6.3.6 Transmission gain

The Transmission gain was calculated to obtain a measure of the amount of internal coupling of both ears. Transmission gain was measured for each owl (n = 5, Figure 6.7). The average maximum averaged gain was about -1 dB (mean) at 0.4 kHz (Figure 6.7b). Note that this was also the starting frequency of the test sweeps, and the average coherence of laser measurements for the five owls was below 0.8 at this frequency. Figure 6.7a shows the transmission gain of the same owl of which directionality data was presented in the previous sections. The transmission gain decreased from 0 dB at 1.6 kHz to -22 dB at 6.5 kHz. Remember that frequencies > 6.5 kHz should be considered carefully due to low coherence. The transmission gain between 2 kHz and 3 kHz can account for the directionality measured in the frequency region (-6 to -12 dB, see simulations below). At higher frequencies the transmission gain is too low to produce considerable directionality of the eardrum. That is congruent with the directionality observations shown in the previous section. The phase offset between the sound directly incoming at the eardrum and the sound transmission through the interaural cavity is the main mechanism producing directionality. The offset is generated by the additional travel time of the sound wave through the interaural canal. The travel time or transmission delay can be calculated from the difference between the phase spectra of ipsi- and contralateral eardrum transfer functions. Here the difference between the spectra is referred to as transmission phase spectra. Transmission phase spectra of five owls are shown in Figure 6.7c. The transmission delays were calculated from a section of the phase spectra around 3 kHz. In these sections the unwrapped transmission phase depended linearly on the frequency. The transmission delay equals the slope divided by $2\pi$. Delays ranged from 224 $\mu$s to 361 $\mu$s for measurements with un-attenuated test signals in 4 owls. The measurement of the fifth owl was performed with 30 dB attenuated signals. Thus the signal to noise ratio was lower in the recordings. The transmission delay of the fifth owl was 452 $\mu$s at 3 kHz. Additionally, directionality was simulated from the data presented in Figure 6.7a and compared with directionality measured in the same owl (Figure 6.7d). The reduction of directionality above 3 kHz can be well explained by the interaural transmission gain. However, the reason for the increase of directionality above 6 kHz remains unclear. The even higher directionality above 7 kHz can be again explained by poor coherence (<0.8).
6 Eardrum directionality induced by interaural transmission

6.4 Discussion

The data on the properties of the barn owls eardrum and the interaural transmission presented here show that the ear is directional in a frequency range where eardrum sensitivity is low (< 3kHz), while directionality is low in a frequency range where sensitivity is high. Likewise, transmission is high (> -12 dB) for low frequencies (<3 kHz), but lower than this value for higher frequencies. In the following these data will be discussed with respect to methodological considerations, interaural transmission, to sound-localization and to mechanisms underlying interaural coupling.

6.4.1 Methodological considerations

Larsen and co-workers (1997) showed that the directionality of the ears of small birds may be impaired during anesthesia, because the Eustachian tube then does not become vented to the same extent as in the awake bird. Consequently, outer and inner ear pressures are not equalized. The overpressure from the outer ear then leads to a decreased motility, and to an inwards-movement of the eardrum, respectively. As a consequence, the low-frequency sound can no longer interact with the tympanum from both sides. A similar effect can occur due to a decrease in temperature caused by anesthesia (Larsen et al. 1997), where the changes in atmospheric pressure exceed the pressure variations that sound waves induce. In both cases, the hampered motility of the tympanum acts as a high-pass filter, and sound localization in the low-frequency range would be impaired. Larsen et al
(1997) observed that in small birds such as budgerigars, the vibrational amplitude of the eardrum strongly increased when the skull was vented (as usually performed via the Eustachian tube in the awake bird) compared to the non-vented situation. As a consequence, intracranial air pressure corresponded to ipsilateral sound pressure, meaning that there was almost no interaction of indirect sound with the eardrum. If venting of the middle-ear cavity was also relevant for the directivity of the owl’s ears, that is, if the eardrum can expand, or vibrate, to a lesser extent in the anesthetized than in the un-anesthetized owl, then the gain (in dB) should be reduced for any frequency. Larsen reported that directivity of the eardrum was lost over all frequencies (1, 2, 4 and 8 kHz) in the non-vented budgerigar skull (Larsen et al. 1997). Such an effect, however, was not observed in the monaural directivity pattern of barn owls suggesting that a reduction of venting did not take place in the recordings examined here.

6.4.2 Relation to earlier interaural transmission data
Moiseff and Konishi (1981b) reported that the ears in the barn owl function as independent pressure receivers for frequencies above 3.5 kHz. The authors measured sound attenuation by the barn owl’s interaural pathway by comparing firing thresholds of cochlear nucleus neurons during ipsi- and contralateral stimulation and measured sound pressure close to the eardrum. Sound attenuation at 7 kHz (63 dB) by the interaural canal was too high, for the ears to act as pressure-gradient receivers, since this would require indirect sound to reach the tympanum. The attenuation was 13 dB at 3.5 kHz frequency. Thus the interaural pathway has low-pass filtering properties. These findings are similar to the transmission gains measured with the laser vibrometry technique in the present study. The findings of Moiseff and Konishi (1981b) thus showed that the interaural pathway attenuates sound coming from the contralateral ear, which speaks in favor of the owl’s ear cavities being uncoupled pressure receivers. Note that the barn owl’s ears are asymmetrically arranged; hence the owl’s ear is not directly comparable to the ears of most other birds such as, for example, the kestrel (Calford 1988).

Analysis of barn owl HRTFs showed that the ITD range does not increase significantly in the tested low-frequency range (500 Hz to 1 kHz) compared to the high frequency range (5-9 kHz) (Hausmann et al 2010). This stands somewhat in contrast to what Calford and Piddington (1988) found in the grass owl (Tyto longimembris), which is a close relative of the American barn owl (Tyto furcata pratincola). Calford and Piddington (1988) used cochlear microphonics (CM) instead of HRTF-measurements and HRTF-measurements might fail in detecting the influence of the interaural canal. However, Cazettes et al. (2014) came to the same conclusion as Hausmann et al., 2010, taking into
account, that the range of best ITDs in the central nucleus of the inferior colliculus did not exceed the range of ITD measured by Hausmann et al. (2010). These were indications that the ears of barn owls do not work as pressure-difference receivers, because the ITD range would then be expected to increase up to threefold (Vossen et al. 2010).

### 6.4.3 Relation to sound localization

The barn owl (*Tyto furcata pratincola*) is an exception compared to most other birds in concerns of sensory processing. These birds have a facial ruff. They hear exceptionally well. In the barn owl, for example, thresholds are very low (down to -15 dB SPL Konishi 1973; Dyson et al. 1998) with a remarkably broad audible frequency spectrum (compared with other birds) that ranges 200 Hz to 10 kHz (Dyson et al. 1998). However, the barn owl does not use the entire frequency range for precise sound localization: Konishi (1973) and Singheiser et al. (2010) reported a decline of precision of sound localization for frequencies below 5 kHz suggesting that the directionality I observed in these frequency regions has no influence on sound localization behavior and, in reverse, the eardrum directionality is low in the frequency range important for sound localization (5-9 kHz). Knudsen and Konishi (1979) also reported a decreased ability to localize low frequency sounds (<5 kHz) compared to frequencies between 5 and 8 kHz. However, this was mainly due to decreased performance in the localization of elevational cues; whereas azimuthal performance was highest between 3 and 8 kHz. The main reason for the poor performance in localizing low frequency sounds in elevation lies in the absence of ILDs at low frequencies and ILDs are a major cue for elevational sound localization in barn owls (Poganiatz and Wagner 2001). Frequencies below 3 kHz are, therefore, not considered to be crucial for sound localization in barn owls, which prompts that the low frequency sound transmission (-10 dB transmission gain below 3 kHz) as reported here and by Moiseff and Konishi (1981b) has no influence on sound localization in barn owls. This conclusion is supported by the fact that responsiveness to low frequencies in the midbrain pathway, which converges at the optic tectum, is missing.

### 6.4.4 Mechanisms underlying interaural coupling

The ears of birds are connected by an interaural canal that opens to the mouth cavity via the Eustachian tube (see Christensen-Dalsgaard 2011 for a review). If the tube is closed, each ear cavity forms a separated, enclosed space, as is the case in mammals. Such ears may act as (non-directional) pressure receivers, and directionality is gained by binaural evaluation of interaural time differences (ITDs), respectively phase differences (IPDs), and level differences (ILDs) in higher brain centres.
6.4 Discussion

At low frequencies, IPDs are small, and the ear, therefore, exhibits little directivity in a pressure-gradient receiver. It is only the tympanal interaction and its amplifying or damping effect on IPDs that yields sufficient cues for sound-localization. This interaction can result in up to three-fold larger ITDs in the low-frequency range (Calford and Piddington 1988, Klump 2000, Vossen, 2010). The interaural transmission delay (~300 μs) was about 3 times longer, then expected by the distance between the two eardrums (3.5 cm, Figure 6.3c) which indicates that the delay is boosted by resonances (compare Vossen 2010). The prolonged transmission delay increases the possible phase difference at low frequencies and thus directionality.

Eardrum directionality can also be an additional source of directionality in auditory neurons, even in neurons insensitive to binaural cues (ITD and ILD). The auditory nerve of lizards, for instance, is sensitive to both ILD and ITD, because the eardrum vibration amplitude varies with these spatial parameters (Christensen-Dalsgaard et al. 2011). Lizards, in contrast to most birds including barn owls, possess nearly loss-free acoustic coupled eardrums resulting in large directionality up to 40 dB and a large frequency spectrum with significant directionality ranging from 2-6 kHz (directionality bandwidth, Christensen-Dalsgaard and Manley 2005, 2008). The large directionality and directionality bandwidth provides Lizards with a powerful mechanism to neurally represent the acoustic space, even in early stations of the auditory pathway. Barn owls, in contrast, use a place code of downstream neurons in the midbrain ICX and optic tectum which maximally respond to preferred ITDs and ILDs. Both nuclei are not tuned to frequencies at which the eardrum of the barn owl is directional (> 3kHz, Vonderschen and Wagner 2009). Nevertheless, an influence of low frequency interaural coupling could still be present in forebrain responses. The forebrain pathway, ending at the acoustic arcopallium (AAr), is also sensitive to low frequencies. Aar neurons often show a bimodal frequency-tuning, with maximum responses up to 3 kHz in addition to a high frequency maximum (Vonderschen and Wagner 2009, 2012). Thus, low-frequency coding of spatial cues in the forebrain might be influenced by eardrum directionality. However, the influence of pressure gradient receiving ears on low frequency auditory processing in the barn owl remains unclear. It is not yet tested whether interaural coupling up to 3 kHz suffices to produce significant ITD-tuning in the early monaural stations of the auditory pathway of barn owls. Therefore, further investigations with respect to low frequency responses to ITD in early stages of the auditory pathway are necessary to uncover a possible influence of low frequency eardrum directionality on sound localization in the barn owl.
7 General discussion

In the preceding chapters I used two specific stimulus paradigms to examine adaptations of the sound-localization system in the barn owl that allow this bird to improve its hunting success. I showed in chapter 4 that precision and accuracy of the fixation of the sound source were decreased, if the stimulus was preceded by a masker (or adapter) stimulus. The experiments presented in chapter 5 demonstrated that frequency modulation helps in disambiguating the sound source due to storage of frequency information within time windows. Last, but not least morphological experiments suggest that transmission through the interaural canal induces eardrum directionality for low sound frequencies (chapter 6). The results of these experiments will be discussed in the following with respect to mechanisms underlying optimization of sound-source localization. Furthermore, the neural basis of sound-localization behavior shall be discussed. Finally, I will conclude the thesis and give a brief outlook on possible future research directions.

7.1 Mechanisms underlying optimization of sound-source localization

Barn owls are efficient hunters that possess several adaptations of the auditory system. Aside from the most obvious adaptation, the prominent facial ruff, the asymmetry of the ears provides the possibility to use ILD and ITD for two different spatial coordinates. In addition, the frequency range used for precise sound localization is extended compared to other birds. Nevertheless, immense selective pressure weighs on the individual bird because of the lack of large amounts of energy reserves, of which the owl cannot afford to waste too much due to unsuccessful hunting. To increase the hunting success, the owls have mechanisms to remember previous locations of sound sources, and should be able to disambiguate the true location of a sound source. These mechanisms are constrained by the physical properties of sound waves and the neuronal process of detecting sound source locations, for example by detecting the ITD. My studies showed that ITD carries ambiguous information about the sound source if the stimulus is narrowband (see also Mazer 1998; Saberi et al. 1999). As a result the owls turn their head toward phantom sound sources with narrowband noise stimuli with a bandwidth $\leq 3$ kHz in this and earlier studies (Saberi et al. 1999). The head-turning angles corresponded to the expected location of phantom sounds which could be determined from the period at the center frequency of the signal and a factor converting ITD into space (von Campenhausen and Wagner 2006). Localization of phantom sound sources, as presented in chapter 5, arises from ambiguous responses of auditory midbrain neurons that are tuned to narrow frequency bands (Wagner et al. 1987, Saberi et al. 1999). Ambiguity is eliminated by integration across frequency channels in the external nucleus of the inferior colliculus (ICX) (Takahashi and Konishi 1986). This system comes with the
7.1 Mechanisms underlying optimization of sound-source localization

drawback of still responding ambiguously if it is driven by narrowband sounds and thus elicits phantom-source localization. The increase of head-turning responses toward the real sound source by increasing the bandwidth up to 3 to 4 kHz can be well explained by the decrease of side-peak suppression in the ICX or optic tectum in the same range of bandwidths (Mazer 1998; Saberi et al. 1999). A similar effect of bandwidth on behavior has been found in humans (Trahiotis and Stern 1989; Buell et al. 1994) who also use ITD as a cue for azimuthal sound-source localization but only for frequencies up to 1500 Hz if only the carrier of the sound signal is considered (Blauert 1997). Since natural sounds are typically broadband the detriment of ambiguity does not become important. However, some gurgling and chirps produced by owl hatchlings are narrowband frequency modulated. The overall bandwidth of the fundamental of these sounds is about 1 kHz to 2 kHz (Bühler and Epple 1980). Chirps elicited head turns toward phantom sound sources but to a lower extent than pure tones in my experiments. This shows that frequency modulation indeed increases the probability of responding to the real sound source. However, this finding also suggests that the information about the frequency content is stored in a certain time interval. Storing frequency information for a certain time is a strategy to counteract the loss of accuracy with sounds that have insufficient frequency content. Storing information in brief time intervals would, therefore, increase the probability of correct responses to significant sounds within the natural environment and thus increase hunting success and recognition of intraspecific vocalizations. Another strategy of the barn owls to avoid unnecessary energy consumption is waiting for a second sound, presumably to confirm the kind and origin of the sound (Payne 1971). This behavior might well be influenced by neuronal adaptation, i.e. the reduction of neural responses in consecutive stimulation. Hereby too, the stimulus history is an important parameter that impacts neural coding and sound-localization behavior. Adaptation is assumed to play a role in several behavioral tasks that includes forward masking (Smith 1977; Nelson et al. 2009), precedence effect (Litovsky and Yin 1998; Spitzer et al. 2004), novelty detection (Ulanovsky et al. 2003; Ulanovsky et al. 2004), contrast enhancement (Furukawa et al. 2005), gain control mechanisms (Laughlin 1989), dynamic range adjustment (Dean et al. 2005), auditory stream segregation (Bee and Klump 2004). Thus, the importance of neural response adaptation for sound detection and localization is evident and assuredly influences nocturnal hunting by listening to faint sounds, and especially waiting for a second sound before actually striking the prey.

The two aspects mentioned above (storing frequency information and response adaptation) can be seen as adaptations within the frequency region that is most important for sound localization (> 3 kHz). The third adaptation of the auditory system (interaural transmission) that was investigated, however, influences sound processing in the low frequency regions up to about 3 kHz. Indeed, the
barn owl’s eardrum was directional at low frequency up to 3 kHz. That means, that the eardrum serves as pressure-gradient receiver, which creates the possibility of a different coding strategy of the auditory space in these frequency regions (compare Christensen-Dalsgaard et al. 2011). However, models of optimal spatial coding in the brain also predicted that 3 kHz is a critical frequency at which coding of interaural time differences switches between two different strategies in owls (Harper and McAlpine 2004). Above 3 kHz, it was proposed that ITD coding is best established by a spatial map of peak coding neurons in a fashion as it was already found in the ICX and optic tectum of the owl’s midbrain. Below 3 kHz the optimal coding strategy was proposed to change into a slope code; that is coding of the ITD by the relative activity of co-occurring populations of neurons. Low frequencies, however, have presumably less influence on sound localization than frequencies above 3 kHz, for the accuracy with which low frequency tones are localized drops considerably (Konishi 1973; Knudsen and Konishi 1979). The prominent interaural cavity in barn owls, therefore, rather serves weight loss than the enhancement of sound-localization abilities. Nonetheless, pressure gradient receiving eardrums might still influence sound processing, especially in the forebrain pathway, in which response to low frequencies is conserved (Cohen and Knudsen 1995; Vonderschern and Wagner 2009) in contrast to the midbrain pathway where low frequency information is discarded (Wagner et al. 2007) which is congruent with the model of Harper and McAlpine (2004). Interaural transmission could influence head orientation based on memory (Knudsen and Knudsen 1996), or top-down control (Winkowski and Knudsen 2006). These aspects are both assumed to be mediated by the forebrain. It is most important to note that eventually optimal coding of the auditory space is based on the decoding of the sensory information by the motor system, and thus on behavior (Salinas 2006). Different behavior controlled by forebrain and midbrain in the barn owl, thus likely result in different coding strategies and tuning properties (Vonderschern and Wagner 2012). In the following section I will further elaborate on the behavior regulated by the different brain regions of barn owls.

7.2 Suggestions for auditory processing underlying sound-localization behavior

Manipulating the stimulus in a specific form allowed answering questions about the contribution of specific physiological mechanisms to sound-localization behavior. However, in the case of the double stimulus paradigm one must take into account several mechanisms, for example, response adaptation (Singheiser et al. 2012), attention (Posner 1980; Johnen et al. 2001), or sub-attentional top-down mechanisms (Winkowski and Knudsen 2006; Winkowski and Knudsen 2007) that might influence the behavior in such a paradigm. Both the double stimulus as well as the frequency modulated stimuli led to increased response latency by about 50 ms. Similar effects on response latency have been found
7.3 Conclusion and outlook

using other paradigms. (Poganiatz et al. 2001) for example observed a 60 ms increase of latencies in response to virtual acoustic stimuli which consisted of ITDs and ILDs pointing to different locations. A similar increase of response latencies (+40 ms) was observed in paradigms with cueing stimuli (Johnen et al. 2001). The additional delay in the head turning response indicates that additional synapses are involved before triggering the actual motor response. This might be due to attentional modulation or similar top-down regulation of the behavior. For example, the sensory responses of the optic tectum are modulated by the forebrain arcopallial gaze fields (Winkowski and Knudsen 2006). The acoustic information might also be undergoing additional processing in the acoustic arcopallium of the forebrain. The spatial parameters ITD and ILD are represented in the AAr in a different way than they are in the midbrain. The coding of these parameters in the AAr resembles the coding in the midbrain of small rodents (McAlpine et al. 2001; Vonderschen and Wagner 2009) in contrast to the place code in the barn owl midbrain nuclei (Knudsen et al. 1977; Olsen et al. 1989). The AAr, as final relay on the forebrain pathway, projects back to the ICX, ICC and optic tectum but also to motor nuclei of the tegmentum (Cohen et al. 1998). Looping the processes through the forebrain nuclei mentioned above might then lead to increased response latency. Behavioral experiments revealed that the forebrain pathway also computes spatial information independently from the midbrain, and therefore contributes to sound localization parallel to the optic tectum (Knudsen et al. 1993). Similar to neurons of the optic tectum AAr neurons do not only code the location of an acoustic signal but also influence motor structures. The AAr is part of the frontal gaze field in which electric stimulation induces head saccades (Masino and Knudsen 1993) independent of the optic tectum. On the contrary, the optic tectum directly projects into motor regions of the spinal cord (Masino and Knudsen 1992). Sound localization via the midbrain optic tectum resembles a reflex loop that does not involve the forebrain. Therefore, localization of sounds that does not include further processing, for example, decision making when hearing two distinct sound sources like the phantom and the real sound source, is presumably accompanied with shorter response latencies.

7.3 Conclusion and outlook

The barn owl has been as model organism for sound localization for nearly half a century. As a specialist for nocturnal hunting and thus possessing a highly adapted auditory system the barn owl is predestined for such investigations. Many adaptations of brain structures underlying auditory-related behavior are developed in the barn owl. These adaptations include enlarged areas within the auditory pathway connected with detecting the locus of a sound source. The mechanisms discribed and discussed above contribute to precise sound-localization behavior on several stages of the auditory pathway. To draw further conclusions from my findings it is essential to probe the physiological basis
underlying the behavior which I observed under well defined conditions. For example, it would be interesting to find out how midbrain neurons respond to frequency modulated tones, and whether side-peak suppression in the ICX or optic tectum neurons is influenced by the past frequency content of a stimulus. Does response adaptation influence the response to different stimulus locations (i.e. ITDs or ILDs), and therefore influence sound-source localization or rather only sound-source detection? This question is currently considered with electrophysiological experiments by Roland Ferger in the Wagner lab. And last but not least, can one find tuning to spatial cues in monaural neurons at the early stages of the auditory pathway? These are some unanswered questions showing that although much is already known about sound-source localization and its neural basis in barn owls a lot of questions still remain open. That also includes questions of genetics and proteomics underlying the generation of behavior which is made accessible for non-generalists in the past decade.

Exploring the evolutionary adaptations of specialist provides the researchers with insight to optimal solutions to specific issues. These solutions and mechanism can then in turn be adapted to technical applications. Findings about the mechanisms of the auditory system can be applied to enhance the performance of cochlear implants and, thus, help hearing impaired persons, or autonomic agents can be equipped with algorithms mimicking brain processes. First attempts in the latter direction are under way in the Wagner lab (Calmes et al. 2007; Göckel et al. 2014). Applications could be autonomous robots that orient by localization of sounds by using an algorithm of a specialist in sound localization like the barn owl (Calmes et al. 2007).
8 References


References


9 Acknowledgements

Hermann gebührt großer Dank; für das Vertrauen, dass er schon während des Studiums und später während der Promotion in mich gesetzt hat, für die Möglichkeit zu vielen Tagungen zu reisen und auf diese Weise Kontakt zu vielen anderen Wissenschaftlern knüpfen zu können, und für den hilfreichen Rat in vielen Situation. Bei Jörg, bedanke ich mich, für die Übernahme des Zweitberichters, aber auch für viele lehrreiche Stunden, neben der Biologie auch beim Segeln und dem ein oder anderen Whisky. Ole und Jakob, many thanks to you. It was my great pleasure to meet you. To have had the opportunity to work with the two of you was terrific.


Für finanzielle Unterstützung bedanke ich mich bei der Deutschen Forschungsgemeinschaft (DFG).