

# **Measuring Primate Gaze Shifts to Study Audio-Visual Evoked Orienting**

Een wetenschappelijke proeve op het gebied van de  
Natuurwetenschappen, Wiskunde en Informatica

## **Proefschrift**

ter verkrijging van de graad van doctor  
aan de Radboud Universiteit Nijmegen,  
op gezag van de rector magnificus prof. mr. S.C.J.J. Kortmann,  
volgens besluit van het college van decanen  
in het openbaar te verdedigen op maandag 12 april 2010  
om 13.30 uur precies

door

**Peter Bremen**

geboren op 29 november 1979  
te Stolberg (Rhd.), Duitsland

Promotor: prof. dr. AJ. van Opstal  
Copromotor: dr. RF van der Willigen

Manuscriptcommissie: prof. dr. H Bekkering  
dr. LJ Bour (AMC Amsterdam)  
prof. dr. MA Frens (Erasmus Universiteit Rotterdam)  
prof. dr. L Goffart (CNRS-INCM Marseille)  
dr. HHLM Goossens

© 2010 Peter Bremen

ISBN-10: 90-9025174-5

ISBN-13: 978-90-9025174-5

This research was supported by a Marie Curie Early Stage Training Fellowship of the European Community's Sixth Framework Program (MEST- CT-2004-007825)

# Contents

Chapter 1	9-26
General introduction	
1.1 Measuring eye movements: A historical perspective	
<i>Robinson's scleral search coil technique</i>	
<i>The alternative: The double magnetic induction method</i>	
<i>Topics addressed in this thesis</i>	
1.2 Mechanisms of Sound Localisation	
<i>Localisation of one sound source in the horizontal plane</i>	
<i>Localisation of two concurrent sound sources in the horizontal plane</i>	
<i>Localisation of one sound source in the vertical plane</i>	
<i>Localisation of two concurrent sound sources in the vertical plane</i>	
1.3 A short primer on audiovisual integration	
<i>Audiovisual integration:</i>	
<i>ventriloquism and calibration</i>	
<i>Audiovisual integration:</i>	
<i>complex environments, congruent stimuli</i>	
<i>Audiovisual integration:</i>	
<i>complex environments, incongruent stimuli</i>	
1.4 References	
Chapter 2	27-46
Using Double-Magnetic Induction to Measure Head-Unrestrained Gaze Shifts – I Theory and Validation –	
2.1 Introduction	
2.2 Methods	
<i>Theoretical considerations for 1D eye rotations in the horizontal plane</i>	
<i>Non-Idealized situation</i>	
<i>Practical application</i>	
<i>Test on a gimbal system</i>	
<i>Calibration</i>	
2.3 Results	
<i>Calibration of the secondary induction signals</i>	
<i>Influence of Vnet</i>	
2.4 Discussion	
2.5 References	

Chapter 3	47-72
Using Double-Magnetic Induction to Measure Head-Unrestrained Gaze Shifts – II Application in Human Subjects –	
3.1 Introduction	
3.2 Methods	
<i>Subjects</i>	
<i>Apparatus</i>	
<i>Stimuli</i>	
<i>Paradigms</i>	
<i>Control Experiment</i>	
<i>Characterization of Vnet</i>	
<i>Extended DMI method</i>	
<i>Calibration</i>	
<i>Characterization of Vnet</i>	
<i>Extended DMI method.</i>	
<i>Data Analysis</i>	
3.3 Results	
<i>Characterization of Vnet</i>	
<i>DMI/SSC Measurement</i>	
<i>Uncalibrated Data</i>	
<i>Calibration</i>	
<i>Saccades during the test paradigm</i>	
3.4 Discussion	
<i>Stability of the DMI system</i>	
<i>Measurement range and calibration</i>	
<i>Recording hardware</i>	
<i>Comparison with video techniques</i>	
<i>Binocular recordings</i>	
<i>Experimental animals</i>	
3.5 References	

Chapter 4	73-102
Using Double-Magnetic Induction to Measure Head-Unrestrained Gaze Shifts – III Calibration & Validation in the Monkey –	
4.1 Introduction	
4.2 Methods	
4.2.1 Simulations	
<i>Gaze shift and head saccade generation</i>	
<i>Simulation of head-unrestrained</i>	
<i>double magnetic induction signals</i>	
<i>Calibration of double-magnetic induction data</i>	



- 4.2.2 Experiments
  - Subjects*
  - Eye-head recording apparatus*
  - Stimuli*
  - Paradigms*
  - DMI calibration paradigm*
  - Head coil calibration paradigm*
  - Calibration*
  - Data Analysis*

4.3 Results

- 4.3.1 Simulations
  - Raw DMI signals*
  - Calibrated gaze signals*
- 4.3.2 Measurements
  - Raw DMI signals*
  - Calibrated gaze signals*

4.4 Discussion

- Training of Neuronal Networks*
- Stability of the assembly*
- Conclusion*

4.5 References

Chapter 5

103-128

Pinna Cues Determine Orienting to Double Sound Sources

5.1 Introduction

5.2 Methods

- Listeners*
- Apparatus and sound generation*
- Stimuli*
- Behavioral testing and paradigms*
- Data analysis*
- Response Normalization*
- Directional transfer functions (DTFs)*
- Similarity model of sound localization*
- Correction procedure*

5.3 Results

- Single-sound localization*
- Double-sound localization*
- DTF similarity-model*

5.4 Discussion

- Averaging and bimodality in the auditory system*
- Comparison to other studies*
- DTF model*

5.5 References

Chapter 6	129-146
Acquired Prior Knowledge Modulates Audiovisual Integration	
6.1 Introduction	
6.2 Methods	
<i>Listeners</i>	
<i>Apparatus and sound generation</i>	
<i>Experiments</i>	
<i>Calibration experiment</i>	
<i>Visual experiment</i>	
<i>Auditory experiment</i>	
<i>Audiovisual 100%-aligned/0%-distractor (AV-100/0)</i>	
<i>Audiovisual 10%-aligned/90%-distractor (AV-10/90)</i>	
<i>Audiovisual 50%-aligned/50%-distractor (AV-50/50)</i>	
<i>Distributions of spatial disparities</i>	
<i>Data Analysis</i>	
<i>Data calibration</i>	
<i>Statistics</i>	
6.3 Results	
<i>AV integration of aligned stimuli in aligned experiment</i>	
<i>Breakdown of AV integration by spatial disparity</i>	
<i>Effect of congruence-disparity distribution</i>	
6.4 Discussion	
<i>Model for Reaction Time Modulation</i>	
<i>by Prior Likelihood Estimation</i>	
6.5 References	
Summary	147-150
Samenvatting	151-156
Publications/Published Abstracts	157-158
Acknowledgements	159-160
Curriculum Vitae	161





## **Chapter 1**

# **General Introduction**

## 1.1 Measuring eye movements: A historical perspective

Historically, new findings are tightly coupled to technical developments. Take for instance the development of the patch clamp technique by Neher and Sakmann for which both were awarded one-half of the Nobel Prize in Physiology or Medicine in 1991. Their method led to new discoveries concerning the function of single ion channels in cells. Needless to say that these developments revolutionized our view on basic functions of nerve cells.

Earlier, in the 1960s, technical advances took place that opened a new avenue for motor physiology and oculomotor research in particular (Robinson, 1987). These advances were on the one hand conceptual in nature, and on the other hand technological. The idea to approach the brain from a systems analysis point of view although at that time controversially debated, is now fully accepted and has proven to be successful in modeling and understanding circuits of the oculomotor system. When applying systems theory to the analysis of neuronal circuits one develops a hypothesis (usually called model) and collects data to test it (scientific method). The collected data either verify or falsify the model. Either way, the cycle of proposing or modifying the model is repeated until the hypothesis agrees with all data collected so far. At this point the hypothesis becomes a theory, like for example Darwin's theory of evolution (Robinson, 1987).

Obviously, when applying systems theory to the oculomotor system in accordance with the scientific method it is imperative to accurately measure the eye movements themselves. The possibility to record from single neurons in awake behaving animals (Evarts, 1968), and to trace the corresponding neuronal circuits for example with horseradish peroxidase (Graybiel and Hartweg, 1974) became available in the late 1960s, early 1970s. In addition, the advent of affordable computers helped to quickly analyse and interpret the collected data. Although with the electro-oculogram (EOG) a technique for measuring eye movements was available prior to the 1960s, its drift and noisiness made it somewhat tedious to use. Thus, when Robinson (1963) introduced a technique based on electromagnetic induction in a scleral search coil that could measure eye movements with unprecedented accuracy, he provided an easy-to-use, straightforward means to monitor eye movements that has dominated the field ever since.

However, as discussed in the following, the scleral search coil technique has some important drawbacks that continue to trouble oculomotor scientists up to the present day. The work described in this thesis tries to overcome this downside making the measurement of eye movements more comfortable for the experimenter and the subject. Chapters 2, 3 and 4 of this thesis describe a novel method to precisely measure eye movements of head-unrestrained human and animal subjects. Before introducing this alternative method we will first describe Robinson's scleral search coil technique in the following section.

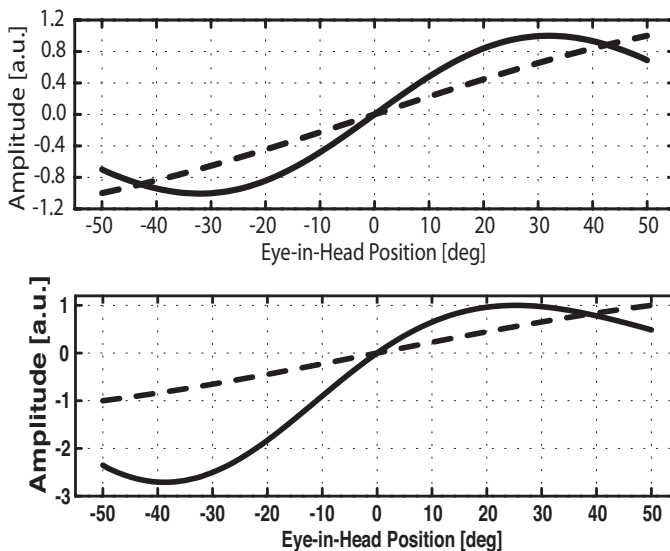
*Robinson's scleral search coil technique.* The golden standard for measuring eye movements in the oculomotor field is still Robinson's scleral search coil (SSC) technique (Robinson, 1963). In the version developed for laboratory animals a small copper-wire coil is implanted beneath the conjunctiva of the eye. The coil's lead wire needed to connect it to the recording apparatus is led through the orbita, beneath the skin up to the top of the head to a connector that is embedded in dental cement (Fuchs and Robinson, 1966; Judge et al., 1979). For human subjects the coil is embedded in a silicon annulus, which can be worn as a contact lens (Collewijn et al., 1975). In that case the connecting wire is led via the eyelashes where it is fixed to the head with tape. For recording the eye movements the subject is seated in an oscillating magnetic field. The amplitude of the voltage induced in the search coil is proportional to the sine of the angle between the axis of the search coil and the magnetic field. Typically, two sets of field coils are used, one producing a vertical field and the other a horizontal field. The two fields are tuned to different frequencies, e.g. 30 kHz and 40 kHz. Lock-in amplifiers are used to demodulate the horizontal and vertical components of the eye movements. The measured signal is roughly linear over a range of about  $\pm 40$  deg, but becomes ambiguous beyond 90 deg. With the addition of a third, frontal magnetic field the measurement range can be readily extended to the full 360 deg. Additionally, the frontal field allows for three-dimensional measurements of eye rotation by having two coils with perpendicular planes embedded on the silicon ring (Robinson, 1963; Collewijn et al., 1975).

The large linear measurement range and accordingly its easy calibration together with the possibility to measure head-unrestrained gaze shifts led to the widespread use of the SSC technique. However, an important drawback of the SSC is the vulnerable lead wire needed to connect the coil to the recording apparatus. This thin wire is prone to breakage. A fact not too surprising when one considers the strain to which the wire is exposed in an implanted animal that makes a few hundred thousand eye movements every day. If the wire breaks a new coil needs to be implanted. In the worst case an animal needs to be implanted several times before the research question at hand can be answered. Obviously this does not only delay the progress of a research project but does also pose an additional burden to the animal. In human subjects the wire easily breaks due to blinks. If this happens at the very beginning of a recording session a new coil can be inserted. However, if the wire breaks later, the whole session needs to be aborted. This problem is particular cumbersome when it concerns measurements in patients. Additionally, recording time is limited to maximally 40 min due to the irritating connecting wires that cause discomfort for the subject when the local anesthetic on the sclera wears out.

*The alternative: The double magnetic induction method.* The hazardous connecting wires of the SSC technique prompted the development of an alternative method for eye movement recordings. In the 1980s the double magnetic induction (DMI) method for head-restrained conditions was developed (Allik et al., 1982;

Reulen and Bakker, 1982; Bour et al. 1984). Instead of a scleral coil, the DMI method makes use of a thin golden ring placed on the subject's eye. Each oscillating magnetic field induces an oscillating current in the ring that in turn induces a (secondary) voltage in a pickup coil placed in front of the eye. The pickup coil is connected to the recording hardware so that there is no strain on the connecting wire, which abolishes the risk of breakage. The ring is easily implanted in animals and can be worn like a contact lens on the eye of human subjects. The ring will not break and thus needs to be implanted only once in a laboratory animal. Due to the absence of connecting wires the ring is more comfortable to wear for human subjects, which prolongs recording time to about an hour. It is also possible to record two eyes simultaneously (Chaturverdi and Van Gisbergen, 2004), since the ring's signal decays exponentially with increasing distance from the ring, i.e. the two ring signals will not interfere with each other.

Because of this rapid decay the measured signal is quite small and highly non-monotonic due to the strong dependence of the ring-pickup coil geometry (Fig. 1). The linear range is about  $\pm 20$  deg, but depends heavily and nontrivially on the exact geometry of ring and pickup coil. If the axes of ring and pickup coil are not perfectly aligned the linear range is shifted to the opposite direction of misalignment. Moreover, the pickup coil does not only measure the small ring signal but also the much larger primary field component, which therefore needs to be cancelled with a second coil connected in anti-parallel to the pickup coil. To keep the primary component small so that the tiny ring signal can be amplified maximally



**Figure 1:** Example of the of simulations of the DMI signal. (**top**): The nonlinear relation of the predicted signals in the pickup coil as function of  $\alpha$  (Eqn. (1); **solid line**) reaches a maximum near  $\alpha = 30^\circ$ . (**bottom**): Simulation of a condition in which the axes of ring and pickup coil are not collinear, instead the ring's center is moved by  $\delta = 0.8$  cm toward the right relative to the pickup coil's center. This misalignment causes the geometrical nonlinearity to become asymmetric re.  $\alpha = 0^\circ$ . The **dashed line** in both plots corresponds to the strength of the induction current in the ring (or to the signal as obtained from the traditional search-coil method and is proportional to  $\sin(\alpha)$ . Taken from: Bremen et al. 2007a.



some researchers employed micromanipulators to optimally align pickup coil and anti-coil (Malpeli, 1998).

The tedious practical implementation, the nonlinear and limited measurement range, its vulnerability to mechanical vibrations, together with its restriction to head-restrained paradigms have so far prevented wide spread acceptance of the DMI method (see however Bour et al., 2008 for a successful application in patients). Clearly, if calibration of the DMI signals could be simplified and if the method could be extended to head-unrestrained preparations, it could prove to be a valuable alternative to the SSC technique.

*Topics addressed in this thesis.* This thesis describes an extension of the DMI method for head-unrestrained preparations. The goal is to allow precise and reliable eye-movement recordings in laboratory animals, like for example monkeys and cats, which are also used for electrophysiological recordings. Due to the absence of vulnerable wires on the eye the DMI method is an excellent candidate. The development of the extended method is characterized by a combination of theoretical and experimental approaches. Chapter two introduces the theoretical background needed to understand the complex DMI signal in head-unrestrained paradigms. Special emphasis is given to the mathematical description of the DMI signal. These theoretical considerations are then compared with measurements obtained with a horizontally rotatable dummy platform simulating the subject's head and eye movements. This chapter demonstrates that in principal the calibration of the DMI signal is possible over the full 360 deg range.

Based on these results, the third chapter describes the application of the DMI method to measure two-dimensional human gaze shifts. The equipment needed and a calibration routine involving coordinated eye-head movements are described in detail. This chapter shows that the implementation of the DMI method is technically feasible and that the achieved spatial resolution of the method is comparable to that of the SSC technique.

Chapter four summarizes the steps needed to setup the DMI method in laboratory animals. To demonstrate the basic design principals of the DMI assembly for use with animals the rhesus macaque monkey has been chosen. These animals perform readily in visual fixation tasks needed for calibration. Compared to the routine employed with human subjects the calibration routine for animal subjects is streamlined and only involves the fixation of randomly presented visual targets across the frontal hemifield, instead of precisely instructed eye-head movements. For a successful calibration a large database of gaze shifts with different eye-head combinations is needed. In laboratory animals this is possible since the DMI assembly is attached to rigid implants on the animal's head. The assembly can thus be fixed in the same way on each recording day. This chapter demonstrates the feasibility of this approach by employing a combination of simulations and real measurements.

## 1.2 Mechanisms of sound localisation

Chapters five and six of this thesis deal with sound localisation behaviour of human subjects (chapter five), and with the integration of auditory and visual information in order to control goal-directed gaze shifts (chapter six). Both chapters highlight localisation performance toward multiple sensory sources within the vertical (mid-sagittal) plane of the head: auditory-only in chapter five, and audio-visual in chapter six. As will be outlined this plane is particularly special for sound localisation, because all binaural sound-localisation cues within this plane are equal. In what follows we first discuss the mechanisms underlying sound localisation, for both single- and for multiple sources. We also provide a brief outline of the major aspects of audiovisual integration that are relevant for this thesis.

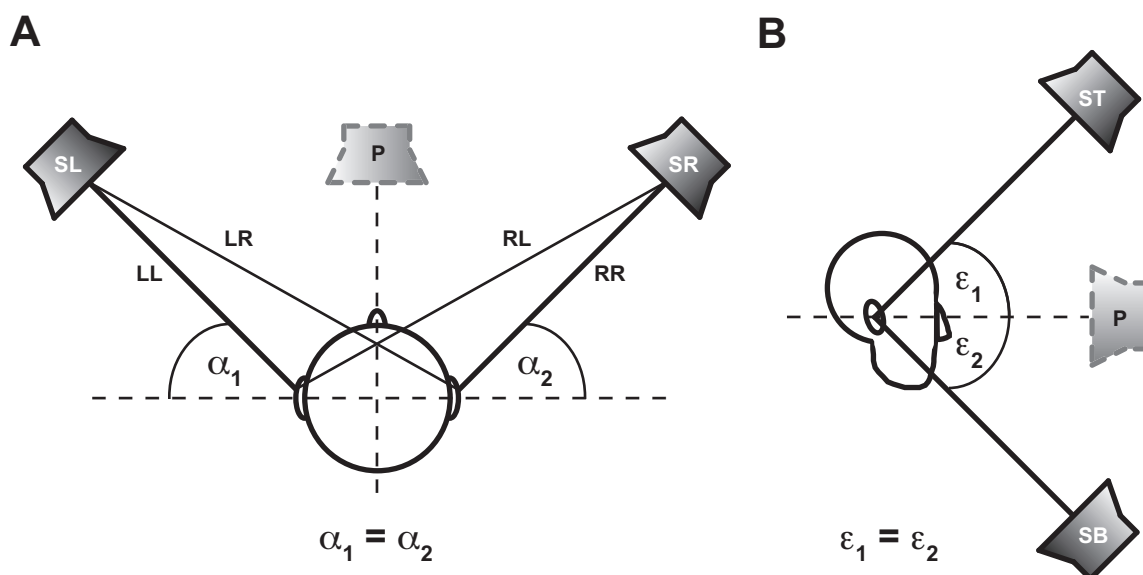
*Localisation of one sound source in the horizontal plane.* Localising, identifying and selecting objects in the environment is crucial for the survival of any animal species. The ability to hear helps to localise and identify objects in the entire space around the head based on their acoustic features. Several acoustic properties of a sound contribute to this process. For example, a sound source that is presented at different locations in space induces different acoustic patterns at the listener's ears. Through experience, the auditory system has learned to use these space-specific acoustic patterns to reliably localise the sound source. Sound-localisation studies in our laboratory have shown that humans can localise single sound sources with great accuracy (e.g. Hofman and Van Opstal, 1998a,b).

Sound localisation relies on three different mechanisms: two types of binaural differences are uniquely related to sound-source directions within the horizontal plane of the head (quantified by the azimuth angle). In addition, specific patterns of amplifications and attenuations at different sound frequencies relate in a unique (albeit complex) way to the location of a sound within the midsagittal plane of the head (the elevation angle; see below). Each of these mechanisms is processed by independent neural pathways within the ascending auditory pathway (Yin, 2002; Young, 2002, for reviews). Adequate sound localisation requires the integration of information from these different pathways.

Sound directions within the horizontal plane impose differences between the ears that vary in a systematic way with azimuth angle. Due to the limited speed of sound (~340 m/s) small but detectable interaural timing differences (ITDs) arise that correspond to unique ongoing phase differences for low frequency tones (<1.3 kHz). Neurons in the auditory brainstem (medial superior olive) are sensitive to these phase differences. Phase differences become ambiguous at higher frequencies, as the path-length difference from the source to the two ears may, or may not contain multiple wavelengths. However, the attenuating effect of the head (the acoustic "head-shadow") at higher frequencies introduces noticeable interaural level differences (ILDs) that also vary in a systematic way with sound-source azimuth and frequency. Neurons in the lateral superior olive are specifically sensitive to these ILDs. This dual principle for the encoding of sound-source

locations in the horizontal plane was already proposed by Lord Rayleigh (1907).

*Localisation of two concurrent sound sources in the horizontal plane.* If two speakers within the frontal horizontal plane of the listener simultaneously emit the same sound (say, a pure tone) the listener will perceive a phantom source at the so-called summing location, which is in between the two physical sources (Blauert, 1997). If the speakers are arranged symmetrically with respect to the listener's midline the phantom source seems to originate from straight ahead (Fig. 2A). This is the well-known stereophonic effect. The summing location for a pure tone in the free field can be readily predicted from interference of the pressure waves at the eardrum. In the current example the path lengths (ITDs) and the attenuation by the head shadow (ILDs) for each of the speakers at the two ears cancel each other, which results in net ITDs and ILDs that point straight ahead. Thus, in such conditions the phantom source is due to the peripheral acoustics at the ears, and requires no explanation in terms of neural integration mechanisms within the auditory system.



**Figure 2:** Schematic of the physical situation in double speaker experiments with symmetric speaker arrangements in azimuth (**A**) and in elevation (**B**). (**A**) The phantom percept in azimuth can be understood by the interactions of the pressure waves at the ear drum. In the case that  $\alpha_1 = \alpha_2$  interaural differences in time of arrival (ITD) and level (ILD) cancel each other. The path lengths LL and RR as well as LR and RL are the same this results in an ITD of 0  $\mu$ s. Likewise, attenuation by the head is the same for both speakers resulting in an ILD of 0 dB. Both cues point towards 0 deg azimuth, i.e. straight ahead. (**B**) Single sound source localisation in elevation is based on monaural spectral cues (HRTFs). Can the phantom source (middle) perceived at zero elevation be explained by the sum of the HRTFs corresponding to speaker location ST (**top**) and SB (**bottom**), i.e. is  $P = ST + SB$ ?

SL = Left speaker, SR = Right speaker, ST = Top speaker, SB = Bottom speaker, P = Phantom source,  $\alpha_1$  = Angle between SL and the interaural axis,  $\alpha_2$  = Angle between SR and the interaural axis, LL = Path of sound from SL to the left ear, LR = Path of sound from SL to the right ear, RR = Path of the sound from SR to the right ear, RL = Path of the sound from SR to the left ear.

Yet, neural integration does exist for horizontal localisation of multiple sources. This can be readily appreciated from headphone stimulation experiments with independently manipulated tones at each ear. These manipulation leads to similar phantom percepts within the head.

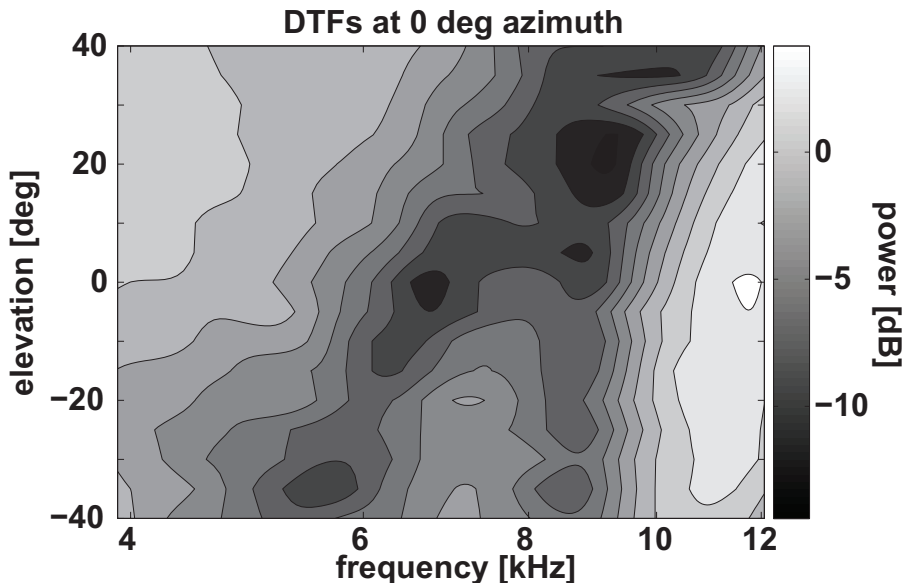
The neural integration of spatial information arising from different sensory inputs is a crucial task for the brain, as it needs to decide whether inputs do, or do not, belong to the same source. Indeed, many studies of visuomotor control have shown that the processing of two visual target locations leads to eye-movement localisation responses that are directed to an average location (just like the auditory phantom source; Ottes et al., 1984). This behaviour is generally attributed to neural interactions of the original source locations within spatially organised neural maps, like e.g. the midbrain superior colliculus. It is therefore an interesting question whether similar mechanisms could underlie the processing of multiple sound sources.

This question has not yet been studied thoroughly for multiple sound sources in the midsagittal plane. As a first attempt, chapter five describes experiments in which we presented synchronous sounds in the midsagittal plane while recording the rapid-orienting responses of subjects to the perceived sound location. In what follows, we briefly outline the idea behind these experiments.

*Localisation of one sound source in the vertical plane.* To localise sounds in the vertical (mid-sagittal) plane humans use spectral shape cues that arise from the direction-dependent filter characteristics of the torso, head and pinna. These filters are known as head-related transfer functions (HRTFs). They can be measured by inserting a small probe microphone into the ear canal of the subject and record the microphone response to broad-band sounds presented at different elevation angles. HRTFs obtained from a typical male subject are shown in Fig. 3. In this plot the power of the signal is grey-coded with brighter (darker) shades indicating an amplification (attenuation) at a given frequency. Note the prominent dark streak running from about 5 kHz at -40 deg elevation up to 10 kHz at +40 deg. This streak represents a “spectral notch” and is thought to be an important cue to localise sounds in elevation. The exact location, depth and angle of the notch in elevation-frequency space, depends on idiosyncratic differences in the geometry of the subject’s pinna. Note that the pinna restricts the spectral shape cues to frequencies above 3-4 kHz, for which the wavelengths become comparable to the size of the pinna folds. Sources containing no spectral power above 3 kHz (like buzzing mosquitoes) can thus not reliably be localized in the elevation direction (Middlebrooks, 1992; Hofman and Van Opstal, 2002; Macpherson and Middlebrooks, 2003).

According to linear systems theory, the sensory sound spectrum, as it arrives at the eardrum can be described in the frequency domain by:

$$P_{ED}(f, \epsilon) = H(f, \epsilon) P_{ST}(f) \quad (1)$$



**Figure 3:** Contour plot of directional transfer functions (DTFs) from the right ear of the author with power in dB being coded in gray scale. Note the prominent notch (dark streak) called notch running from -40 deg elevation at 5 kHz up to +40 deg elevation at 10 kHz. DTFs are obtained from HRTFs by subtracting the average spectrum calculated over all recorded HRTFs. In this way the direction independent stimulus spectrum, the resonance of the ear canal and the response of the recording apparatus are removed so that only the unique direction-dependent contribution of a given location remains.

with  $P_{ED}(f, \varepsilon)$ , the frequency and elevation direction-dependent spectrum at the eardrum,  $H(f, \varepsilon)$  the direction-dependent HRTF and  $P_{ST}(f)$  the spectrum of the stimulus. Note that it follows from Eqn. 1, that the auditory system is faced with an ill-posed problem. Without a priori knowledge or assumptions of the stimulus spectrum ( $P_{ST}(f)$ ) there are infinitely many combinations of source spectra and HRTFs (and hence elevation directions) that give rise to the same spectrum. Nevertheless, studies have demonstrated that spectral shape cues are crucial for the elevation percept and disambiguation of the front-back angle and that human subjects can use these cues with high fidelity (Batteau, 1967; Blauert, 1969; Gardner and Gardner, 1973; Hofman et al., 1998).

Several models propose a mechanism for how the auditory system may get around the ill-posed problem and compute a reliable and accurate location estimate from the sensory spectrum. The most prominent of these models is based on the idea that the auditory system somehow compares the perceived sound spectrum to a learned internal representation of the spectral shape cues (Middlebrooks 1992; Hofman and Van Opstal, 1998; Langendijk and Bronkhorst, 2002). This correlation model assumes (i) that the sets of stored HRTFs do not correlate with each other, i.e. they are all unique, and (ii) that a given stimulus spectrum,  $P_{ST}(f)$ , does not resemble any of the stored HRTFs. It can be mathematically shown that such a model can successfully localise sounds for a large variety of broad-band spectra. The model also predicts specific illusory locations, in case the source spectrum does resemble one of the stored HRTFs.

Note that although the spectral cues are frequently referred to as being “monaural localisation cues”, experiments with perturbed spectral cues have revealed that the cues from both ears are centrally weighted to create the actual elevation percept (Hofman and Van Opstal, 2002; Macpherson and Sabin, 2007; Van Wanrooij and Van Opstal, 2007).

*Localisation of two concurrent sound sources in the vertical plane.* Compared to the situation of two simultaneous sources in azimuth (see above) the analysis in elevation is different since the interactions now take place in the spectral domain. Suppose that at the eardrum the spectrum of a sound at elevation,  $\varepsilon_1$ , is  $H_1(f)$  and that of a sound at elevation,  $\varepsilon_2$ , is given by  $H_2(f)$ . Since at the ears acoustic waves follow the linear superposition principle, Fourier analysis of two spectra yields:

$$H_{12}(f) = H_1(f) + H_2(f) \quad (2)$$

It is important to note that the frequency spectrum is a complex variable and can be written as:

$$H(f) = \text{Re}(H(f)) + i \times \text{Im}(H(f)) = R(f)e^{i\Phi(f)}$$

$$R(f) = \sqrt{\text{Re}^2(H(f)) + \text{Im}^2(H(f))} \quad \Phi(f) = \arctan \frac{\text{Im}(H(f))}{\text{Re}(H(f))} \quad (3)$$

with  $R(f)$  and  $\Phi(f)$  the amplitude and phase spectra, respectively. Although it is generally assumed that the auditory system solely relies on the amplitude spectra, and is insensitive to phase information, the latter nevertheless becomes important when considering the amplitude spectrum for multiple sound sources. Constructive or destructive interference between the two sources may significantly alter the shape of  $H_{12}(f)$ , as can be seen when considering the amplitude spectrum of Eq. 2:

$$R_{12}(f) = \sqrt{R_1^2(f) + R_1(f)R_2(f) \times [\cos(\Phi_1(f))\cos(\Phi_2(f)) + \sin(\Phi_1(f))\sin(\Phi_2(f))] + R_2^2(f)} \quad (4)$$

The result is clearly different from a mere linear summation of the individual amplitude spectra of  $R_1(f)$  and  $R_2(f)$ . The question arises whether or not the summed HRTF,  $H_{12}(f)$ , suffices to explain a subject’s localisation behaviour (Fig. 2B).

If true, the elevation percept would be determined by peripheral acoustic interactions, which would be crucially different from the results obtained in visuomotor studies (see above). If false, however, the elevation percept would arise from neural interactions of the two source representations within the auditory system, possibly like those observed in visuomotor behaviour. For free-field sound sources this would clearly be different from the mechanisms described for the azimuth direction. Chapter five deals with this problem.



### 1.3 A short primer on audiovisual integration

*Audiovisual integration: ventriloquism and calibration.* While chapter five deals with the localisation of multiple unimodal stimuli, chapter six addresses the problem of audiovisual integration, in particular the combination of an auditory and a visual stimulus presented at the same or at different locations in the mid-sagittal plane.

Having multiple senses increases the likelihood of detecting and identifying objects in the environment. Each sensory modality (e.g. audition and vision) is optimally equipped to encode, decode and interpret information from different physical events (Stein and Meredith, 1993). However, to integrate the information from different sensory modalities, two fundamental problems arise: (1) how to determine whether the activity of the different modalities arises from one physical object, or from distinct objects? And (2) how to best integrate the information provided by the different sensory modalities, in case of one physical object (and not to integrate for distinct objects)?

In general the brain might assume that events that are close in time and space (congruent sensory events) belong to one object and should therefore be fused into the percept of one object. Such integration occurs e.g. when looking at a ventriloquist, who moves the lips of the dummy, while uttering speech without moving his own lips. In this case, the visual and auditory signals are temporally synchronized, but they arise from (slightly) different locations in space. Yet, the brain assumes that the dummy's moving lips produced the sounds. Apparently, the visual information can dominate the auditory percept, in case of a conflict (spatial non-alignment).

A second important function for audiovisual integration is spatial calibration of the sound-localisation cues. Seminal experiments in the barn owl (Knudsen and Knudsen, 1985) have shown that vision guides the development of sound-localisation behaviour. In their experiments, Knudsen and colleagues mounted shifting prisms on young barn owl chicks and showed that the visual shift induced by the prisms transferred to the sound-localisation responses, even though the acoustic cues of the chicks had not altered. Zwiers and colleagues (2003) demonstrated a similar effect in humans who were equipped with minifying glasses that compressed the visual scene. After two days of altered visual exposure the horizontal sound-localisation responses adapted to the visual compression, thus following the altered visual input rather than the unchanged auditory inputs.

To study audiovisual integration in a quantitative way, actual localization measurements have to be performed under precisely controlled conditions. The gaze-control system is especially suited to study audiovisual integration, as the orienting response is a natural behaviour to these stimuli. Indeed, oculomotor experiments employing relatively simple audiovisual conditions (e.g. one visual and one auditory stimulus, emitted from a limited number of spatial configurations) have demonstrated that audiovisual integration takes place, and is advantageous for the organism. In such settings the saccade reaction times for temporally and

spatially congruent audiovisual stimuli are shorter than the unimodal reaction times, and the benefit depends in a systematic way on the spatial and temporal alignment of the two modalities (e.g. Frens et al., 1995).

*Audiovisual integration: complex environments, congruent stimuli.* Given the complexity of natural environments it is not always straightforward, as in the example of the ventriloquist, to determine whether events arise from an auditory, a visual or an audio-visual object. To study audiovisual integration under more natural conditions, more complex stimulus situations need to be employed. However, such studies are rare. In a study from our lab, we used a complex two-dimensional audiovisual background scene in which a target (either auditory, visual, or audiovisual) could appear at an unpredictable location. In the audiovisual case the stimuli were always spatially aligned. In such a condition, the brain cannot draw a preprogrammed response from a limited repertoire, but needs to continuously evaluate the properties of the scene to find the target. It was shown that despite the complexity, also in such environments audiovisual integration occurs. Indeed, the benefit of integration turned out to be strongest when the uncertainties of the two modalities were largest (i.e. for the weakest auditory stimuli, and very dim visual stimuli; Corneil et al., 2002; Van Wanrooij et al, 2009). In the experiment, the fact that auditory saccades can be more inaccurate than visual saccades, but they are elicited at much shorter reaction times, was exploited. Interestingly, audiovisual responses in this complex environment could be characterized as: auditory reaction times at visual accuracy. Hence audiovisual integration appeared to follow a “best of both worlds” principle, when auditory and visual target stimuli are perceived to be congruent in space-time. In a follow-up study, Van Wanrooij et al. (2009) showed that the best-of-both worlds principle breaks down when the perceived stimuli are not aligned. In that condition, early responses are elicited by the auditory stimulus (in that case, considered to be the distractor), late responses are visually driven, whereas saccades with intermediate reaction times follow a bimodal distribution (either auditory, or visual). In other words, the system cannot ignore the auditory distractor when it is forced to react fast.

*Audiovisual integration: complex environments, incongruent stimuli.* In the Corneil et al. (2002) experiments subjects were told that the auditory, visual and audiovisual stimuli were always the target, as they would be spatially aligned when presented together. In the real world, however, such knowledge is not available, and the brain has to decide on the basis of current evidence whether or not to integrate the audiovisual inputs. So how does the brain know whether or not a given audiovisual event emanates from the same object in space-time? This question is the topic of chapter six.

To study this question, we presented visual and audiovisual stimuli in the midsagittal plane from a large variety of locations and spatial disparities. The subject was told that only the visual stimulus was the target, that the auditory accessory did not provide a spatial cue for the visual stimulus, and should therefore be



ignored. In the experiments we manipulated the probability that the audiovisual stimuli would be spatially aligned. In a given block of trials this probability could be 100% (always aligned), only 16%, or 50%. In the latter two cases we also varied the distributions of the misalignment (uniform vs. peaked). Interestingly, because of the statistics, different random stimulus sequences of aligned and non-aligned stimuli will occur, but some of these sequences will contain one or more repetitions of aligned conditions. We wondered whether the brain keeps track of the history of recent stimulus configurations to estimate the probability that in the current trial the audiovisual stimuli will or will not be aligned. If so, audiovisual integration within complex, unpredictable environments continuously assesses the current statistics of that environment, to infer the most likely situation. Our results indicate that this may indeed be the case.

## 1.4 References

Allik J, Rauk M, Luuk A. Control and sense of eye movement behind closed eyelids. (1981) *Perception* 10:39–51.

Batteau DW. The role of pinna in human localization. (1967) *Proc R Soc London, Ser B* 168:158–80.

Blauert J. Sound localization in the median plane. (1969) *Acustica* 22:205–13.

Blauert J. Spatial hearing: the psychophysics of human sound localization. Revised Edition. (1997) MIT Press, Cambridge, MA, USA.

Bour LJ, Van Gisbergen JAM, Buijns J, Ottes FP. The double magnetic induction method for measuring eye movement—results in monkey and man. (1984) *IEEE Trans Biomed Eng* 31:419–27.

Bour LJ, Van Rootselaar AF, Koelman JH, Tijssen MA. Oculomotor abnormalities in myoclonic tremor: a comparison with spinocerebellar ataxia 6. (2008) *Brain* 131:2295–303.

Chaturvedi V, Van Gisbergen JA. Stimulation in the rostral pole of monkey superior colliculus: effects on vergence eye movements. (2000) *Exp Brain Res* 132:72–8.

Collewijn H, Van der Mark F, Jansen TC. Precise recording of human eye movements. (1975) *Vision Res* 15:447–50.

Corneil BD, Van Wanrooij MM, Munoz DP, Van Opstal AJ. Auditory-visual interactions subserving goal-directed saccades in a complex scene. *J Neurophys* 88:438–54.

- Evarts EV. A technique for recording activity of subcortical neurons in moving animals. (1968) *J Electroenceph Clin Neurophysiol* 24:83–6.
- Fecteau JH, Munoz DP. Exploring the consequences of the previous trial. (2003) *Nature Rev Neurosci* 4:1–9.
- Findlay JM. Global visual processing for saccadic eye movements. (1982) *Vision Res* 22:1033–45.
- Frens MA, Van Opstal AJ, Van de Willigen RF. Spatial and temporal factors determine auditory-visual interactions in human saccadic eye movements. (1995) *Percept & Psychophysics* 57:802-16.
- Fuchs AF, Robinson DA. A method for measuring horizontal and vertical eye movement chronically in the monkey. (1966) *J Appl Physiol* 21:1068–70.
- Gardner MB, Gardner RS. Problem of localization in the median plane: Effect of pinna occlusion. (1973) *J Acoust Soc Am* 53: 400–8.
- Graybiel AM, Hartweg EA. Some afferent connections of the oculomotor complex in the cat: An experimental study with tracer techniques. (1974) *Brain Res* 81:543–51.
- Hofman PM, Van Opstal AJ. Spectro-temporal factors in two-dimensional human sound localization. (1998a) *J Acoust Soc Am* 103:2634–48.
- Hofman PM, Van Riswick JG, Van Opstal AJ. Relearning sound localization with new ears. (1998b) *Nat Neurosci* 1: 417–21.
- Hofman PM, Van Opstal. Bayesian Reconstruction of sound localization cues from responses to random spectra. (2002) *Biol Cybern* 86:304–16.
- Judge SJ, Richmond BJ, Chu FC. Implantation of magnetic search coils for measurement of eye position: An improved method. (1979) *Vision Res* 20:535–38.
- Knudsen EI, Knudsen PF. The adjustment of auditory localization in young barn owls. (1985) *Science* 230:545–8.
- Langendijk EHA, Bronkhorst AW. Contribution of spectral cues to human sound location. (2002) *J Acoust Soc Am* 112:1583–96.
- Macpherson EA, Middlebrooks JC. Vertical-plane sound localization probed with ripple-spectrum noise. (2003) *J Acoust Soc Am* 114:430–45.

Macpherson EA, Sabin AT. Binaural weighting of monaural spectral cues for sound localization. (2007) *J Acoust Soc Am* 121:3677–88.

Malpeli JG. Measuring eye position with the double magnetic induction method. (1998) *J Neurosci Methods* 86:55–61.

Middlebrooks J. Narrow-band sound localization related to external ear acoustics. (1992) *J. Acoust Soc Am* 92:2607–24.

Ottes FP, Van Gisbergen JAM, Eggermont JJ. Metrics of saccade responses to visual double stimuli: two different modes. (1984) *Vision Res* 24:1169–79.

Reulen JPH, Bakker L. The measurement of eye movement using double magnetic induction. (1982) *IEEE Trans Biomed Eng* 29:740–4.

Rayleigh, Lord. On our perception of sound direction. (1907) *Philos Mag* 13:214–32

Robinson DA. A Method of measuring eye movement using a scleral search coil in a magnetic field. (1963) *IEEE Trans Biomed Eng* 10:137–45.

Robinson DA. The windfalls of technology in the oculomotor system. Proctor lecture. (1987) *IOVS* 28:1912–24.

Stein BE, Meredith MA. *The merging of the senses*. (1993) MIT Press, Cambridge, MA, USA

Stein BE, Stanford TR. Multisensory integration: current issues from the perspective of the single neuron. *Nature Rev Neurosci* (2008) 9:255–266.

Van Wanrooij MM, Van Opstal AJ. Sound localization under perturbed binaural hearing. (2007) *J Neurophys* 97:715–26.

Van Wanrooij MM, Bell AH, Munoz DP, Van Opstal AJ. The effect of spatial-temporal audiovisual disparities on saccades in a complex scene. (2009) *Exp Brain Res* 198:425–37.

Yin TC. Neural mechanisms of encoding binaural localization cues in the auditory brainstem. (2002) In: *Integrative functions in the mammalian auditory pathway* (Oertel D, Fay RR, Popper AN, eds), pp 99 –159. Heidelberg: Springer.

Young ED, Davis KA. Circuitry and function of the dorsal cochlear nucleus. (2002) In: *Integrative functions in the mammalian auditory pathway* (Oertel D, Fay RR, Popper AN, eds), pp 160 –206. Heidelberg: Springer.

Zwiers MP, Van Opstal AJ, Page GD. Plasticity in human sound localization induced by compressed spatial vision. (2003) *Nat Neurosci* 6:175–81.





## Chapter 2

# Using Double-Magnetic Induction to Measure Head-Unrestrained Gaze Shifts – I Theory and Validation –

### Abstract

So far, the double-magnetic induction (DMI) method has been successfully applied to record eye movements from head-restrained humans, monkeys and cats. An advantage of the DMI method, compared to the more widely used scleral search coil technique, is the absence of vulnerable lead wires on the eye. A disadvantage, however, is that the relationship between the eye-in-head orientation and the secondary induction signal is highly nonlinear and non-monotonic. This limits the effective measuring range to maximum eye orientations of about  $\pm 30^\circ$ .

Here, we analyze and test two extensions required to record the full eye-head orienting range, well exceeding  $90^\circ$  from straight-ahead in all directions. (1) The use of mutually perpendicular magnetic fields allows for the disambiguation of the non-monotonic signal from the ring. (2) The application of an artificial neural network for offline calibration of the signals. The theoretical predictions are tested for horizontal rotations with a gimbal system. Our results show that the method is a promising alternative to the search-coil technique.

## 2.1 Introduction

Currently, two invasive electromagnetic methods for monitoring eye movements are being used in oculomotor research: the scleral search coil (SSC) technique, and the double magnetic induction (DMI) method. The SSC technique has become the gold standard because (1) it allows for high spatial (down to a few minutes of arc) and high temporal (typically one ms, or better) resolution, (2) it is approximately linear across the entire human (monkey and cat) oculomotor range ( $\pm 40^\circ$ ), (3) it does not limit the field of view, (4) it can be applied in complete darkness, even with closed eye lids, and (5) it can be readily used to measure head-free gaze shifts over the full motor range. To measure eye movements with the SSC technique in human subjects, a silicon annulus is placed on the eye (Robinson, 1963; Collewijn, 1975). This annulus contains an embedded coil made of thin copper wire with protruding lead wires connected to a detection device. When a subject is placed in the center of a periodically changing uniform magnetic field, eye orientation in space (gaze) can be determined from the amplitude of the induction voltage in the SSC.

The SSC method, however, has a number of drawbacks mostly linked to the lead wires (Bour et al. 1984; Bos et al. 1988; Malpeli, 1998). (1) Dragging of the wires through the eyelashes causes irritation, which limits experimental time to about 30 minutes. (2) Visual acuity deteriorates in the course of the experiment. (3) Breakage of the wires can occur during the experiments. (4) The lead wires, combined with the visco-elastic coupling between the SSC and the conjunctiva, have been hypothesized to introduce extra variability in eye movement kinematics, which is not observed with non-invasive, video-based eye-tracking methods (Van der Geest and Frens, 2001; Smeets, 2003; see, however, Houben et al., 2006). These complications become increasingly significant when SSCs are implanted chronically in laboratory animals. There, the SSC is implanted beneath the conjunctiva. A small loop of wire is fashioned to give enough lead way for the eye to move. The connecting wire is then led subcutaneously up to the top of the head, where it is embedded in dental cement (Judge, 1980). The exit point of the wire introduces an additional risk for infections. However, the most important drawback, of the SSC technique is the considerable risk of wire breakage, making re-implantation of a new coil inevitable.

The lead wire complications of the SSC method have prompted efforts to develop the so-called double magnetic induction (DMI) method (Allik, 1981; Reulen and Bakker, 1982; Bour et al., 1984; Bos et al., 1988; Malpeli, 1998). In this method, the SSC is replaced by a short-circuited metal ring that is comparable in size to a SSC. Apart from prolonging experimental time in human subjects this also improves the comfort, as the need of connecting wires between the eye and recording equipment is avoided altogether. The head-restrained subject is seated in an alternating magnetic field that induces a current in the eye ring. The current strength depends on the orientation of the eye with respect to the magnetic field. This current, in turn, generates an additional secondary magnetic field that is



picked up by a coil placed in close proximity to the ring and that remains stationary relative to the head. The secondary induction voltage in this, so-called, pickup coil depends on the current strength, and on the precise geometrical relationship between pickup coil and ring. Because the signal in the pickup coil also contains a contribution from the primary magnetic field, an additional coil, the so-called compensation coil is placed in the magnetic field and connected in anti-phase with the pickup coil. The primary component can thus be subtracted from the pickup coil's signal, and what remains, is the contribution of the ring.

Apart from requiring a head-restrained subject, the main drawback of the DMI method, however, is that the effective measurement range is limited to about  $30^\circ$ , due to complicated geometrical relationships and a non-monotonic, non-linear dependency between eye orientation and signal. The linear range, being only about  $\pm 10^\circ$ , is even more restricted (Reulen and Bakker, 1982). To deal with this problem, Bour et al. (1984) proposed a calibration procedure that employed a quadratic polynomial fit of the data. Later, Bos et al. (1988) improved the description of the inherent nonlinearity by using the electrodynamical concept of mutual induction that for a fixed geometrical relationship between ring and pickup coil was derived from the Biot-Savart law. So far, the DMI method has been used successfully in head-restrained human subjects (e.g., Reulen and Bakker, 1982, Bour et al., 2000) and patients (e.g. Bour et al., 2008), monkeys (e.g., Bour et al., 1984) and cats (Malpeli, 1998). The spatial and temporal resolution of the DMI method is comparable to that of the SSC. Still, the limited effective measurement range and the complex setup, together with the high susceptibility to noise due to insufficient compensation of the primary field, has so far restricted the design of experiments with human subjects. However, the major advantage of the DMI method, compared to the SSC technique, is the easy implantation of the ring in animal subjects and the lack of connecting wires, which not only makes re-implantations unnecessary but also minimizes the risk of infections.

In this study we investigated how to overcome the inherent limitations of the DMI method so that it may be used in a greater variety of experimental setups. Our main aim is to use the DMI-method in head-unrestrained subjects, and apply it over the full gaze-motor range, well exceeding  $90^\circ$  in all directions. To understand the basic underlying principles in this situation, we have extended the theoretical approach of Bos et al. (1988) to a head-unrestrained condition, and to a more realistic situation in practice, in which neither the ring/pickup coil assembly, nor the pickup coil and compensation coil system are perfectly aligned.

To test the theoretical predictions of our analysis, we employed a dummy platform to simulate the subjects' head and eye movements over the  $180^\circ$  left/right range from straight-ahead in the horizontal plane. To calibrate these data two extensions to the original DMI method (Reulen et al., 1982; Bour et al., 1984; Malpeli, 1998) were made. (1) Three mutually perpendicular, primary magnetic fields were used instead of two, enabling for a disambiguation of the non-monotonic ring signal. And (2) a model-free artificial neural network was employed to calibrate this signal. Our results indicate that the extended DMI method may be a promising alternative for the SSC technique.

## 2.2 Methods

### Theoretical considerations for 1D eye rotations in the horizontal plane

*Idealized situation.* For the head-restrained situation, Bos et al. (1988) derived on the basis of mutual induction between the ring and the pickup coil and the Biot-Savart law that the voltage induced in the pickup coil by the ring is given by:

$$V_h(\alpha, t) = K_h(t) \cdot \sin(\alpha) \cdot L[\cos(\alpha)] \quad (1)$$

with  $\alpha$  the horizontal eye-in-head orientation (in deg),  $K_h(t)$  is a proportionality constant that depends on the geometry of the system and the time-varying horizontal magnetic field strength, and  $L[\cos(\alpha)]$  is a shape factor that expresses the geometrical nonlinearity of the DMI method (see below).

Using the parameter notation introduced by Reulen and Bakker (1982) the coefficient  $K_h(t)$ , can be expressed as:

$$K_h(t) = -N_{coil} \cdot N_{ring} \cdot \omega_h^2 \cdot B_h \cdot \sin(\omega_h t) \cdot \frac{\pi^2 \cdot \mu_o \cdot R_{ring}^3 \cdot R_{coil}}{b \cdot Z_{elec}} \quad (2)$$

with

$$b = \sqrt{(R_{eye} + d)^2 + R_{coil}^2} \quad (3)$$

In Eq. (2),  $N_{coil}$  and  $N_{ring}$  are the number of turns of the coil and the ring ( $N_{ring} = 1$ ), respectively.  $\omega_h$  and  $B_h$  are the angular frequency and strength of the horizontal primary magnetic field, respectively. Time is denoted as  $t$ , and  $\mu_o$  denotes the magnetic permeability.  $R_{coil}$ ,  $R_{ring}$  and  $R_{eye}$  are the radii of the pickup coil, ring and eye, respectively.  $Z_{elec}$  is the electrical impedance of the ring and  $d$  in Eq. (3) denotes the coaxial distance between the ring and the pickup coil.

In Eq. (1),  $\sin(\alpha)$  gives the strength of the magnetic flux through the ring (Fig. 1) as a function of horizontal eye rotation over azimuth angle  $\alpha$ . Since in this case the head is fixed,  $\alpha$  represents the angle of the ring with respect to the horizontal primary magnetic field, hereafter referred to as the horizontal field. Most notably, apart from the right-hand factor,  $L[\cos(\alpha)]$ , and details in the proportionality constant  $K_h(t)$ ,  $V_h$  is identical as determined for the SSC technique (Robinson, 1963; Collewijn, 1975). The shape factor  $L[\cos(\alpha)]$  distinguishes the DMI method from the SSC technique by introducing an additional non-linearity that limits the measurement range between  $\pm 30^\circ$ . In particular,  $L[\cos(\alpha)]$  expresses the geometric dependence of the ring and the pickup coil by:

$$L[\cos(\alpha)] = \sum_{n=1}^{\infty} \frac{1}{n(n+1)} \cdot \left(\frac{c}{b}\right)^n \cdot P_n\left(\frac{R_{eye}}{c}\right) \cdot P_n\left(\frac{R_{eye} + d}{b}\right) \cdot P_n(\cos(\alpha)) \quad (4)$$

with

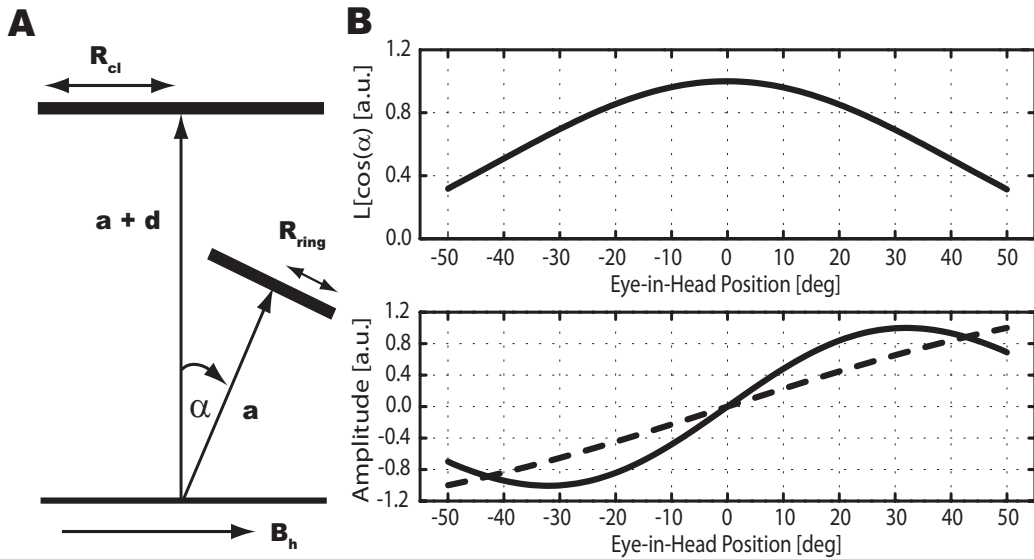
$$C = \sqrt{R_{eye}^2 + R_{ring}^2} \quad (5)$$

Here,  $P_n$  are Legendre polynomials and  $Q_n$  are associated Legendre polynomials of the first kind. In practice, the first six polynomials ( $n = 6$ ) suffice to adequately approximate this series to within 0.05% (Bos et al., 1988).

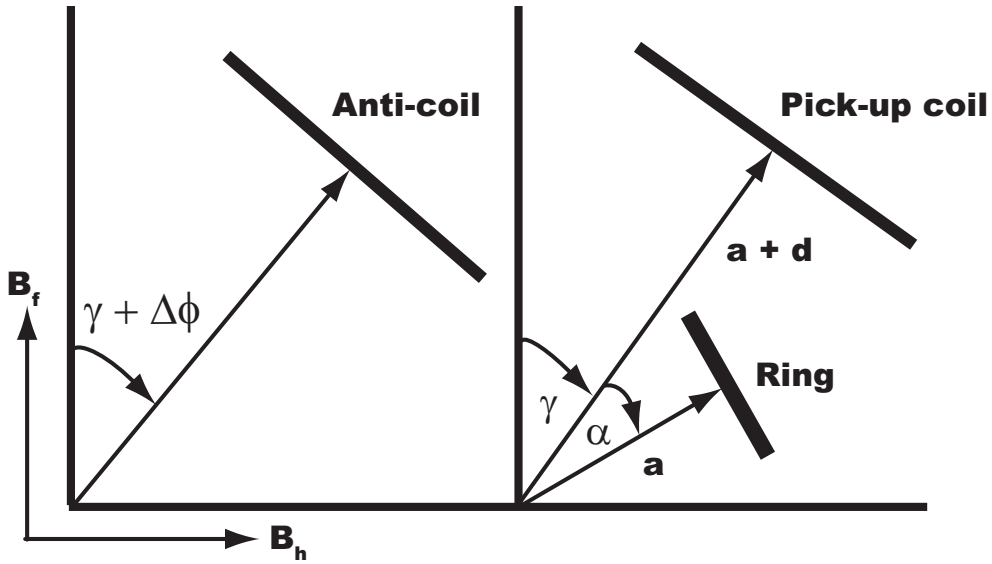
The DMI method can be readily extended to two dimensions, by adding a vertical primary magnetic field, with its own driving frequency,  $\omega v$  and magnetic field strength,  $B_v$  (Reulen and Bakker, 1982; Bour et al., 1984; Bos et al., 1988).

It should be noted that Eq. (1) is only valid for the idealized situation in which the axes of the ring and pickup coil are aligned, i.e. they intersect at the common origin of rotation (see Fig. 1A). A second approximation in Eq. (1) is that the primary magnetic field does not induce a net voltage in the pickup coil, i.e., the coil's orientation is exactly parallel to the magnetic field. In practice, these assumptions may not be exactly met, in which case Eqs. (1) and (2) will have to be modified. We will deal with this non-ideal situation later. First, we proceed with the idealized coaxial condition, and perfect primary field cancellation.

Fig.1A shows the coaxial 1D situation, with  $B_h$  the horizontal field, and  $\alpha$  the eye-in-head orientation. Fig. 1B (top) shows the geometrical factor  $L[\cos(\alpha)]$  as function of  $\alpha$  (Eq. (2), up to  $n = 6$ ), together with



**Figure 1:** (A) Geometry of the head-restrained DMI method, in which the pickup coil (radius  $R_{cl}$ , at a minimum distance  $d$  from the ring) is parallel to the magnetic field ( $B_h$ ). Here, the eye (radius  $a$ ) with the ring (radius  $R_{ring}$ ) has been rotated by angle  $\alpha$  with respect to the normal vectors of the coil and the field. The axes of coil and ring intersect at a common rotation center. We refer to this as the ring/pickup coil assembly. (B, top): The ring-coil shape factor  $L[\cos(\alpha)]$  as function of the eye-in-head orientation (Eq. (2)). (B, bottom): The nonlinear relation of the predicted signals in the pickup coil as function of  $\alpha$  (Eqn. (1); solid line) reaches a maximum near  $\alpha = 30^\circ$  (for fixed  $a = 1.2$  cm,  $d = 2$  cm,  $R_{ring} = 0.8$  cm,  $R_{cl} = 2.5$  cm). The dashed line corresponds to the strength of the induction current in the ring (or to the signal as obtained from the traditional search-coil method) and is proportional to  $\sin(\alpha)$ .



**Figure 2:** Hypothetical DMI geometry for head-free gaze shifts. The anti-coil is positioned at a sufficiently large distance from the ring (e.g., on the head). Its electrical properties are identical to those of the pickup coil, but may make a small, fixed, angle  $\Delta\phi$  with it. Angle  $\alpha$  is the orientation of the ring re. the pickup coil;  $\gamma$  is the angle of the pickup coil with the magnetic field,  $B_h$ , and corresponds to the head orientation in space. Their sum,  $\alpha + \gamma$ , is the eye orientation in space (gaze).  $B_f$  is an additional frontal magnetic field.

the prediction of the pickup voltage according to Eq. (1) (bottom). As can be seen from Fig. 1B (bottom), although the current strength in the ring varies as  $\sin(\alpha)$  (dotted line), which is nearly linear for angles within the oculomotor range (up to about  $\pm 40^\circ$ ), the coil's secondary induction voltage depends in a highly nonlinear way on the ring's orientation (solid line). A more severe problem with the DMI method, however, is its limited measurement range, as the signal reaches its extremes at about  $\alpha = \pm 30^\circ$ .

Therefore, in its present form the method is useless when eye orientations would exceed the  $\pm 30$ - $35^\circ$  range, because the secondary induction voltage then becomes ambiguous. For head-restrained conditions, however, this potential ambiguity is not problematic, since the available measurement range covers a major part of the human (cat and monkey) oculomotor range. Under natural conditions ocular excursions beyond  $30^\circ$  eccentricity are rare and can be readily constrained within an experiment. In head-unrestrained subjects, this is of course not feasible. We will address the issue of head-unrestrained gaze shifts in the following paragraphs.

Note, that in Eq. (1) the  $L[\cos(\alpha)]$  term is only determined by the geometry of the coil-ring assembly, and does not depend on the orientation of that assembly with respect to the primary magnetic field. In contrast, the  $\sin(\alpha)$  factor is only determined by the flux of the primary magnetic field through the ring. Therefore, in the situation that the entire assembly of Fig. 1A is rotated by angle  $\gamma$  with respect to the horizontal magnetic field (this simulates a head turn; see Fig. 2), only the  $\sin(\alpha)$  term in Eq. (1) will change and  $V_h$  becomes:

$$V_h(\alpha, \gamma, t) = K_h(t) \cdot \sin(\alpha + \gamma) \cdot L[\cos(\alpha)] \quad (6)$$

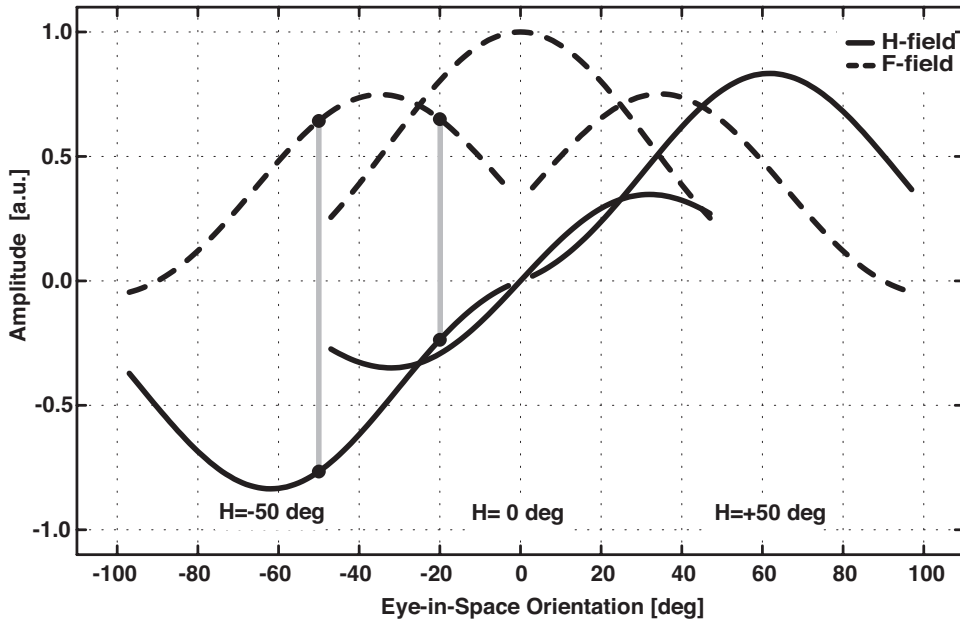
In Eq. (6), angle  $\alpha$  is the orientation of the eye in the head,  $\gamma$  is the rotation of the head in space. Thus, their sum  $\alpha + \gamma$  represents the direction of gaze (eye-in-space). The inherent nonlinearity of the DMI method (due to  $L[\cos(\alpha)]$ , see Fig. 1B, top) is therefore not influenced by head-unrestrained gaze shifts.

Instead, for eye-in-head orientations in the  $30\text{-}35^\circ$  range the curve in Fig. 1B (bottom) will now shift as a function of head orientation, as determined by the factor  $\sin(\alpha + \gamma)$ . Accordingly, the ambiguity of the eye-in-head signal will shift by the same amount, considerably restricting the monotonic measurement range in the direction of head rotation (Fig. 3, solid curves).

However, the introduction of an additional primary magnetic field that is perpendicular to the horizontal field, will fully resolve this ambiguity, and in addition, drastically improve the resolution of the system near the extremes. A frontal primary field, with strength  $B_f$  and frequency  $\omega_f$  induces the following secondary voltage in the pickup coil:

$$V_f(\alpha, \gamma, t) = K_f(t) \cdot \cos(\alpha + \gamma) \cdot L[\cos(\alpha)] \quad (7)$$

Eq. 7 is identical to Eq. 6 except that  $\sin(\alpha + \gamma)$  now becomes  $\cos(\alpha + \gamma)$ .



**Figure 3:** Predicted Pickup coil signals for the horizontal (solid line, H-field) and frontal fields (dashed line, F-field) as function of the eye-in-space orientation, shown for three different head orientations ( $50^\circ$  leftward, straight-ahead, and  $50^\circ$  rightward) assuming an ideal geometry as depicted in Fig. 2. The eye-in-head orientation is restricted to  $\pm 45^\circ$ . Note that, the shape of the DMI nonlinearity varies with the head orientation (see Eq. (6)). Because the voltages induced by the two magnetic fields have a  $90^\circ$  phase difference, a peak or valley in one field's signal coincides with a monotonic relation for the other field (e.g., between the points connected by the gray vertical lines). Thus, even though the signal from the horizontal field may be ambiguous, the eye-in-head orientation can be uniquely retrieved, once the head-in-space orientation is known.

In Fig. 3, simulated signals of the frontal induction voltages for three head orientations (straight ahead, and  $\pm 50^\circ$  horizontal) are plotted as function of gaze (with the eye-in-head orientation  $-45 < \alpha < +45^\circ$ ; dashed lines), for the idealized situation that the primary signals in the anti-coil and pickup coil fully cancel. With the help of Eq. (7) it is possible to disambiguate the non-monotonic input-output relation of the ring/pickup coil assembly as will be discussed later. Note here, however, that for each head orientation a unique combination of voltages for the horizontal (H-field) and frontal (F-field) fields exists, as indicated by the black dots connected by the grey vertical lines. As an example, the signal of the frontal field for  $50^\circ$  and  $20^\circ$  eye orientation is ambiguous as indicated by the black dots. Nevertheless, the signals for the same eye orientations of the horizontal field are not ambiguous (black dots). In combination with the head orientation there will always be a unique combination of signals that exactly determines the eye-in-head orientation.

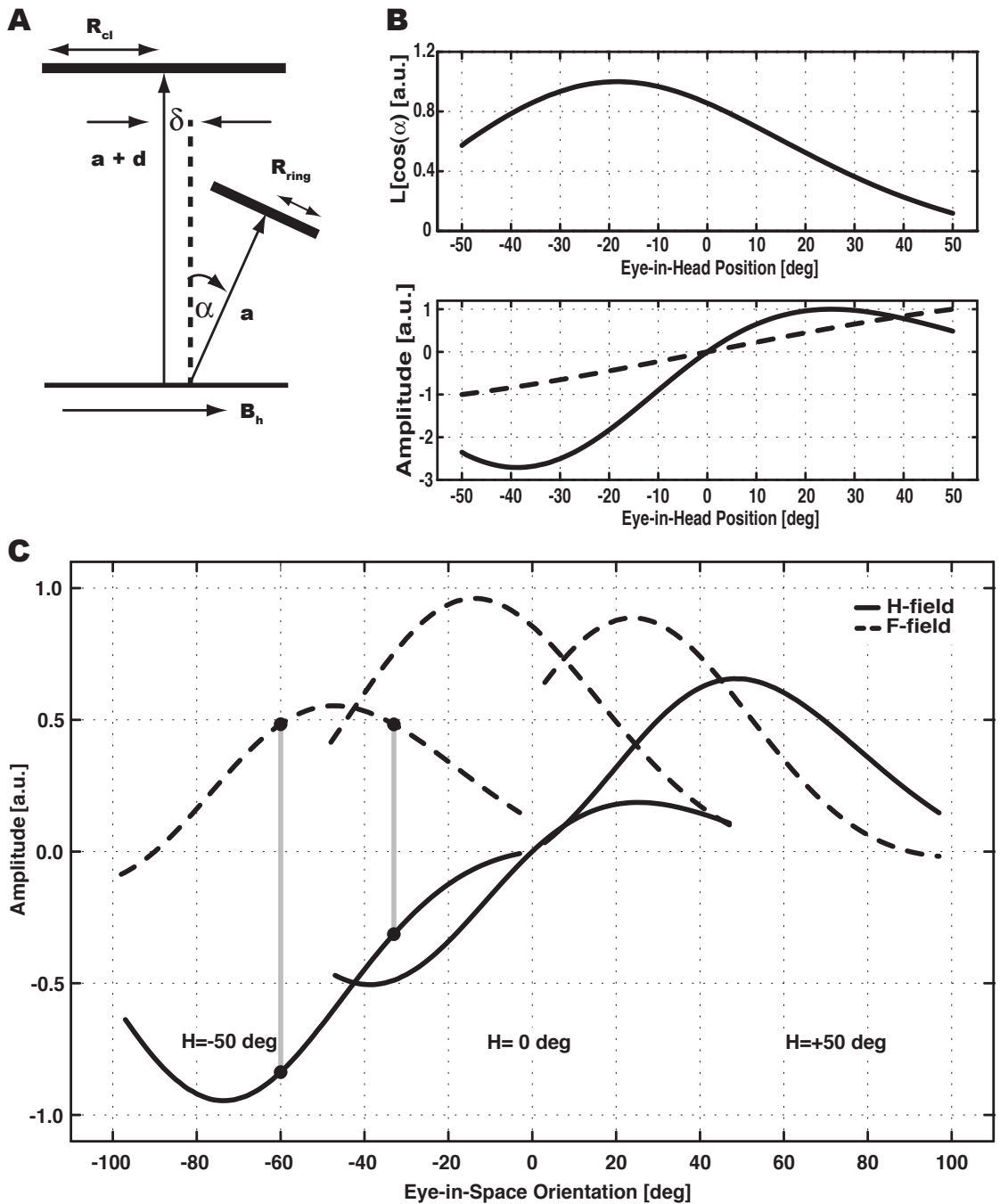
Finally, it is important to note, that Eqs. (6) and (7) assume that the primary induction voltage in the pickup coil is fully canceled by a compensation coil (see Fig 2 anti-coil). For this to work, however, the two coils need to be identical in their electrical properties, the compensation coil should not be influenced by the ring, should be exactly parallel to the pickup coil, and the magnetic field should be perfectly homogeneous. In practice, however, even though the coils may be identical, minor inhomogeneities of the field will be present, and the two coils will not be exactly aligned. Below, we will deal with this situation.

*Non-Idealized situation.* In case of imperfect cancellation, there will be an additional net primary-field voltage in the pickup coil's signal. Theoretically this voltage would be determined by the geometrical difference of the two otherwise identical coils and would depend only on the head orientation,  $\gamma$ . As noted in Eq. (8) for the horizontal field:

$$V_{net,h}(\gamma, t) = L_{cl,h} \cdot \sin(\omega_h t) \cdot [\sin(\gamma) - \sin(\gamma + \Delta\phi)] \quad (8)$$

with  $L_{cl,h}$  a constant, proportional to the coils' self-induction (here taken to be identical for the two coils), and  $\Delta\phi$ , the (small) difference in angle between the coils and the magnetic field lines (Fig. 2). In practice, however, the electrical and geometrical properties of the two coils will not be identical, rendering the dependency of the net voltage on head movement slightly more complicated (e.g. Fig 8 C, D). A quantitative description of the exact relation, however, is not necessary for the calibration of the signals. Since (1) it is possible to measure  $V_{net}$  before the experiment (Humans: Measurement without ring; Animals: Coils are usually fixed to an assembly that can be measured without the animal).  $V_{net}$  can then be subtracted from the experimental data offline. And (2) in the neuronal network calibration approach the network will not be hampered by a complex  $V_{net}$  function, since the unique relationship between the secondary induction voltages and gaze remains unchanged as is explained below.

Since perfect cancellation for all head orientations is not attainable in prac-



**Figure 4:** Simulation of a condition in which the axes of ring and pickup coil are not collinear, instead the ring's center is moved by  $\delta = 0.8$  cm toward the right relative to the pickup coil's center (**A**). This misalignment causes the geometrical nonlinearity to become asymmetric re.  $\alpha = 0^\circ$ . (**B, top**) The monotonic measurement range has shifted towards the left, at the expense of a smaller range to the right. (**B, bottom**) Nevertheless, together with the frontal field, the eye-in-head position can still be retrieved over the full oculomotor range (**C**, compare the points connected by the gray vertical lines).



Since perfect cancellation for all head orientations is not attainable in practice, the induction voltages measured by the pickup coil are described as

$$V_h(\alpha, \gamma, t) = K_h(t) \cdot \sin(\alpha + \gamma) \cdot L(\cos \alpha) + V_{net,h}(\gamma, t) \quad (9)$$

$$V_f(\alpha, \gamma, t) = K_f(t) \cdot \cos(\alpha + \gamma) \cdot L(\cos \alpha) + V_{net,f}(\gamma, t)$$

Thus, imperfect cancellation of the primary field contributions in the pickup coil will merely add an offset to the signals that varies as a function of the eye-head orientation, without affecting the shape of the signals.

It is also to be expected that the assumption of a perfectly aligned ring/pickup coil assembly is violated in practice. That is, the rotation axes of the ring and the pickup coil will often not coincide (Fig. 4A). For example their axes may be shifted with respect to each other ( $\delta$ , Fig. 4A). This shift will remain constant for all eye orientations (the ring on the eye rotates about angle  $\alpha$ , Fig. 4A). It is then important to realize that this will only affect the  $L[\cos(\alpha)]$  term in Eqs. (1), (6), (7) and (9), but not the induction current in the ring, and the primary induction voltages in the coils.

Yet, in general the analytical solution of this problem is not straight forward, and has to be done by numerical approximations of the Biot-Savart law. Our calculations show that the main effect, of such a misalignment is the introduction of an asymmetry in the shape factor  $L[\cos(\alpha)]$ , with respect to the straight-ahead eye-in-head fixation direction ( $\alpha = 0^\circ$ ). In other words, the peak of this function will shift leftward when the ring's axis is shifted to the right, or rightward when it is shifted to the left of the center of the pickup coil. As a result, the shape factor  $L[\cos(\alpha)]$  can no longer be described analytically by simple Legendre polynomials. Fig. 4B (top) shows the result of a numerical simulation of this effect when the ring was shifted rightward with respect to the coil's center by a full radius  $\delta = 0.8$  cm. As a result, the linear range of the pickup coil signal has shifted to the left and ambiguity in the signal already arises for relatively small positive angles of eye orientation (Fig. 4B, bottom). This can be seen in greater detail in Fig. 4C. Although an asymmetry of the signals as function of eye-in-head orientation is now apparent, frontal and horizontal field signals still form unique combinations as a function of the gaze angle over the full range (Fig. 4C, black dots connected by vertical gray lines). Thus, the combination of head orientation  $\gamma$ ,  $V_h$  and  $V_f$  constitutes a triplet of values that uniquely defines the eye orientation in the head.

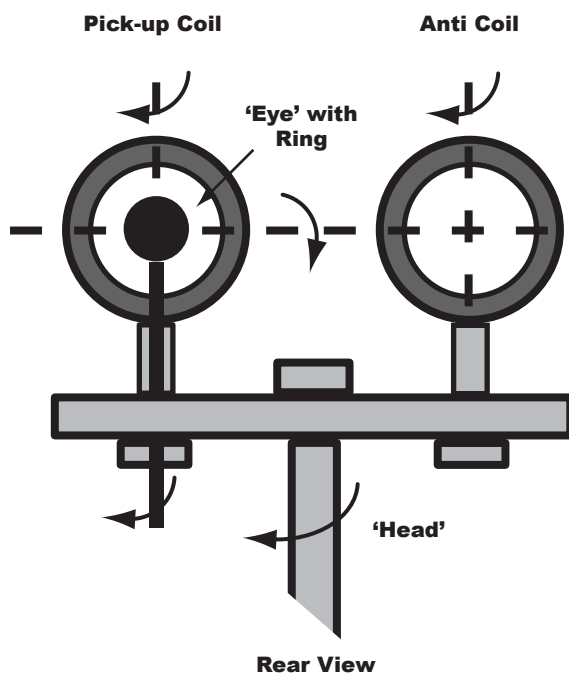
### *Practical application*

So far, we have extended the theoretical analysis of the DMI method by Bos et al. (1988) to head-unrestrained gaze shifts. In what follows, we describe an experimental test of the method on a gimbal system. We also propose a non-parametric calibration procedure that relies on a feed forward neural network to reconstruct the eye-in-head orientation,  $\alpha$ , from the signals measured by the pick-



up coil and anti-coil over a large  $\pm 120^\circ$  range of gaze shifts from the straight ahead direction.

*Test on a gimbal system.* A custom-made gimbal system was used to simulate combined eye/head movements in the horizontal plane. This system consisted of a horizontal plateau that could manually be rotated about a vertical axis to simulate horizontal head rotations over the full  $360^\circ$  range. Two coils were mounted on the plateau inside two Fick gimbals (e.g., Crawford et al., 1999) that allowed independent alignment of the coils both horizontally and vertically with respect to the magnetic fields, and to each other (Fig. 5). The anti-coil and pickup coil were connected in anti-phase, such as to produce cancellation of the primary magnetic fields. The whole system was positioned in the center of three orthogonal, square-wave pulsed magnetic fields consisting of three pairs of Helmholtz coils of  $2.5 \times 2.5$  m and a Rempel field-generating system (EM 5 Rempel Labs., Katy, TX, USA; frontal field: 48 kHz 25 mV rms at a gain of 50, horizontal field: 60 kHz 25 mV rms at a gain of 50). The pickup coil and anti-coil consisted of 50 turns of 0.1 mm insulated wire, with a diameter of 5.0 cm, and had a self-induction of 0.3 mH, and a d.c. resistance of 17 Ohm (Malpeli 1998). The diameter was chosen such that it would not occlude the field of vision of a human subject. For animal subjects the diameter may be slightly smaller. The ring was made of gold-plated copper, shaped such that it would fit around the cornea of a human eye (Bour et al., 1984). It had a central diameter of 1.3 cm, and a width of 4 mm. It could be placed behind the pickup coil on an artificial globe, at several distances and positions with respect to the coil's center, and could be manually rotated about the vertical axis to simulate horizontal eye rotations in the head.



**Figure 5:** Gimbal system used to test the head-unrestrained DMI method. The pickup coil with a ring on an artificial globe placed behind it (**left**) and the anti-coil (**right**) are each mounted inside a Fick gimbal on top of a horizontal plateau that can be rotated around a vertical axis. In this way the 'eye' and 'head' can be rotated independently in the horizontal plane, and the two coils can be optimally aligned with respect to the fields and each other. The distance of 1.8 cm of the ring relative to the pickup coil, as well as, the size (diameter: 5 cm) of the pickup coil itself was chosen to mimic the situation attainable in subjects. Notice also that the absolute position of the ring on the artificial globe determines  $\delta$  of Fig. 1.

To minimize the primary field component, the two coils were manually aligned. The system (without the ring) was rotated in the horizontal plane, and the anti-coil's orientation was adjusted to produce a minimal peak amplitude. Due to minor inhomogeneities in the fields, and to small differences between the two coils, a full cancellation across the entire measurement range was not achieved. Typically, the peak amplitude of the difference signal (about 2.0 V at a gain of 300) was about twice the peak-to-peak amplitude of the secondary induction voltage generated by the ring.

Once the two coils were aligned, the ring was positioned in the assembly, at a fixed distance of 1.8 cm behind the center of the pickup coil. The offsets of the anti-coil and pickup/anti-coil differences from the horizontal field were set to zero for the condition in which the eye and the head were both facing the straight-ahead direction. Since the signal of the frontal field follows a cosine behavior (Eq. (7)), this field was zeroed at  $-90^\circ$  relative to the horizontal field direction. Then, the following set of measurements was performed: for each head position, the eye was rotated between  $-45^\circ$  and  $+45^\circ$ , in  $5^\circ$  steps. The Gimbal system was rotated between  $\pm 180^\circ$  in  $10^\circ$  steps; thus resulting in a total of  $19 \times 36 = 684$  measurements.

The system was designed such that a large number of variable settings was possible. Here we only concentrate on a setting that would typically resemble a recording from human or animal subjects. That is, a distance of 1.8 cm between ring and pickup coil and a radius of 2.5 cm for the pickup coil. In the case of human subjects a precisely fitting bite-board could be employed on which both pickup coil and anti-coil can be mounted. In animal subjects both coils can be mounted on implanted headposts. Note however, that the anti-coil does not have to be placed in the immediate vicinity of the pickup coil, as long as the magnetic field is sufficiently homogeneous.

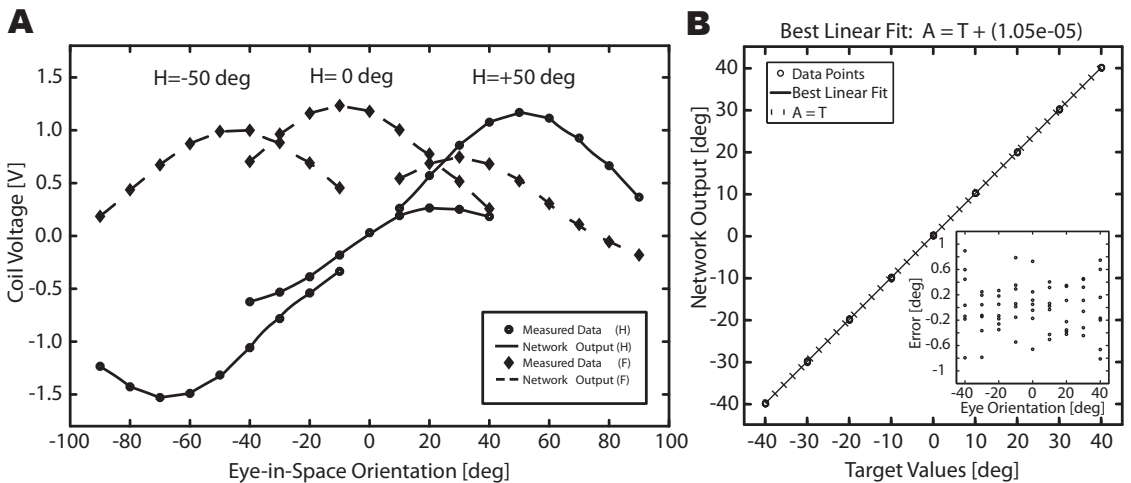
*Calibration.* The recorded coil difference voltages,  $V_h^D$ , and  $V_f^D$ , were calibrated by applying a three-layer feed forward neural network. The three input units of the network were (1)  $V_h^D$ , (2)  $V_f^D$ , and (3)  $\gamma$  (as set on the gimbal system), while the desired, single output of the network was the eye-in-head angle,  $\alpha$ . The network, which contained one layer of 8 hidden units, was trained using the Bayesian-Regularization implementation of the back-propagation algorithm (Matlab 7.0, Neural Networks Toolbox, The Mathworks, Inc.) in order to avoid over-fitting (MacKay, 1992). The network was trained on a restricted subset of the sampled data: Only 63 of the 684 recorded data points were used. These included head orientations from  $-90^\circ$  to  $+90^\circ$  in steps of  $30^\circ$  and eye orientations of  $-40^\circ$  to  $+40^\circ$  in steps of  $10^\circ$ . This set was designed to mimic a data set as would be typically obtained during an actual calibration session with a human subject. After training, the network's performance was tested on a set of interpolated values (interpolation factor 40), of which the training data were a subset.

## 2.3 Results

*Calibration of the secondary induction signals.* The recorded signals could be calibrated reliably. Fig. 6A shows the results of gimbal eye movement measurements and calibration for the same three head orientations as shown in Figs. 3 and 4. To achieve a better comparison with those two figures, Vnet was subtracted. Note, however, that this subtraction is not necessary to achieve a good calibration. Markers denote data points (circles: Horizontal field, diamonds: Frontal field) and lines (solid: Horizontal field, dashed: Frontal field) the network's output to the interpolated data set.

Note, that compared to the theoretical data of Figure 3, the measured curves of Figure 6 are shifted leftward; creating the predicted asymmetry as shown in Figure 4. Thus, one can conclude that the ring relative to the pickup coil was shifted rightwards during the recording. In practice it is not possible to align ring and coil perfectly. There will always be a slight misalignment ( $\delta$ , Fig 4A) between the center of rotation of the pickup coil and that of the ring. But this does not pose a problem for the calibration procedure, since the neuronal network will not be hampered by a misalignment. These data represent actual DMI measurements not unlike the ones that would be obtained from a human subject "wearing" a stable bite-board.

The performance of the network was assessed by a linear regression of the network's output values as function of the target values (Fig. 6B). It is apparent that the network is very well capable of fitting the data. As can be seen in the

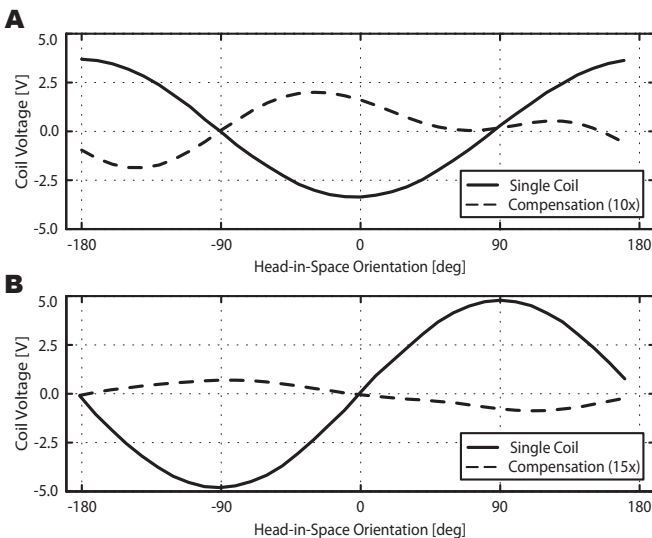


**Figure 6:** (A) Measured voltages in the pickup coil for three different head-in-space orientations ( $-50^\circ$ ,  $0^\circ$ ,  $+50^\circ$ ). The net primary signal has been subtracted. Eye-in-head positions range from  $-40^\circ$  to  $+40^\circ$  in steps of  $10^\circ$  for each head orientation. Circles denote signals from the horizontal field and diamonds from the frontal field. The calibrated network output is plotted as solid lines for the horizontal field and dashed lines for the frontal field, respectively. The network was trained on data limited to a head range of  $-90^\circ$  to  $+90^\circ$  in steps of  $30^\circ$  and an eye range from  $-40^\circ$  to  $+40^\circ$  in steps of  $10^\circ$  ( $\Sigma$  63). (B) Network output versus target values for seven different head positions. Data is fitted with a linear function ( $r^2 = 0.998$ ,  $N = 63$ ). Inset: Difference between network output and target values for each eye orientation at seven different head orientations.

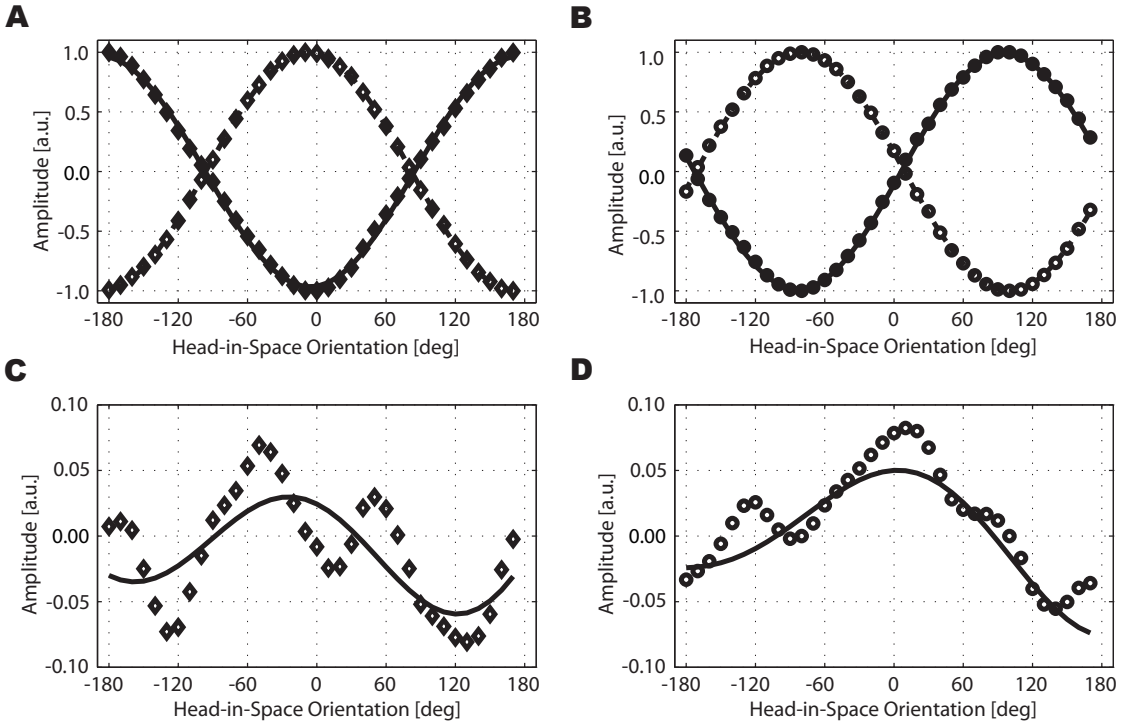
inset, the difference between network and target value, i.e. the error, varies within  $\pm 1^\circ$  for different head and eye orientations. The largest errors per eye orientation occur at the most extreme head orientations ( $\pm 90^\circ$ ) and eye orientations ( $\pm 40^\circ$ ). In conclusion, the network yielded an acceptable fit even with such a restricted training set and a limited number of hidden units ( $N=8$ ). This will be important for the application of this calibration method in subjects. In this case the dataset used for calibration has to be limited to a minimum, since time-consuming calibration measurements will considerably reduce experimental time. In addition, calibration of the entire dataset is possible demonstrating that the network can generalize the problem (data not shown).

*Influence of  $V_{net}$ .* We recorded the pickup/anti-coil difference signals –the difference signals for short– without the ring for the horizontal and frontal magnetic fields, as well as the primary field signals from the pickup coil and the anti-coil alone. Note, here, that the secondary voltage induced by the ring (about 1 V) is tenfold smaller than the primary field component (approximately 8–10 V peak-to-peak). Therefore, the gain of the system needs to be increased in order to obtain a sufficiently large eye orientation signal. In Fig. 7, the primary field component for the frontal (A) and the horizontal (B) field is plotted as a function of head orientation. The single coil signal (solid line) reaches a peak-to-peak amplitude of about 8 V whereas the difference signal (dashed line) reaches a peak-to-peak amplitude of about 0.5 V for the frontal and 0.2 V for the horizontal channel. The gain for the difference signal was thus set tenfold higher than for the anti-coil. This amplification was also used to record the gimbale eye orientations. It is apparent that in this case  $V_{net}$  is not negligible. Thus, cancellation of the primary field is not perfect for this setup. Nevertheless, manual alignment of pickup coil and anti-coil is sufficient to ensure a secondary-to-primary component ratio such that the ring signal can be calibrated. Recording hardware saturation is avoided as well.

Note, that the resulting difference coil signals are not readily predictable. We fitted sinusoids to the pickup coil and the anti-coil signals for both the



**Figure 7:** Measured primary field component for a single coil (solid line) and the pickup/anti-coil compensation assembly (dashed line) for the frontal field (A) and the horizontal field (B). The gain for the single coil is 30 and 20 for the horizontal and the frontal field, respectively, and the gain for the pickup/anti-coil compensation assembly is 300 for both fields.



**Figure 8:** Measured primary field component for the pickup coil (black symbols) and the anti-coil (open symbols) for the frontal field (**A**) and the horizontal field (**B**). The data are fitted with sinusoidal functions. (**A**) pickup coil:  $SSE = 0.0025$  and  $r^2 = 0.99$  ( $N = 36$ ), anti-coil  $SSE = 0.0102$  and  $r^2 = 0.99$  ( $N = 36$ ) (**B**) pickup coil  $SSE = 0.0307$  and  $r^2 = 0.99$  ( $N = 36$ ), anti-coil  $SSE = 0.0025$  and  $r^2 = 0.99$  ( $N = 36$ ). (**C**) and (**D**) Mathematical subtraction (markers) of the primary field component for the frontal (**C**) and horizontal (**D**) fields. The data are fitted with the difference of the two sinusoidal fits of the corresponding pickup coil and anti-coil.

horizontal and frontal fields. The results of these fits are shown in Fig. 8A and B. Only small deviations from the data can be observed. To assess if the compensation signals for both fields can be described by the difference between two sinusoids (see Eq. (8)), we subtracted the independent signals from each other and compared them with the difference of the original two sinusoidal fits. As can be seen in Fig. 8C and D, the compensation signal cannot simply be described as the difference of two sinusoids.

## 2.4 Discussion

We have described an extension of the DMI method to record head-unrestrained gaze shifts. By including an additional perpendicular magnetic field, the inherent limited measurement range of the method (maximally  $30^\circ$  from straight ahead) could be overcome. At the same time, the measurement range could be extended to the full eye-head motor range. In particular, we have shown that by using a simple, three-layer feed forward neural network with only a few hidden units, in

combination with knowledge about the head orientation in space, the small ring signal can be extracted from the difference coil signal. This is even possible, when the coils are not exactly aligned, or when small inhomogeneities are present in the fields that induce a complex head-position dependent offset. The neural network parameters can be determined within a minute after the calibration experiment, and can be readily used to calibrate eye movements in real time. For this to work, it is not important whether the primary field component is present in the coils or not. The obtained calibration precision with the neural network (8 hidden units) is  $\pm 1^\circ$  over the entire measurement range, which is comparable to results obtained with the head-free SSC technique (Goossens and Van Opstal, 1997).

The data presented in Fig. 6 yielded an angular resolution of about  $0.3^\circ$ , which is worse than the value obtainable with the SSC technique and with the head-restrained DMI method (down to  $0.1^\circ$ , or better). For these latter two methods, the resolution is only limited by the recording system's noise. One should bear in mind, however, that the optimal resolution of the (head-restrained) human oculomotor system is about  $0.5^\circ$ .

The following factors influence the resolution of the head-unrestrained DMI technique: (1) the peak-to-peak variation of the Vnet signal, as function of head rotation, (2) the system's inherent noise, (3) the peak-to-peak variation of the ring-induced secondary signal, and (4) mechanical vibrations of the coil assembly. It is advisable to reduce the Vnet signal to a minimum, so that an amplification of the ring signal is possible without saturating the recording hardware at more extreme head positions. As a result, the resolution of the system increases, since the ring signal gets bigger while the noise stays the same.

Mechanical vibrations of the assembly will introduce additional low-frequency noise to the measurements that further reduce the system's resolution. These vibrations could be minimized e.g. by having human subjects wear a bite-board on which the two coils are rigidly attached. Preliminary 2D experiments conducted with human subjects indicate a resolution of about  $0.6^\circ$  (recorded with a EM 7 eye movement monitor, Remmel Labs, manuscript in preparation). With experimental animals, these vibrations could be virtually eliminated by attaching the coils to head posts that are fixed to the skull.

Note here, that in experiments with head-unrestrained subjects the orientation of the head will not be known a priori and will have to be measured and calibrated. For example a head mounted coil could be used to measure the head movements independently. To calibrate these signals, a simple minimizing mse fitting routine could be employed (Goossens and Van Opstal, 1997) using

$$\begin{aligned} V_h^A(\gamma) &= A_h \cdot \sin(\gamma + \Phi_h) \\ V_f^A(\gamma) &= A_f \cdot \cos(\gamma + \Phi_f) \end{aligned} \quad (10)$$

with  $A_h$ ,  $A_f$ ,  $\Phi_h$ , and  $\Phi_f$  parameters. These parameters can be easily constrained.



For perpendicular fields the offset angles,  $\Phi_h$  and  $\Phi_f$  will be the same; the same holds true for  $A_h$  and  $A_f$  by tuning the amplifiers of the magnetic fields. This leaves only two free parameters,  $A$  and  $\Phi$ , for rotations in the horizontal plane. Once these are known, the absolute head-in-space orientation,  $\gamma$ , can be readily extracted from the primary induction voltages in the anti-coil over the full  $360^\circ$  range.

A critical requirement of the extended method is a rigid fixation of the pickup coil and anti-coil to the subject's head. If not fixated rigidly, vibrations of the assembly will introduce additional low-frequency noise to the measurements. In humans, the use of a precisely fitting bite-board to which the two coils can be rigidly attached, is thus essential.

#### Bour technique

In experimental animals rigid fixation can be achieved more easily by attaching both coils to the skull with bone screws and dental cement.

Our extended method is thus a promising alternative for the search coil technique. The latter one is expensive (coils may break already after a few recording sessions), while the lead wires often irritate the eye, restricting experimental time to about 40 min. The DMI method could therefore also better be suited for use on patients (provided they can use a bite-board), especially when head movements are to be incorporated (vestibular studies, blinking studies, etc.). Note also, that wire breakage is especially a problem in experimental animals. The method can also be employed for binocular recordings: as the ring's signal strength falls off rapidly with distance, the danger of interference of a ring's signal in the pickup coil of the other eye is negligible (e.g. Chaturvedi and Van Gisbergen, 2000).

In conclusion, the theory and data presented here demonstrate the feasibility of the extended DMI method for recordings in head-unrestrained subjects. The next step will be its full application in two-dimensional gaze control studies with human and non-human subjects.

## Acknowledgements

We thank Stijn Martens and Hans Kleijnen for valuable technical assistance. Jeremiah Wanga is thanked for his simulations of the Biot-Savart law. This research project has been supported by a Marie Curie Early Stage Research Training Fellowship of the European Community's Sixth Framework Program under the contract number MEST-CT-2004-007825 (PB), the Neurocognition program of the Netherlands Organization for Scientific Research (NWO, project number. 051.04.022 RFVDW), and the Radboud University Nijmegen (AJVO)

## 2.5 References

Allik J, Rauk M, Luuk A. Control and sense of eye movement behind closed eyelids. (1981) *Perception* 10:39–51.

Bos JE, Reulen JPH, Boersma HJ, Ditters BJ. Theory of double magnetic induction (DMI) for measuring eye movements: Correction for nonlinearity and simple calibration in two dimensions. (1988) *IEEE Trans Biomed Eng* 35:733–9.

Bour LJ, Van Gisbergen JAM, Bruijns J, Ottes FP. The double magnetic induction method for measuring eye movement—results in monkey and man. (1984) *IEEE Trans Biomed Eng* 31:419–27.

Bour LJ, Aramideh M, de Visser BW. Neurophysiological aspects of eye and eyelid movements during blinking in humans. (2000) *J Neurophysiol* 54:866–71.

Bour LJ, Van Rootselaar AF, Koelman JH, Tijssen MA. Oculomotor abnormalities in myoclonic tremor: a comparison with spinocerebellar ataxia type 6. (2008) *Brain* 131:2295–303.

Chaturvedi V, Van Gisbergen JA. Stimulation in the rostral pole of monkey superior colliculus: effects on vergence eye movements. (2000) *Exp Brain Res* 132(1):72–8.

Crawford JD, Ceylan MZ, Klier EM, Guitton D. Three-Dimensional eye-head coordination during gaze saccades in the primate. (1999) *J Neurophysiol* 81:1760–82.

Collewijn H, Van der Mark M. F, Jansen TC. Precise recordings of human eye movements. (1975) *Vision Res* 15:447–50.

Frens MA, Van der Geest JN. Scleral search coils influence saccade dynamics. *J. Neurophysiol.*, 2002; 88: 692–8.

Van der Geest JN, Frens MA. Recording eye movements with video-oculography and scleral search coils: A direct comparison of two methods. *J. Neurosci. Methods*, 2001; 114: 185–95.

Goossens HHL, Van Opstal AJ. Human eye-head coordination in two dimensions under different sensorimotor conditions. *Exp. Brain. Res.*, 1997; 114: 542–60.



Houben MM, Goumans J, Van der Steen J. Recording three-dimensional eye movements: scleral search coils versus video oculography. (2006) *IOVS* 47:179–87.

Judge SJ, Richmond BJ, Chu FC. Implantation of magnetic search coils for measurement of eye position: An improved method. (1979) *Vision Res* 20:535–8.

MackKay DJC. Bayesian interpolation. (1992) *Neural Comput* 4:415–47.

Malpeli JG. Measuring eye position with the double magnetic induction method. (1998) *J Neurosci Methods* 86:55–61.

Rommel RS. An inexpensive eye movement monitor using the scleral search coil technique. (1984) *IEEE Trans Biomed Eng* 31:388–90.

Reulen JPH, Bakker L. The measurement of eye movement using double magnetic induction. (1982) *IEEE Trans Biomed Eng* 29:740–4.

Robinson DA. A Method of measuring eye movement using a scleral search coil in a magnetic field. (1963) *IEEE Trans Biomed Eng* 10:137–45.

Smeets JBJ and Hooge ITC. Nature of variability in saccades. (2003) *J Neurophysiol* 90:12–20.



## Chapter 3

# Using Double-Magnetic Induction to Measure Head-Unrestrained Gaze Shifts – II Application in Human Subjects –

### Abstract

This study compares the performance of a newly developed gaze (eye-in-space) measurement technique based on double magnetic induction (DMI) by a custom-made gold-plated copper ring on the eye with the classical scleral search coil (SSC) technique to record two-dimensional (2D) head-unrestrained gaze shifts. We tested both systems simultaneously during head-free saccades towards light emitting diodes (LEDs) within the entire oculomotor range ( $\pm 35$  deg). The absence of irritating lead wires in the case of the DMI method leads to a higher guarantee of success (no coil breakage), and to less irritation on the subject's eye, which results in a longer and more comfortable measurement time. Correlations between DMI and SSC signals for horizontal and vertical eye position, velocity and acceleration were close to 1.0. The difference between the SSC signal and the DMI signal remains within a few degrees. In our current setup the resolution was about 0.3 deg for the DMI method, vs. 0.2 deg for the SSC technique. The DMI method is an especially good alternative in the case of patient and laboratory animal gaze control studies where breakage of the SSC lead wires is particularly cumbersome.

### 3.1 Introduction

At present, two invasive electromagnetic methods for monitoring eye movements are being used in oculomotor research: the scleral search coil (SSC) technique (Robinson 1963; Collewijn et al. 1975), and the double magnetic induction (DMI) method (Allik et al. 1981; Reulen and Bakker 1982; Bour et al. 1984; Bos et al. 1988; Malpelli 1998). Temporal and spatial resolutions of both methods are comparable when recording head-restrained eye movements. However, while the SSC technique can be readily used to measure head-unrestrained two dimensional (2D) gaze shifts, the DMI method so far has been restricted to head-fixed experimental settings only. Two problems inherent to the DMI method make its use somewhat tedious. To understand these it is necessary to briefly review the theoretical and practical background of DMI.

The DMI method makes use of a gold-plated short-circuited copper ring that is placed on the subject's eye instead of a silicon annulus with an embedded copper coil, as is the case for the SSC technique. Since the ring is not connected to the recording apparatus, a coil that picks up the tiny secondary magnetic field generated by the primary currents in the ring (hence double magnetic induction) has to be placed in front of the eye.

Note that the pickup coil also records the much stronger signal of the primary magnetic fields generated by the field coils, which are not related to the eye-in-head orientation. An anti-coil that is connected anti-parallel to the pickup coil should therefore cancel these components. In practice, however, it is not possible to perfectly cancel the primary field contributions, resulting in dc-components that depend in a complicated way on the head orientation and on small electromagnetic differences between pickup coil and anti-coil. This dc, together with the eye signal's non-monotonic nonlinearity, have been considered a serious problem for the application of DMI to head-unrestrained situations.

In a recent paper (Bremen et al. 2007) we described the theoretical background needed to extend the classical DMI method to head-unrestrained conditions. For example the signal induced by the horizontal magnetic field, as recorded by the pickup coil, is described by the following equation:

$$V_h(\alpha, \gamma, t) = K(t) \cdot \sin(\alpha + \gamma) \cdot L(\cos \alpha) + V_{net}(\gamma, t) \quad (1)$$

with  $V_h$  the voltage recorded by the pickup coil,  $\alpha$  the horizontal eye-orientation in the head,  $\gamma$  the horizontal head-orientation in space and  $t$  time.  $K(t)$  is a coefficient that includes the characteristics of the time varying primary magnetic field, the ring geometry and the electromagnetic properties of the pickup coil.  $L[\cos(\alpha)]$  is a non-monotonic geometrical shape factor that is only determined by the ring/pickup coil geometry. This shape factor distinguishes the DMI method from the SSC technique, for which  $L[\cos(\alpha)] = 1$  (Bos et al. 1988).  $V_{net}$  describes the head dependent dc that is due to imperfect cancellation of the primary field component (Bremen et al. 2007). Note that the shape factor only

depends on eye-in-head orientation, while Vnet solely depends on head-orientation in space.

As a consequence of the complicated three-dimensional (3D) geometrical relationship between ring and pickup coil, the effective measurement range is limited to about 30 deg, since the non-linear relation between gaze orientation and signal strength is non-monotonic. The linear range, being only about  $\pm 10$  deg, is even more restricted (Reulen and Bakker 1982; Bos et al. 1988). Additionally, the exact ring-pickup coil geometry strongly influences the measurement range and the signal/ring-orientation relation. If the ring is positioned slightly off-center relative to the pickup coil the shape factor changes too, thereby shifting the non-monotonic part of the signal. This creates a pronounced signal asymmetry, which may further reduce the effective measurement range (Bremen et al. 2007).

For these reasons the SSC technique is still widely used and is considered the gold standard. However, the vulnerable lead wires needed to connect the SSC to the recording hardware have proven to pose a considerable risk – as they are prone to breakage –, especially in experimental animals and patients. Since the DMI method uses a gold-plated copper ring without connecting wires this risk is absent. Therefore, should it be possible to extend the DMI method to head-unrestrained conditions, it could serve as a valuable alternative to the widely used SSC technique.

In order to use DMI in head-unrestrained subjects, the measurement range should not be restricted by the non-monotonic, and hence ambiguous, dependence of the signal on gaze angle  $\alpha + \gamma$  (Eq. 1). In our previous paper we demonstrated that introducing an additional, frontal, magnetic field might suffice to resolve the signal ambiguity. In addition, a model-free artificial neuronal network could solve the relatively complex, and in principle unknown signal asymmetries that are introduced by imperfections in the ring-pickup coil geometry and the magnetic fields, and can account for the head-orientation dependent offsets (Bremen et al. 2007).

Yet, the system has so far not been tested on real subjects performing in 2D gaze-orienting tasks. A serious problem of DMI is the strong dependence of the shape factor  $L[\cos(\alpha)]$  on the ring-pickup coil geometry. Any shift of either of the two coils relative to the induction ring will strongly influence the signal. In addition, mechanical vibrations of the coils relative to the head (and hence the ring) will introduce unwanted low-frequency noise to the signals, which strongly affect the system's resolution and signal-to-noise ratio. Preventing these problems necessitates a rigid and vibration-free pickup coil/anti-coil assembly, which so far has only been achieved in head-restrained subjects (Bour et al. 1984; Malpeli 1998). The present study demonstrates that all these potential problems can be overcome with relatively simple means. We show experimental data obtained with a low-cost custom-made pickup coil/anti-coil assembly applied to human subjects who made rapid 2D head-unrestrained gaze shifts. We also present a simple routine to calibrate the system. We tested our method over a range of  $\pm 35$  deg in all directions, and compared the calibrated signals to simultaneous recordings

obtained from the classical SSC technique that was applied to the other eye. Our results show that, despite a slightly lower resolution (0.3 deg vs. 0.2 deg), the extended DMI method is a good and low-cost alternative to the search-coil technique for measuring 2D gaze shifts, in principle over the entire motor range.

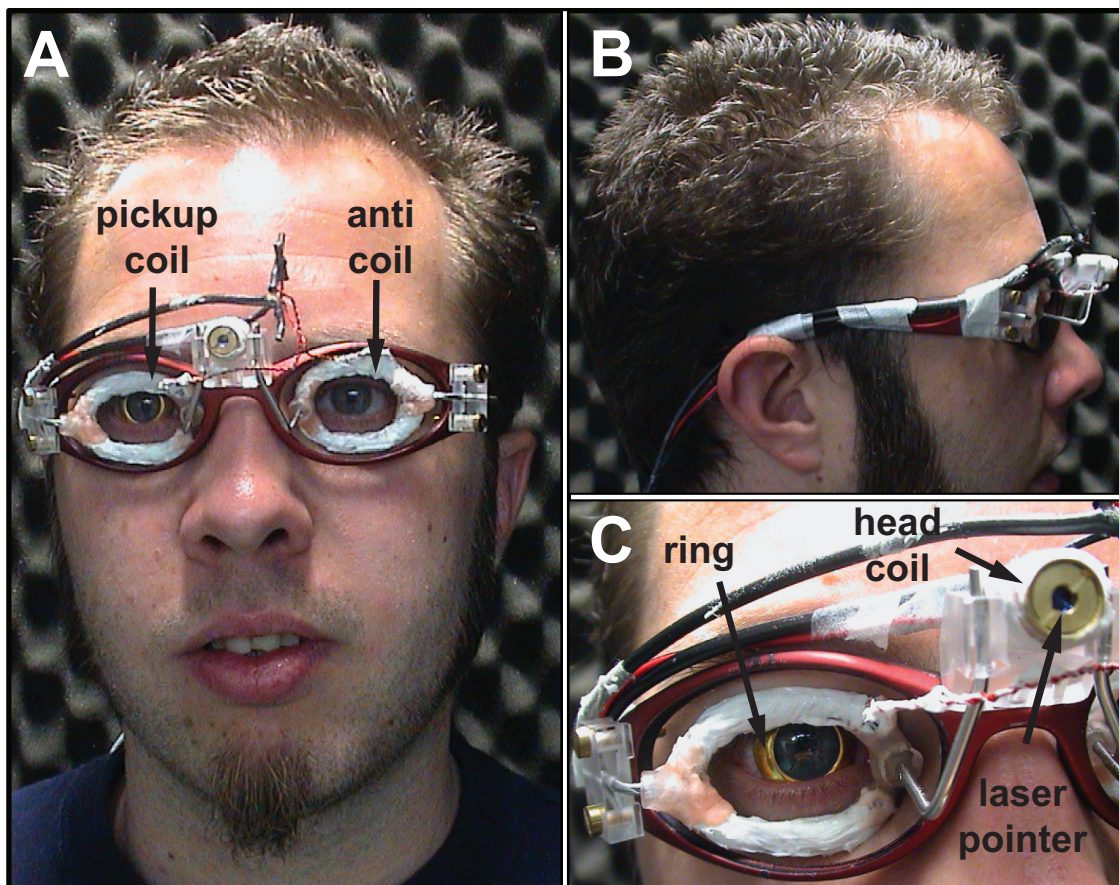
## 3.2 Methods

*Subjects.* Three male subjects participated in these experiments (JG, PB and RW). Ages ranged from 27 to 40 years. In all three subjects simultaneous recordings with the SSC technique and DMI method were performed. To familiarize the subjects with the paradigms, a short practice session was performed prior to the first experimental session. Subjects PB and RW are authors of this paper. Subjects JG, PB and RW had corrected to normal vision (JG: -3.5D right, -3.5D left; PB: -3.5D right, -3.0D left; RW: -2.8D right, -2.0D left) but did not wear their glasses during the experiments. All subjects were experienced in oculomotor studies and participated on a voluntary basis. Experiments were performed in accordance with the Declaration of Helsinki.

*Apparatus.* Experiments were conducted in a completely dark room (3 x 3 x 3 m<sup>3</sup>). Three orthogonal pairs of single turn field coils (surface area 6 mm<sup>2</sup>) used to generate the primary magnetic field were mounted alongside the edges of the four walls, ceiling and floor. The coils were driven by a Remmel field-generating system (Remmel 1984, EM 7 Remmel Labs, Katy, TX, USA; horizontal field: 48 kHz, vertical field: 60 kHz, frontal field: 80 kHz). To monitor eye movements and to have a direct comparison of the DMI method with the SSC technique, the subjects wore a gold-plated copper ring (custom made, thickness 0.3 mm, inner diameter 11.3 mm, top angle 108 deg, weight 0.46 g; see also Bour et al. 1984, their fig. 4) on the right eye, and a SSC (Scalar Instruments, Delft, The Netherlands) on their left eye. Prior to the insertion of the ring and the coil both eyes were anesthetized with two drops of oxybuprocain (0.4%, in HCl (pH 4.0), Thea Pharma S.A., Wetteren, Belgium). Both ring and coil were manually inserted without the aid of a suction device. To that end, the subject was requested to look downward while the upper eyelid was lifted, so that the ring/coil could be gently inserted. Subsequently, the ring/coil was positioned in the frontal plane around the cornea, with the aid of a wetted cotton swab, if needed (Fig. 1A&C). The coil's lead wires were secured to the nose and forehead with tape. After the experiments the ring/coil were removed with a blunt forceps.

The ring's secondary magnetic field was registered with a pickup coil placed in front of the eye. The pickup coil and the anti-coil (both: surface area of the wire 0.13 mm<sup>2</sup>, 35 turns,  $\varnothing$  of the coil 2.5 cm, L = 120  $\mu$ H, DC resistance 1.8  $\Omega$ , connected in series) were mounted on a lightweight plastic glasses frame (total weight: 148 g; hereafter referred to as the DMI assembly) that was worn by the subjects (Fig. 1A-C). To reduce noise that appeared to be caused by skin contact, the coils were further insulated with a thin layer of silicon. The DMI





**Figure 1.** Subject wearing coil assembly and ring. **(A)** Frontal view. Note that the coils are elliptical and isolated from the skin with white silicon rubber. The elliptical shape optimized the subject's horizontal field of view, and allowed putting the pickup coil closely in front of the ring. The signal strength could thus be optimized. A small laser pointer was mounted on the nose bridge. A coil to measure head orientation is wound around it (not visible). The ring on the subject's right eye is made from thin gold plated copper and fits around the subject's cornea. The ring is not embedded in silicon, as is the case with the SSC. **(B)** Side view. **(C)** Close-up of the pickup coil, ring and laser pointer.

assembly was constructed such that both coils remained in the same plane and thus had the same orientation relative to the primary magnetic fields. A rubber strap around the subject's head connected the temples of the glasses frame to further tighten the assembly. The DMI assembly was sufficiently rigid so that it did not slip or distort when the subject made rapid head-unrestrained gaze shifts. The distance between pickup coil and eye varied from subject to subject but ranged between 1.5 cm and 2.0 cm. In some subjects play dough (Play-Doh®, Hasbro) was molded around the nose bridge to roughly center the pickup coil in front of the ring, to optimize the field of view ( $\pm 30$ - $35$  deg in both directions) and to achieve a comfortable fit of the glasses frame. Horizontal and vertical head movements were measured with the search-coil technique. For that purpose a small custom-made coil was wound around a miniature laser device (LQB-1-650, World Star Tech, Toronto Ont., Canada, see section Stimuli) that was mounted on the nose

bridge of the glasses frame (Fig. 1A,C).

The head coil (horizontal & vertical channel) and gaze signals (DMI & SSC: horizontal, vertical & frontal channel) were amplified (total of eight channels), demodulated and low-pass filtered at 150 Hz (4th order Butterworth, custom built) before being stored on hard disk at a sampling rate of 500 Hz/channel for subsequent off-line analysis.

*Stimuli.* Subjects were seated comfortably in the center of the experimental chamber facing a spherical wooden frame (radius: 85 cm) with 12 spokes and five concentric rings (LED sky; described in detail in other studies from our lab, e.g. Hofman and Van Opstal 1998). A polar coordinate system with its origin at the straight-ahead position was defined to describe target positions. Target eccentricity,  $R$ , is measured as the angle with respect to the straight-ahead position, whereas target direction,  $\varphi$ , is measured relative to the horizontal meridian. LEDs (0.2 deg diameter as viewed by the subject, intensity 0.15 cd·m<sup>-2</sup>) were mounted at  $R = \{0, 2, 5, 9, 14, 20, 27, 35\}$  deg (rings) and  $\varphi = \{0, 30, 60, \dots, 330\}$  deg (spokes) on the frame. For example, the position of the center LED is described by  $(R, \varphi) = (0, 0)$  deg. A target in the upper right quadrant can be described by  $(R, \varphi) = (14, 60)$  deg and one in the lower left quadrant as  $(R, \varphi) = (9, 240)$  deg. For offline analysis these polar coordinates were transformed into their corresponding azimuth ( $\alpha$ ) and elevation ( $\varepsilon$ ) angles (Hofman and Van Opstal 1998) by:

$$\alpha = \arcsin(\sin R \cdot \cos \Phi) \quad \varepsilon = \arcsin(\sin R \cdot \sin \Phi) \quad (2)$$

respectively. In addition to these LEDs the small laser device mounted on the nose bridge of the glasses frame was used to align the subject's head to the target (see section Paradigms).

### Paradigms

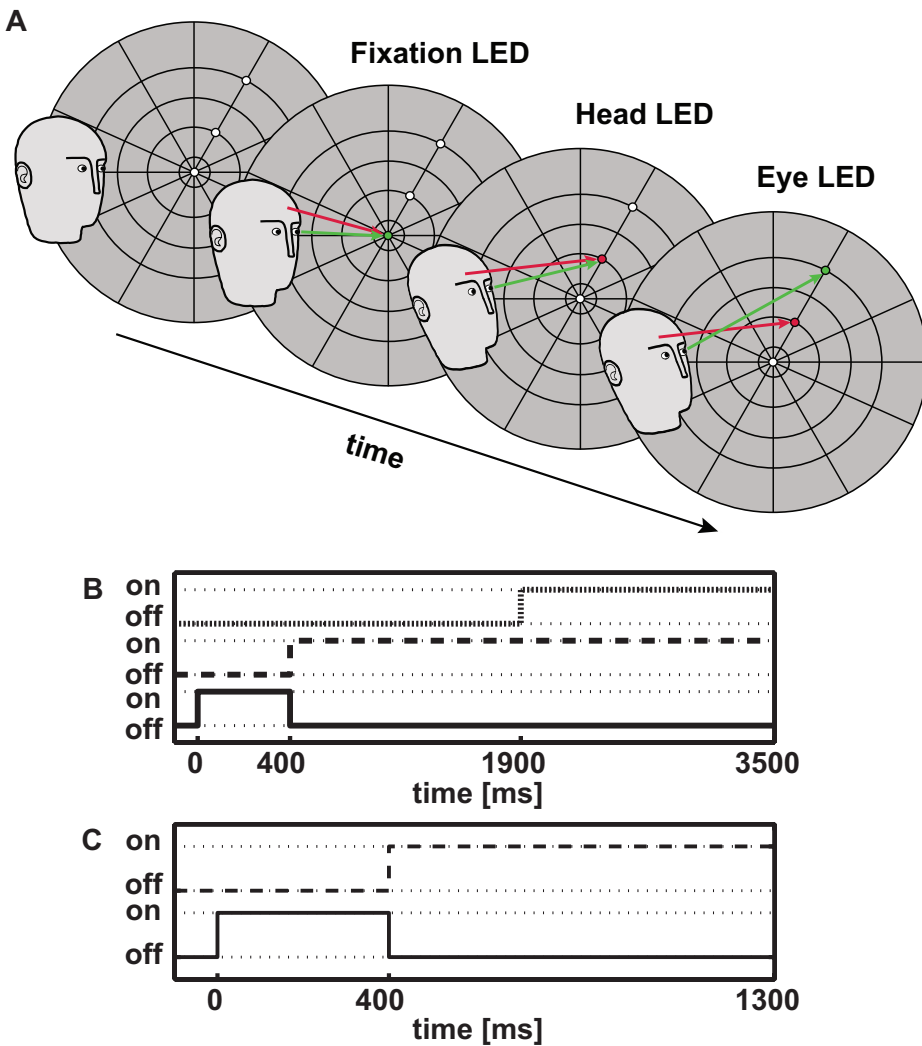
All experiments were performed in darkness.

*Control experiment.* To ensure that there was no interference of the SSC on the DMI assembly and vice versa two control experiments were performed with subject JG (SSC and DMI assembly without ring; DMI assembly and ring without SSC) that showed that the respective signal was not disturbed (data not shown).

*Characterization of  $V_{net}$ .* We conducted measurements with subjects wearing the assembly without ring and SSC to characterize the head-orientation dependent remnant primary field component (referred to as  $V_{net}$  in Eq. 1) picked up by the DMI assembly and to assess whether it remains stable during the course of an experimental session. In addition, we recorded the signal from the head coil to have a standard with which to compare the  $V_{net}$  signal with. The subjects had to align the head fixed laser pointer to targets presented in pseudo-random order on the LED sky (see below). Each target location was presented four times to test the stability and reproducibility of the assembly.



*Extended DMI method.* In the actual recording experiments we conducted measurements while subjects wore the DMI assembly, ring and SSC. Two paradigms were used: (1) a closed-loop calibration experiment (Fig. 2A, B), and (2) a test experiment (Fig. 2C). Each trial in the calibration paradigm started with a green fixation LED at  $(R, \varphi) = (2, 0)$  deg, which was presented for 800 ms. The subject had to align the head-fixed pointer to that fixation LED. This procedure assured that the head was at the same starting position for every trial. The fixation LED was then first followed by the so-called head target (red LED) that remained lit until the end of the trial. The subject had to align the head-fixed pointer, and thus his head, to this target LED and keep it there for the remainder of the trial. After 1500 ms a second target LED (green, the so-called eye target) was lit for 1600 ms. The subject had to make an eye saccade to this new target *without moving the head*.



**Figure 2.** Paradigms used for calibration and testing. **(A)** Cartoon of one trial of the calibration paradigm. **(B)** Time sequence of the calibration paradigm. **(C)** Time sequence of the test paradigm. F = Fixation LED, T = Target LED, TH = Head target LED, TE = Eye target LED.

A series of computer simulations in one dimension had been performed (Bremen et al. 2007) to determine the minimal number of different head-eye combinations needed for a successful calibration. Typically, only five eye positions along the same direction  $\varphi$  as the head (two at the extreme opposite ends of the oculomotor range, and one at the center) for each given head position sufficed. This assured that the majority of points were sampled at the non-linear and non-monotonic part of the ring's eye-in-head signal (see Fig. 7). Each target pair was repeated three times.

In the test paradigm the offset of the central fixation LED was immediately followed by the appearance of a peripheral target that remained lit for 900 ms. The subject was instructed to make a natural orienting gaze shift with both the head and eye towards this target, as quickly and as accurately as possible. To elicit large gaze shifts the head-fixed pointer remained on during the test paradigm. For all experiments, the order of stimulus conditions and positions was randomized throughout a session. One recording session lasted for 30 to 40 minutes of which about 10 min were used for calibration. Note that the maximum recording time was restricted by the SSC. The lead wires typically become increasingly irritating during the course of an experiment eliciting disturbing blinks. Since the ring is less irritating due to the lack of lead wires it can be worn for up to 60 min (Bour et al. 1984).

*Calibration.* Prior to off-line calibration, the data were digitally low-pass filtered at 75 Hz using a 50-point FIR-filter. Fixation endpoints were determined off-line. A window with duration of 100 ms around saccade offset determined the average endpoint values of the eye- and head position signals. The window was adjusted after visual inspection to correct for errors. Endpoints were averaged (Vnet: four repeats per target location; DMI/SSC: three repeats per target combination) after removing outliers.

For both the Vnet and the DMI/SSC measurements, the head coil signal was calibrated using two three-layer feed forward neuronal networks for the horizontal and vertical component of the head movement. The input to the networks consisted of the horizontal and vertical voltages as recorded with the head coil. Input units were connected to 4 hidden units that projected to the output unit of the networks. The networks were either trained on the whole dataset (Vnet) or on the set of data points collected during the head target epoch of the calibration paradigm (DMI/SSC). The networks' output was the horizontal or vertical head orientations in space,  $H_\alpha$  and  $H_\varepsilon$ , respectively.

Networks were trained by the Bayesian-Regularization implementation of the back-propagation algorithm (Matlab 7.0, Neural Networks Toolbox, The Mathworks, Inc.) in order to avoid over-fitting (MacKay 1992).

*Characterization of Vnet.* Although  $V_{net}$  is only a head-dependent dc we mapped it on the target positions in space. This allowed us to better assess the stability and reproducibility of the DMI assembly and to compare the amplitude

of Vnet with the calibrated head coil signal.

For this, the pickup coil signals were calibrated with two three-layer feed-forward neuronal networks for the horizontal and the vertical component, respectively. Three input units (1) voltage of the horizontal field  $V_h$ , (2) voltage of the vertical field  $V_v$ , (3) voltage of the frontal field  $V_f$  were connected to four hidden units that projected to one output unit yielding horizontal ( $H_\alpha$ ) and vertical ( $H_\epsilon$ ) head orientation, respectively. The networks were trained on a total of 67 averaged endpoints with the teacher value being the target location in space. Subsequently the thus trained networks were used to calibrate the raw endpoints for all target repetitions.

*Extended DMI method.* DMI and SSC signals were both calibrated with three-layer feed-forward neuronal networks. Note that for the calibration of the ring signal it is not necessary to subtract the complicated Vnet contribution. The networks used for calibration automatically take  $V_{net}$  into account.

Horizontal and vertical components of saccades were calibrated independently. The four input units of the networks were (1) voltage of the horizontal field  $V_h$ , (2) vertical voltage of the vertical field  $V_v$ , (3) voltage of the frontal field  $V_f$ , and (4) either  $H_\alpha$ , or  $H_\epsilon$  (calibrated with the above mentioned neuronal networks), while the desired, single output of the network was the gaze angle,  $G_\alpha$  or  $G_\epsilon$ . The networks, which contained one layer of 4 hidden units, were trained on a total of 109 head-eye combinations (27 head-eye combinations x 4 axes + mean fixation LED value) collected during the eye target epoch of the calibration paradigm. Networks were trained within one minute. After training, the networks were used to calibrate the test data.

### Data Analysis

To assess how well the networks approximated the data, linear fits of target position versus network output were performed. If not stated otherwise all linear fits were made employing the least-squares criterion.

*Characterization of  $V_{net}$ .* The four calibrated target repeats were used to assess whether the DMI assembly remained stable during the recording session. The mean endpoint for a given target location was calculated and the four corresponding repetitions were subtracted from this mean to obtain an estimate of the variability. These procedures were performed for the DMI assembly and the head coil signal.

*Extended DMI method.* We performed a spectral analysis of all trials recorded during the test experiment for each subject. For this, we applied a 512-point fast fourier transform (FFT) to the horizontal and vertical components of both the DMI method and the SSC technique. Subsequently the spectra were averaged for each subject. This analysis would show for example, whether the DMI assembly introduces characteristic vibrations to the ring signal.

To provide an estimate of the resolution of the extended DMI method and the SSC technique we calculated the standard deviation over 100 calibrated

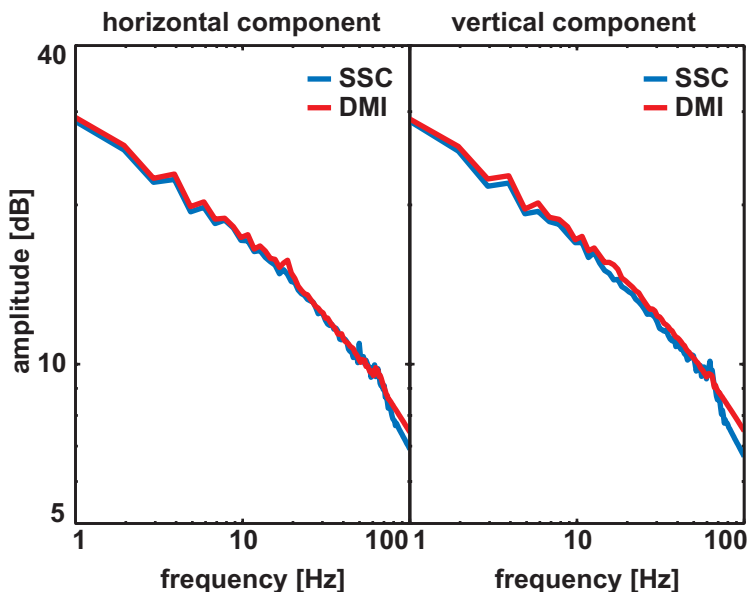
samples taken from the fixation epoch and all target positions of the test experiment. This was done independently for the horizontal and vertical components and for each subject. Subsequently, the values obtained for each subject were averaged. We argue that the standard deviation provides a measure of the resolution since it is the root mean square (RMS) deviation from the arithmetic mean. Thus, it provides information about the magnitude of the noise inherent to the system. Since the noise level determines the lower detection limit, a signal must be bigger than this limit to be detected.

Additionally, we calculated the track velocity and track acceleration for all saccades and subjects by applying a two-point difference differentiation algorithm on the radial gaze position data.

### 3.3 Results

We performed a power spectrum analysis of the horizontal and vertical gaze-position data of the DMI and SSC signals of each subject (see Methods) and for all trials. The spectra did not show any peculiarities on the horizontal and vertical signal components for neither the DMI method, nor the SSC technique nor between the horizontal and vertical components of the signals. In figure 3 the average spectra across all trials of the test experiment from subject RW are shown. Horizontal and vertical components are depicted in the left- and right-hand column, respectively. The DMI signal is drawn in red and the SSC signal in blue. All signals are in good agreement. These data ascertain that the DMI assembly did not introduce spurious mechanical vibrations to the ring signal.

Table 1 lists the resolutions obtained in our current setup for the DMI method and the SSC technique for all subjects. The values are in good agreement with the resolution of the DMI method ( $\sim 0.3$  deg) being slightly inferior to the SSC technique ( $\sim 0.2$  deg). Similar resolution values were obtained for fixations



**Figure 3.** Average power spectra for the horizontal (**left**) and vertical component of the DMI method (**red**) and the SSC technique (**blue**). Spectra for all trials of the test experiment from subject RW were calculated and subsequently averaged.

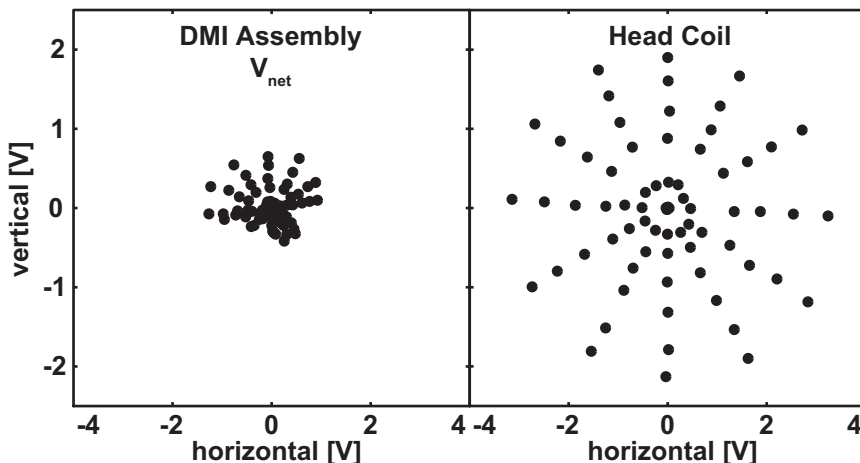
**Table 1.** Resolution in degree calculated as the standard deviation over 100 samples taken from the fixation epoch of the test experiment for all subjects.

Subject Method & Component	JG	RW	PB	Mean all subjects
DMI horizontal	0.12	0.37	0.37	0.29
SSC horizontal	0.10	0.22	0.15	0.16
DMI vertical	0.23	0.29	0.49	0.34
SSC vertical	0.28	0.24	0.17	0.23

across the whole tested range. Only extreme gaze fixations, i.e. beyond  $R = 30$  deg, tended to yield somewhat lower resolutions (between 0.3 and 0.6 deg) for both the DMI method and the SSC technique. Yet, the DMI resolution is of the same order, or better, as the optimal resolution of the oculomotor system, which is limited by the fovea to about 0.6 deg.

In what follows, we will first present the results of the Vnet characterization before we will describe the results of the DMI/SSC measurements.

*Characterization of Vnet.* Figure 4 depicts the mean raw endpoints of both the DMI assembly (left) and the head coil (right) for subject PB. The head coil endpoints clearly mirror the stimulus array, indicating that the coil's signal varies almost linearly with target position for the range tested. The net DMI assembly signal, in contrast, is highly asymmetric. Note that an ideal assembly would yield a  $V_{net}$  of zero for all head positions. In practice the pattern is more complicated. Here, the vertical signal is strongly compressed in the lower hemisphere. This is



**Figure 4.** Uncalibrated endpoints for the DMI assembly ( $V_{net}$ ) (left column) and the head coil (right column) for subject PB. Note that an ideal DMI assembly would yield a  $V_{net}$  of zero for all head orientations. Here, the vertical endpoints for the DMI assembly are strongly compressed in the lower hemisphere.

a feature of the current assembly. Different pickup coil/anti-coil geometries will produce different  $V_{net}$  shapes, and it is apparent that it would be difficult to describe  $V_{net}$  analytically. The compensation of the anti-coil is far from perfect, as the remnant signal's amplitude reaches about 30% of the head coil signal's maximum amplitude. This remnant signal (maximum value of  $V_{net}$ ) is, however, small enough to allow for a sufficiently large ring signal (e.g. figs. 7 & 8).

The calibration networks were nevertheless able to linearize the data and thus could account for the head-induced dc (fig. 6) as well as the complex changes in  $V_{net}$  (fig. 4). Table 2 lists the best-fit parameters of the linear regressions performed on network output versus target position for both the pickup coil and the head coil. The coefficient of determination,  $r^2$ , exceeded .99 for horizontal and vertical components in both cases. Correspondingly, the rmse values were small. The values obtained for the pickup coil were slightly higher than for the head coil. This holds especially for the vertical component with its pronounced asymmetry. However, all gains were close to one and the offsets close to zero deg also indicating that the networks were able to fit the data well.

In figure 5 the variation around the mean endpoint of the four fixations of a given target is plotted as a function of horizontal (left) and vertical (right) target position, for both the DMI assembly (red circles) and the head coil (blue diamonds). The DMI- and head-coil distributions overlap considerably. In addition, it is clear that the endpoint variability does not vary in a systematic way with target eccentricity. The standard deviations are smaller than 2 deg for both methods and components. The distributions for the DMI assembly and the head coil do not significantly differ from each other (student's t-test,  $p < 0.01$ ). From these data we conclude that the DMI assembly remained stable during a recording session.

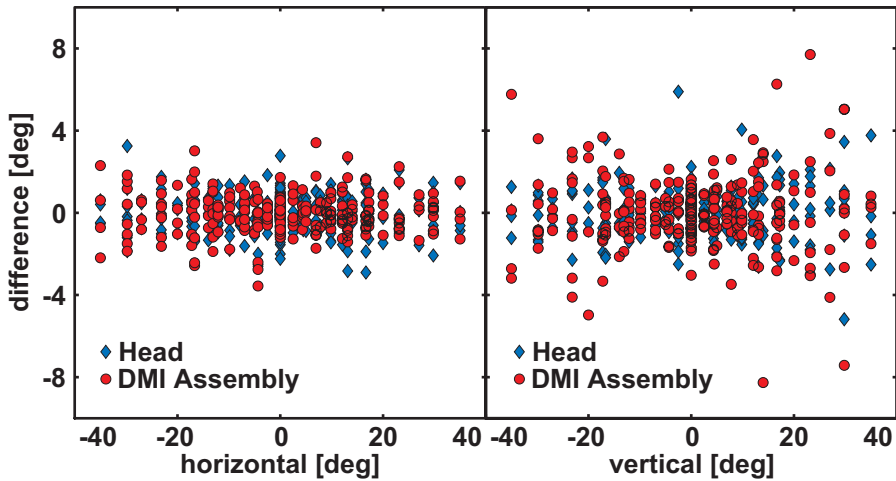
### *DMI/SSC Measurement.*

*Uncalibrated Data.* Uncalibrated ring signals for the endpoint fixations on the horizontal meridian, as measured during the calibration paradigm are shown

**Table 2.** Coefficients of the fits between the neuronal network output and the teacher value for horizontal and vertical component of the DMI assembly and the head coil for subject PB.

Slope (95% confidence intervals)	Intercept (95% confidence intervals)	Adj. $r^2$	rmse	N	Technique/Component
0.997 (0.987 1.007)	-0.033 (-0.182 0.116)	0.99	0.61	67	DMI assembly horizontal
0.995 (0.977 1.015)	0.027 (-0.258 0.312)	0.99	1.17	67	DMI assembly vertical
1.000 (0.993 1.007)	0.002 (-0.101 0.104)	0.99	0.42	67	Head horizontal
0.997 (0.991 1.004)	-0.027 (-0.130 0.073)	0.99	0.41	67	Head vertical

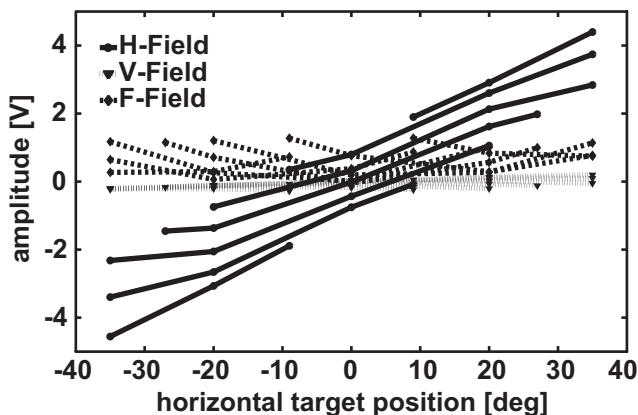




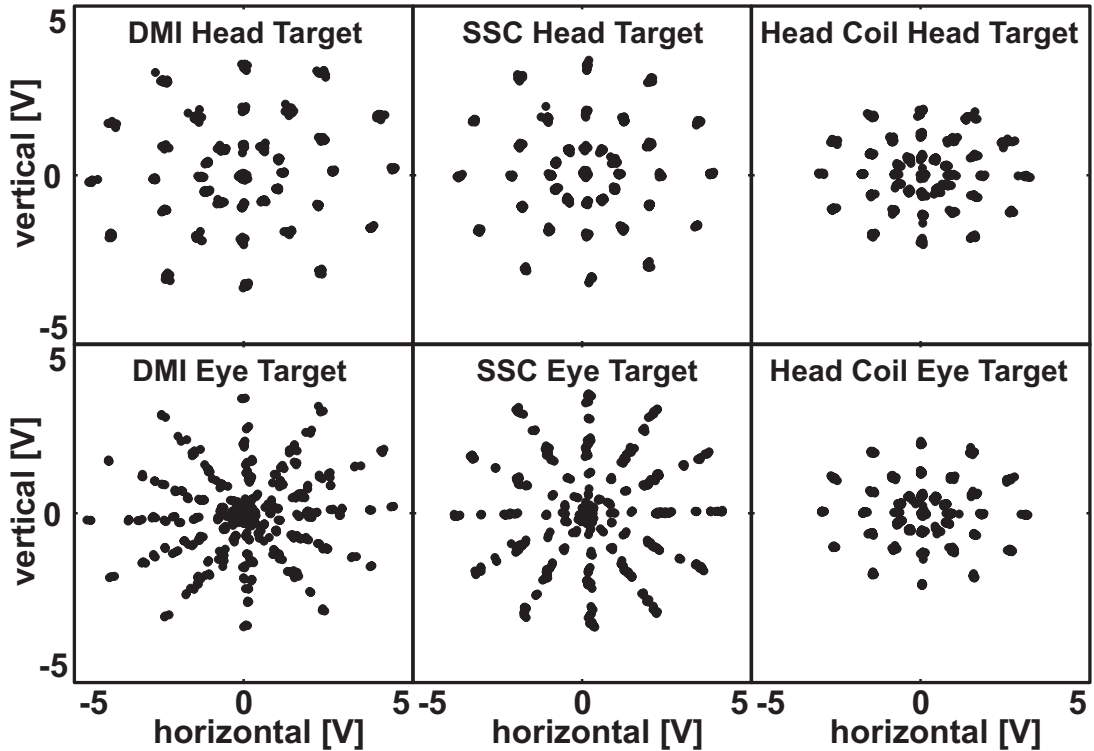
**Figure 5.** Difference of calibrated mean endpoint of a given target and four repetitions made to the same target as a function of horizontal (**left**) and vertical (**right**) target position for Vnet of the DMI assembly (**red circles**) and the head coil (**blue diamonds**). Subject PB.

in figure 6 (subject PB). These signals were in good agreement with the theoretical predictions described in our previous paper (Bremen et al., 2007). Note that as each line represents one head orientation, a head orientation-dependent offset can be seen, which is due to two factors: (1) a head orientation related component in the ring signal (the  $\sin(\alpha+\gamma)$  term in Eq. 1), and (2) a dc resulting from non-perfect cancellation of the primary field signals ( $V_{\text{net}}$ ). In addition, it is apparent that the combination of horizontal, vertical and frontal field signals together with the known head orientation uniquely defines the 2D gaze orientation.

Figure 7 depicts the uncalibrated raw endpoints of subject PB recorded during the calibration paradigm. DMI signals are shown in the left-hand column, SSC signals are shown in the center column and head-coil signals, used for both methods, in the right-hand column. Recall that in the calibration paradigm a head target is followed by an eye target. Endpoints for the former are shown in the top row and those for the latter in the bottom row. Note that the signals for all three techniques are somewhat asymmetric, as the vertical range is smaller than



**Figure 6.** Raw ring signals for the horizontal (H-field, solid lines, black dots), vertical (V-field, dotted lines, black triangles) and frontal (F-field, dashed lines, black diamonds) fields as function of horizontal gaze position for head orientations  $R = \{\pm 35 \text{ deg}, \pm 20 \text{ deg}, \pm 9 \text{ deg}, 0 \text{ deg}\}$  and  $\phi = \{0 \text{ deg}, 180 \text{ deg}\}$ . Subject PB.

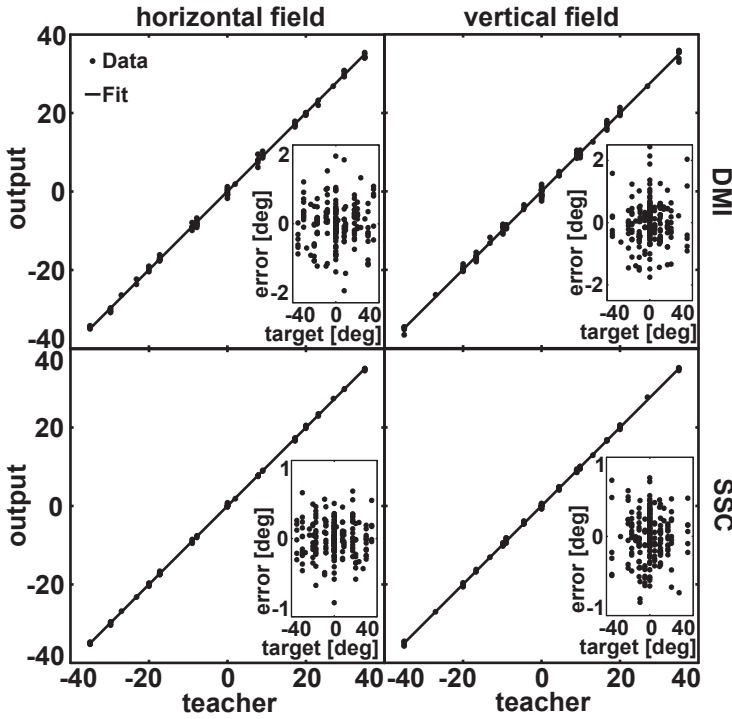


**Figure 7.** Raw endpoints for the DMI method (**left column**), the SSC technique (**middle column**) and the head coil (**right column**) for subject PB. Upper row: Endpoints extracted during presentation of the head target. Lower Row: Endpoints extracted during presentation of the eye target.

the horizontal range. This is due to the Remmel recording hardware, for which field-driving frequency and signal strength are inversely related. i.e. the lower the frequency the larger the signal. The asymmetry between horizontal (field driven at 48 kHz) and vertical (field driven at 60 kHz) components is most prominent for the head coil signal, for which the vertical amplitude reaches about 65% of the horizontal amplitude. The asymmetry is less pronounced for the DMI signal (74%) and least pronounced for the SSC signal (86%). However, the head-target endpoints for both the DMI method (left) and the SSC technique (center) are quite similar. While it is possible to recognize the spatial arrangement of the eye targets by looking at the raw SSC endpoints (center bottom), this is less obvious for the DMI endpoints. Here, the more complex geometric dependence of the DMI method (Eq. 1) becomes apparent. Nevertheless, as described below, it is possible to calibrate the more complex ring signal, such that the differences between the DMI method and the SSC technique disappear. Note also that the head-coil signal during the head-target epoch (right top) and the eye-target epoch (right bottom) are virtually identical, indicating that the subject indeed held his head stable during the two epochs.

*Calibration.* Figure 8 shows linear regressions that compare the goodness of fit for the network calibrations of the DMI method and the SSC technique for





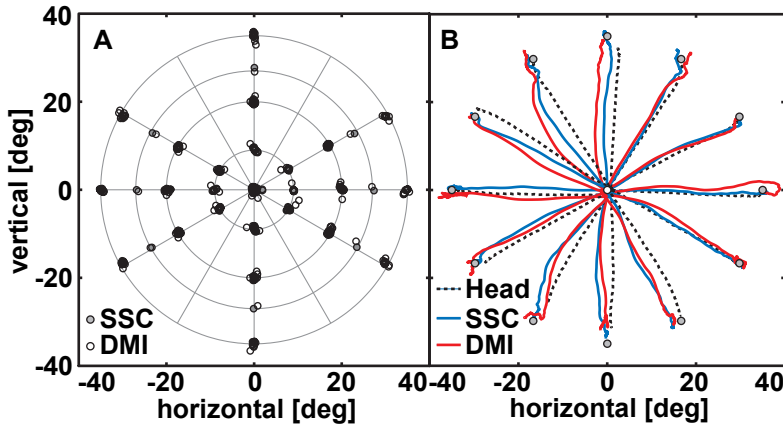
**Figure 8.** Linear regressions of the calibrated neuronal network output and the teacher values for the horizontal (left column) and vertical (right column) component of the DMI method (upper row) and the SSC technique (lower row) for subject PB. The insets in each panel depict the corresponding errors, i.e. teacher - output.

the horizontal and vertical components for subject PB. The corresponding coefficients are listed in table 3. Insets depict the network errors, i.e. teacher value minus network output value. The linear regressions for the two systems were in good agreement, indicating that the networks were well able to fit the DMI data. However, the network errors for the DMI method were slightly larger than for the SSC technique. Standard deviations  $\sigma$  for the DMI method were 0.72 deg for the horizontal and 0.62 deg for the vertical component. For the SSC technique the standard deviations were 0.30 deg and 0.32 deg for horizontal and vertical components, respectively.

A direct comparison of the calibrated endpoints for the DMI method and the SSC technique can be seen in figure 9A. Open circles denote DMI endpoints and

**Table 3.** Coefficients of the fits between the neuronal network output and the teacher value for horizontal and vertical component of the DMI method and the SSC technique for subject PB.

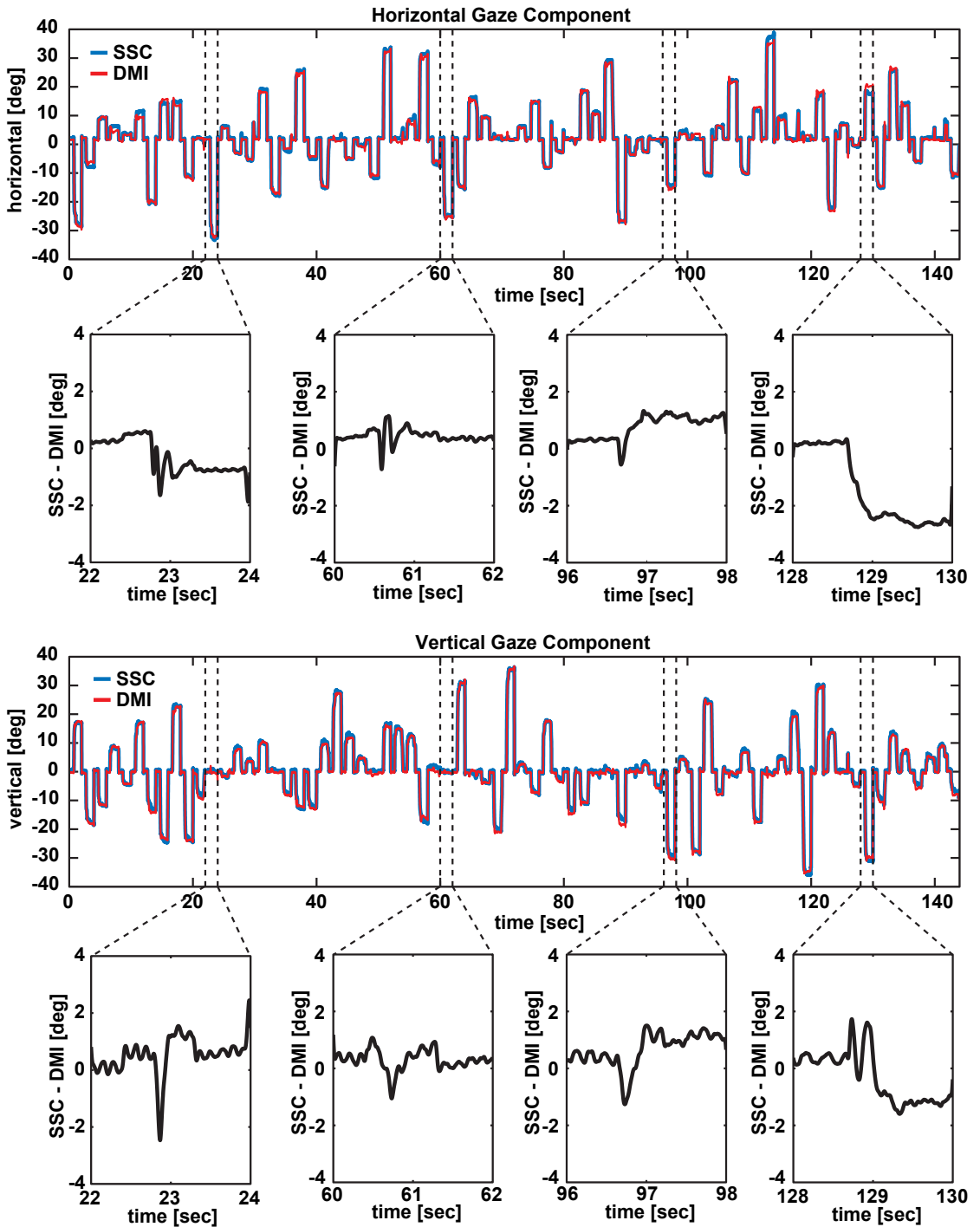
Slope (95% confidence intervals)	Intercept (95% confidence intervals)	Adj. $r^2$	rmse	N	Technique/Component
0.999 (0.995 1.003)	-0.006 (-0.068 0.056)	0.99	0.46	109	DMI horizontal
0.998 (0.992 1.005)	0.017 (-0.062 0.097)	0.99	0.59	109	DMI vertical
0.999 (0.998 1.002)	-0.002 (-0.031 0.028)	0.99	0.22	109	SSC horizontal
0.999 (0.996 1.002)	-0.007 (-0.047 0.034)	0.99	0.30	109	SSC vertical



**Figure 9.** (A) Calibrated fixation points for both the DMI method (**open circles**) and the SSC technique (**grey circles**) for subject PB. The light grey rings and spokes denote the LED sky. On each intersection of the rings and spokes a stimulus LED is installed. (B) Calibrated spatial traces for the DMI method (**red lines**), the SSC technique (**blue lines**) and the head (**dashed line**) for subject PB. The fixation point is indicated by the white circle and target positions by grey circles. Note small differences between SSC and DMI gaze trajectories. Note also that the SSC and the DMI ring were worn on different eyes.

grey circles SSC endpoints. The intersections of the light grey rings and spokes correspond to the targets on the LED sky. Note that the multiple endpoints at one LED location result from different gaze shifts to the same spatial location, not from repetitions of the same head-eye combination. The mean difference between the SSC and DMI endpoints for both the horizontal and the vertical component is 0.00 deg and the standard deviation is 0.68 deg for the horizontal component and 0.67 deg for the vertical component. Thus, despite the pronounced differences for the raw eye target endpoints of the DMI and SSC signals, as depicted in figure 7, the network is able to calibrate the data.

*Saccades during the test paradigm.* After training, the networks were tested on an independent data set obtained during the test paradigm, in which the subject made head-unrestrained gaze shifts towards peripheral targets. Figure 9B shows spatial trajectories for subject PB to the most eccentric circle of the LED sky ( $R=35$  deg). Red lines indicate DMI data, blue lines SSC gaze data, and black dashed lines are the head-movement trajectories. Grey symbols denote target positions, whereas the white central symbol is the fixation LED. Note that the ring and the SSC are worn on different eyes. Thus, it is expected that the eye-movement trajectories will be similar, but not necessarily identical. Indeed, small differences may be observed in the gaze trajectories of Fig. 9B. It is important to note that these differences were not systematic, and were not reproduced between trials. For example, toward the end of the rightward horizontal gaze shift the DMI and SSC signals seem to diverge. This feature was not due to a calibration artifact, of either recording method, as it did not occur on other trials. Rather, it might be due to small differences in binocular gaze control, and differences in the amplitude of blinks by the left and right eye.

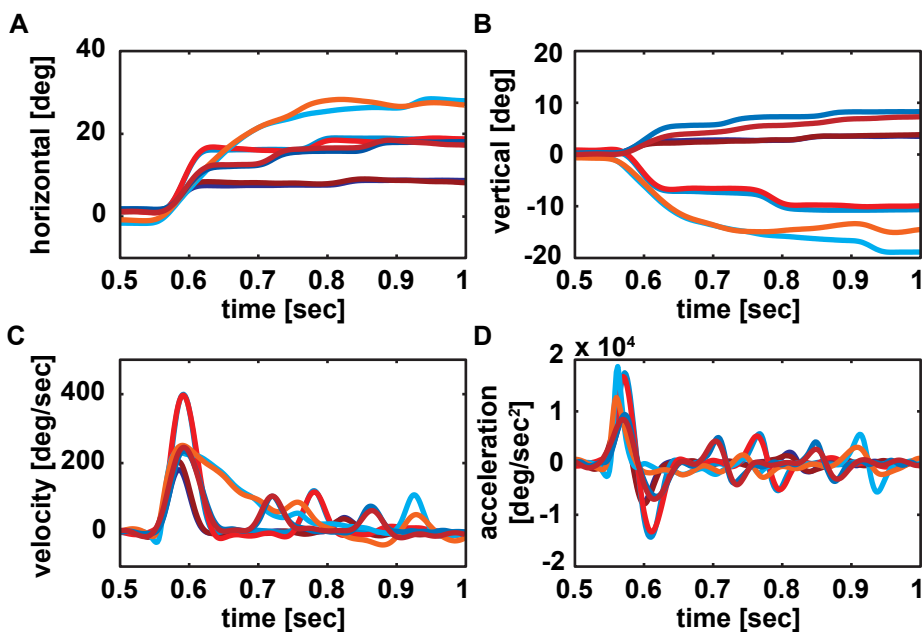


**Figure 10.** Calibrated temporal traces of the test paradigm for the DMI method (red lines) and the SSC technique (thick blue lines) for subject JG. **Upper panel:** Horizontal gaze component. **Lower panel:** Vertical gaze component. Some of the largest differences between the two methods are highlighted in the subplots for four trials. Differences were not systematic and typically small (less than 1 deg).

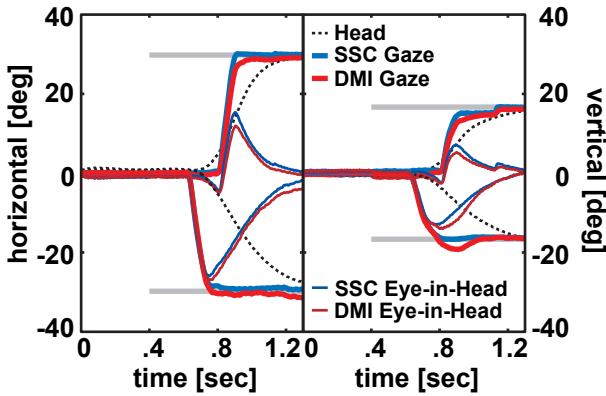
Figure 10 depicts in an analogous manner the temporal traces for subject JG. All 72 trials of the head-unrestrained test paradigm are shown as one continuous trace for both the horizontal and vertical gaze component. The insets highlight instants with the largest differences between SSC and DMI signals. We found no systematic relation between these differences and target location. The mean difference for the horizontal component is 0.30 deg with a standard deviation of 0.64 deg and 0.34 deg with a standard deviation of 0.72 deg for the vertical component.

Temporal gaze component, velocity and acceleration plots of subject RW to four selected targets are shown in figure 11. The DMI signals are plotted as red lines, the SSC signals as blue lines. Different hues denote the different trials. The traces have been temporally aligned such that the gaze saccade onsets coincide. Note that the DMI and SSC signals of the individual trials are in good agreement, even for the acceleration profiles. Again, since ring and SSC were worn on different eyes small differences in the traces are most likely due to differences in binocular oculomotor control and not due to the calibration of either signal. The head-free DMI method could thus be used for detailed trajectory and gaze kinematics analysis as has been done in head-restrained paradigms (for example Van Opstal and Van Gisbergen 1987).

In figure 12 the gaze and reconstructed eye-in-head traces of the DMI method (red lines) and the SSC technique (blue lines) for two target locations (grey lines) are shown. The head trace is depicted as a black dashed line (data from subject PB). The signals from both methods are very similar. Note the small



**Figure 11.** Temporal, velocity and acceleration plots for the DMI method (red lines) and the SSC technique (blue lines) of four trials (different line hues) from subject RW. Traces are plotted from stimulus onset to stimulus offset and are temporally shifted so that saccade onsets are aligned. (A) Horizontal gaze components. (B) Vertical gaze components. (C) Velocity profiles. (D) Acceleration profiles.



**Figure 12.** Calibrated temporal gaze and eye-in-head traces of the DMI method (red lines) and the SSC technique (blue lines) for subject PB. The head signal is denoted with a dashed line. Light grey horizontal bars indicate target positions ( $R=[35,30]$ ,  $\Phi=[35, 210]$ ). **Left panel:** Horizontal component. **Right panel:** Vertical component.

correcting saccade in the gaze and eye-in-head traces towards the vertical 20 deg target for both the DMI and the SSC signal. Note also that the head traces reach the target, since the subject has visual feedback via the laser pointer. These large head movements elicit counter rotation of the eyes, such that the eyes keep fixating the target while the head is still moving. The operation of the vestibular ocular reflex can thus be seen in both signals.

### 3.4 Discussion

We have shown that the extended DMI method is a valuable tool for measuring head-unrestrained 2D gaze shifts. The simultaneous recordings of the SSC technique and the extended DMI method are in good agreement. Although in our current setup the resolution of 0.3 deg obtained with the extended DMI method is somewhat inferior to the one of the SSC technique (0.2 deg) it remains well within the resolution of the oculomotor system. We believe that the signal-to-noise ratio can still be improved substantially. Most of the noise in the DMI method is due to electronic limitations imposed by the Rimmel system (see below).

Although we regarded the SSC signal as the perfect gold standard for comparison with the DMI method, this assumption need not be valid at all times. For example, coil slippage and minor pulling forces induced by the lead wires could have an influence on the trajectories and kinematics of saccades, or lead to an increase of blinking activity, and could thus be partly responsible for small trial-to-trial differences between the DMI and SSC signals, like in Figs. 9B and 10. On the other hand, since we recorded movements from both eyes simultaneously, small differences in binocular control signals, most notably due to differences in oculomotor plant characteristics (e.g. medial rectus vs. lateral rectus eye muscles), likely contribute to these differences too.

*Stability of the DMI system.* The extended DMI method as presented here with the DMI assembly mounted on a plastic glasses frame is a low-cost solution that is easy to implement. The frame's position remains stable relative to the ring and no disturbances of the signal were observed during highly accelerated gaze shifts. However, special care should be taken to assure that the pickup coil and

the anti-coil are fixed in the same plane (not necessarily close to each other if the magnetic fields are reasonably homogeneous) to assure a minimal head-orientation dependent residue of the three primary field components. Furthermore, depending on the exact size of the glasses frame the field of view can be somewhat limited. In our case some subjects reported to have difficulties seeing the far upward  $(R, \varphi) = (35, 90)$  deg target. Moreover, in its current form, subjects with corrected to normal vision will not be able to wear their own glasses. However, the glasses frame could be easily modified so that exchangeable lenses could be attached.

The glasses frame turned out to be a far better solution than a bite-board. Pilot measurements with a bite-board yielded substantial vibration artifacts, especially in the vertical channel that severely disturbed the measurements. In addition to this technical issue the glasses frame is much better suited for the use in patients, for whom a bite-board is considered a serious burden in addition to the invasive nature of techniques based on electromagnetism.

We also tested the inter-session stability of the assembly, by applying the Vnet calibration of a given session to the data obtained on a different recording day. However, the differences, which on average exceeded two deg, were found to be too large (data not shown). Probably, day-to-day variations in the play dough used to roughly center the assembly in front of the ring may have caused these differences. As a result, it was deemed necessary to calibrate the system prior to each use.

*Measurement range and calibration.* Note that the extended DMI method is not limited to the 2D range of  $\pm 35$  deg tested here. Provided the calibration subset is chosen appropriately the range can be readily extended to the full 360 deg. For the range tested here, a calibration dataset of 109 unique eye-head combinations already sufficed to obtain a good calibration. It is important to choose the set of unique eye-head combinations in such way that most points fall into the non-linear and non-monotonic part of the ring's signal. These are usually the most extreme eye positions. In this context it is also interesting to note that the used networks (4 input units, 4 hidden units, 1 output) could theoretically use 40 (20 weights and 20 biases) parameters. In practice, however, the number of effective parameters used was typically between 20 and 25. This order of magnitude seems surprisingly small when compared to the calibration routine for the head-restrained DMI method described by Bour et al. (1984) that needed 36 parameters to correct for the direction-dependent signal non-linearity. Note that Bos et al. (1988) described an analytical method to reduce the number of parameters to only 6, their calibration routine was based on the theoretical assumption that the axes of ring and pickup coil would exactly align, a requirement that is hard to realize in practice. In the head-unrestrained setting the non-monotonic ring signal with its pickup coil induced asymmetry and the complex primary field remnants can be calibrated with roughly the same number of parameters using the calibration routine presented here. More importantly, by using a simple model-free feed-forward neuronal net.



work, no special care needs to be taken to account for the complex Vnet signal. In addition, there is no need for the use of micromanipulators to exactly center the pickup coil in front of the ring (Malpeli 1998). As long as pickup coil and ring are roughly the network will be able to disambiguate the non-monotonic ring signal.

As we have shown in our previous study (Bremen et al. 2007) the extended DMI method can in principle be used to record over the entire 360 deg range. Due to limitations in the target arrays in our current setup the recording range was restricted to 35 deg in all directions. As a consequence of how the complex (and unknown) geometry of the assembly enters the equations for the DMI method (Eq. 1) it is not possible to simply extrapolate recorded signals beyond the range for which the system is actually calibrated. Thus, when measuring gaze shifts over a larger range a calibration procedure is required that covers this extended range. Although the resolution within the 35 deg range remained fairly constant, and while the application of three orthogonal magnetic fields theoretically allows for a constant resolution within the entire hemifield, the behavior of the DMI method for a larger measurement range will have to be established by future experiments.

*Recording hardware.* It is noteworthy that the head-unrestrained DMI method is not restricted to the use with the Rempel recording hardware. As long as three perpendicular primary magnetic fields are provided any recording hardware can be used. For example the lock-in amplifiers of Princeton Applied Research (model PAR 128A) have been used successfully in the head-restrained DMI method (Bour et al. 1984; Ottes et al. 1984; Van Opstal and Van Gisbergen 1987; Chaturverdi and Van Gisbergen 2004), and could also be applied with head-free DMI, as long as the input sensitivity is adapted to the signal strength of Vnet plus ring. Note that when using the Rempel eye monitor system it is important to use low impedance primary field coils since this will drastically increase the S/N ratio of the ring signal. The use of aluminum bars for field generation could also have a beneficial effect on the S/N ratio and thus on the system's resolution.

*Comparison with video techniques.* When comparing the extended DMI method to currently available video eye tracking systems basically the same points that hold true for the SSC technique can be noted as well. The invasive nature of the DMI method can be considered a disadvantage especially when working with patients or children. However, with DMI head-unrestrained measurements are possible, measurements can be performed in total darkness, even with closed eyelids, and at arbitrarily high sampling rates. The glasses frame holding the DMI assembly is lightweight in comparison with video helmets so that the glasses frame will not disturb the head movement kinematics.

When comparing the costs of the DMI, SSC and video systems, current video-based techniques are by far the most of expensive. However, even the SSCs cost well over US\$ 100,- per coil, so that with an average lifetime of 4-6 experiments an SSC experiment will typically cost about US\$ 25,-. These running expenses do not apply to the DMI method (a single gold-plated copper ring

suffices), which renders it an accurate, efficient, simple, and relatively low-cost recording technique.

*Binocular recordings.* With minor modifications of the current assembly, simultaneous recordings of both eyes are also possible, thus allowing the study of head-free vergence eye movements (see e.g. Chaturvedi and van Gisbergen, 2000, for head-fixed binocular DMI). However, torsional eye movements (Hess et al. 1992; Houben et al. 2006) cannot be measured with the DMI method.

*Experimental animals.* Especially in experimental animals the head-unrestrained DMI method will be a promising alternative to the SSC technique. The implantation of the ring is less invasive than of the SSC, as one need not guide the lead wire subcutaneously to a connector on top of the animal's head (Judge et al. 1979, Bour et al. 1984). Instead, the ring is implanted by simply opening and closing of the conjunctiva, which heals within a few days. More importantly, however, the absence of a lead wire renders the DMI method less vulnerable to malfunction than the SSC technique. In comparison to its use in human subjects, as described in this paper, the DMI method can be applied even more robustly in laboratory animals since the DMI assembly can be rigidly fixed to the animal's skull. This will ensure that the day-to-day geometry of the DMI assembly and ring is fixed for laboratory animals, and that therefore calibration of the ring signal will progressively improve as more data accumulates with every new recording day. Once a good calibration network has been determined it can be applied to the data recorded on different days, without the need for extensive calibration sessions.

However, since it is time consuming, and often impossible, to train animals to perform the type of coordinated eye-head movements as applied in the present study, a different calibration approach may have to be employed. Note that in principal two signals are needed for the calibration (1) the head-in-space signal (head coil) and (2) the eye-in-head signal (DMI ring). The head coil can easily be calibrated beforehand, for example by determining the gains of the coil as it is rotated without the subject in the experimental chamber. This is a standard calibration routine used for example when implanting eye coils in cats or monkeys. Then, the animal can be trained in a standard paradigm to make head-free gaze shifts to peripheral visual targets. The information of the calibrated head position can then be related to the measured (and known) eye-orientation in space from which the eye-in-head orientation can readily be reconstructed. In this way a catalogue of eye orientations for a given head orientation is collected. After a number of sessions sufficient data could be collected to calibrate the ring's signal. It is, however, to be expected that the collection of such a catalogue will take longer than the procedure described here for human subjects.



## Acknowledgements

Dr Ron Rimmel is thanked for providing assistance related to technical issues concerning the EM7 and the adjustment of the field coils. We are especially grateful to Stijn Martens and Hans Kleijnen for their valuable technical assistance. This work was supported by a Marie Curie Early Stage Research Training Fellowship of the European Community's Sixth Framework Program (MEST-CT-2004-007825 PB), the Neurocognition program of the Netherlands Organization for Scientific Research (NWO 051.04.022 RFVDW), a VICI grant within the Earth and Life Sciences of NWO (ALW 865.05.003 AJVO), and the Radboud University Nijmegen (AJVO).

### 3.5 References

Allik J, Rauk M, Luuk A. Control and sense of eye movement behind closed eyelids. (1981) *Perception* 10:39–51.

Bos JE, Reulen JPH, Boersma HJ, Ditters BJ. Theory of double magnetic induction (DMI) for measuring eye movements: Correction for nonlinearity and simple calibration in two dimensions. (1988) *IEEE Trans Biomed Eng* 35:733–9.

Bour LJ, Van Gisbergen JAM, Bruijns J, Ottes FP. The double magnetic induction method for measuring eye movement—results in monkey and man. (1984) *IEEE Trans Biomed Eng* 31:419–27.

Bremen P, Van der Willigen RF, Van Opstal AJ. Using double-magnetic induction to measure head-unrestrained gaze shifts I. Theory and validation. (2007) *J Neurosci Methods* 160: 75–84.

Chaturvedi V, Van Gisbergen JA. Stimulation in the rostral pole of monkey superior colliculus: effects on vergence eye movements. (2000) *Exp Brain Res* 132: 72–8.

Collewijn H, Van der Mark F, Jansen TC. Precise recording of human eye movements. (1975) *Vision Res* 15: 447–50.

Hess BJ, Van Opstal AJ, Straumann D, Hepp K. Calibration of three-dimensional eye position using search coil signals in the rhesus monkey. (1992) *Vision Res* 32: 1647–54.

Hofman P, Van Opstal AJ. Spectro-temporal factors in two-dimensional human sound localization. (1998) *J Acoust Soc Am* 103: 2634–48.

Houben MMJ, Goumans J, Van der Steen J. Recording three-dimensional eye movements: Scleral search coils versus video oculography. (2006) *IOVS* 47: 179–87.

Judge SJ, Richmond BJ, Chu FC. Implantation of magnetic search coils for measurement of eye position: An improved method. (1979) *Vision Res* 20: 535–38.

MacKay DJC. Bayesian interpolation. (1992) *Neural Comput* 4: 415–47.

Malpeli JG. Measuring eye position with the double magnetic induction method. (1998) *J Neurosci Methods* 86: 55–61.

Ottes FP, Van Gisbergen JAM, Eggermont JJ. Metrics of saccade responses to visual double stimuli: two different modes. (1984) *Vision Res* 24: 1169–79.

Rommel RS. An inexpensive eye movement monitor using the scleral search coil technique. (1984) *IEEE Trans Biomed Eng* 31: 388–90.

Reulen JP, Bakker L. The measurement of eye movement using double magnetic induction. (1982) *IEEE Trans Biomed Eng* 29: 740–4.

Robinson DA. A method of measuring eye movement using a scleral search coil in a magnetic field. (1963) *IEEE Trans Biomed Eng* 10: 137–45.

Van Opstal AJ, Van Gisbergen JAM. Skewness of saccadic velocity profiles: a unifying parameter for normal and slow saccades. (1987) *Vision Res* 27: 731–45.



## Chapter 4

# Using Double-Magnetic Induction to Measure Head-Unrestrained Gaze Shifts – III Calibration and Validation in the Monkey –

### Abstract

The double magnetic induction (DMI) method has successfully been used to record head-unrestrained gaze shifts in human subjects (Bremen et al. 2007b). This method employs a small golden ring placed on the eye that, when placed within an oscillating magnetic field, induces a voltage in a pickup coil in front of the eye. Here we develop and test a streamlined calibration routine for use with experimental animals, in particular monkeys. The calibration routine requires the animal solely to accurately follow visual targets presented at random locations in the visual field. This task is readily learned by the animals. Additionally, we use the fact that the pickup coil can be fixed rigidly and reproducibly on implants on the animal's skull. Therefore, the accumulation of calibration data leads to increasing accuracy.

As a first step we simulated gaze shifts and the resulting DMI signals. Our simulations showed that the complex DMI signals can be calibrated with random target sequences. Subsequently, we tested our paradigm with one monkey. Our results show that the data for a successful calibration can be collected in a single recording session, in which the monkey makes about 1500-2000 goal-directed saccades. We obtained a resolution of 30 arc minutes that is comparable to the fixation resolution of the monkey's oculomotor system.

---

*Adapted from:* Bremen P, Van der Willigen RF, Van Wanrooij MM,  
Van Grootel TJ, Van Opstal AJ *in preparation*

## 4.1 Introduction

The double magnetic induction (DMI) method has been successfully applied to measure eye movements in head-restrained subjects and laboratory animals (Allik et al. 1981; Reulen and Bakker 1982; Bour et al. 1984; Ottes et al., 1987; Bos et al. 1988; Malpeli 1998). For this method the subject's eye is positioned in the center of two perpendicular (horizontal and vertical) oscillating primary magnetic fields that induce a superposition of ac-currents in a thin, golden ring that is implanted on the sclera of the eye. The currents in the ring in turn generate weak secondary magnetic fields that can be picked up by a coil placed in front of the eye. These secondary magnetic fields vary in a non-linear way as a function of eye-in-head orientation, whereby the measurement range strongly depends on the exact geometrical relationships (i.e. size, distance, orientation and alignment) between ring and pickup coil (Bos et al. 1988; Bremen et al. 2007a). In addition, the pickup coil also carries the much stronger primary field components. The contribution of the primary fields can be reduced by adding a second coil that is connected in anti-phase to the pickup coil. If the subject's head is fixed the anti-coil can be positioned within the primary fields in such a way that the primary-field cancellation is nearly perfect (Bour et al. 1984; Ottes et al. 1987; Malpeli 1998). However, minor changes in geometry of the ring-coil/anti-coil assembly (e.g. through mechanical vibrations) can appreciably distort the measured signals. Moreover, when the head is unrestrained cancellation will be far from perfect due to small electromagnetic and geometrical differences between pickup coil and anti-coil as well as to inhomogeneities of the primary fields, all resulting in head-orientation dependent dc-offsets (Bremen et al. 2007a).

Yet, the major advantage of the DMI method compared to the classical scleral search coil (SSC) technique (Robinson 1963; Collewijn et al. 1975) is the absence of vulnerable connecting wires. The ring is easily implanted in laboratory animals (DMI: Bour et al. 1984; SSC: Judge et al. 1979) and does not break. It is therefore also more comfortable to wear on the eye, e.g. of human patients. However, the restriction to head-restrained preparations, the necessity for a bite-board, the strong input-output nonlinearity, which, in combination with the limited measurement range, leads to ambiguities in the signals for angles exceeding about 20 deg, as well as the method's sensitivity to mechanical vibrations and artifacts have so far prevented its widespread acceptance.

Recently, we have shown that the DMI method can be readily applied to measure two-dimensional (2D) head-unrestrained gaze shifts in human subjects (Bremen et al. 2007b). We introduced two critical extensions to the classical head-restrained DMI method. First, adding a third, frontal primary magnetic field that is perpendicular to the horizontal and vertical fields disambiguates the non-linear DMI signal. By combining the signals from all three fields with the associated head orientation (that can easily be measured e.g. with the search coil technique) the ring signal can in principle be calibrated over the full 360 deg range. The second extension is calibration of the signals with artificial neural networks that make

analytical description of the complicated ring signals unnecessary. The networks account for the unknown inhomogeneities in the primary fields, for cross-talk as the fields are not exactly perpendicular, for the unknown ring/pickup coil geometry in 3D space as well as for the head-orientation dependent dc-offsets that result from imperfect cancellation of primary field components.

Our goal is to apply the DMI method to head-unrestrained laboratory animals, like cats and monkeys, which requires a simple calibration paradigm. In our previous study we employed an intricate calibration routine in which human subjects were required to generate a number of precisely coordinated eye-head movements (Bremen et al. 2007b). Although such an instructed calibration paradigm is feasible for human subjects, it is quite cumbersome to apply to laboratory animals, as it would require extensive training. Here we propose and test a simple calibration paradigm that overcomes this problem. The gaze-orienting task solely requires accurate fixation of randomly presented visual targets that are distributed over the range of interest. Such a paradigm can be readily learned by cats and monkeys, and is already used in numerous gaze control studies. Here we test the applicability of our method to gaze shifts of rhesus monkeys.

## 4.2 Methods

### 4.2.1 Simulations

*Gaze shift and head saccade generation.* The purpose of the simulations was to create a simple tool that allowed quantitative assessment of different potential calibration paradigms for later use in the experiments with real subjects. We therefore implemented a simplified two-dimensional (2D) version of Robinson's local feedback model (Van Gisbergen et al. 1981; Van Weter and Van Opstal 2008) that could simulate pseudo-realistic head- and eye saccades for which the main parameters were randomly varied over a wide range: eye- and head saccade-component kinematics, initial eye-in-head position, and the head-saccade gain (head amplitude divided by desired gaze amplitude). We generated the horizontal and vertical saccade components independently to create various amounts of curvature in the gaze- and head trajectories. We also introduced different amounts of white noise to the simulated signals. The details of this implementation are given in Appendix A.

*Simulation of head-unrestrained double magnetic induction signals.* The simulated gaze shifts and head saccade signals were subsequently used to calculate the resulting DMI signals from the horizontal, vertical and frontal fields. We hence extended the theoretical description of the DMI method (Bremen et al. 2007a) to 2D (i.e., azimuth and elevation angles) in the following way:



$$\begin{aligned}
 V_h(\alpha, \gamma) &= K_h \cdot \sin(\alpha + \gamma) \cdot L(\cos(\alpha + \psi_h)) + V_{NET,h}(\gamma) \\
 V_v(\varepsilon, \delta) &= K_v \cdot \sin(\varepsilon + \delta) \cdot L(\cos(\varepsilon + \psi_v)) + V_{NET,v}(\delta) \\
 V_f(\alpha, \varepsilon, \gamma, \delta) &= K_f \cdot \cos(\sqrt{(\alpha + \gamma)^2 + (\varepsilon + \delta)^2}) \cdot L(\cos(\alpha + \psi_h)) \cdot L(\cos(\varepsilon + \psi_v)) + V_{NET,f}(\gamma, \delta)
 \end{aligned} \tag{1}$$

for the horizontal ( $V_h$ ), vertical ( $V_v$ ) and frontal ( $V_f$ ) field signals, respectively. Here,  $\alpha$  and  $\varepsilon$  denote the azimuth and elevation components of the eye-in-head orientation and  $\gamma$  and  $\delta$  are the azimuth and elevation components of the head-in-space orientation. Note that the frontal field signal depends on both the azimuth and the elevation components of eye and head and the signal's amplitude is phase shifted by 90 deg with respect to the other two fields. For simplicity, we did not incorporate crosstalk between the three channels.  $K$  is a proportionality constant that depends on the geometry of the system (ring-coil distance, coil self-inductance etc.) and on the magnetic field strengths and frequencies (Reulen and Bakker 1982; Bremen et al. 2007a). To simplify the implementation we assumed the same constant for all three fields, which essentially means that the amplifiers were tuned such that they generated the same signal strengths.

$L[\cos(x)]$  is the so-called shape factor that expresses the geometrical non-linearity of the DMI method (Bos et al. 1988; Bremen et al. 2007a). It only depends on the azimuth ( $\alpha$ ) and the elevation ( $\varepsilon$ ) components of the eye-in-head orientation (Bos et al. 1988). In the physical derivation of this coefficient, the geometry was idealized by letting the normal vectors of the pick-up coil and the ring precisely coincide. In that case  $L[\cos(x)]$  can be well approximated by a linear superposition of the first six Legendre polynomials of the first kind (Bos et al. 1988).

In a realistic application, however, the coil will typically be placed off-center relative to the eye, and the planes of coil and ring will not be exactly parallel (Bremen et al. 2007b). Because an analytical treatment of such a situation is impractical, and application of the Biot-Savart law to this is computationally expensive (Bremen et al. 2007a), we approximated the effect of geometrical misalignment by simply adding an offset angle,  $\psi$ , to the eye-in-head orientation of the  $L[\cos(x)]$  term. Our experiments verified that this correction leads to a realistic approximation (see Results). To simplify the situation we took this offset angle to be the same for the horizontal and vertical fields.

In practice, cancellation of the primary field component by the anti-coil will be imperfect, which results in a prominent head-orientation dependent dc-offset called  $V_{NET}$  (Bremen et al. 2007a,b). Since this offset depends on field inhomogeneities and small geometrical and electromagnetic differences between the coils, it cannot be described analytically in a straightforward manner (Bremen et al. 2007a). We therefore approximated  $V_{NET}$  by taking the difference of two phase-shifted sinusoids:

$$\begin{aligned}
 V_{NET,h}(\gamma) &= H_C \cdot [\sin(\gamma) - \sin(\gamma + \Delta\Phi)] \\
 V_{NET,v}(\delta) &= H_C \cdot [\sin(\delta) - \sin(\delta + \Delta\Phi)]
 \end{aligned} \tag{2}$$

with  $H_C$  proportional to each coil's self-induction (taken to be the same for the two

coils). For simplicity, we also took the phase shift to be identical for the horizontal and vertical primary fields. Values of all parameters used are given in table 1.

Target generation for the calibration of simulated gaze shifts. In the actual experiments we made use of a LED board attached to the wall of the experimental chamber (see Section 2.2.3). To enable a quantitative comparison between simulations and experimental results we simulated gaze shifts towards targets on this LED board. Target locations were defined in a polar coordinate system. Eccentricity,  $R$ , is measured as the angle with respect to 0 deg (straight ahead), whereas direction,  $\phi$ , is measured relative to the horizontal meridian. Targets could be at  $R = \{0, 5, 9, 14, 20, 27, 35, 43\}$  deg (rings) and  $\phi = \{0, 30, 60, \dots, 330\}$  deg (spokes). For simulations and easy visualization polar coordinates were transformed into the corresponding azimuth ( $\alpha$ ) and elevation ( $\varepsilon$ ) angles (Hofman and Van Opstal 1998) by:

$$\alpha = \arcsin(\sin R x \cos \phi) \text{ and } \varepsilon = \arcsin(\sin R x \sin \phi) \quad (2)$$

respectively. Targets were randomly drawn from these locations with the restriction that the target location did not exceed the field-of-view (<50 deg) relative

**Table 1.** Parameters for the simulation of the double-magnetic induction signals.

Parameter	Value	Description
$N_{coil}$	100	Number of turns of the pickup coil
$N_{ring}$	1	Number of turns of the ring
$\omega$	$75 \cdot 10^3 \cdot 2 \cdot \pi$	Angular frequency of the driving primary magnetic field in rad
$B$	$10^{-4}$	Magnetic field strength in T
$\mu_0$	$1.26 \cdot 10^{-6}$	Vacuum permeability/magnetic constant in H/m
$R_{coil}$	$2.5 \cdot 10^{-2}$	Radius of the pickup coil in m
$R_{ring}$	$0.8 \cdot 10^{-2}$	Radius of the ring in m
$R_{eye}$	$1.2 \cdot 10^{-2}$	Radius of the eye in m
$Z_{elec}$	$1.26 \cdot 10^{-3}$	Electric impedance of the ring in $\Omega$
$d$	$2 \cdot 10^{-2}$	Coaxial distance between ring and pickup coil in m
$L_{cl}$	2.5	Constant that is proportional to the coils' self-induction in a.u.
$\Delta\phi$	250	Phase shift representing the angle between the pickup coil and anti-coil for the horizontal & vertical fields in deg
$\psi$	2	Misalignment between center of rotation of the ring and pickup coil for the horizontal & vertical fields in deg.

to the fixation LED. The fixation location of the first trial was chosen randomly within the field-of-view from straight ahead. For all subsequent trials the target location of the previous trial was taken as the new fixation location, i.e. endpoint and starting point of subsequent saccades were connected. This scheme yielded a combination of saccades with different start- and end positions and different amplitudes and directions that sampled the whole frontal hemisphere between  $\pm 43$  deg.

*Calibration of double-magnetic induction data.* For a successful calibration of the DMI signals endpoints are needed that are related to known target locations. We aimed to extract the maximum amount of unique eye-head combinations from any given trial. To achieve this we used the fact that the head typically lags the eye: i.e. the head is still moving when gaze has already reached the target. We selected all simulated samples after head saccade had reached peak velocity. In the simulated data set this amounted to about 183,000 unique eye-head combinations from a total of 400,000 samples (45%).

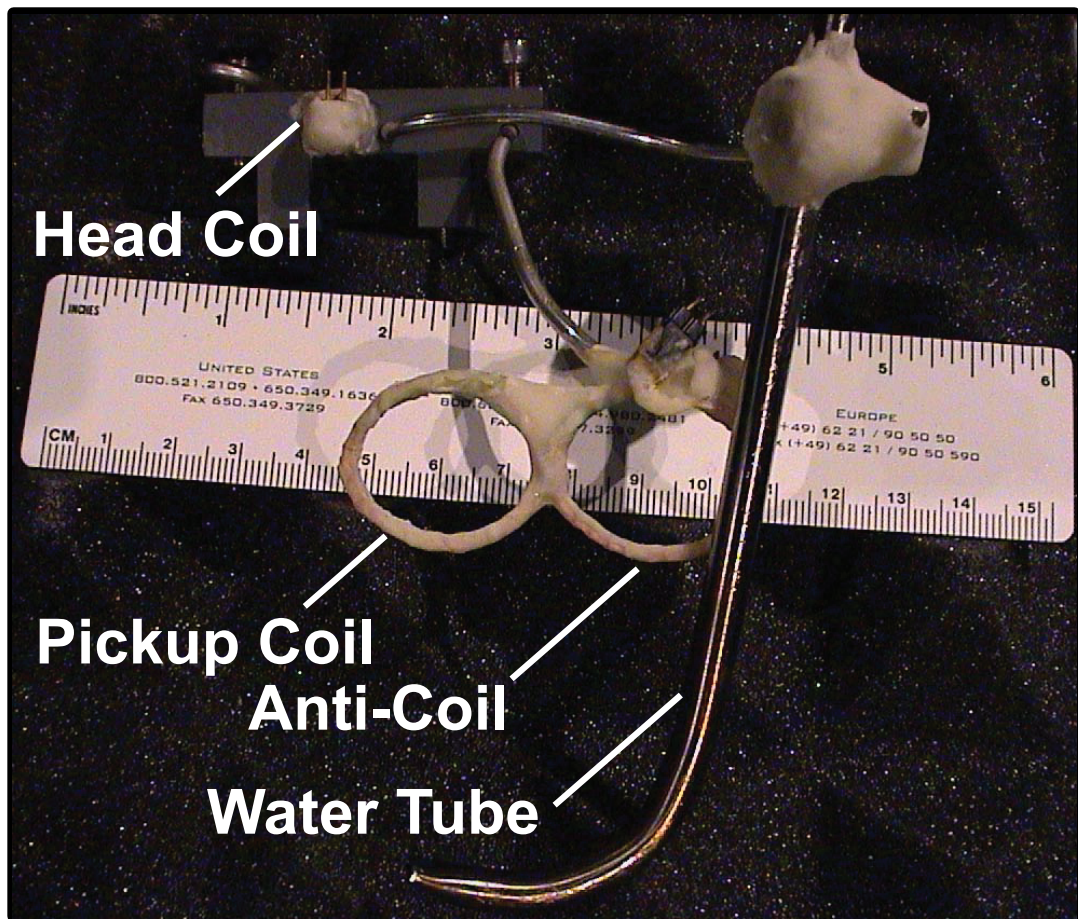
The selected gaze endpoints were calibrated by applying two three-layer feed forward neural networks, one for the azimuth component and one for the elevation component. The five input units of the network were (1)  $V_h$ , (2)  $V_v$ , (3)  $V_f$ , (4) simulated  $\gamma$  and (5) simulated  $\delta$ , while the desired, single output of the network was the gaze angle,  $\alpha$  or  $\epsilon$ . The networks were trained using the Bayesian-Regularization implementation of the back-propagation algorithm (Matlab 7.0, Neural Networks Toolbox, The Mathworks, Inc.) in order to avoid over-fitting (MacKay, 1992). The trained networks were used to calibrate the remaining gaze shift samples, here called the test set (55% of the data set). To assess the performance of the networks the calibrated test set was compared to the simulated gaze signal by calculating the error as the difference between simulated and calibrated sample. The mean error and the mean SD over all samples and trials were used to assess the influence of (1) the number of units in the hidden layer, (2) the number of trials, i.e. the number of unique eye-head combination samples and (3) the reproducibility of calibration (see Fig. 4).

#### 4.2.2 Experiments

*Subjects.* Behavioral experiments were conducted with one adult male rhesus monkey (*Macaca mulatta*; weight  $\sim 7$  kg)<sup>1</sup>. All experiments were conducted in accordance with the European Communities Council Directive of November 24, 1986 (86/609/EEC) and were approved by the local ethics committee (dier experimenten commissie, DEC) of the Radboud University Nijmegen for the use of laboratory animals.

<sup>1</sup> Additional experimental data from two other monkeys are currently collected in the laboratory and will become available before the thesis will go to print.

*Eye-head recording apparatus.* The monkey was seated in a completely dark room (2.5 m x 2.5 m x 2.5 m) facing a light-emitting diode (LED) board used for stimulus presentation. Three orthogonal pairs of single-turn field coils (surface area 6 mm<sup>2</sup>) used to generate the primary magnetic fields and were mounted alongside the edges of the four walls, ceiling and floor in a Helmholtz configuration. The coils were driven by custom-built amplifiers (horizontal field: 48 kHz, vertical field: 60 kHz, frontal field: 80 kHz). To monitor the monkey's eye movements a golden ring that fitted around the cornea without obstructing the extraocular muscles was placed onto the sclera underneath the conjunctiva in an aseptic surgery (Judge et al. 1979; Bour et al., 1984). In order to pick up the secondary magnetic field from the ring a DMI assembly (Bremen et al. 2007b) was attached to a head holder that had been implanted onto the skull in an earlier surgery (Fig. 1). The assembly consisted of a pickup coil connected antiparallel to the anti-coil (both 100 turns, wire diameter 0.1 mm, inner diameter 2.5 cm,  $R = 37 \Omega$ ,  $L = 983 \mu\text{H}$  @ 1 kHz). The two coils, shaped roughly as an ellipse, were cemented



**Figure 1.** Frontal view of the DMI assembly. Pickup coil, anti-coil and head coil are embedded in dental cement for protection. The water tube used to deliver the liquid reward, the pickup coil and anti-coil are fixed via two small steel wires to a plastics bridge that can be screwed to bolts implanted on the monkey's skull. The head coil is glued to this bridge and is also embedded in dental cement. Note that the small laser pointer is not mounted. Its mounting point is above the head coil.



to each other in a figure-eight shape so that both lied in exactly the same plane. The pickup coil was positioned in front of the implanted eye while the anti-coil protruded temporally (like a monocle). Note that in this way the other eye was completely free, which differs from the DMI assembly we have previously used with human subjects (Bremen et al. 2007b), and it will be possible to record binocular eye movements as well. The distance between eye and pickup coil was approx. 1 cm. To monitor head movements we attached an additional small search coil to the DMI assembly. Finally, we mounted a small laser pointer (LQB-1-650, World Star Tech, Toronto, Ontario, Canada) to the assembly. This laser could be used to align the monkey's head with the LEDs; a procedure that we employed to calibrate the head coil. For liquid reward delivery a thin stainless steel tube was mounted on the DMI assembly to which a rubber tube could be attached once the primate chair was in place. The rubber tube was connected to a water reservoir and a latch outside the experimental chamber. All wires and the water tubing were led from the top of the monkey's head to a custom-built tether system consisting of a metal rod attached to the monkey chair and a plastic ring (diameter 5 cm) through which the cables were led. The whole assembly without the tether system weighed about 50 grams.

The outputs of the DMI assembly and the head coil were connected to six lock-in amplifiers (Princeton Applied Research, Model 128A). Three were used for the horizontal, vertical and frontal field signal of the DMI assembly and the other three for the head coil. These signals were low-pass filtered at 150 Hz (4th order Butterworth, custom-build), digitized at a rate of 1000 samples/s (RX6 System 3, Tucker Davis Technologies; Alachua, Florida, USA) and finally stored on a computer's hard disk for subsequent off-line analysis.

*Stimuli.* The monkey faced a frame (radius: 149 cm) with twelve spokes and seven concentric rings (LED board) mounted on the wall of the experimental chamber. LEDs (0.2 deg diameter as viewed by the subject, intensity 0.5 cd·m<sup>-2</sup>,  $\lambda = 565$  nm) were mounted at  $R = \{0, 5, 9, 14, 20, 27, 35, 43\}$  deg (rings) and  $\varphi = \{0, 30, 60, \dots, 330\}$  deg (spokes) on the frame.

Target selection and data acquisition was done with an in-house program written in C++ running on a PC (2.8 GHz Intel Pentium D, Dell, Windows XP). To assure microsecond timing precision trial information was send from the PC to a micro-controller that worked as a latch initiating a trial when all hardware components were ready.

*Paradigms.* All experiments were conducted in darkness. One recording session lasted for several hours depending on the monkey's motivational level. At the beginning and at the end of an experimental session the monkey's head was fixed to the primate chair via a custom-built stainless steel rod that could be attached to the head holder on the monkey's skull. After head fixation the DMI assembly with head coil, laser pointer and reward system was mounted/removed. Since all components were attached to skull-fixed anchor points the rigid DMI

assembly could be placed reproducibly between consecutive recording sessions.

*DMI calibration paradigm.* We trained the monkey on a simple visual fixation task by instrumental conditioning. To initiate a trial the head-unrestrained monkey had to press a handle bar. A randomly selected LED was lit, extinguished after 600-1100 ms, upon which a different LED was illuminated. This sequence was repeated for a random number of LEDs (up to six). The last LED in a sequence changed its intensity after a randomly selected duration. The monkey had to indicate this intensity change as quickly as possible by releasing the handle bar within 600 ms. All LEDs in the trial were lit for a randomized period of time (between 600 ms and 1100 ms). The trial was aborted when the monkey released the bar too early, i.e. before the change in intensity occurred.

In case of a correct trial the monkey was rewarded with approx. 0.2-0.4 ml of fresh water. We increased (decreased) task difficulty by decreasing (increasing) the LED intensity change. To perform well the monkey needed to closely follow and foveate the LEDs. Additionally, we monitored the monkey's behavior via an infrared video camera, latencies and overall performance. This was especially important in the initial stages of training when we could not rely on calibrated eye and head signals.

Note that the definition of a trial differs between the simulations and the measurements. While in the simulations one trial contains one target and thus one saccade, in the measurements one trial can contain up to six targets and thus more than one saccade. On a typical recording day the monkey would make about 1500-2000 gaze shifts.

*Head coil calibration paradigm.* Note that the calibration of the DMI signal (gaze in space) can be done with the uncalibrated head-coil signals. However, to determine eye-in-head orientations from the calibrated gaze traces the head coil needs to be calibrated, too. Since the head-coil signals have a simple sinusoidal relationship with head orientation, this procedure is relatively straightforward. We manually aligned the monkey's head with the head-fixed laser pointer to the LEDs on the board. Due to physical constraints, it was not possible to align the head with the most eccentric LED ring ( $R = 43$  deg) and with the upward  $R = 35$  deg,  $\varphi = 90$  deg LED. We sampled 500 msec of fixation for a given LED and the monkey was rewarded after each LED with approx. 0.2-0.4 ml water via the head-fixed drinking tube. The monkey rapidly adapted to the procedure and after having fixated a few LEDs in this way he actively cooperated with the experimenter. Collection of one calibration set took about 3-5 min.

*Calibration.* Combinations of raw data (AD values horizontal, vertical and frontal channels) and known LED locations (azimuth and elevation in deg) were used to train two three-layer neuronal networks for azimuth and elevation. The networks comprised of two input units, four hidden units and one output unit and

were trained in the same way as the gaze networks. In addition to a nearly linear mapping from AD values to degrees the networks also accounted for small inhomogeneities in the fields and cross-talk between the three channels. The trained networks were used to calibrate the experimental head data yielding the  $\gamma$  and  $\delta$  angles.

The recorded ring signals were calibrated employing the same network architecture as described for the simulated data. Prior to calibration data were digitally low-pass filtered at 75 Hz using a 50-point FIR (finite impulse response) filter. The networks were trained on a subset of the sampled data. Gaze shift and head saccade endpoints were extracted as described in section 2.1.3. However, the velocity criteria were applied to the raw ring and head coil data. Note that raw head saccade endpoints were used in the input layer.

*Data Analysis.* We calculated the spatial resolution of the DMI method by determining the SD over the fixation epoch for each target location individually and converted the obtained values to minutes of arc. We argue that in order to resolve a movement the signal must be larger than the SD. Note that the recorded and calibrated signals do not only contain system noise but also minute fixational movements of the eye.

If not stated otherwise all fits were made employing the least-squares criterion.

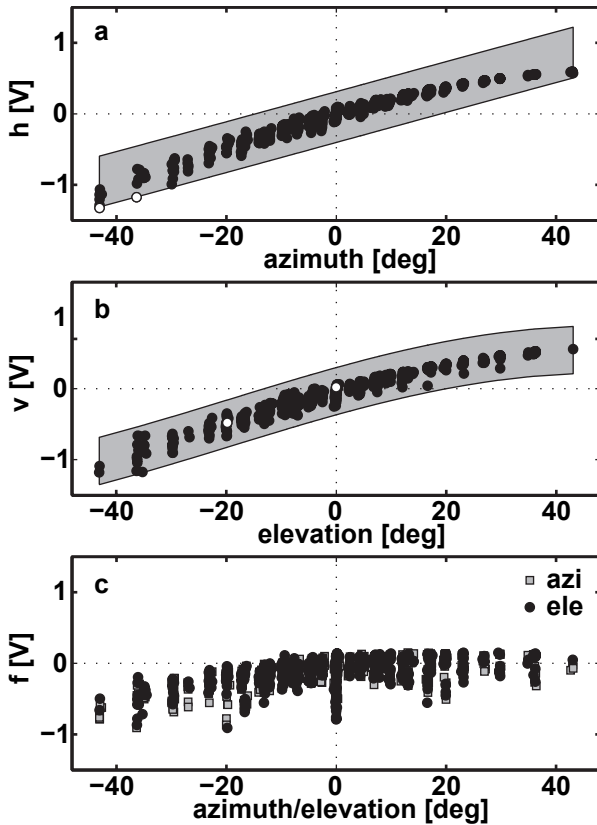
## 4.3 Results

### 4.3.1 Simulations

*Raw DMI signals.* Figure 2 depicts simulated DMI signals of the horizontal (Fig. 2a), vertical (Fig. 2b) and frontal (Fig. 2c) field as a function of azimuth and elevation (N=492). The simulations accounted for a 2 deg (see Tab. 1) misalignment of ring and pickup coil as well as for a head-orientation dependent  $V_{NET}$  function into account (see Methods and Tab. 1). The pronounced nonlinearity of the signals is particularly obvious for the horizontal and vertical fields. While the linear range is shifted to negative azimuth and elevation angles the nonlinear range starts already around 0 deg. The strong reduction in measurement range for rightward gaze positions is caused by misalignment of the centers of ring and pickup coil. The frontal field signals are plotted as a function of azimuth (gray squares) and of elevation (black circles). Note that the emerging pattern is more complex. Due to the influence of the complex  $V_{NET}$  function data points are tilted around zero degree azimuth/elevation. For a comparison with measured data compare Fig. 2 to Fig 8.

In order to check for endpoints that were extracted incorrectly we fitted a cosine through the data and discarded all data points that fell outside  $\pm 1$  SD. The gray patches in Fig. 2a and b indicate this range. Two data points were discarded (plotted as open circles) as they failed to meet the criterion (Fig. 2a). This procedure





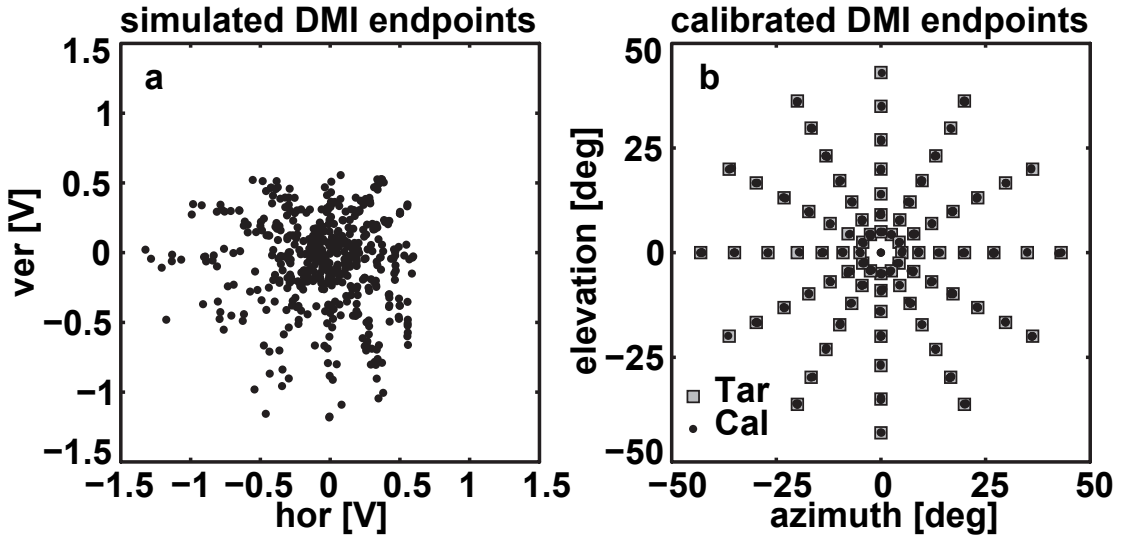
**Figure 2.** Simulated DMI endpoint signals for the horizontal (a), vertical (b) and frontal (c) field components. (a) Horizontal voltages of simulated endpoints as a function of azimuth. The **black circles** indicate data points. A sine was fit to the data, in order to remove outliers. The gray patch depicts the area of one SD around the fit. **Open circles** indicate data points that fell outside this selection criterion. Note that in the simulations all endpoints were correct by definition. The employed criterion can thus be considered conservative. (b) the same as in (a) but for the vertical component. (c) Frontal voltages as a function of either azimuth (**gray squares**) or elevation (**black circles**). For a detailed description see text.

was designed to minimize the inclusion of wrong endpoints for the training of the artificial neuronal networks. As can be seen this approach is conservative since in the simulations all endpoints were extracted correctly.

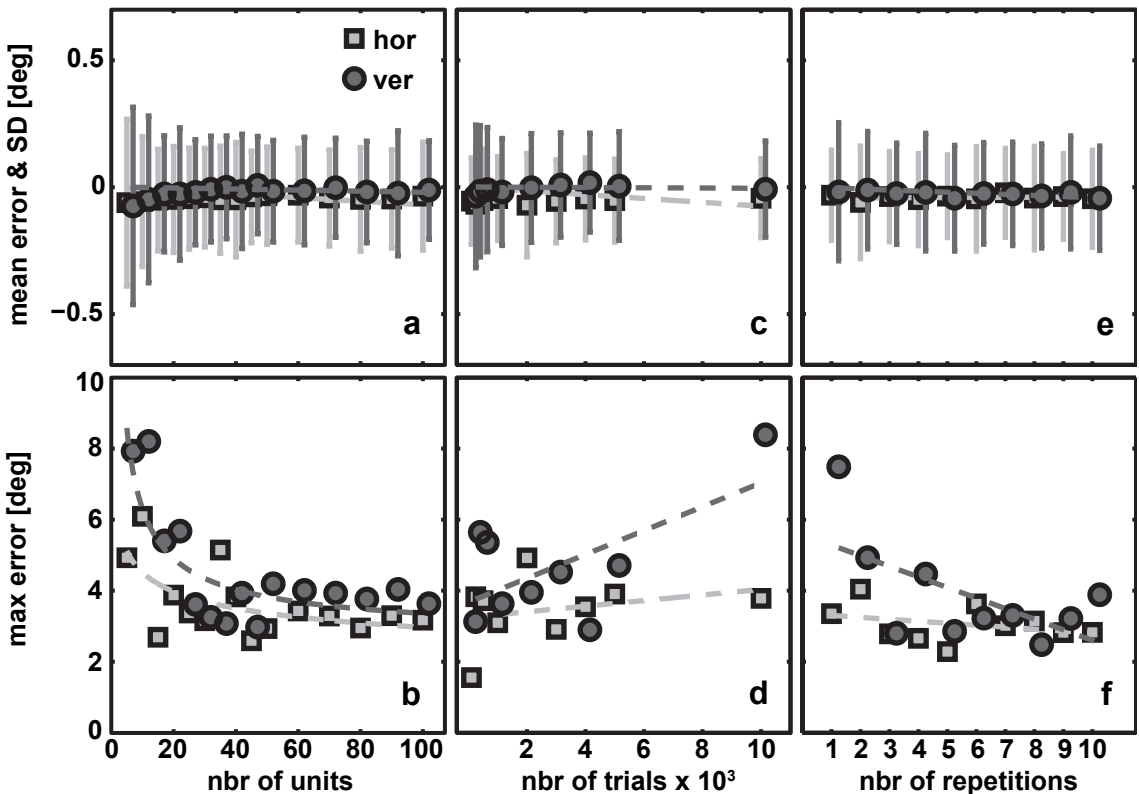
In Fig. 3a we have plotted uncalibrated endpoints of the vertical field versus those of the horizontal field. The pronounced non-linearity and asymmetry in the data can be clearly seen as the data points for positive horizontal and vertical voltages are squeezed while the negative range is spread out. A similar pattern is observed for the measured data (Fig. 9a).

*Calibrated gaze signals.* Figure 3b shows the gaze endpoints of Fig. 3a after calibration with two three layer artificial neuronal networks with 40 hidden units each. Target locations are indicated by gray squares and calibrated endpoints as black circles connected with black lines to their respective target. The horizontal and vertical networks were well able to map the complex, non-monotonic DMI signals.

In order to determine the best settings for the artificial neuronal networks we performed several optimization simulations on the entire data set. First, the optimal number of hidden units was determined by running the calibration routine with different numbers of hidden units (5 to 50 in steps of 5, 60 to 100 in steps of 10) for 1000 simulated trials. This procedure was repeated three times and the resulting mean error and the SD are shown in figure 4a. The maximal error is shown in figure 4d. The data of the vertical component (dark grey circles) were shifted



**Figure 3.** Simulated DMI endpoints before (a) and after (b) calibration. (a) Two-dimensional plot of the simulated raw horizontal and vertical voltages. Note the pronounced signal asymmetry. (b) After calibration the pronounced signal asymmetry is gone. The endpoints can be mapped to the target locations accurately.



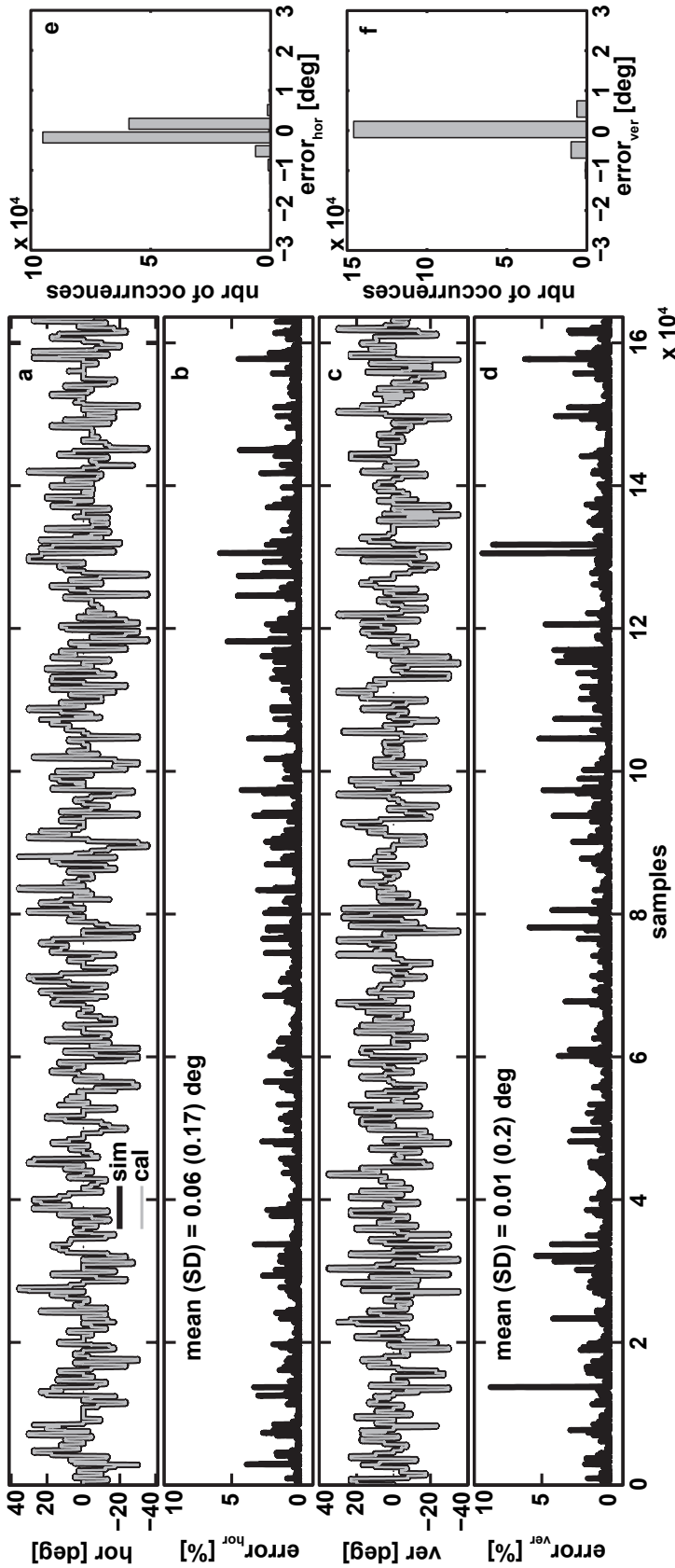
**Figure 4.** Simulated DMI endpoints before (a) and after (b) calibration. (a) Two-dimensional plot of the simulated raw horizontal and vertical voltages. Note the pronounced signal asymmetry. (b) After calibration the pronounced signal asymmetry is gone. The endpoints can be mapped to the target locations accurately.

to the right with respect to the horizontal data (light grey squares) for graphical purposes only. The mean error is smaller than 0.05 deg for both horizontal and vertical components and does not vary with the number of units. In fact, linear fits for both components (dashed grey lines, vertical fit occludes horizontal fit) have a slope and an offset of 0 deg/unit and 0 deg, respectively. The SD of the error, however, decreases exponentially with the number of units and reaches its minimum at about 30-40 units. We therefore chose 40 units for all subsequently used networks.

Second, an important aspect of the simulations is to assess how many trials are needed for a good calibration of the DMI signals. We therefore simulated a varying number of gaze shifts and calibrated the extracted endpoints. To assess the influence of the number of training trials - and therefore the number of unique eye-head combinations - on performance we simulated and calibrated experiments with 100, 500, 1,000 to 5,000 (steps of 1,000) and 10,000 trials. Fig. 4b and e show the mean error, SD of the error and maximal error as a function of number of trials (gaze shifts). Mean error (close to 0 deg) and SD ( $< .25$  deg) do not vary with number of trials. Note that the number of trials does not affect the mean error. The SD decreases slightly and reaches its minimum at 3,000 trials. Further increasing the number of trials does not improve the performance. The maximal error as expected does jitter somewhat more and is always smaller than 6 deg and larger than 1.5 deg. The data point for the vertical network at 10000 trials is due to an outlier in one of the three repetitions. In the other two runs the maximal error was close to 4 deg. This demonstrates an issue one has to deal with when working with artificial neuronal networks. A local minimum can be reached during training leading to an early abortion of the process that results in a badly trained network. For the simulations we chose 500 trials in order to limit computation time.

Third, in order to test the robustness/reproducibility of the networks we repeated the calibration of 500 simulated trials with 40 hidden units 10 times. As can be seen in figure 4c the mean error and the SD do not vary appreciably with the number of repetitions. However, the SD between repetitions did vary to some degree (compare for example repetition 1 with 3 of the vertical component). Clearly, large SD's correspond to larger maximal errors and smaller SD's to smaller maximal errors (Fig. 4f). It is therefore advisable to select networks in the training process with a small SD of the error for both components.

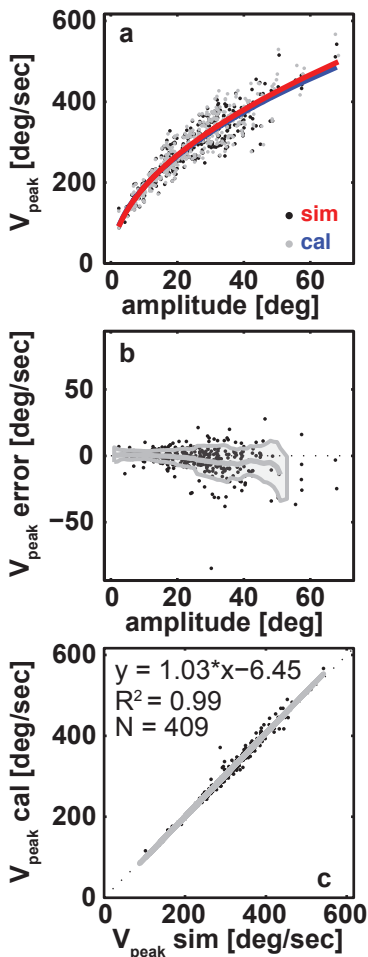
So far we have only discussed calibrated endpoints obtained after training the networks. Figure 5a and d depict 500 simulated (thick black line) and calibrated (gray lines) gaze shifts. As can be seen the calibrated gaze shifts nicely follow the simulated horizontal (Fig. 5a) and vertical components (Fig. 5c). A more quantitative view is provided by Fig. 5b and d, which show the error, i.e. the difference between simulation and calibration, in percent for the horizontal (Fig. 5b) and the vertical component (Fig. 5d). The error never exceeded 10%, but note that large errors were very rare, as can be better appreciated from the histograms of Fig. 5e (horizontal) and f (vertical). Note that in these subplots the error is



**Figure 5.** Overview of all simulated (**black line**) and calibrated (**thin gray line**) saccades in one continuous trace for the horizontal (**a**) and vertical (**c**) components. Relative errors, i.e. simulation - calibration in percent, are shown in (**b**) for the horizontal and in (**d**) for the vertical components. The error for both components is always smaller than 10%. Histograms of the error in degrees are depicted in (**e**) and (**f**) for the horizontal and vertical components, respectively. The majority of errors are in the order of  $\pm 1$  deg.

given in degrees. The far majority of errors were at or close to zero degree. The number of occurrences of errors larger than  $\pm 1$  deg are not visible on the scale of the histograms (10\*104 samples horizontal; 15\*104 samples vertical). Overall this shows that the networks were very well able to calibrate not only endpoints but also the intermediate samples of simulated simulated head-unrestrained gaze shifts that were not used in the training set.

As a further test of calibration we extracted saccade parameters from the simulated gaze shifts (prior to feeding them to the DMI conversion) and compared them to the calibrated gaze shifts. We extracted saccade on- and offset, amplitude and peak velocity. Figure 6a depicts the peak velocity as a function of amplitude. The original simulated data points are shown as black circles, calibrated data points as gray circles. The peak velocity versus amplitude relationship was fit with an exponential (red: simulated; blue calibrated). Data and fits for simulated and calibrated gaze shifts are in good agreement. At amplitudes above 40 deg the calibrated data tend to slightly underestimate the velocity slightly, but this may be due to the small number of large saccades. Figure 6b shows the error in peak velocity between simulated and calibrated data as a function of amplitude (black circles). The running average is plotted through the data as a gray line and



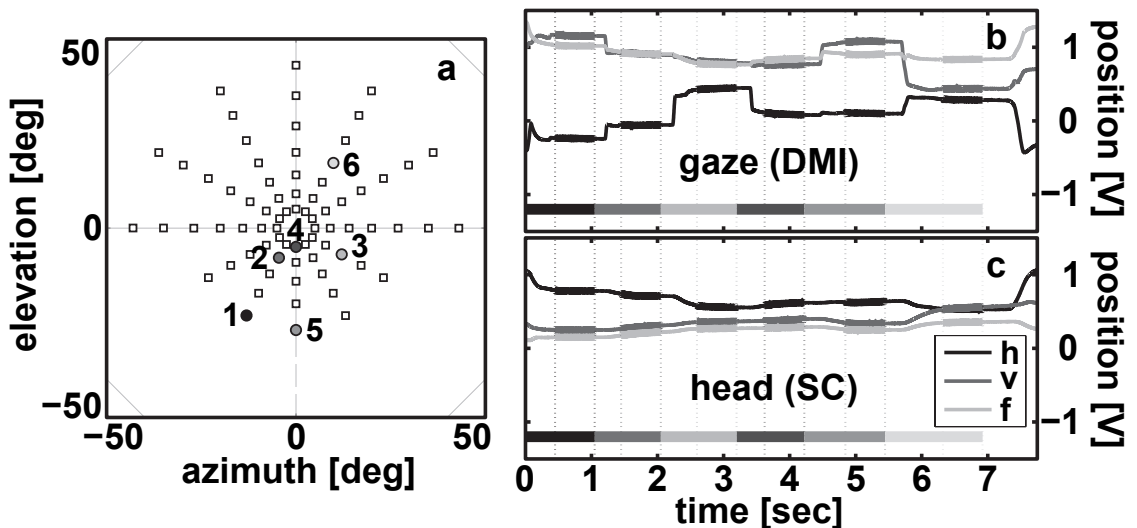
**Figure 6.** Comparison of the main sequence velocity-amplitude relation for simulated (**black circles**) and calibrated (**gray circles**) gaze shifts. (a) Velocity plotted as a function of amplitude. Velocity increases exponentially with amplitude. **Red line:** exponential fit to the simulated data. **Blue line:** fit to the calibrated data. (b) Difference between simulated and calibrated data as a function of amplitude. the **gray line** indicated the running average through the data and the patch covers one SD around the running average. (c) Linear regression analysis by plotting the calibrated peak velocity as a function of simulated peak velocity. Note the high correlation.

the patch indicates the SD of the error. Although the SD slightly increased with amplitude, the overall error remained very small. To further assess how well the calibrated data reflect the simulated gaze shifts we plot peak velocity obtained from the calibration against peak velocity from the simulations in Fig. 5c. As can be seen the two highly correlate ( $r^2=0.99$ ). From these data we conclude that the calibration preserves the saccade kinematics well.

### 4.3.2 Measurements

Data presented in this section were collected from one monkey at two successive days. One session lasted for about 4 hours in which the monkey completed a total of about 2000 trials.

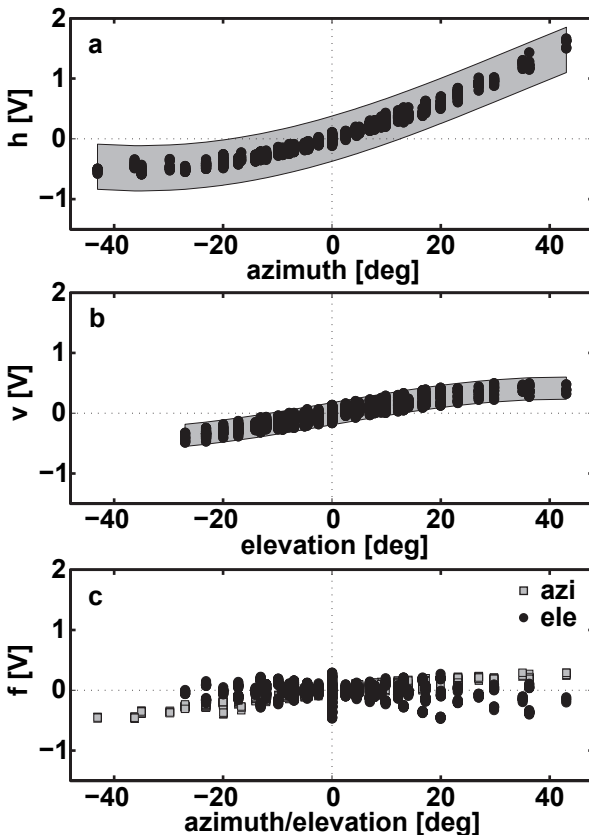
*Raw DMI signals.* Next we present measured gaze shifts and head saccades from a rhesus monkey equipped with the DMI assembly shown in Fig. 1. Figure 7a shows the location and order of six targets presented to the monkey in a representative trial. Note that in the lower hemisphere two rings of the setup with the eccentricity  $R = 35$  and  $43$  deg could not be used since the monkey was not able to see them. In this particular trial the monkey had to follow 6 targets before the last target changed intensity and he had to release the bar. The whole trial lasted for more than 7 sec before the monkey released the bar and got rewarded. The monkey's uncalibrated gaze shifts and head saccades can be seen in Fig. 7b and c, respectively. The horizontal component is depicted as a black line, the



**Figure 7.** Uncalibrated measured DMI and head coil signals. (a) Stimulus array (open squares) and targets (circles in different shades of gray). The numbers close to the targets indicate the order of presentation. (b) Recorded DMI position signals of the horizontal (black line), vertical (dark gray line) and frontal (light gray line) field. The thick gray bars at the bottom of the plot indicate the presentation of a target. The shades of gray correspond to the ones used in (a). (c) Same as (b) but for the head coil signal.

vertical as a dark gray line and the frontal as a light gray line. The thick lines in different shades of gray at the bottom of each subplot indicate the on- and offset of a target. The colors correspond to the colors used in Fig. 7a. Note that the traces both for the DMI method and the search coil (SC) are smooth with little noise. Fig. 7b and c can also be used to explain how endpoints for calibration were extracted. First, the offset of each target was determined than a velocity criterion was applied to the preceding 600 samples, i.e. gaze velocity had to be smaller than 0.3 V/sec, a value determined after visual inspection of the uncalibrated velocity traces. Since neighboring samples have similar values we selected every tenth sample from the original 600 samples in order to decrease computational time needed to train the networks. The reason for not selecting for example the mean over a given range of samples can be seen when looking at head saccade made towards the third and fourth target. The head still moves while gaze is on target, i.e. the eye counteracts the head movement. These additional eye-head combinations are essential for a successful calibration (see Bremen et al. 2007b). Extracted samples are indicated as thick lines in the color corresponding to the respective field.

Figure 8 depicts the extracted endpoints in the same format as Fig. 2. In Fig. 8a the horizontal DMI signal is plotted as a function of azimuth, the vertical component as a function of elevation is shown in Fig. 8b and the frontal component plotted versus azimuth (gray squares) and elevation (black circles) is shown in Fig. 8c. The endpoints were aligned such that the signal recorded at 0 deg



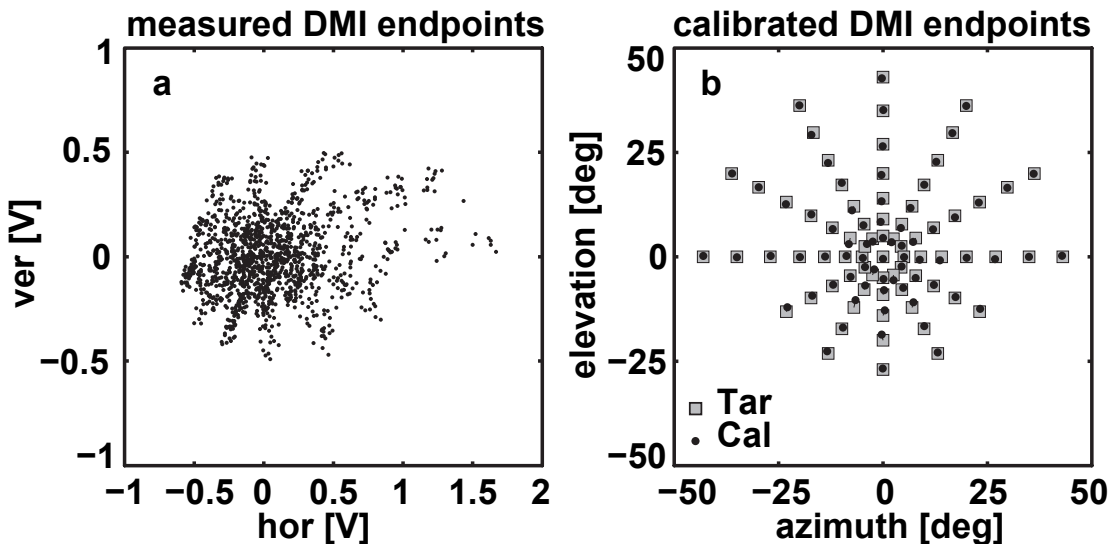
**Figure 8.** Measured endpoints after extraction from the saccades. The convention is the same as used in Fig. 2.



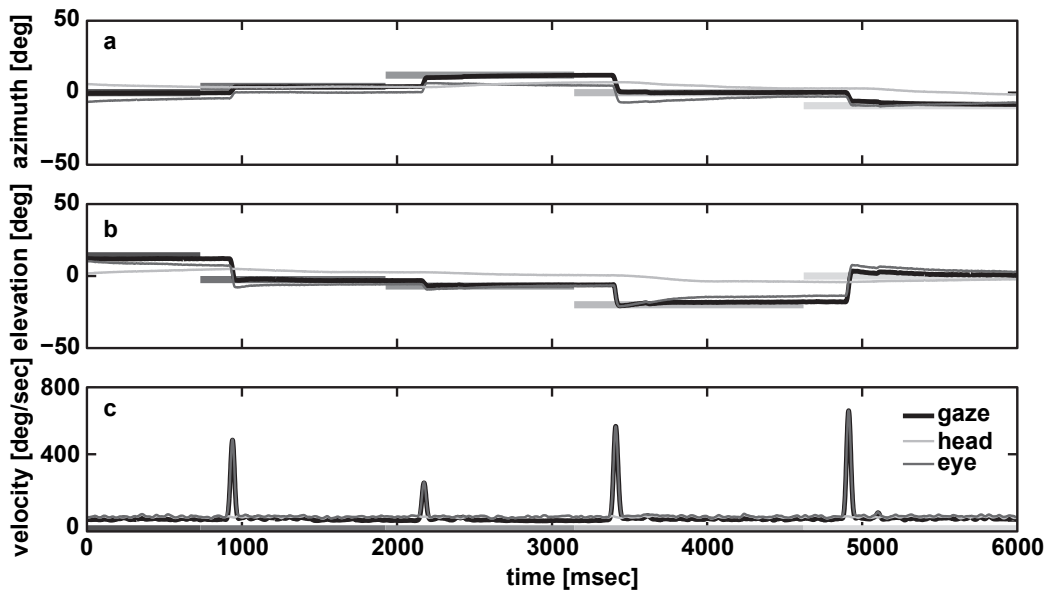
azimuth and elevation coincided with 0 V. A comparison with Fig. 2 shows that the simulations nicely captured the main features of the recorded signals. The misalignment of ring and pickup coil can be seen in Fig. 8a as a shift of the linear range towards positive azimuth angles while the signal is ambiguous for negative azimuth locations. For the vertical component some target locations are missing compared to the simulations since the monkey was not able to see these low locations (see also Fig. 7a). The frontal field nicely exhibits the tilt due to the influence of the complex  $V_{NET}$ .

Another view of the raw recorded endpoints is provided in Fig. 9a. There the vertical component is plotted against the horizontal one. Again the pronounced non-monotonicity is seen for negative voltages, i.e. azimuth locations. The shape of the target LED board can be distinguished to some extent, but it is clear that the signals provide a highly distorted representation of the actual gaze shifts. Nevertheless, the rather complex signals can be successfully mapped onto the target array as shown in Fig. 9b. Target locations are indicated by gray squares while the calibrated endpoints are shown as black circles. In total 82,655 of 3,311,057 (2.5%) recorded samples of one session were used for calibration.

*Calibrated gaze signals.* Fig 10a and b depict the calibrated gaze shifts (black lines) and head movements (light gray lines) and additionally the reconstructed eye-in-head traces (dark gray lines). The monkey performed the task of fixating the LEDs well especially when considering the long duration of the trials (78% correct trials). Gaze shifts end on the presented targets (thick lines in different shades of gray). As expected, the head does not reach the target location and lags the gaze shifts. The eye-in-head traces nicely demonstrate the VOR with the



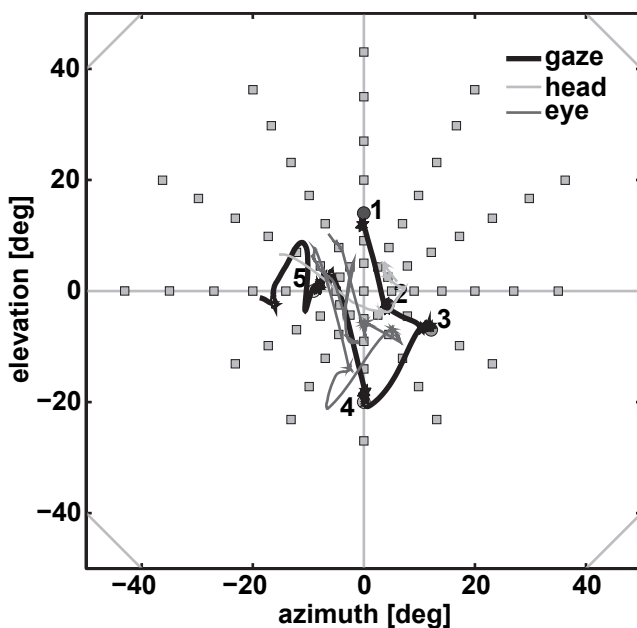
**Figure 9.** Measured DMI endpoints before (a) and after (b) calibration. (a) Two-dimensional plot of the simulated raw horizontal and vertical voltages. Note the pronounced signal asymmetry. (b) After calibration the pronounced signal asymmetry is gone. The endpoints can be mapped to the target locations accurately. Compare with Fig. 3.



**Figure 10.** Calibrated gaze (black line), head (light gray line) and eye-in-head (dark gray line) traces of the azimuth (a) and elevation (b) components. Targets are indicated by thick lines in shades of gray. (c) Velocity profiles for the traces shown in a and b. Color conventions as in a and b.

eye counteracting the slow head movement. Velocity profiles of the gaze shifts, head and eye-in-head movements are shown in Fig. 10c. Note that the profiles are smooth with little noise. The gaze shifts (black thick line) are faster than the head movements (light gray line) with larger gaze shifts exhibiting higher peak velocities. Eye-in-head velocity follows the gaze velocity closely since the contribution of the head movement to the gaze shift is extremely small.

Figure 11 depicts the calibrated gaze shifts (black line), head movements (light gray line) and eye-in-head-traces (dark gray line) in 2D. Target locations are



**Figure 11.** Two dimensional display of gaze (black line), head (light gray line) and eye-in-head (dark gray line) traces. The circles in different shades of grays indicated target positions and the number close to the circles indicate the presentation order.

indicated by circles in different shades of gray and numbers close to the circles indicate the order of presentation (compare with Fig. 10a and b). In our lab, data as shown in figures 10 and 11 are used as the online display for experimenter to check the performance of the monkey during experiment.

An important aspect of the head-unrestrained DMI method for the use with laboratory animals is reproducible fixation of the DMI assembly session to session. To test whether this could be guaranteed we calibrated a data set from a different recording day (set A) with the networks that were obtained on a previous recording day (set B). Fig. 12 compares raw data (left column) and calibrated data (right column) of set A versus set B. The data were obtained by averaging endpoints of each set directed to the 74 unique LED locations.

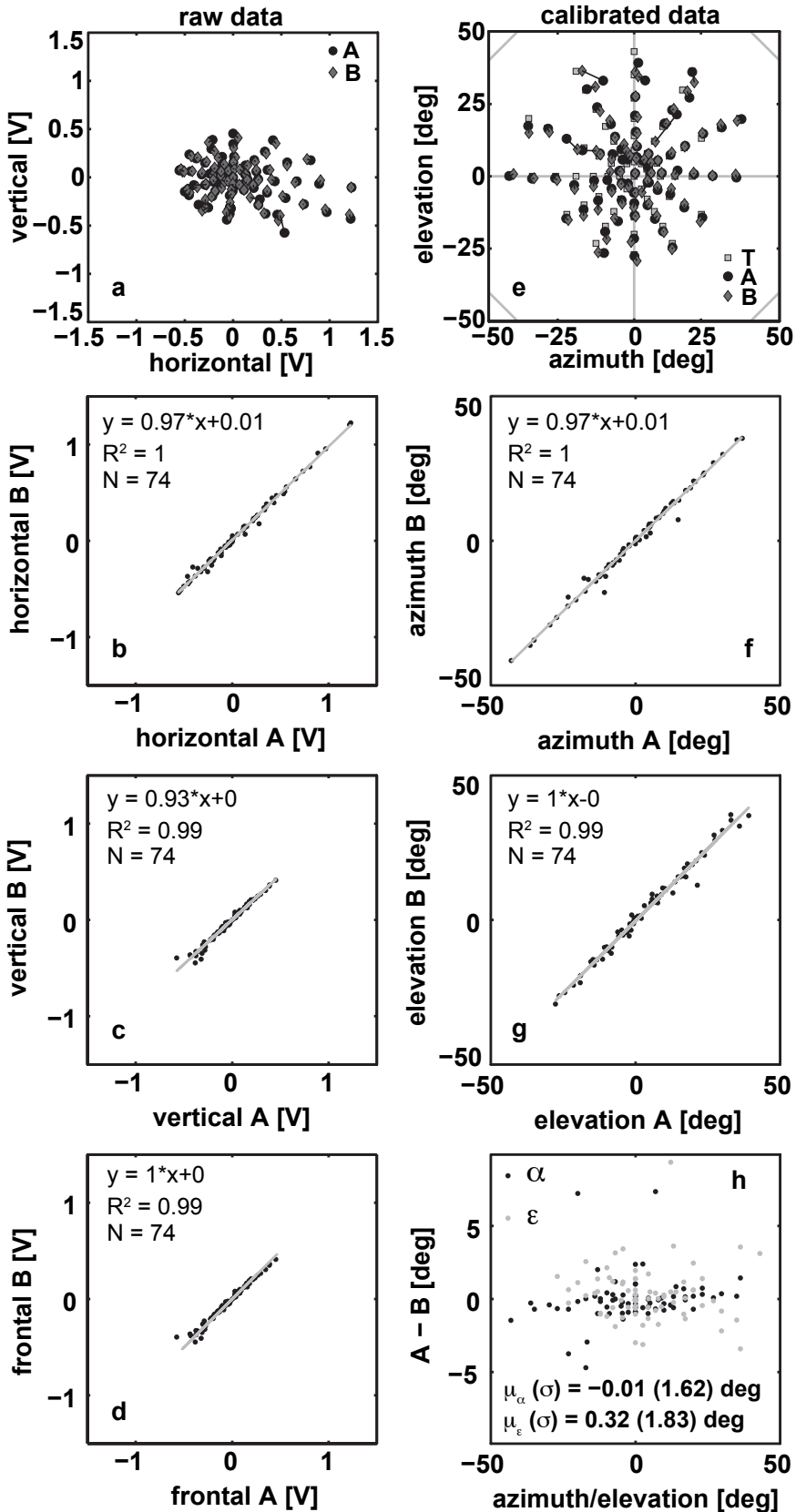
Raw data endpoints for set A (black circles) and set B (dark gray diamonds) are shown in Fig. 12a. Note the close correspondence between the two sets collected on two consecutive days. This can also be seen when looking at the individual fields. Fig. 12b plots the horizontal component of set B as a function of the horizontal component of set A. Figures 12c-d plot the data of the vertical and frontal field in the same manner. Data of all fields could be fit with a straight line. The correlation coefficients  $R^2$  for all three components were larger than 0.9 with slopes close to 1 and offsets close to zero.

In Fig. 12e the calibrated saccadic endpoints of the set A (black circles) and set B (dark gray diamonds) are plotted in 2D together with the target location (light gray squares). Endpoints made towards the same target are connected with black lines. The two sets correspond well with only small differences. There were only three outliers with a difference between set A and B larger than 5 deg (see also Fig. 12h). These are likely attributable to the monkey fixating the target incorrectly. Since the endpoint of one set is close to the target location while the endpoint of other set is far off target.

A yet better comparison of the two sets is achieved when looking at Fig. 12f-g. There the azimuth (Fig. 12f) / elevation (Fig. 12g) component of set B is plotted against the azimuth/elevation component of set A. A regression analysis for the azimuth component yielded a slope of 0.97 and an offset of 0.01 deg with a  $r^2$  of 1.00. For the elevation component the values were 1.00 for the slope, 0.00 deg for the offset and 0.99 for  $r^2$ . The correlation is thus high with fits falling closely on the unity line.

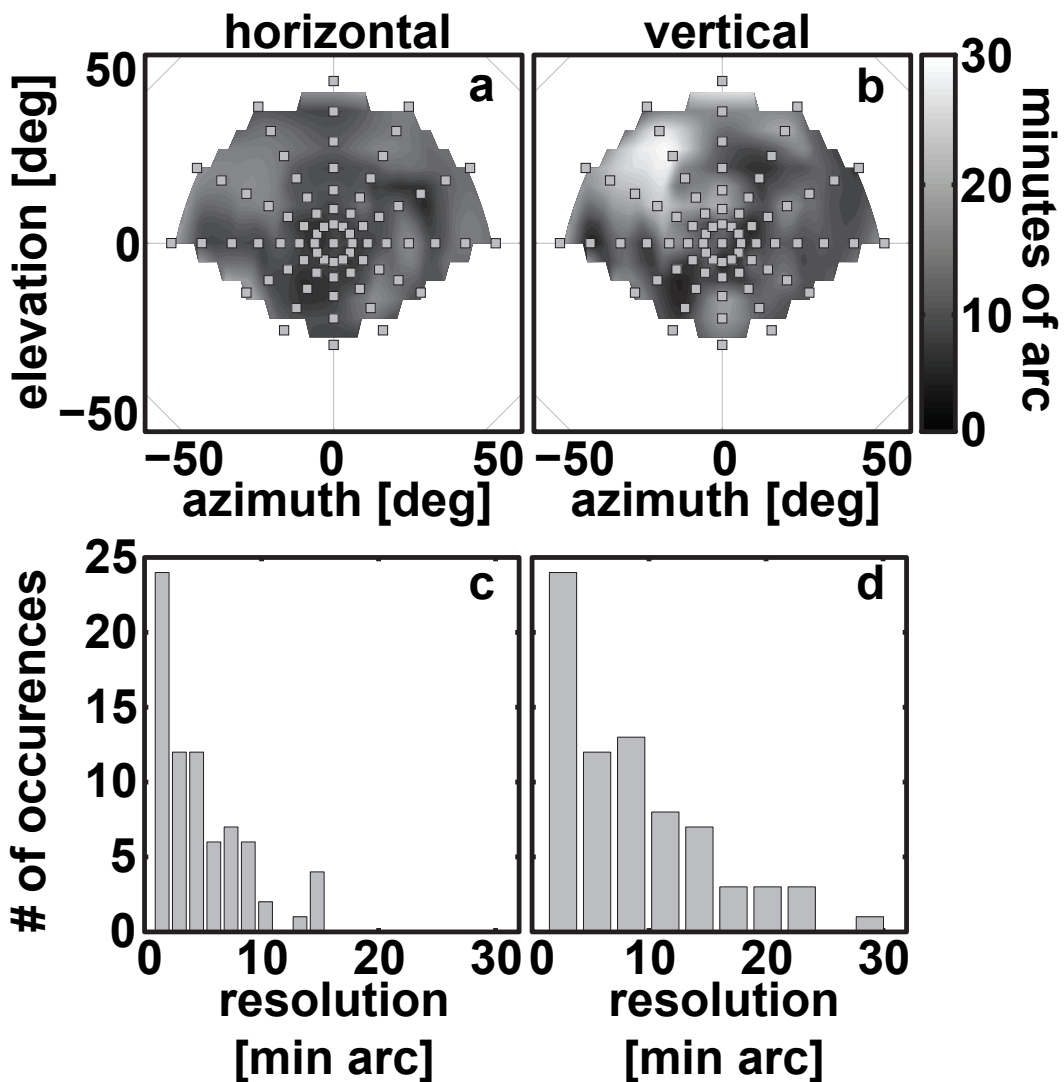
---

**Figure 12.** Analysis of the stability of the DMI assembly. **Left column:** raw data. **Right column:** calibrated data. Data from set A (**black circles**) and B (**dark gray diamonds**) were collected on two consecutive days. **(a)** Raw endpoints of set A and B. Corresponding endpoints are connected with black lines. **(b)** Horizontal field set A as a function of horizontal field B (**black circles**). The gray line is a linear fit through the data. **(c)** Same as **(b)** for the vertical field. **(d)** Same as **(b)** for the frontal field. **(e)** Calibrated endpoints. Networks were trained on set A and used to calibrate both data sets. Target locations are shown as light gray squares. **(f)** Azimuth component of set B plotted against azimuth component of set A. The gray line is a linear fit through the data. **(g)** Same as **(f)** for elevation. **(h)** Difference between set A and B of the azimuth (**black circles**) and elevation (**gray circles**) components as function of location. (see next page)



The error, i.e. set A - set B, of the azimuth (black circles) and elevation (light gray circles) components is plotted as a function of location in Fig. 12h. It can be seen that the majority of errors is smaller than  $\pm 5$  deg. The mean error for both the azimuth and the elevation components are close to 0 deg with SD's smaller than 2 deg. Taken together this demonstrates that the DMI assembly is fixed rigidly and reproducibly between successive recording sessions.

The resolution in minutes of arc of the head-unrestrained DMI method is assessed in Fig. 13 for the horizontal (Fig. 13a) and vertical (Fig. 13b) components. The worst resolution found is 30 arc minutes in the upper left quadrant of the vertical component. Typical values were smaller than 10 minutes for both components (Fig. 13c,d). These values are comparable to previously reported ones for the DMI method and the standard SSC technique (Bremen et al., 2007b).



**Figure 13.** Resolution of the head-unrestrained DMI method for the horizontal (a) and vertical (b) component in minutes of arc (shades of gray). (c) and (d) histograms of the horizontal and vertical resolution, respectively.

## 4.4 Discussion

The data presented in this paper demonstrate that the DMI method can be successfully applied to measure head-unrestrained gaze shifts in laboratory animals. Although we tested our method on trained rhesus macaques, there is no reason why the method could not be used in other experimental animals, like cats, or even mice. The only requirement for the method to work is the ability of the animal to make instructed goal-directed gaze shifts toward visual target locations.

An important aspect of the current study is the introduction of a simple calibration routine that solely relies on the accurate fixation of randomly presented visual targets in the range of interest. Before applying this calibration routine to real measurements obtained with a monkey, we performed realistic computer simulations that mimicked natural head-unrestrained gaze shifts with a physical model of the resulting DMI signals (Fig. 2). We showed that artificial neuronal networks could be efficiently trained to calibrate fixation endpoints with great accuracy (Fig. 3), but also that the networks were able to interpolate the remaining data points to precisely capture the whole saccadic gaze traces (Fig. 5), while even preserving the details of the main sequence relations, for which the most sensitive parameter is peak gaze velocity (Fig. 6).

The experimental data were calibrated successfully (Figs. 9, 10, 11). In accordance with previous data (Bremen et al., 2007b) the worst spatial resolution was about 30 minutes of arc which is still within the range of the highest precision of the oculomotor system. The typical resolution of the gaze shifts, however, was below 20' and was therefore comparable to values values that have been reported with the SSC technique (e.g. Bour et al.).

*Training of Neuronal Networks.* The networks were trained on gaze endpoints and their inputs were the raw head orientation and the recorded voltages of the horizontal, vertical and frontal components induced by the ring. Note that it is not necessary to use calibrated head orientations as input to the networks. But in order to make full use of all recorded signals, including a reconstruction of eye-in-head orientation during gaze shifts, it is essential to calibrate the head coil as well. This is especially important for example in studies concerned with eye-head coordination, or with the study of the vestibulo-ocular reflex (VOR).

For a successful calibration of the data it is important to feed endpoints into the networks with different gaze-head combinations. It is noteworthy that each subject has a idiosyncratic way of generating head-unrestrained gaze shifts, i.e. the contribution of the head to the total gaze shift varies from subject to subject. The calibration routine deals with this idiosyncratic behavior in a natural way. Since the subject will produce the same combinations during calibration and experiment there is no need for a guided calibration paradigm, in which subjects have to make coordinated eye-head movements (Bremen et al. 2007b).

Note that the head coil data could also be calibrated in a different way.

The signals of the three magnetic fields (horizontal, vertical and frontal), and thus each head orientation ( $\alpha$  and  $\epsilon$ ), falls on the surface of an ellipsoid. The axes of this ellipsoid are fully determined by the properties of the coils and the magnetic fields. The field strengths and amplifiers of the recording system could even be tuned such that the data fall on the surface of a sphere. Calibration of the head coil data could then be achieved by fitting the offset-corrected data to the surface of a sphere. However, determining the optimal fit parameters for the surface of an ellipsoid or even sphere is not straightforward (Gander et al., 1994; Turner et al., 1999) and crosstalk between the fields can lead to serious distortions of the data, which has to be accounted for. Although for the present study it was not necessary to implement this calibration scheme it could be necessary for smaller laboratory animals, like mice, in which it is not so easy to align the head with the calibration target.

*Stability of the assembly.* Due to anatomical differences around the nose and eyes, and to differences of the head implants, the DMI assembly needs to be separately manufactured for each individual animal. However, an assembly can be used indefinitely, as it will fit on the animal in exactly the same way for every recording session (Fig. 12). In this way, the acquisition of calibration data in the beginning of an animal's training is a cumulative process that yields better calibration as the number of sessions increases. From this point on the experimenter only has to quickly check the quality of calibration at the beginning of a recording session.

*Conclusions.* Studies employing animal subjects will have a direct benefit of the DMI method as presented here. Due to the absence of vulnerable lead wires the DMI method is more reliable. At the same time the temporal and spatial resolution is comparable to the resolution of the SSC technique. The calibration routine presented here is easy and straightforward to implement. We therefore think that the DMI method is a valuable alternative to the classical SSC technique.

## Acknowledgements

We thank dr. Eric Schaling, MD, for performing the implantation of the DMI ring in our monkeys. We are grateful to Hans Kleijnen, Dick Heeren and Stijn Martens for valuable technical assistance. This research was supported by a Marie Curie Early Stage Training Fellowship of the European Community's Sixth Framework Program (MEST- CT-2004-007825 PB), a VICI grant within the Earth and Life Sciences of NWO (ALW 865.05.003; AJVO, RFVDW, MMVW, TJVG), and the Radboud University Nijmegen (AJVO).



## Appendix A

*Simulation of eye and head saccades.* Eye and head saccades were generated by a simplified version of the feedback model of Van Gisbergen et al. (1981), in which the horizontal and vertical oculomotor and head subsystems were kept independent (to allow for substantial curvature in simulated trajectories). Parameters of the horizontal and vertical burst generators were different for eye and head (see Table 1A). We added random noise to these parameters to allow for a substantial amount of inter-trial variability, both in eye-head coupling, as well as in eye-head kinematics and trajectories. Here, only one basic module of the 2D eye-head system (consisting of independent horizontal and vertical eye, and horizontal and vertical head modules) is described.

We begin with defining the instantaneous motor error of the module  $m_e(t)$  as

$$m_e(t) = L - S(t), \quad (\text{A1})$$

**Table A1.** Parameters for gaze shift and head saccade generation.

Parameter	Value	Description
$t$	400	Trial duration in samples
$n$		Gaussian noise with $\mu = 0$ deg and $\sigma = 0.05$ deg
$G_h V_{pk}$	600	Horizontal gaze peak velocity in deg/sample of the burst generator
$G_v V_{pk}$	800	Vertical gaze peak velocity in deg/sample of the burst generator
$Gm_0$	7	Time to rise to 70% of peak velocity in sample of the burst generator
$GT_1$	150	Long time constant of the oculomotor plant → elastic properties of the muscles
$G_{on}$	80	Gaze shift onset re. begin of trial in samples
$H_h V_{pk}$	400	Horizontal head peak velocity in deg/sample of the burst generator
$H_v V_{pk}$	300	Horizontal head peak velocity in deg/sample of the burst generator
$Hm_0$	15	Time to rise to 70% of peak velocity in sample of the burst generator
$HT_1$	300	Long time constant of the head plant → elastic properties of the muscles
$H_{on}$	20	Head saccade onset re. gaze onset in samples
$H_{gain}$	0.5-1	Head gain drawn from a Gaussian distribution with $\mu = 0.75$ and $\sigma = 0.05$
$T_2$	20	Short time constant of oculomotor and head plant → plant's viscosity

with  $S(t)$  being the instantaneous saccade displacement evoked by a step of the target location,  $L$ . The brainstem burst generator (BG) encodes saccade velocity  $V(t)$  by the following saturating nonlinearity:

$$V(t) = v_{pk} \left( 1 - \exp\left(\frac{-m_e(t)}{m_0}\right) \right) \quad (A2)$$

This main sequence nonlinearity contains two free parameters: the angular constant  $m_0$  (in deg) and the asymptotic peak velocity  $v_{pk}$ . The output of the BG is fed into the neural integrator, which is implemented simply by integrating Eq. (A2) and solving for  $S(t)$  by substituting A(1):

$$S(t) = m_0 \cdot \ln \frac{A \cdot \exp\left(\frac{v_{pk} t}{m_0}\right)}{1 + A \cdot \exp\left(\frac{v_{pk} t - L}{m_0}\right)} \quad (A3)$$

$$\text{with } A = \frac{1}{1 - \exp\left(-\frac{L}{m_0}\right)}. \quad (A4)$$

This signal (the step) is added with the scaled burst from the BG ( $T_1^* V(t)$ ), and the resulting pulse-step signal is subsequently convolved with the oculomotor plant impulse response given by

$$h(t) = \frac{1}{T_1 - T_2} \left( \exp\left(-\frac{t}{T_1}\right) \right) - \exp\left(-\frac{t}{T_2}\right), \quad (A5)$$

with  $T_1$  the long time constant describing the elastic properties of the plant and  $T_2$  being the short time constant which is related to the plant's viscosity. Note that since in our simulations vertical and horizontal components are generated independently, oblique saccades do not exhibit component stretching. This detail is, however, not essential for the purpose of our study. It is more important to create a realistic interaction between the eye and head in order to account for subject idiosyncrasy. Therefore, the gain of the head saccades was varied randomly between 0.5 and 1 from trial to trial. At the same time all generated saccades obeyed the restriction that that eye-in-head orientation cannot exceed  $\pm 30$  deg in azimuth and elevation due to physical limitations. In order to mimic the experimental data better Gaussian noise  $n$  with  $\mu = 0$  deg and  $\sigma = 0.05$  deg was added to each sample of both the gaze shifts and head saccades. We obtained an estimate of  $\sigma$  by determining the SD of the fixation epoch over all trials of calibrated experimental gaze data (Bremen et al. 2007b). Note, however, that the experimentally measured SD of the gaze signal was only 0.03 deg and of the head signal only 0.02 deg. Table A1 lists the values of all parameters for both the simulated gaze shifts and head saccades. Parameters were chosen such that they roughly mimicked human saccade kinematics.

We simulated gaze shifts and head saccades to random target locations distributed between  $\alpha = \pm 80$  deg and  $\varepsilon = -55$  to 85 deg specified in double-polar coordinates (see section 2.2.3). We restricted ourselves to this range since we

wanted to run the exact same series of trials in our experimental setup with human subjects (see section 2.2). For the same reason fixation and target location were chosen such that their difference did not exceed  $\pm 50$  deg in  $\alpha$  and  $\epsilon$  since the field of view of a subject wearing the DMI assembly needed to pickup the ring's signal is restricted to about this range. Each simulated trial had a duration of 400 samples and started with a fixation epoch followed by a target epoch that always started at 80 samples and ended at 400 samples.

## 4.5 References

Allik J, Rauk M, Luuk A. Control and sense of eye movement behind closed eyelids. (1981) *Perception* 10:39–51.

Bos JE, Reulen JPH, Boersma HJ, Ditters BJ. Theory of double magnetic induction (DMI) for measuring eye movements: Correction for nonlinearity and simple calibration in two dimensions. (1988) *IEEE Trans Biomed Eng* 35:733–9.

Bour LJ, Van Gisbergen JAM, Bruijns J, Ottes FP. The double magnetic induction method for measuring eye movement—results in monkey and man. (1984) *IEEE Trans Biomed Eng* 31:419–27.

Bremen P, Van der Willigen RF, Van Opstal AJ. Using double-magnetic induction to measure head-unrestrained gaze shifts I. Theory and validation. (2007a) *J Neurosci Methods* 160: 75–84.

Bremen P, Van der Willigen RF, Van Opstal AJ. Applying double magnetic induction to measure two-dimensional head-unrestrained gaze shifts in human subjects. (2007b) *J Neurophysiol* 98:3759–69.

Collewijn H, Van der Mark F, Jansen TC. Precise recording of human eye movements. (1975) *Vision Res* 15: 447–450.

Gander W, Golob GH, Strebel R. Least squares fitting of circles and ellipses. (1994) *BIT Num Math* 34:558–578.

Irving EL, Zacher JE, Allison RS, Callender MG. Effects of scleral search coil wear on visual function. (2003) *Invest Ophthalmol Vis Sci* 44: 1933–1938.

Judge SJ, Richmond BJ, Chu FC. Implantation of magnetic search coils for measurement of eye position: An improved method. (1979) *Vision Res* 20: 535–38.

MacKay DJC. Bayesian interpolation. (1992) *Neural Comput* 4: 415–47.

Malpeli JG. Measuring eye position with the double magnetic induction method. (1998) *J Neurosci Methods* 86: 55–61.

Murphy PJ, Duncan AL, Glennue AJ, Knox PC. The effect of scleral coil lens wear on the eye. (2001) *Br J Ophthalmol* 85:332–335.

Ottes FP, Van Gisbergen JAM, Eggermont JJ. Collicular involvement in a saccadic colour discrimination task. (1987) *Exp Brain Res* 66:465–478.

Reulen JP, Bakker L. The measurement of eye movement using double magnetic induction. (1982) *IEEE Trans Biomed Eng* 29:740–744.

Robinson DA. A method of measuring eye movement using a scleral search coil in a magnetic field. (1963) *IEEE Trans Biomed Eng* 10:137-145.

Turner DA, Anderson IJ, Mason JC, Cox MG. An algorithm for fitting an ellipsoid to data. (1999) <http://www.scientificcommons.org/42871609>, DOI: 10.1.1.36.2773.

Van Gisbergen JA, Robinson DA, Gielen S. A quantitative analysis of generation of saccadic eye movements by burst neurons. (1981) *J Neurophysiol* 45:417–442.

Van Wetter SMCI, Van Opstal AJ. Experimental test of visuomotor updating models that explain perisaccadic mislocalization. (2008) *J Vision* 8:1–12.





## Chapter 5

# Pinna Cues Determine Orienting To Double Sound Sources

### Abstract

To program a goal-directed orienting response toward a sound source embedded in an acoustic scene the audiomotor system should detect and select the target against a background. Here we focus on whether the system can segregate synchronous sound sources in the mid-sagittal plane (elevation), a task requiring the auditory system to dissociate the pinna-induced spectral localization cues. Human listeners made rapid head-orienting responses towards either a single sound source (broad-band buzzer, or Gaussian noise), or towards two simultaneously presented sounds (buzzer and noise) at a wide variety of locations in the mid-sagittal plane. In the latter case, listeners had to orient to the buzzer (target), and ignore the noise (non-target). In the single-sound condition localization was accurate. However, in the double-sound condition response endpoints depended on relative sound level and spatial disparity. The loudest sound dominated the responses, regardless of whether it was the target or the non-target. When the sounds had about equal intensities and their spatial disparity was sufficiently small, endpoint distributions were well described by weighted averaging. However, when spatial disparities exceeded about 45 deg, response endpoint distributions became bimodal. Similar response behavior has been reported for visuomotor experiments, for which averaging and bimodal endpoint distributions are thought to arise from neural interactions within retinotopically organized visuomotor maps. We show, however, that the auditory-evoked responses can be well explained by the idiosyncratic acoustics of the pinnae. Hence, basic principles of target representation and selection for audition and vision appear to differ profoundly.

---

*Adapted from:* Bremen P, Van Wanrooij MM, Van Opstal AJ (2010)  
J Neurosci 30:194-204.



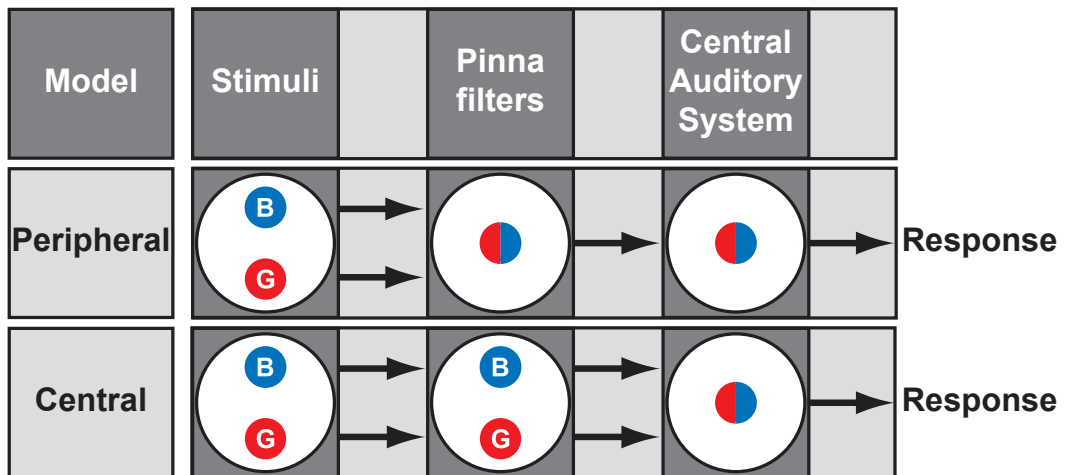
## 5.1. Introduction

Target selection has been addressed extensively in the visuomotor literature (Becker and Jürgens, 1979; Findlay, 1982; Ottes et al., 1984, 1985; Chou et al., 1999; Aitsebaomo and Bedell, 2000; Watanabe, 2001; Arai et al., 2004; Nelson and Hughes, 2007). When two visual targets are presented in spatial-temporal proximity, saccadic eye movements often terminate between the two target locations (averaging). Varying target features (saliency, onset asynchrony, spatial disparity, size), task constraints (instruction), or saccade-reaction times, systematically affect response endpoint distributions (Findlay, 1982; Ottes et al., 1984, 1985). This behavior is thought to arise from neural interactions within spatially organized maps, like in midbrain superior colliculus (Ottes et al., 1984; Lee et al., 1988; Van Opstal and Van Gisbergen, 1990; Glimcher and Sparks, 1993; Kim and Basso, 2008).

In contrast to the visual system's retinotopic organization, the auditory system is tonotopic. Hence, sound locations are derived from implicit acoustic cues. Interaural time and level differences determine sound-source azimuth (Middlebrooks and Green, 1991; Blauert, 1997); pinna-related spectral-shape cues encode elevation (Shaw, 1966; Blauert, 1969; Wightman and Kistler, 1989; Middlebrooks, 1992; Hofman and Van Opstal, 1998, 2002; Kulkarni and Colburn, 1998; Langendijk and Bronkhorst, 2002). Two simultaneous sounds at different azimuth locations induce the percept of one phantom source at the so-called summing location (Blauert, 1997). For example, identical sounds symmetrically presented left and right of the listener are perceived as a single sound-source at straight-ahead. Its location systematically varies with relative intensities and timings of the speakers, which can be understood from sound-wave interference at the ear canals.

Here we test whether similar principles apply to elevation, as the effect of spatial summation for complex spectral-shape cues is far from obvious. The only study addressing this issue focused on discriminating sounds in virtual acoustic space, and noted that listeners could not segregate sounds in the mid-sagittal plane (Best et al., 2004). As previous research did not assess sound-localization performance, we took a different approach, by characterizing free-field sound-localization with head saccades evoked by two simultaneous, but different, sounds in the mid-sagittal plane: a broadband buzzer (BZZ, target), and a Gaussian white noise burst (GWN, non-target). Both sounds were perceptually well distinguishable and localizable when presented alone. We systematically varied their locations, relative level and spatial disparities over a large range. Figure 1 considers two hypotheses to explain localization percepts: peripheral-acoustics, which holds, unlike vision, that stimulus interactions at the sensory periphery determine percepts, vs. central-analysis, in which, like vision, localization percepts rely on neural interactions.

Our data indicate that, except when BZZ intensity well exceeded GWN, listeners were unable to ignore the non-target, regardless of response-reaction times and spatial disparity. Interestingly, simulations with a pinna-based



**Figure 1.** Two competing accounts for perceiving a phantom sound-source at a weighted average location in the mid-sagittal plane. **(Top)** In the peripheral-interaction model the double-sound spectra of BZZ (B) and GWN (G) yield an amplitude spectrum corresponding to the weighted-averaged location (mixed circle). **(Bottom)** In the neural-interaction scheme the periphery preserves spectral shape of either source. A weighted-averaged percept emerges through neural interactions in an audio-spatial representation.

similarity model indicate that observed response patterns may be fully understood from idiosyncratic pinna acoustics. These results suggest a marked difference between target-selection mechanisms for vision and audition.

## 5.2. Methods

*Listeners.* Four listeners (ages 29 to 32, three male, one female) with normal hearing as indicated by their audiometric curves (hearing threshold  $\leq 20$  dB between 250 and 11300 Hz, eleven 0.5 octave steps) participated in the experiments. Two of the authors (PB & MW) served as listeners, the other two listeners were naive about the purpose of the study.

*Apparatus and sound generation.* All experiments were performed in a dark 3 m x 3 m x 3 m room lined with acoustic foam (UXEM Flexible Foams B.V.; Lelystad, The Netherlands) that attenuated sound reflections above 500 Hz. Background noise level was 30 dBA. Sounds were presented from a total of 29 speakers (SC 5.9, Visaton; Art. No. 8006) mounted on a vertical motorized circular hoop, 2.5 m in diameter. Loudspeakers were mounted in 5 deg increments from -55 to 85 deg on the front (double-polar coordinates, see below). The listener was seated in the center of this hoop (head-centered) on a straight-back chair. In the present study the hoop was always aligned with the mid-sagittal plane.

For easier visualization, target locations and head-movement responses were transformed to double-polar coordinates (Knudsen and Konishi, 1978). The vertical location is given by the elevation coordinate  $\varepsilon$ , i.e. the angle formed by the center of the hoop (listener's head), sound source/response location, and the

horizontal plane. Positive (negative) elevation values indicate locations above (below) the listener's interaural axis. The horizontal location is given by the azimuth coordinate  $\alpha$ , which is the angle formed by the center of the hoop, the sound source, and the median plane. In the current study the speaker azimuth was always at  $\alpha = 0$  deg.

Speaker selection was done with a custom-made program written in C++ running on a PC (2.8 GHz Intel Pentium D, Dell). The same program was used to record head saccades (see section Behavioral testing and paradigms), directional transfer functions (see section Directional transfer functions) and sound playback. For the presentation of stimuli a stored wav file made offline with MatLab (The MathWorks, Natick, MA, USA) was sent to a real-time processor (RP2.1 System3, Tucker Davis Technologies; Alachua, Florida, USA) at a sampling rate of 48,828 kHz. After attenuation by custom-built amplifiers the audio signal was sent to the selected speaker. In a given trial either one or two speakers could be selected to play the stimuli. Transfer characteristics of the speakers differed by less than 0.3 dB (root-mean-square) from 1 to 15 kHz. Accordingly no attempt was made to correct for these small speaker differences.

To ensure microsecond timing precision relevant trial information was sent from the PC to a custom-made micro-controller that initiated and controlled events in the trial when all hardware components were stand-by.

*Stimuli.* All sounds had 50 ms duration, including 5 ms smooth sine/cosine-squared onset/offset ramps. We used two different sounds in the localization experiments: a Gaussian white-noise burst (GWN, 0.5-20 kHz bandwidth) and a periodic quasi-noise burst (BZZ) that had the same amplitude spectrum as the GWN but differed in its temporal structure (Zwiers et al., 2001). The quasi-noise had a fixed periodicity of 20 msec (making it sound like a 50 Hz buzzer). Sound levels were varied between 35 and 55 dBA in 5dB steps for the GWN and from 32 to 52 dBA in 5 dB steps for the BZZ (measured with Brüel & Kjær BK2610 sound amplifier, and Brüel & Kjær BK4144 microphone; at the location of the listener's head). In double sound trials (see section Behavioral testing and paradigms) we always held either the BZZ or the GWN constant at 42 or 45 dBA, respectively, while varying the level of the other sound. This resulted in a total of nine different combinations, for which the level difference,  $\Delta L$ , could assume the values -13, -8, -3, +2 and +7 dB. Positive/negative differences indicate that the BZZ/GWN is louder. In additional trials we created summed sounds that were constructed by linearly adding (in the temporal domain) the GWN and the BZZ wave forms at all different level combinations. These combined sounds were always played from one randomly selected speaker and served as a control (see Fig. 2C,D). The different combinations of these stimuli were ordered as indicated by the following example: B32G45, contains a BZZ at 32 dBA and GWN at 45 dBA, i.e.  $\Delta L = -13$  dB. It's numerical order on the scale between BZZ-only (nr. 1) to GWN-only (nr. 11) is 10. B52G45 is nr. 2, etc.

*Behavioral testing and paradigms.* The behavioral testing procedure required the listeners to orient a head-fixed laser pointer towards the perceived location of the BZZ “as quickly and as accurately as possible”. If no BZZ was perceived the listener simply had to localize the presented sound. The laser pointer (LQB-1-650, World Star Tech; Toronto, Ontario, Canada) was attached to a modified lightweight sunglasses frame (glasses were removed) and projected its red beam onto a small, frame-attached disk (diameter 1 cm) at about 30 cm in front of the listener’s nose. This assured that no visual cues (e.g. reflections on the wall and hoop) influenced the localization behavior of the listeners. Head movements were measured with the magnetic search-coil technique (Robinson, 1963). To that end a small custom-made coil was wound around the laser pointer and connected to an eye monitor (EM7, Remmel Labs; Katy, Texas, USA) that was also used to drive three pairs (horizontal, vertical, frontal) of field coils mounted alongside the edges of the experimental chamber. The demodulated data were A/D-converted at 1000 Hz (RA16 System3, TDT) and stored on disk for further offline analysis.

To calibrate the head coil a calibration session was performed prior to each experiment. The listener was asked to orient the laser pointer towards light-emitting diodes (LEDs) mounted in front of the loudspeakers on the hoop. 56 points distributed in the frontal hemisphere were sufficient to calibrate the horizontal and vertical components of head movements to within 0.6 deg.

In the experimental session a trial started with a green fixation LED at  $\alpha = 0$  deg and  $\varepsilon = 0$  deg. The listener aligned his head with this fixation LED after which he initiated the trial sequence by a button press. In a trial one of the following three stimulus types could be presented:

1) *Single sound trials:* The sound - either the GWN or the BZZ - was emitted from one single speaker with varying level (see Stimuli section) at one of nine locations ( $\varepsilon = [-50, -35, -20, \dots, 70]$  deg) on the frontal mid-sagittal plane. This amounted to a total of 45 single-target trials per stimulus type. The responses towards these stimuli were used to assess the listener’s standard localization behavior.

2) *Double sound test trials:* Two speakers emitted the test sounds simultaneously. Speaker locations were identical to the single-sound locations. The separation between the two speakers could thus vary between 15 and 120 deg elevation in steps of 15 deg. For example holding speaker 1 at -50 deg the following eight locations were used for speaker 2: -35, -20, -5, 10, 25, 40, 55, 70 deg. Additionally, we varied the level difference as described in the Stimuli section. All possible spatial and level configurations were tested, leading to a total of 720 trials per listener.

3) *Double-sound control trials:* A single speaker emitted the linear sum of BZZ and GWN at the same level differences as in the double speaker trials. Ninety trials were performed with these stimuli.

In total a measurement session contained 900 localization trials and lasted for about one hour.

Note that the three largest spatial disparities in double-sound trials had only a limited amount of possible spatial combination per level difference

(90 deg = 6 trials; 105 deg = 4 trials; 120 deg = 2 trials). Therefore, listeners MW and PB performed an additional experiment in which the level difference in double-sound trials was fixed at -3 dB and spatial disparity for the double-sound trials was limited to 90, 105 and 120 deg (total of 528 trials per listener).

*Data analysis.* Head orientation was calibrated using the data obtained in the calibration experiment. Combinations of raw data (AD values horizontal, vertical and frontal components) and known LED locations (azimuth and elevation in deg) were used to train two three-layer neuronal networks for azimuth and elevation, respectively. The networks were trained by the Bayesian-Regularization implementation of the back-propagation algorithm (MatLab, Neural Networks Toolbox) in order to avoid over-fitting (MacKay 1992). In addition to a linear mapping from AD values to degrees the networks also accounted for small inhomogeneities in the fields and crosstalk between the three channels. The thus trained networks were then used to calibrate the experimental data.

A custom-written MatLab script was used to automatically detect saccades in the calibrated data by employing a preset velocity criterion (15 deg/s) to saccade onset and offset. Detected saccades were visually inspected for errors and corrected if necessary. Saccades with a reaction time shorter than 150 ms were discarded as anticipatory responses.

*Response normalization.* For the data analysis shown in Figs. 7 and 8, the head-saccade endpoints in double-sound trials were normalized by:

$$\hat{R} = \frac{R - (B + G)/2}{(B - G)/2} \quad (1)$$

with  $B$ ,  $G$  the elevation of BZZ and GWN stimuli, respectively, and  $R$  the head-movement response elevation. For a response directed to the GWN non-target ( $R=G$ ),  $\hat{R} = -1$ , while for a response to the target BZZ ( $R=B$ ),  $\hat{R} = +1$ .

Regression analysis and statistics. In all regression analyses parameters were found by minimizing the mean squared error (Press et al., 1992).

We performed a linear regression analysis on the stimulus-response relation to quantify localization behavior in elevation:

$$R = a \cdot T + b \quad (2)$$

with  $R$  and  $T$  response elevation and target elevation, respectively. The slope,  $a$ , is the response gain, and offset,  $b$  (in deg), the response bias. To determine the goodness of fit we calculated the correlation coefficient between fit and data and the mean absolute error.

To test whether responses could be described by a weighted average of the locations of BZZ, TBZZ, and GWN, TGWN, we first determined the optimal

weights separately for each  $\Delta L$  as described by the weighted average formula:

$$\tilde{R}_{AVG} = w_B \cdot \tilde{T}_{BZZ} + (1 - w_B) \cdot \tilde{T}_{GWN} \quad (3)$$

with  $\tilde{X} \equiv (X - \mu_X) / \sigma_X$  the dimensionless z-score of variable  $X$  ( $\mu_X$  mean,  $\sigma_X$  standard deviation), and  $w_B$  and  $1 - w_B$  are the dimensionless weights (partial correlation coefficients) of BZZ and GWN, respectively.  $R_{AVG}$  is the weighted-average prediction for the response. Then, we evaluated to what extent the measured responses in the entire data set (pooled over all  $\Delta L$  per listener) could be best described by either the BZZ location (i.e. the target), the GWN location (i.e. the non-target), the weighted average prediction,  $R_{AVG}$ , of Eq. 3, or the level difference by employing a normalized (z-transformed) multiple linear regression analysis according to

$$\tilde{R} = p \cdot \tilde{T}_{BZZ} + q \cdot \tilde{T}_{GWN} + w \cdot \tilde{R}_{AVG} + m \cdot \Delta \tilde{L} \quad (4)$$

To obtain confidence limits of the coefficients we used a bootstrap method. To that end 1000 data sets, randomly drawn from the responses (with replacement) were generated and subjected to the regression analysis of Eq. 4. The SD of the resulting set of 1000 coefficients was used to estimate the confidence levels of the partial correlation coefficients. with  $R$  the response elevation,  $R_{AVG}$  the weighted-average response prediction of Eq. 3,  $\Delta L$  the level difference, and  $p$ ,  $q$ ,  $w$  and  $m$  the partial correlations coefficients.

*Directional transfer functions (DTFs).* For all listeners we measured their DTFs, simultaneously from both ears. To that end, we sampled 360 locations in the frontal hemisphere. A small probe microphone (Knowles EA1842) connected to a small rubber tube (1.5 mm outer diameter, length 5.5 cm) was positioned at the entrance of the external auditory meatus of each ear, and fixed with tape without obstructing or deforming the pinnae. The listener's interaural axis was aligned with  $\alpha = 0$  deg and  $\varepsilon = 0$  deg and the head was supported in this position by a neck rest. As a probe stimulus a periodic flat spectrum Schroeder-phase signal (Schroeder 1970) was used that consisted of 20 periods with a duration of 20.5 ms each (total duration 410 ms). The probe was presented at a sound pressure level of 50 dBA. The first and last periods were  $\sin^2/\cos^2$ -ramped (5 ms) and were discarded in the analysis. The same measurements were also performed without the listener in place. In that case the two microphones were positioned at the location of the listener's interaural axis. These latter measurements were used to correct for speaker- and microphone characteristics as well as location-specific reflections from the measured transfer functions.

The recorded microphone signal was amplified (custom-built pre-amplifiers), band-pass filtered (Krohn-Hite 3343, passband 0.2-20 kHz) and sampled at 48.828 kHz (RP2.1 System 3, TDT). The subsequent offline analysis was from



performed in MatLab. First, the average signal over Schroeder periods 2 to 19 was calculated (1024 samples) for listener and microphone measurements. Subsequently the magnitude spectra were computed by means of the fast Fourier transform (512 bins). The obtained spectra were smoothed using a simple Gaussian filter with a constant Q-factor of 8 (Algazi et al., 2001) and converted to sound level. Then the microphone measurements were subtracted for all locations. Finally, the DTFs were obtained by subtracting the mean spectrum of the whole data set from each measurement. In that way, only the direction dependent information of the pinnae remained (Middlebrooks, 1992).

*Similarity model of sound localization.* Since acoustic pressure waves add linearly at the ear, we reconstructed the sensory spectrum for a given BZZ/GWN combination by first adding the corresponding measured sensory spectra of the same single targets. To that end, the two selected DTFs (on linear magnitude scale) were filtered (i.e. multiplied) with the magnitude spectrum of BZZ and GWN and corrected for the actual levels used in the experiment. These operations were performed in the frequency domain before log-(dB) transformation. After summation we performed a log-transformation on the combined amplitude spectrum to approximate the neural input spectrum for the double sound (here termed 'double-DTF'). For simplicity, in this analysis we only used the right-ear DTFs measured at azimuth 0 deg.

We then calculated the standard deviation over a frequency range from 3 to 12 kHz of the difference between the double-DTF (considered as the sensory input), and all measured single-DTFs (presumed to be stored in the brain as neural templates). This SD is taken as a measure of similarity (Langendijk and Bronkhorst, 2002), rather than the correlation coefficient between sensory spectrum and DTFs (Middlebrooks, 1992; Hofman and Van Opstal, 1998). The resulting similarity index (SI) was scaled so that it ranged from zero (no similarity) to one (identical). The idea behind this procedure is that a conceptually similar analysis is thought to be performed by the ascending auditory pathway to estimate the most likely elevation angle, given the sensory spectrum (Middlebrooks, 1992; Hofman and Van Opstal, 1998, 2003; Kulkarni and Colburn, 1998; Langendijk and Bronkhorst, 2002). According to this model, the DTF template (with its associated location) that yields the highest similarity index has also the highest probability for being the source's real location.

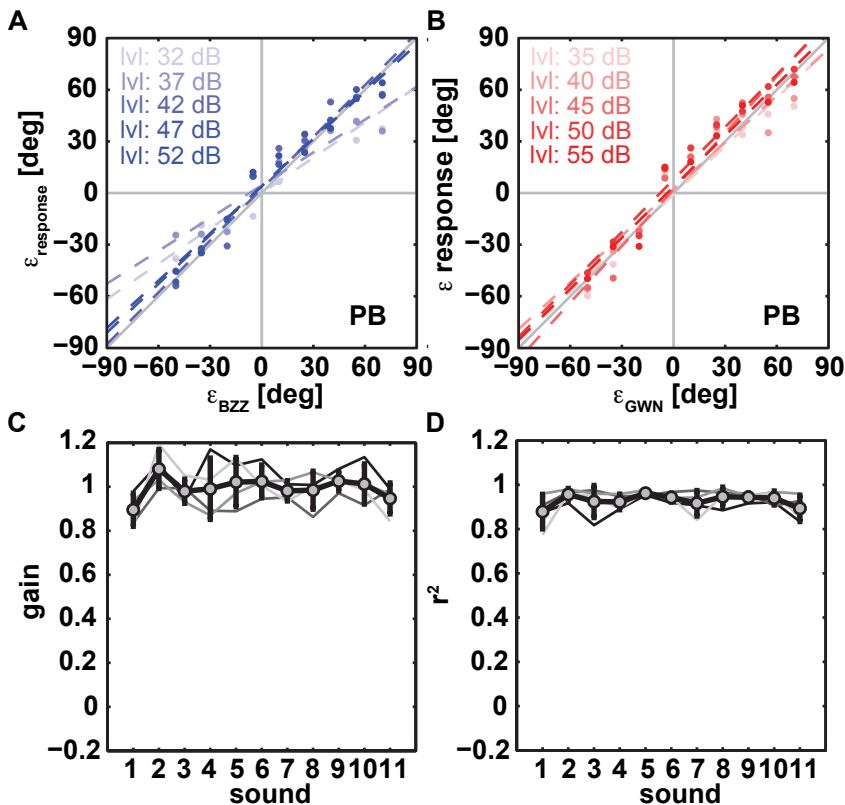
*Correction procedure.* We assessed the quality of the SI as a predictor of behavior by determining the SI that corresponded to a given response. Note that the SI analysis is based entirely on the acoustics and does not account for a potential response bias in the head-motor response (the gain,  $a$ , and offset,  $b$  in Eq. 2). Since some listeners could display a significant response bias, we corrected responses (for both single- and double-sound trials) by using the mean gain and bias obtained from all single-sound trials (data shown in Fig. 2). In this way,



the double sound responses were normalized with respect to the single-target localization results. Since our SI maps had a finite resolution we calculated the distance between response location and all sampled DTF locations and selected the closest location. The thus obtained SI's were binned (bin size 0.1) and the resulting histograms were normalized to the maximal number of occurrences (Figs. 9 and 10).

### 5.3 Results

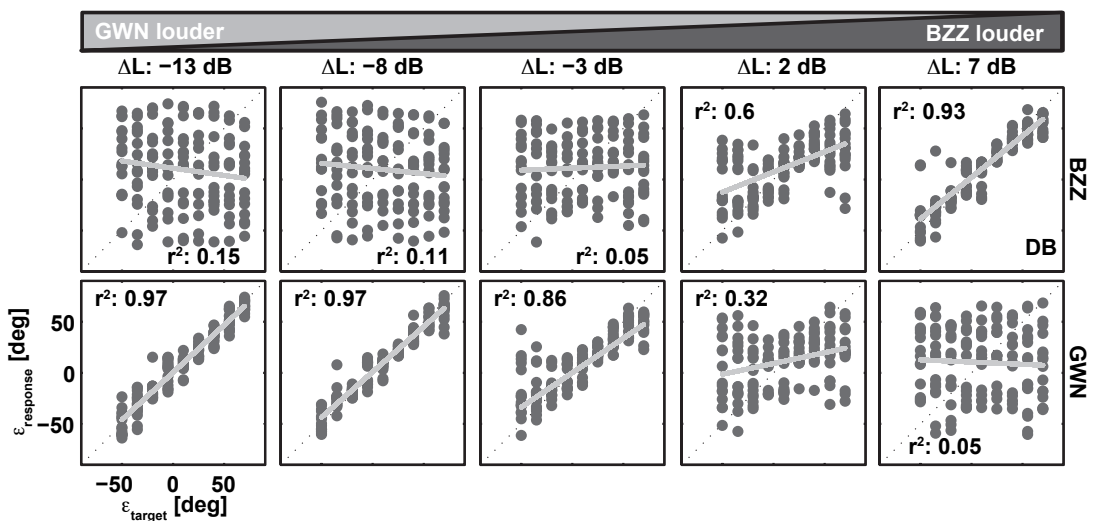
*Single-sound localization.* We first assessed localizability of BZZ and GWN by analyzing the stimulus-response relations for all single-sound trials. In these trials either the BZZ or the GWN were presented alone, or superimposed, at all nine combinations used in the double sound trials (see Methods). Figure 2 shows representative stimulus-response plots for listener PB. In this figure response



**Figure 2.** Standard localization behavior of listener PB to single-speaker sounds at five levels (BZZ: 32, 37, 42, 47, 52 dBA; GWN: 35, 40, 45, 50, 55 dBA). Stimulus response plots for BZZ (A) and GWN (B). Different color shades indicate different levels. Gains and correlation coefficients are close to one, biases close to 0 deg, indicating good localization performance. Dashed lines indicate linear regression lines. (C) Gains for all single speaker sounds. Subscript numbers indicate the level of BZZ (1) and GWN (11) and the summed sounds (see Methods, for details). Lines in different shades of grays are from different listeners. Thick black line with gray circles is the average over all four listeners. Error bars denote one SD. (D) Correlation coefficients obtained for single-sound stimulus-response plots. Both gains and correlations are close to one, indicating high localization accuracy and precision, respectively.

elevation for the BZZ (Fig. 2A) and for the GWN (Fig. 2B) are plotted against target elevation for all levels tested (different shades). The gains for the BZZ at soft levels (32 and 37 dBA) were somewhat lower (0.69 and 0.64) than at higher sound levels. These levels were close to the background noise of about 30 dBA that was always present in the experimental chamber, so that the signal-to-noise ratio (NSR) was low too. The effect of low SNR on elevation localization has been reported before (Zwiers et al., 2001). Note that at the other levels elevation gain was close to one and the biases were close to zero deg for both BZZ and GWN indicating excellent localization performance. Figure 2C and D show gains and correlation coefficients for BZZ, GWN and the summed control sounds for the four listeners (different shades of gray) and the average across listeners (thick black line with gray circles). Both gain and  $r^2$  are close to one, and do not depend systematically on stimulus level. These data demonstrate that all listeners accurately localized BZZ and GWN, as well as the summed control sounds when these were presented alone.

*Double-sound localization.* The listener's task in the double-sound trials was to localize the BZZ and to ignore the GWN. To test how well listeners performed in this task we plotted the listener's localization response of all double sound-trials as a function of the actual BZZ location. The results of this analysis for listener DB are shown in Fig. 3 (top row). Each column shows the data for a given  $\Delta L$  (left: -13 dB, GWN is much louder than BZZ; right (+7 dB, BZZ louder than the GWN). It is apparent that for negative  $\Delta L$  the responses do not correlate with the actual BZZ location. However, when BZZ levels start to exceed the GWN ( $\Delta L +2$ dB) responses start to be directed toward the BZZ location. Only at a  $\Delta L$  of +7 dB 93% of the variance in the response data can be explained by the BZZ location. In other words, the listener could not perform the task when the BZZ



**Figure 3.** Localization behavior of listener DB in double-speaker trials. Each column shows stimulus-response data for one level difference between BZZ and GWN. **Top row.** Responses plotted against BZZ location. **Bottom row.** Responses plotted versus GWN location.

sound level was comparable to, or lower than, the GWN level.

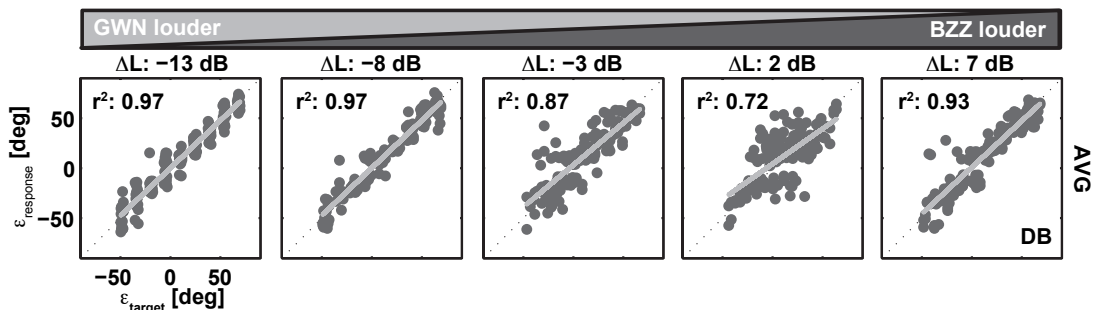
One may therefore wonder whether in these cases the listener instead of localizing the BZZ actually localized the GWN. The bottom row of Fig. 3 re-plots the responses as a function of GWN location, which shows that the response pattern is now reversed. When the GWN stimulus was much louder than the BZZ ( $\Delta L = -13$  and  $-8$  dB), the non-target GWN location explained 97% of the responses variance. Only for the case of a much louder BZZ target ( $\Delta L = +7$  dB), responses can not be accounted for by the GWN location.

Interestingly, although at the extreme intensity differences localization was guided to either the GWN (large negative  $\Delta L$ ), or the BZZ (large positive  $\Delta L$ ), at near equal loudness values of  $\Delta L = -3$  dB and  $+2$  dB the correlation coefficient decreased appreciably. We therefore wondered whether, like in the visuomotor system, a weighted average of the stimulus locations (Eq. 3) could perhaps better explain the entire set of responses.

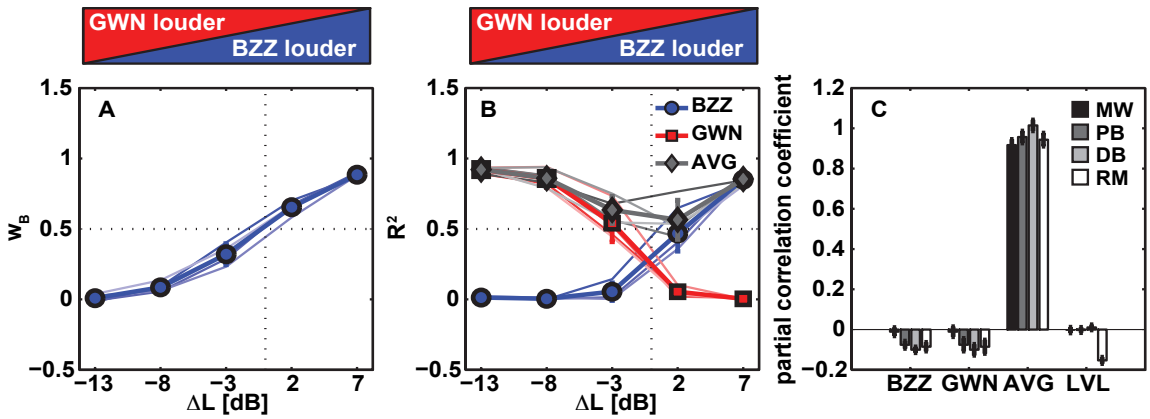
To obtain the weights for the two targets at each  $\Delta L$  we first performed a multiple linear regression analysis for each listener (Eq. 3). The variables in the regression were the respective locations of BZZ and GWN. Note that because of the large range of stimulus disparities used in the experiments, a potential spurious correlation between BZZ and GWN locations was eliminated. We then used the optimal weights ( $w_B$  and  $1-w_B$ ) from this regression analysis to calculate the weighted averaged target locations for linear regression (Eq. 2) with the listener's responses.

Figure 4 shows the results of the weighted average predictions for listener DB. Note that the data correlate much better with the averaged locations over the entire range of  $\Delta L$  values than with the actual target positions (cf. Fig. 3). At extreme  $\Delta L$  values the correlation coefficients of the weighted average prediction are identical to the corresponding BZZ/GWN single target result. Importantly, however, the weighted average estimate also accounts significantly better for the data variance at the smaller level differences ( $\Delta L -3$  and  $2$  dB) than either the BZZ or the GWN locations.

The resulting weight of the BZZ,  $w_B$  (Eq. 3), for all listeners are shown in Fig. 5A as function of  $\Delta L$ . The thin lines in different shades of blue are the data from individual listeners; the thick line with markers indicates the mean over four listeners. It is obvious that the contribution of the BZZ to the responses strongly



**Figure 4.** Localization behavior of listener DB in double-speaker trials as a function of a weighted average target prediction (Eqn. 3) of BZZ and GWN location for five level differences.



**Figure 5.** (A) Partial correlation coefficients for BZZ location on localization response obtained with regression analysis of Eqn. 3. The weight of the GWN is  $1-w_B$ . Thin lines indicate data from individual listeners, thick lines with markers show pooled data. The influence of a target depends in a sigmoid fashion on the level difference. Note that the point of equal contribution (weight = 0.5) is at  $\Delta L = 0$  dBA. (B) Correlation coefficients obtained from the linear regression shown in Fig. 2. Color convention as in (A). The coefficients for the weighted-average prediction are depicted in gray colors. For BZZ and GWN stimuli the correlation decreases with decreasing level of the corresponding single target. The correlation coefficient of the weighted-average prediction is equal to the single target values at extreme level differences. At  $\Delta L$  of -3 and +2 dBA, however, the weighted average prediction correlates better with the responses than either BZZ or GWN. (C) Partial correlation coefficients of Eqn. 4 for all four listeners (MW, PB, DB, RM). Most of the data can be explained by the weighted-averaged prediction, with coefficients  $>0.8$  for all four listeners.

depends on  $\Delta L$ . For example at  $\Delta L +7$  dB the BZZ dominates with a weight close to one. The opposite can be seen at  $\Delta L -13$  dB, for which the BZZ has a weight close to zero. The reverse behavior holds for the GWN since  $w_G=1-w_B$ . The crossing point, i.e. the point at which both targets contribute equally (0.5), is at a  $\Delta L$  close to zero.

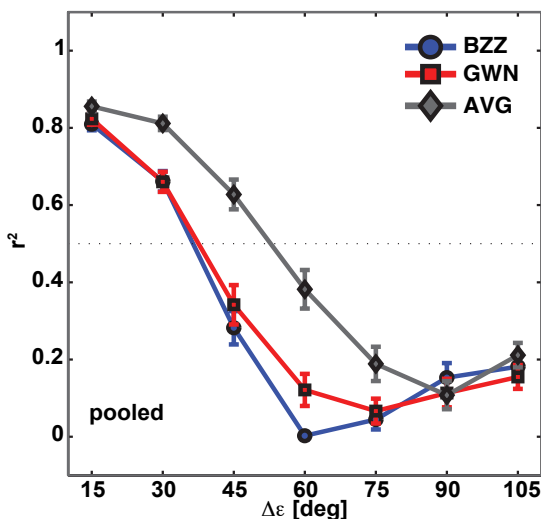
The superiority of the weighted average prediction can also be seen in Fig. 5B, which also shows the  $r^2$  values (variance explained) of the linear regression (Eq. 2) as a function of  $\Delta L$ . Again individual data are depicted as thin lines of different colors and the mean over all listeners as thick lines with markers. Correlation coefficients for the BZZ are shown in blue, those for the GWN are shown in red, and those for the weighted average prediction in gray. It can be seen that the coefficient for the single stimuli decreases towards zero with decreasing level of that respective stimulus. The goodness-of-fit for the weighted-averaged locations drops slightly for the smallest  $\Delta L$  values of -3 and 2 dB (to about 0.6; see below). Yet, these values are significantly larger than the ones obtained for BZZ or GWN locations alone.

To assess that the weighted averaging model captures the essence of the data best, we performed an extended multiple linear regression analysis (Eq. 4), in which the variables were the actual BZZ and GWN locations, the weighted-average target prediction (from Eq. 3), and  $\Delta L$ . Results for each listener are shown as bars in Fig. 5C. This analysis makes clear that the responses can indeed best be explained by the weighted average prediction, since the partial correlation coefficients for this variable exceeded 0.8 for all four listeners.

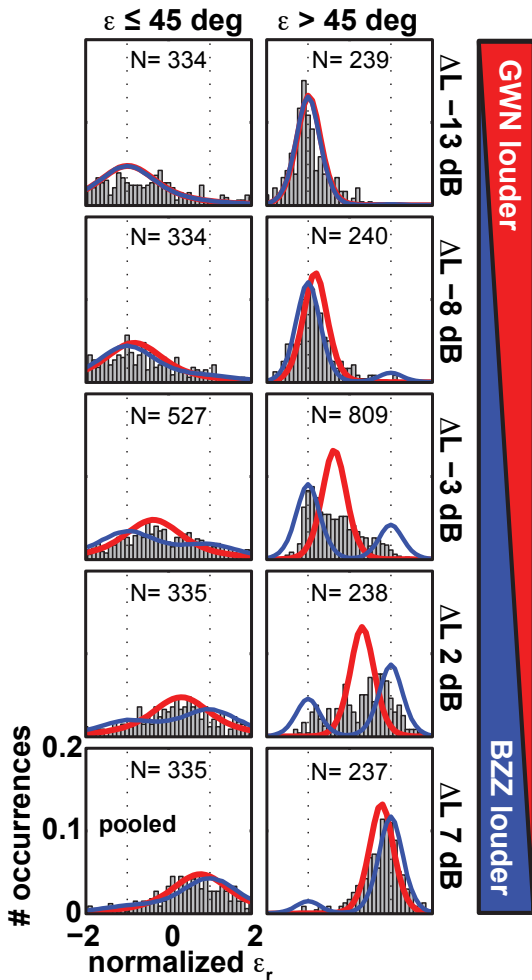
Yet, the partial correlation coefficients for the weighted average predic-

tion shown in Fig. 5A, as well as the goodness-of-fit (Fig. 5B), and coefficients obtained from the extended regression of Fig. 5C deviated significantly from the optimal value of one. Note that these results were based on pooled target conditions, in which we included all spatial disparities,  $\Delta\varepsilon$ . Possibly, target averaging breaks down at large spatial disparities, in which case also the averaging model would provide a poor predictor for the data. To test for this possibility, we re-evaluated the correlation coefficients for BZZ, GWN and weighted average prediction (Eq. 2), but now for each of the tested  $\Delta\varepsilon$  separately. In this analysis we combined the three largest  $\Delta\varepsilon$  values (90, 105 and 120 deg) for the linear regression analysis, as we had only two target configurations for  $\Delta\varepsilon=120$  deg (see Methods). The results pooled across all listeners are shown in Fig. 6. Coefficients for the BZZ are shown in blue, those for the GWN and the average prediction in red and gray, respectively. Values decline rapidly for BZZ and GWN; they drop below 0.4 at  $\Delta\varepsilon$ 's  $> 45$  deg. In contrast, the coefficients obtained with the weighted average prediction are higher than the single-target coefficients for all  $\Delta\varepsilon < 90$  deg. Only for the largest  $\Delta\varepsilon$ 's (90, 105 and 120 deg) BZZ, GWN and average prediction are indistinguishable from each other (all are close to 0.1-0.2). This indicates that at larger  $\Delta\varepsilon$  none of the three models is able to predict the responses.

To further study what happens at the large spatial disparities, we divided the data into two sets:  $\Delta\varepsilon \leq 45$  deg, and  $\Delta\varepsilon > 45$  deg. At this criterion, the weighted average model could explain at least 60-65% of the data variance (Fig. 6). Figure 7 shows the probability distributions (pdfs) of normalized (Eq. 1) head-saccade endpoints pooled across listeners. The data are shown separately by level difference,  $\Delta L$  (rows; GWN loudest top; BZZ loudest bottom) and spatial separation,  $\Delta\varepsilon$  (columns;  $\Delta\varepsilon \leq 45$  deg left;  $\Delta\varepsilon > 45$  deg right). A value of +1.0 indicates a response to BZZ, a value of -1.0 to GWN (see Methods). In addition to the measured endpoints we performed numerical simulations based on the weighted-average prediction (red lines) and on a bimodal prediction (blue lines). In these simulations we used SD = 12 deg of single-sound responses (pooled across all listeners) to generate Gaussian response distributions, and simulated 104 double-sound re-



**Figure 6.** Linear regression gains (Eqn. 2) for BZZ (blue), GWN (red) and weighted-averaged prediction (gray) as a function of  $\Delta\varepsilon$  pooled across all four listeners. Data at  $\Delta\varepsilon = 105$  deg is averaged over  $\Delta\varepsilon = (90, 105 \text{ and } 120 \text{ deg})$ . Error bars denote one SD.



**Figure 7.** Normalized saccade-endpoint distributions for the five  $\Delta L$  values (rows), pooled over all listeners. Data is separated in  $\Delta \epsilon \leq 45$  deg (left column) and  $\Delta \epsilon > 45$  deg (right column). Black dotted lines indicate normalized target locations with 1.0 denoting the BZZ and -1.0 denoting GWN. The red line indicates a simulated weighted-averaged target prediction ( $\mu = \epsilon_{\text{AVG}}$ ,  $\sigma = 12$  deg,  $N = 104$ ). The blue line shows a simulated bimodal prediction ( $\mu_1 = w_{\text{BZZ}} \cdot \epsilon_{\text{BZZ}}$ ,  $\mu_2 = w_{\text{GWN}} \cdot \epsilon_{\text{GWN}}$ ,  $\sigma = 12$  deg,  $N = 104$ ).

sponses for each condition in the experiments. At large  $\Delta L$  values (top and bottom rows), the weighted average simulations and measured data are in good agreement irrespective of  $\Delta \epsilon$ , and show a unimodal distribution. However, a marked difference between measurements and weighted average simulation is seen for  $\Delta L = -3$  and  $+2$  dB data. At  $\Delta \epsilon$  values  $\leq 45$  deg the response distribution is single peaked so that the weighted average model seems to fit the data better than the bimodal model. The pattern observed at  $\Delta \epsilon$  values  $> 45$  deg is more complex, as the responses are clearly not single peaked. Accordingly, the weighted average model is not in good agreement with the data. Interestingly, however, neither is a simple bimodal model: although the main response peak lies closest to the location of the louder sound, and a secondary peak is found near the softer sound, either peak appears to be shifted to lie between the two actual sound locations at  $+1$  and  $-1$ .

As listeners were instructed to ignore the GWN non-target, we wondered whether their localization behavior might have benefited from longer reaction times, which has been shown to be the case in visuomotor experiments (e.g.



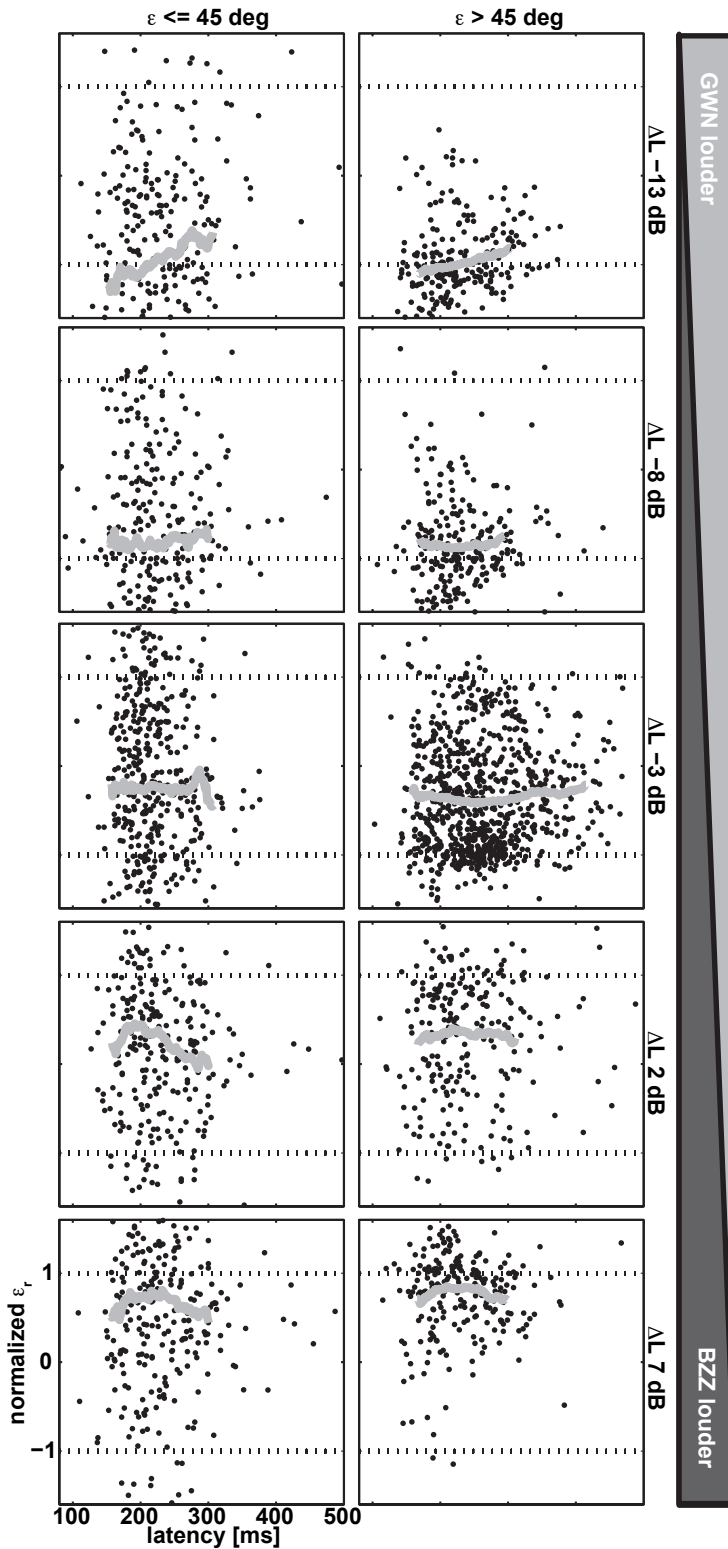
Becker and Jürgens, 1979; Ottes et al., 1985). Figure 8 shows the normalized response location as a function of head-saccade latencies (pooled across listeners). For graphical purposes we restricted the abscissa to range from -1.5 to 1.5. We divided the data set according to  $\Delta L$  (rows) and  $\Delta \epsilon \leq 45$  deg vs.  $\Delta \epsilon > 45$  deg (columns), like in Fig. 7. The gray line in each panel indicates the running average through all data. This analysis demonstrates no systematic influence of response latency on localization behavior. Thus, response accuracy did not improve when listeners postponed their responses, although their auditory system could have accumulated more evidence for the actual target location.

*DTF similarity-model.* The response patterns described so far appear to resemble those found in the visuomotor literature. In what follows, our analysis will be guided by the two opposing hypotheses described in the Introduction (Fig. 1). According to the peripheral hypothesis averaging and bimodal response behavior arise as a result of acoustic interactions at the pinna. In the case of averaging responses, the auditory system has no way of retrieving the original source locations because the averaging location is already encoded by the double DTF at the pinna filtering stage.

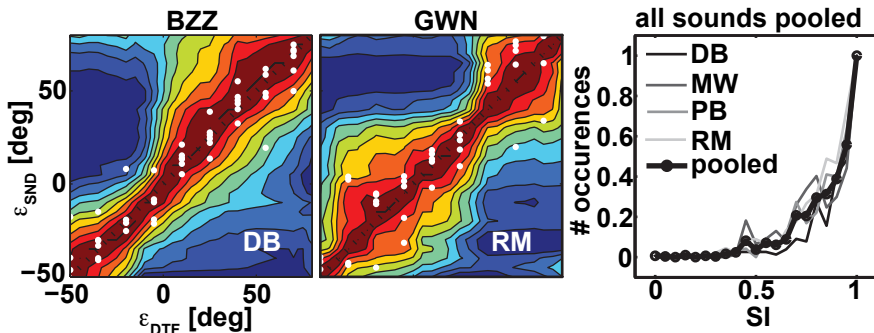
To test for this hypothesis, we applied the similarity analysis of recorded DTFs to the localization data (see Methods). As a consistency check, we first calculated the SI maps of the single-sound trials on which we superimposed the listener's head movement responses (cf. Fig. 2). Figure 9 shows the results obtained with listener DB (left panel) and RM (center panel). The gain and bias-corrected elevation responses of the listener are plotted as white dots (ordinate: target location; abscissa: response location), and the SI of the different DTFs is shown color encoded. Warm colors indicate high similarity between template DTF and filtered single-DTF; cold colors indicate poor similarity. As DTFs were obtained from broadband flat-spectrum Schroeder sweeps, and the single-sounds used in this experiment had similar long-term amplitude spectra, the SI map is expected to show a single strip of high correlation along the main diagonal. The figure shows that this is indeed the case, which indicates that the listener's DTFs are unique for each elevation angle, and therefore contain unambiguous elevation information. The left-hand panel shows data for BZZ (listener DB), the center panel for GWN (listener RM). In line with Fig. 2 listeners were well able to localize the single targets, so that all responses lie close to the black-dashed identity line with little variability. Importantly, as all responses fall on the dark-red patches that indicate high similarity (close to 1.0) between template and sensory DTFs, locations of high similarity thus correlate well with response location. This finding was also obtained without the gain-bias correction of single-sound responses (not shown).

The right-hand panel shows the accumulated histograms of the similarity indices obtained for the single-sound responses of all listeners (gray-coded lines) pooled across all 11 single-target conditions. The thick solid line shows the average for all conditions and listeners. Clearly, the SI peaks sharply at a value 1.0,





**Figure 8.** Normalized head-saccade endpoints as a function of saccade latency for the five  $\Delta L$  values (rows), pooled over all listeners. Data is separated in conditions in which  $\Delta\varepsilon \leq 45 \text{ deg}$  (left column) and  $\Delta\varepsilon > 45 \text{ deg}$  (right column). BZZ and GWN locations are indicated by the dotted black lines at 1 and -1, respectively. The thick gray line indicates the running average.

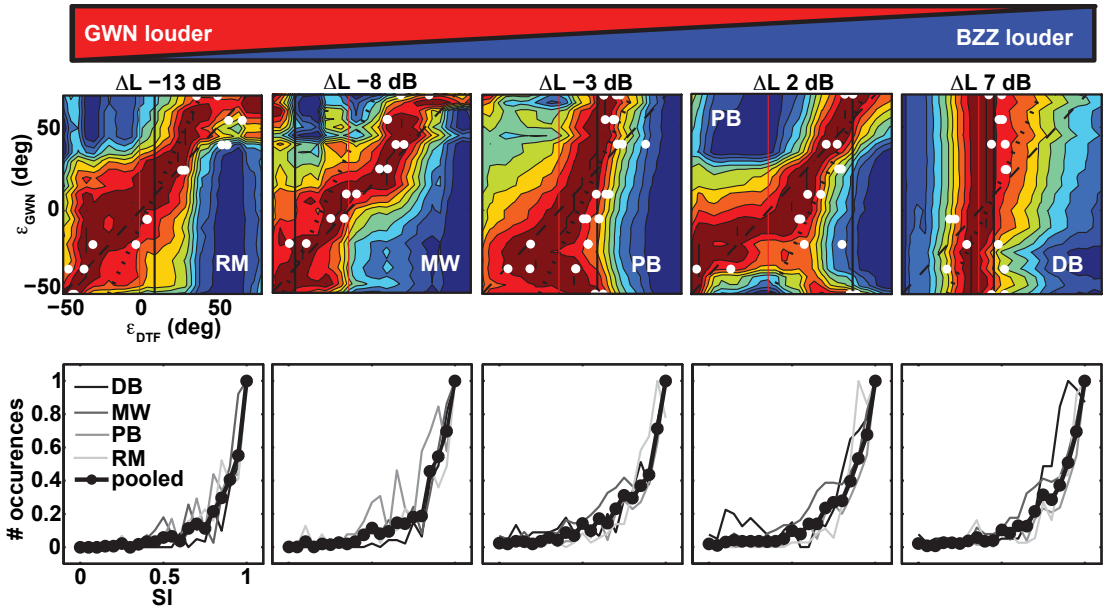


**Figure 9.** Similarity index (SI) of single-sound DTFs and Schröder-DTF templates (colored patches) and normalized single-sound head-movement responses (white dots) for BZZ (listener DB; **left**) and GWN (listener RM; **center**) as a function of target location. Responses and areas of highest similarity (warm colors) between template and target DTF coincide with response location. (**Right**) Normalized histograms of the SI's obtained at the response locations pooled across all 11 single-sounds for all four listeners (different shades of gray). The thick line with markers indicates the average across all four listeners.

with little variability, indicating that sound-localization responses to a single broadband sound in the mid-sagittal plane are highly reproducible and can be well explained by perceived similarity between the sensory spectrum and the internally stored representation of pinna filters.

Figure 10 shows the results of the similarity analysis for the double-sound conditions in the same format as Fig. 9. We selected examples from all four listeners (top row), for each of the applied  $\Delta L$  values (left to right). In order to give an overview of the data we selected one BZZ location for each case (positioned at the black solid line) and plotted the SI (colored patches) and the listener's responses (white dots) as a function of GWN location (corresponding with the black-dashed identity line). Note that the situation is now more complex than for the single-sound conditions of Fig. 9, as the sensory spectra deviate significantly from the single-DTFs obtained with Schroeder-sweeps. In the bottom row of Fig. 10 we show the accumulated histograms of the SI's obtained at the response location for all  $\Delta\epsilon$  values and all four listeners (different shades of gray). The average across the listeners is indicated as a black line with markers.

In the left-hand panel of the top row the GWN is 13 dB louder than the BZZ, and the response endpoints (gain- and bias corrected; see Methods) are distributed close to the dark-red patch parallel to the central diagonal that corresponds to the GWN location (listener RM). The lower-left panel shows the pooled histograms for the individual listeners and the mean for this condition. The histogram peaks close to the highest value of  $SI = 1.0$ . The same is seen for the  $\Delta L = -8$  dB condition of listener MW. On the right-hand side of Fig. 10 we show results of the  $\Delta L = +7$  dB condition for listener DB (top), and the resulting SI-histograms (bottom), which indicate localization responses that are dominated by the BZZ location (which in this figure is positioned at  $\epsilon_{\text{BZZ}} = +10$  deg). More interesting are the near-equal  $\Delta L$  conditions (-3 and +2 dB) shown in the central panels for listener PB (top). Especially at large spatial disparities (lower half of the upper panels) the responses became bimodal, whereas for small spatial disparities (up-



**Figure 10.** Similarity index (SI) of double-sound DTFs and Schröder-DTF templates (colored patches) as a function of GWN location, and the template location for all five  $\Delta L$  values and four different listeners (**top row**). The listener’s responses are indicated as white dots. In each plot the buzzer location was held constant at the location indicated by the solid black line. The dashed-black line indicates the GWN location (unity) and the dotted-black line indicates a prediction based on the weighted average of buzzer and GWN location. Warm colors indicate small differences between template and simulated double-DTF, cold colors indicate large differences. (**Bottom row**) Normalized histograms of the SI’s obtained at the response locations for all five  $\Delta L$  values pooled across all  $\Delta\epsilon$  conditions for all four listeners (different shades of gray). The average across all listeners is denoted as the thick black line with markers.

per half of the panels) we obtained weighted averaging responses ( $\Delta L = +2$  dB data align with the dotted-black line with a slope of approx. 2.0). Again, for both response modes, the head-pointing responses fall on or very near to the dark-red patches of high similarity. Thus, when sounds of equal  $\Delta L$  are spatially well separated, the acoustic periphery preserves major features of individual DTF characteristics. However, when they are closer together (within about 45 deg), highest similarity is found for the weighted averaged elevation. The listener’s responses are guided by these acoustic similarities. This is shown in the histograms (central bottom panels), which demonstrate that these results were obtained for all listeners and for all  $\Delta\epsilon$  values.

## 5.4 Discussion

*Summary.* When confronting the auditory system with two broadband sounds presented synchronously in the mid-sagittal plane, the perceived location is determined by the sound level differences and spatial disparity, but not task requirements (target vs. non-target), or reaction time. Our data reveal three different response modes of the sound-localization system in this situation. When level differences exceeded about 5 dB, the loudest sound determined perceived location, regardless of spatial disparity. However, when

sound levels were about equal, we observed two response modes that depended on spatial disparity. For spatial separations within 45 deg we obtained weighted-averaged responses, with unimodal distributions. Larger spatial separations resulted in bimodal response distributions, in which the listener oriented to either the target location, or the non-target. However, even in this case responses were drawn toward the location of the competing stimulus (Fig. 7).

We show that all observed response patterns may be understood from pinna-induced spectral cues. Our results therefore favor the hypothesis in which auditory spatial percepts to synchronous stimuli are guided by acoustic interactions at the pinnae (Fig. 1).

*Averaging and bimodality in the auditory system.* Response averaging in azimuth is a well-established phenomenon (stereophony) and, for simple sounds like a pure tone presented from two speakers in the free field, can be fully explained by interference of sound waves at each ear canal. A similar percept, however, can be obtained through dichotic stimulation (e.g. binaural beats), in which case sounds cannot interact at the ear canals, and hence a neural cause for the perceived intracranial location is more likely. Two localization cues underlie the percept of sound-source azimuth. Interaural time differences (ITDs) operate at low frequencies (<1.5 kHz), as the relation between ongoing phase difference and azimuth is unambiguous. Interaural level differences (ILDs) arise from the head-shadow effect and functions at high frequencies (>3kHz). Both cues are processed by independent brainstem pathways that involve the medial and lateral superior olive, respectively (Yin, 2002). Interestingly, imposing a conflicting ILD in combination with a given ITD under dichotic listening induces a weighted averaged intracranial location (time-intensity trade-off). Also in this case, the percept must arise from neural integration of binaural inputs. In short, spatial averaging in azimuth may have both a peripheral (for very simple sounds) and a central (for complex sounds) origin.

Note that for synchronous broadband stimuli in the mid-sagittal plane the emergence of response averaging is not obvious. The elevation percept is thought to be determined by the amplitude spectrum of the direction-dependent pinna-filter impulse response, the so-called DTF (Wightman and Kistler, 1989). The direction-dependent phase characteristic of the pinna filter is considered irrelevant. When two sounds are presented simultaneously, sound waves add linearly at the pinnae:

$$h_D(\tau) = h_1(\tau) + h_2(\tau) \quad (5)$$

with  $h_D(\tau)$  the total pressure wave at the eardrum in response to two impulses presented synchronously at locations 1 and 2. The respective DTFs at the locations are described by their amplitude spectra,  $R_1(\omega)$  and  $R_2(\omega)$ . However, since spectra are complex variables,  $H(\omega) = R(\omega) \cdot \exp(-i \cdot \Phi(\omega))$ , with  $\Phi(\omega)$  the direction-dependent phase spectrum, the combined amplitude spectrum is not simply the

linear superposition of the respective amplitude spectra, i.e.  $R_D(\omega) \neq R_1(\omega) + R_2(\omega)$ , but:

$$R_D = \sqrt{R_1^2 + R_2^2 + 2R_1R_2(\cos\Phi_1\cos\Phi_2 + \sin\Phi_1\sin\Phi_2)} \quad (6)$$

which relies on the amplitude and phase spectra of either DTF. Given the complexity of spectral-shape functions, it is not immediately clear from Eq. 6 that the (weighted) combination of two sounds yields an amplitude spectrum that best corresponds to the weighted averaged location (or to two shifted locations, for widely separated stimuli, Fig. 7). Interestingly, the data in Fig. 10 indicate that this is indeed the case.

Note, that listeners were able to perceptually distinguish the conditions in which only one sound (either BZZ, or GWN) was presented, or a double sound (BZZ+GWN). Thus, they could report reliably whether or not the target sound was presented together with a non-target. Yet, they invariably reported to always perceive a single, mixed sound-source coming from one location, even when their responses resulted in bimodal distributions.

Does this all mean that the auditory system is not capable to segregate auditory objects in elevation? We believe the answer is “yes” for synchronous sounds. However, introducing a brief (few and even sub-ms) onset asynchrony in either azimuth or elevation has been shown to immediately shift the localization percept toward the first sound (the precedence effect; Dizon & Litovsky 2004, Litovsky et al., 1999; Yin, 2002). This clearly indicates the involvement of neural processing that weeds out secondary acoustic input to enable localization. Our results hint at the interesting possibility that this temporal filter may be functionally imperative: if the secondary signal were not filtered out, the auditory system cannot segregate different sound sources in the mid-sagittal plane.

*Comparison to other studies.* The different response modes in our experiments are quite reminiscent to earlier reports from visuomotor experiments, in which two visual targets evoked saccadic eye movements (Becker and Jürgens, 1979; Findlay, 1982; Ottes et al., 1984,1985; Chou et al., 1999; Aitsebaomo and Bedell, 2000; Watanabe, 2001; Arai et al, 2004; Nelson and Hughes, 2007). Those studies revealed that the visuomotor system typically responds with averaging saccades when stimuli are presented in spatial-temporal proximity, and when the response reaction time is fast. Responses become bimodal when the spatial separation increases beyond about 30 deg. However, in both cases the saccadic system can optimize accuracy in a target/non-target paradigm by prolonging reaction times (Becker and Jürgens, 1979; Ottes et al., 1985): while early saccades invariably end at averaged locations, or in many cases at the distractor, late saccades can all be directed toward the task-imposed target.

Our auditory-evoked head saccades differ from these visuomotor response properties in two major respects. First, auditory response patterns did not evolve

over time, since localization accuracy did not improve with increasing reaction time (Fig. 8). Second, in visuomotor experiments visual stimuli were well separable, both at the retina, and in early sensory responses of neurons within the visuomotor pathways. This holds also for example for the midbrain superior colliculus (SC), a crucial sensorimotor interface for the programming and generation of eye-head orienting responses (Arai et al., 2004; Kim and Basso, 2008). The actual visual response selection leading to either averaging, or bimodal responses is therefore due to neural processing, rather than to visual peripheral limitations. It has been hypothesized that such response selection could take place within topographically organized neural maps, like in SC, where target locations are mapped onto spatially separated neural populations. Competition between different populations, combined with local-excitatory/global-inhibitory interactions, shapes the population that represents the saccade goal. Task constraints and stimulus saliency help favor neurons that represent the target to win this competition, yet also averaging may be the result of this competition (Van Opstal and Van Gisbergen, 1990; Glimcher and Sparks, 1993; Arai et al., 2004; Kim and Basso, 2008).

Our data indicate that such selective processes do not occur in the audio-motor system for synchronous stimuli, since acoustic interactions at the level of the auditory periphery appear to impose the auditory goal. The spectral shapes of sounds are only preserved when they have about equal loudness and are widely separated in the mid-sagittal plane (Fig. 10), leading to bimodal response distributions. However, even under those conditions listeners do not perceive two segregated auditory objects. In a pilot experiment subjects also listened to the double sounds without making head saccades, but instead indicated whether they perceived two distinct auditory events. In all trials they perceived only one acoustic event (data not shown). The differences in temporal fine structure between BZZ and GWN were therefore not sufficient to separate the sounds. Interestingly, we also obtained weighted-averaging responses when temporal fine structure and spectral content of the stimuli were very different (two male voice utterances of the same duration). Again, subjective reporting indicates the percept of only one (mixed) sound source that emanated from one, averaged, location (data not shown).

*DTF-Model.* Our SI-analysis follows previous studies that correlated DTFs with sound-localization (Middlebrooks, 1992; Hofman and Van Opstal, 1998, 2003; Langendijk and Bronkhorst, 2002). These studies assume that the sensory spectrum is somehow compared to internally stored templates that are related to elevation after a learning process (Hofman et al., 2001; Van Wanrooij and Van Opstal, 2005). The template with highest similarity to the sensory input is selected and mapped to the source location. Our data show that sound-localization responses faithfully follow the DTF similarity map that is based on a representation on single broadband sounds, as head-saccade endpoints are clustered at locations with high similarity values. Similar findings have been reported for cats, who rely on their spectral cues for localization (pinna-related notches) when presented



with illusory localization cues in 2D virtual space (Tollin and Yin, 2003).

Although the neural representations of templates and the associated mapping stages in the auditory pathway are still not known, dichotic and free-field perceptual experiments showed that DTFs can be smoothed substantially without hampering spatial percepts (Kulkarni and Colburn, 1998; Hofman and Van Opstal, 2002). Further study is required to assess the relative contributions of acoustic variables, task constraints, temporal asynchronies, and combined azimuth/elevation cues, to the segregation and selection of multiple sounds. The present experiments provide a first step toward that goal.

## **Acknowledgements**

We thank Robert Hovingh for his help in the initial phase of this project. Dick Heeren, Hans Kleijnen and Stijn Martens are thanked for excellent technical assistance. We acknowledge Arno Engels and Jaap Nieboer for engineering and building the auditory hoop. Supported by Marie Curie Early Stage Training Fellowship of European Community's Sixth Framework Program (MEST-CT-2004-007825 PB), VICI grant within Earth and Life Sciences of NWO (ALW 865.05.003; AJVO, MMVW), and Radboud University Nijmegen (AJVO).



## 5.5 References

Aitsebaomo AP, Bedell HE. Saccadic and psychophysical discrimination of double targets. (2000) *Optom Vis Sci* 77:321–30.

Algazi R, Duda RO, Thompson DM, Avendano C (2001) The CIPIC HRTF database. Proc. 2001 IEEE Workshop on Applications of Signal Processing to Audio and Electroacoustics, pp. 99–102, Mohonk Mountain House, New Paltz, NY, Oct. 21–4.

Arai K, McPeck RM, Keller EL. Properties of saccadic responses in monkey when multiple competing visual stimuli are present. (2004) *J Neurophysiol* 91:890–900.

Becker W, Jürgens R. An analysis of the saccadic system by means of double step stimuli. (1979) *Vision Res* 19:967–83.

Best V, Van Schaik A, Carlile S. Separation of concurrent broadband sound sources by human listeners. (2004) *J Acoust Soc Am* 115:324–36.

Blauert J. Sound localization in the median plane. (1969) *Acustica* 22:205–13.

Blauert J. Spatial hearing: the psychophysics of human sound localization. Revised Edition. (1997) MIT Press, Cambridge, MA, USA.

Chou I, Sommer MA, Schiller PH. Express averaging saccades in monkeys. (1999) *Vision Res* 39:4200–16.

Dizon RM, Litovsky RY. Localization dominance in the mid-sagittal plane: effect of stimulus duration. (2004) *J Acoustic Soc Am* 115: 3142–55.

Findlay JM Global visual processing for saccadic eye movements. (1982) *Vision Res* 22:1033–45.

Glimcher PW, Sparks DL. Representation of averaging saccades in the superior colliculus of the monkey. (1993) *Exp Brain Res* 95:429–35.

Hofman PM, Van Opstal AJ. Spectro-temporal factors in two-dimensional human sound localization. (1998) *J Acoust Soc Am* 103:465–70.

Hofman PM, Van Riswick JRA, Van Opstal AJ. Relearning sound localization with new ears. (2001) *Nature Neurosci* 1:417–21.

Hofman PM, Van Opstal AJ. Bayesian reconstruction of sound localization cues from responses to random spectra. (2002) *Biol Cybern* 86:305–16.

Hofman PM, Van Opstal AJ. Binaural weighting of pinna cues in human sound localization. (2003) *Exp Brain Res* 148:458–70.

Kim B, Basso MA. Saccade target selection in the superior colliculus: a signal detection theory approach. (2008) *J Neurosci* 28:2991–3007

Knudsen EI, Konishi M. Mechanisms of sound localization in the barn owl (*Tyto alba*). (1979) *J Comp Physiol* 133:13–21.

Kulkarni A, Colburn HS. Role of spectral detail in sound-source localization. (1998) *Nature* 396:747–9.

Langendijk EHA, Bronkhorst AW. Contribution of spectral cues to human sound location. (2002) *J Acoust Soc Am* 112:1583–96.

Lee C, Rohrer WH, Sparks DL. Population coding of saccadic eye movements by neurons in the superior colliculus. (1988) *Nature* 332:357–60.

Litovsky RY, Colburn HS, Yost WA, Guzman SJ The precedence effect. (1999) *J Acoust Soc Am* 106: 1633–54.

MackKay DJC. Bayesian interpolation. (1992) *Neural Comput* 4:415–47.

Middlebrooks J. Narrow-band sound localization related to external ear acoustics. (1992) *J. Acoust Soc Am* 92:2607–24.

Middlebrooks JC, Green DM. Sound localization by human listeners. (1991) *Annu Rev Psychol* 42:135–59.

Nelson MD, Hughes HC. Inhibitory processes mediate saccadic target selection. (2007) *Percept Mot Skills* 105:939–58.

Ottes FP, Van Gisbergen JAM, Eggermont JJ. Metrics of saccade responses to visual double stimuli: two different modes. (1984) *Vision Res* 24:1169–79.

Ottes FP, Van Gisbergen JAM, Eggermont JJ. Latency dependence of colour-based target vs nontarget discrimination by the saccadic system. (1985) *Vision Res* 25:849–62.

Press WH, Flannery BP, Teukolsky SA, Vettering WT Numerical recipes in C: the art of scientific computing. (1992) Cambridge University Press. Cam

bridge, MA, USA.

Robinson DA. A method of measuring eye movement using a scleral search coil in a magnetic field. (1963) *IEEE Trans Biomed Electron BME* 40:137–45.

Schroeder MR. Integrated-impulse method measuring sound decay without using impulses. (1979) *J Acoust Soc Am* 66:497–500.

Shaw, EAG. Earcanal pressure generated by free sound field. (1966) *J Acoust Soc Am* 39:465–70.

Tollin DJ, Yin TC. Spectral cues explain illusory elevation effects with stereo sounds in cats. (2003) *J Neurophysiol* 90:525–30

Van Opstal AJ, Van Gisbergen JA. Role of monkey superior colliculus in saccade averaging. (1990) *Exp Brain Res* 79:143–9.

Van Wanrooij MM, Van Opstal AJ. Relearning sound localization with a new ear. (2005) *J Neurosci* 25:5413–24.

Watanabe K. Inhibition of return in averaging saccades. (2001) *Exp Brain Res* 138:330–42

Wightman FL, Kistler DJ. Headphone simulation of free-field listening. II: Psychophysical validation. (1989) *J Acoust Soc Am* 85:858–67.

Yin TC. Neural mechanisms of encoding binaural localization cues in the auditory brainstem. (2002) In: *Integrative functions in the mammalian auditory brainstem* (Oertel D, Fay RR, Popper AN, eds). pp 99–159. Heidelberg, Springer.

Zwiers MP, Van Opstal AJ, Cruysberg JRM. A spatial hearing deficit in early-blind humans. (2001) *J Neurosci* 21:1–5.



## Chapter 6

# Acquired Prior Knowledge Modulates Audiovisual Integration

### Abstract

Orienting responses to audiovisual events in the environment can benefit markedly by the integration of visual and auditory spatial information. However, logically, such integration would only be considered successful for stimuli that are spatially and temporally aligned, as these would be emitted by a single object in space-time. Since humans do not have prior knowledge about whether novel auditory and visual events do indeed emanate from the same object, such information needs to be extracted from a variety of sources. For example, expectation about alignment or misalignment could modulate the strength of multisensory integration. If evidence from previous trials would repeatedly favour aligned audiovisual inputs, the internal state might also assume alignment for the next trial, and hence react to a new audiovisual event as if it were aligned. To test for such a strategy, subjects oriented a head-fixed pointer as fast as possible to a visual flash that was consistently paired, though not always spatially aligned, with a co-occurring broadband sound. We varied the probability of audiovisual alignment between experiments. Reaction times were consistently reduced in blocks containing only aligned audiovisual stimuli, when compared to blocks containing also pseudo-randomly presented spatially disparate stimuli. Results demonstrate dynamic updating of the subject's prior expectation of audiovisual congruency. We discuss a model of prior probability estimation to explain the results.

## 6.1 Introduction

Each of our senses extracts information about events in the environment. Successful integration of these separate information streams can be highly beneficial in numerous tasks, ranging from improved stimulus identification to speeding of orienting responses, and enhanced localisation performance. In the present study we focus on the control of rapid head saccades toward a novel audiovisual stimulus in the peripheral visual field.

A large body of experimental evidence has indicated that audiovisual integration leads to a marked reduction of saccadic reaction times for co-occurring and spatially aligned audiovisual targets. Typically, experiments tested relatively simple conditions in which a single auditory and a single visual stimulus could occupy a limited number of possible configurations (Hughes et al., 1994; Nozawa et al., 1994; Frens et al., 1995; Goldring et al., 1996). Yet, also for more complex audiovisual scenes that contain more uncertainty about upcoming target locations and audiovisual combinations, perceptually aligned audiovisual events consistently produce faster and more accurate orienting responses than their uni-sensory counterparts (Corneil et al., 2002; Van Wanrooij et al., 2009).

The mechanisms and rules that govern audiovisual integration to evoke rapid and accurate orienting behaviour have been explained by neural interactions within spatially organized neural maps, such as in the midbrain Superior Colliculus (Stein & Meredith, 1993; Frens & Van Opstal, 1998; Wallace et al., 1998; Bell et al., 2005). These interactions induce excitatory effects for congruent multisensory inputs, but suppress each other when they fall outside the spatial-temporal integration window (Meredith & Stein, 1986a; Meredith et al., 1987). A theoretical account for such a mechanism was offered by (Anastasio et al., 2000), who proposed that the principle of optimal statistical inference (Bayesian reasoning) underlies multisensory integration.

Strict integration of audiovisual cues (Alais & Burr, 2004), however, is not always the most desirable option (Hillis et al., 2002). In that case, audiovisual events emanating from distinct objects would also be integrated, thus losing their segregation and identities. Obviously, the brain is able to cope with such situations, and can readily distinguish spatially disparate audiovisual stimuli (Wallace et al., 2004; Kording et al., 2007; Sato et al., 2007). The breakdown of multisensory integration is also evidenced by a systematic reduction of speed and accuracy of saccadic eye movements to disparate stimuli (Frens et al., 1995; Harrington & Peck, 1998; Hughes et al., 1998; Colonius & Arndt, 2001; Van Wanrooij et al., 2009).

But how does the brain know which auditory and visual signals to fuse into an integrated percept, and which not? According to Bayesian models of multisensory integration the strength of multisensory fusion is proportional to the amount of coupling between sensory streams, which reflects the prior knowledge that multisensory inputs belong together (Ernst, 2005). Such prior knowledge could be based on experience, and is likely to be adaptive.

Here we address the question whether subjects extract and use information about expected alignment of audiovisual stimuli on the basis of the experimental stimulus statistics, and accordingly adapt multisensory integration when orienting to audiovisual stimuli. To that end, we manipulated the proportion of audiovisual spatial congruency in separate experimental blocks. If subjects dynamically adjust their expectation on audiovisual congruency, one expects enhanced integration effects when the probability of spatial alignment is high and decreased integration for low probabilities. Our results corroborate this hypothesis.

## 6.2 Methods

*Listeners.* Seven subjects, aged 21-33 (mean: 27.7 yrs), participated in this study. Two subjects (MW, PB) are authors of this paper; the remaining five participants were naive about the purpose of the study. All subjects had normal hearing (within 20 dB of audiometric zero) as determined by an audiogram obtained with a standard staircase procedure (10 tone pips, 0.5-octave separation, between 500 Hz and 11.3 kHz) and had normal or corrected (MA, MW) binocular vision, with the exception of subject PB who did not wear his prescription glasses during the experiments. Since the flash was supra-threshold, and his V- and AV-responses were within the normal range, we included his results in this study.

Experiments were conducted after subjects gave their full understanding and written consent. The experimental procedures were approved by the Local Ethics Committee of the Radboud University Nijmegen and adhered to The Code of Ethics of the World Medical Association (Declaration of Helsinki), as printed in the British Medical Journal of July 18, 1964.

*Apparatus and sound generation.* During the experiments, subjects sat comfortably in a chair in the centre of a completely dark, sound-attenuated room (3m x 3m x 3m). The floor, ceiling and walls were covered with sound-attenuating black foam (50 mm thick with 30 mm pyramids, AX2250, Uxem b.v., Lelystad, The Netherlands), effectively eliminating echoes for frequencies exceeding 500 Hz. The room had an ambient background noise level of about 30 dB SPL.

The chair was positioned at the centre of a vertically oriented circular hoop (radius 1.2 m) on which an array of 29 small broad-range loudspeakers (SC5.9; Visaton GmbH, Haan, Germany) was mounted at 5-deg intervals from -55 to +85 deg in the midsagittal plane (elevation angles, with 0 deg at straight ahead). Acoustic stimuli were digitally generated using Tucker-Davis System 3 hardware (Tucker-Davis Technologies, Alachua, Florida, USA), with a real-time processor (RP2.1 System3, 48,828-Hz sampling rate). All acoustic stimuli consisted of 65-dB (A-weighted), 50 ms Gaussian white noise (0.5-20 kHz bandwidth), with 0.5 ms sine-squared on- and cosine-squared offset ramps. Visual stimuli consisted of green (wavelength 565 nm) light-emitting diodes (LEDs) mounted at the centre of each speaker (luminance 0.5 cd/m<sup>2</sup>).

Head movements were recorded with the magnetic search-coil technique



(Robinson, 1963). To that end, the listener wore a lightweight spectacle frame on which a small coil was attached to its nose bridge. Three orthogonal pairs of square coils (6 mm<sup>2</sup> wires, 3m x 3m) were attached to the room's edges to generate the horizontal (80 kHz), vertical (60 kHz), and frontal (48 kHz) magnetic fields, respectively. The head-coil signal was amplified and demodulated (EM7; Rimmel Labs, Katy, TX, USA), low-pass filtered at 150 Hz (custom built, fourth-order Butterworth), and digitized by a Medusa Head Stage and Base Station (TDT3 RA16PA and RA16; Tucker-Davis Technology, Alachua, FL, USA) at a rate of 1017.25 Hz per channel. A custom-written C++ program running on a PC (Precision 380; 2.8 GHz Intel Pentium D; Dell, Limerick, Ireland) controlled data recording and stimulus generation.

*Experiments.* We performed six different experiments: a calibration experiment, a unisensory auditory (A) and visual (V) experiment, and three audiovisual (AV) experiments. The AV experiments differed mainly in their distributions of spatial disparities between auditory and visual stimuli, as outlined below. Apart from the calibration experiment, all stimulus locations in the V, A and AV experiments were presented in the midsagittal plane. Every experimental session began with the visual calibration experiment, followed by four blocks of a single other experiment. Each session was performed on separate days.

In all experiments, except for the calibration experiment, subjects initiated a trial by a button press after first fixating a straight-ahead fixation LED. This button press extinguished the fixation LED 100-200 ms later, immediately followed by the visual target flash (50 ms) and/or a synchronous sound. Subjects were instructed to direct the head-fixed laser pointer as quickly and as accurately as possible to the target location. In the AV-experiments, the target was always the visual flash. Since reaction times typically exceeded 100 ms, all responses were made under complete open-loop conditions.

*Calibration experiment.* To obtain the head-position data for the calibration procedure subjects accurately pointed the head-fixed laser pointer towards 56 LED locations in the 2D frontal hemifield that encompassed the stimulus range of the actual experiments. Each experimental session started with this calibration run.

*Visual experiment.* Visual targets (50 ms duration) were presented at 10 possible locations:  $\pm$  [15, 20, 25, 30, 35] deg in elevation. Each location was presented 80 times, yielding 800 trials. These trials were pseudo-randomly presented in four separate, consecutive blocks of 200 trials. In between those blocks, subjects were allowed a short break (1-3 min) from the experiment, in which the room lights were illuminated.

*Auditory experiment.* The same target locations as in the visual experiment were employed, with targets being auditory.

*Audiovisual 100%-aligned/0%-distractor experiment: AV-100/0.* In the AV-100/0 experiment, the visual targets (same locations as in the visual experiment) were accompanied by spatially and temporally aligned sounds. Subjects were specifically instructed that this was the case, yet that the flash was at the target location.

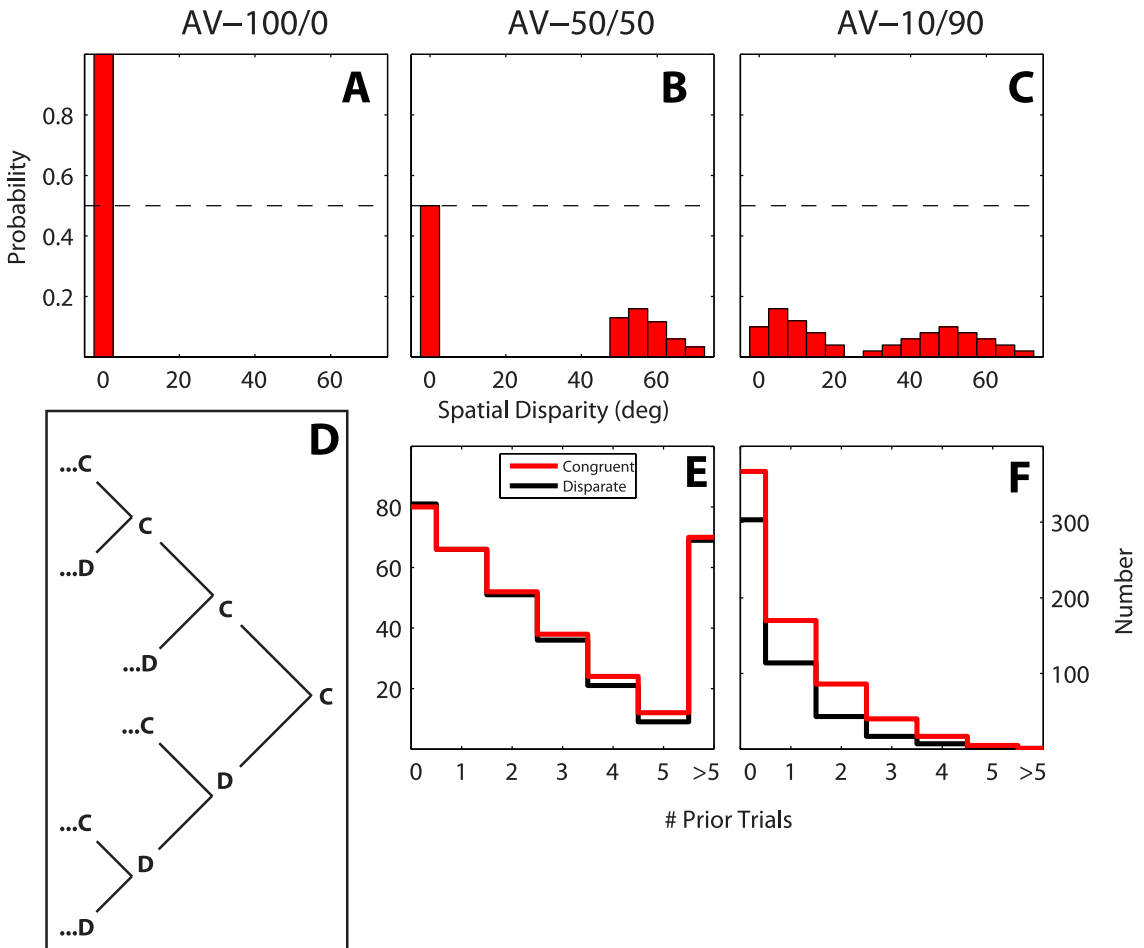
*Audiovisual 10%-aligned/90%-distractor experiment: AV-10/90.* All spatial combinations of flash and sound locations were presented eight times in the AV-10/90 experiment, yielding (10 flash locations x 10 sound locations x 8 repetitions=) 800 trials. As with the visual experiment, these trials were pseudo-randomly presented in four consecutive blocks of 200 trials. In this experiment, the sounds thus provided no a priori knowledge about the visual target. Subjects were instructed beforehand that the sound could be ignored. Only 10% of all trials contained spatially aligned audiovisual stimuli.

*Audiovisual 50%-aligned/50%-distractor experiment: AV-50/50.* In the AV-50/50 experiment, 50% of all trials contained spatially aligned AV stimuli, while the other 50% had a spatial disparity exceeding 45 deg. The (10 flash locations x 2 disparities x 15 repetitions=) 300 trials in this experiment were divided in 2 blocks of 150 trials. Subjects were instructed, as in the AV-10/90-experiment, that the sounds could be ignored.

*Distributions of spatial disparities.* The different spatial distributions of each AV experiment are depicted in figure 1A-C. Due to the pseudorandom presentation of stimuli, each trial could be preceded by various combinations of disparate or congruent trials. This trial order, which is an important aspect of our experimental rationale, is exemplified in figure 1D; a congruent trial (C) might be preceded by either a congruent trial (CC), or a disparate trial (DC). And this preceding trial is again preceded by either a congruent or a disparate trial, leading to four different triplet trial sequences ending in a congruent trial. As an example, the distribution of two uninterrupted series of either congruent or disparate trials is shown in figure 1E&F.

*Data Analysis.* All data analysis was performed off-line in MatLab (r2008a; The Mathworks, Natick, MA, USA).

*Data calibration.* Response data were calibrated by training two three-layer neural networks with the back-propagation algorithm that mapped final head orientations onto the known target positions of the visual calibration experiment (Goossens & Van Opstal, 1997). Head-position data from the other experiments were calibrated off-line using these networks with an absolute accuracy <3% over the entire range. Head movements were automatically detected from calibrated data based on velocity criteria (onset >20 deg/s, offset <15 deg/s). Onset and offset markings were visually checked by the experimenter, and adjusted if necessary.



**Figure 1.** Distribution of AV spatial disparities in the AV-100/0 (A), AV-50/50 (B), and AV-10/90 (C) experiments. (D) The pseudo-random nature of trial presentation leads to different combinations of trial disparity order. The number of trials for uninterrupted series of congruent (disparity < 15 deg) and disparate stimuli (disparity > 45 deg) is shown for the AV-50/50 (E) experiment, and the AV-10/90 (F) experiment. **Red line** – congruent streak; **black line** – disparate streak.

**Performance.** Performance of the subjects was quantified by the reaction time (RT; onset head movement - onset target), and localisation error (see below) of the first goal-directed head movement in a trial. Responses with reaction times shorter than 60 ms, or exceeding 600 ms, or with an amplitude below 5 deg, were discarded from the analysis, as they were deemed to be due to prediction or to inattentiveness of the subject.

**Localisation Error.** We quantified localisation accuracy by linear regression on the stimulus-response relation:

$$\varepsilon_R = a\varepsilon_T + b \quad (1)$$

with  $\varepsilon_R$  and  $\varepsilon_T$  response elevation and target elevation, respectively. Parameters

$a$  and  $b$  were found by minimizing the mean squared error (Press et al., 1992). We took the localisation error as the value of the residuals between data and fit.

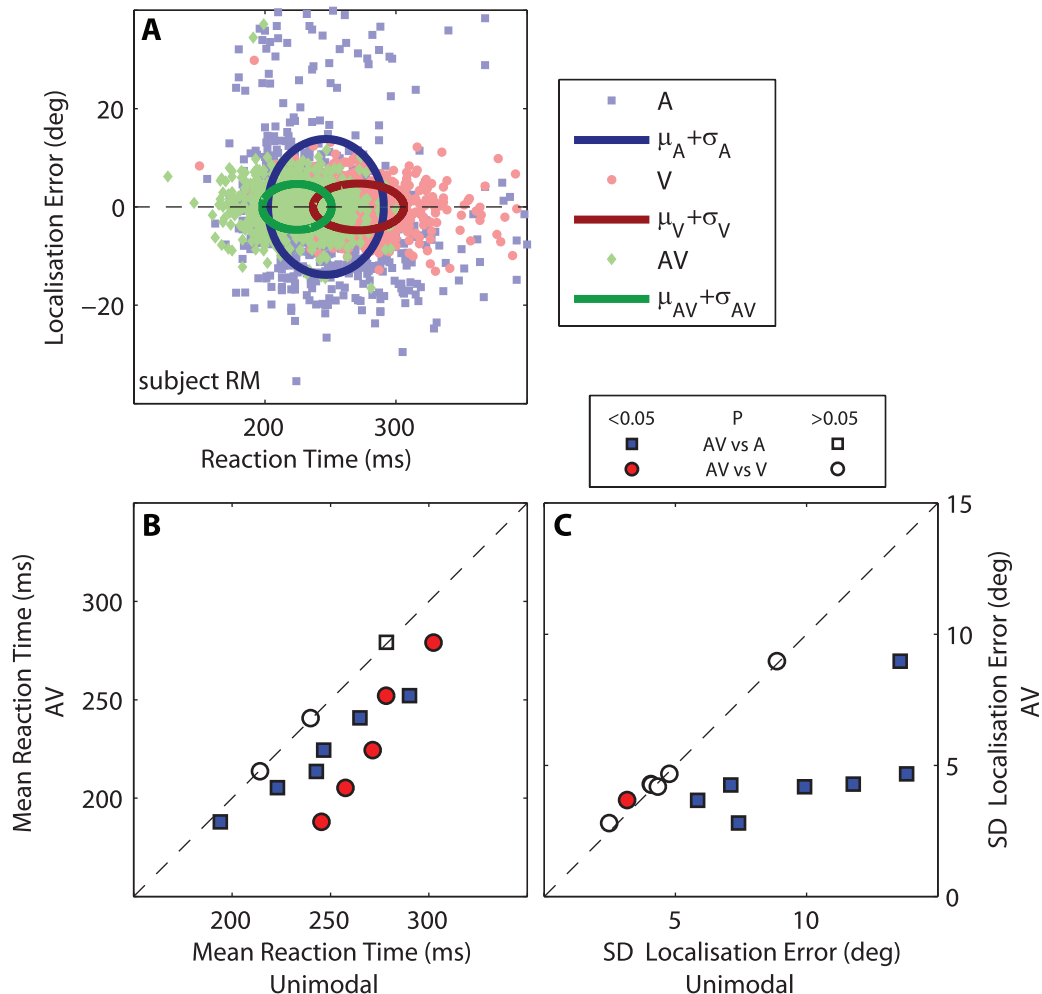
*Statistics.* Statistical significance of a difference between two 2-dimensional (2-D, reaction time versus localisation error) distributions was assessed a 2-D Kolmogorov-Smirnov-test. The means of two reaction time distributions were compared with a t-test, while the variances of localisation errors were compared with an F-test. We took  $P=0.05$  as the accepted level of significance.

## 6.3 Results

*AV integration of aligned stimuli in aligned experiment.* Synchronous presentation of spatially-aligned audiovisual stimuli in the vertical plane led to faster and more accurate AV-evoked head saccades in the AV-100/0 experiment (only AV-aligned stimuli), than to the unisensory stimuli. Figure 2A shows a representative example of the AV integration properties in this AV-100/0 experiment: the two-dimensional (2D) distributions (reaction time versus localisation error) of A (blue), V (red) and aligned AV responses (green) of subject RM are compared to each other. The A-responses are faster than the V-responses, as the A-distribution is systematically shifted to the left of V-responses, but the V-responses are clearly more accurate. Yet, the AV-responses are, on average, the fastest and most accurate ( $K\text{-SA-AV}=0.59$ ,  $P \ll 0.0001$ ;  $K\text{-SV-AV}=0.26$ ,  $P \ll 0.0001$ ). This improvement was generally observed: 5 out of 7 subjects had shorter AV reaction times than both A and V reaction times (Figure 2B, t-test,  $P \ll 0.001$ , exceptions,  $P > 0.05$ : subject MA for A; subjects MM and PB for V). The variance of the localisation error (Figure 2C) was always lower than the variance in the A error (F-test,  $P \ll 0.001$ ), while the AV error variance did not differ from the V error variance (F-test,  $P > 0.05$ ), with the exception of subject PB ( $F=0.74$ ,  $P=2.2 \times 10^{-5}$ ).

We have termed this type of multisensory integration the “best of both worlds”-effect (Corneil et al., 2002), as AV responses appeared to be as fast, or faster, than A-responses, but at visual localisation precision (equal variance). Having established this typical improvement in performance for the basic, spatially aligned AV stimuli also when stimulus locations are confined to the vertical plane and responses are measured with head movements, we next quantified the effect of spatial disparity on AV integration.

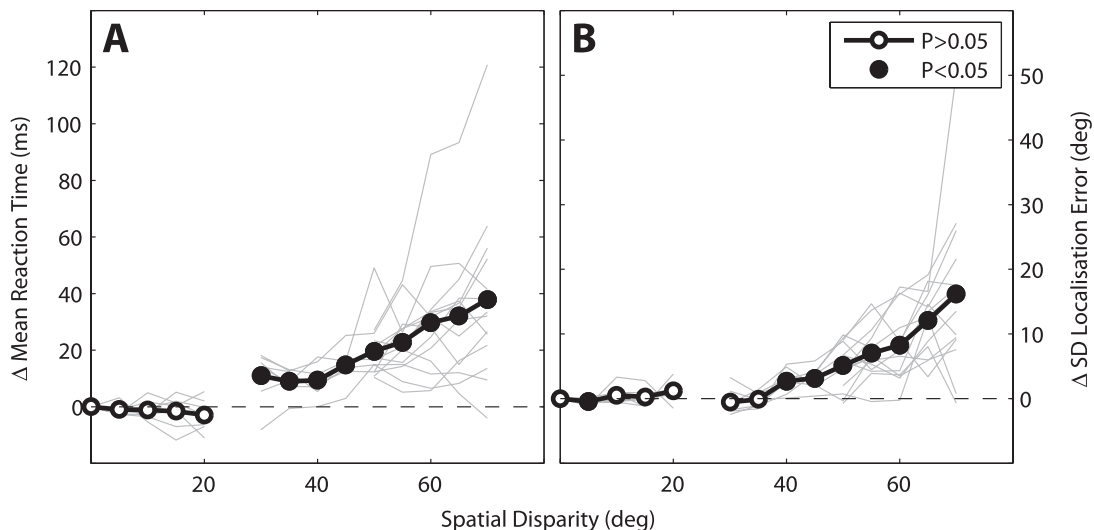
*Breakdown of AV integration by spatial disparity.* A large spatial disparity between the A- and V-stimuli degraded the AV integration effect (Figure 3) in the two experiments that contained auditory distractors (the AV-50/50 [50% distractors at  $> 45$  deg disparity; Fig. 1E] and the AV-10/90 experiment [90% distractors at  $> 15$  deg disparity; Fig. 1F]). For small spatial disparities ( $\leq 20$  deg, for which the V target and A distractor were in the same lower or upper hemifield), AV reaction times (Figure 3A) and localisation errors (figure 3B) were not affected (t-test across subjects,  $P > 0.05$ ). When the auditory distractor was presented in the op-



**Figure 2.** AV integration in AV-100/0 experiment. Responses to unisensory (A – **blue**; V – **red**) and AV-aligned (**green**) responses. **(A)** Signed localisation error versus reaction time for subject RM. Ellipses denote 1 standard deviation around the mean. **(B)** Mean reaction time of AV-aligned responses versus A and V unisensory responses for all subjects. **(C)** Standard deviation of localisation error for AV-aligned responses versus A and V unisensory responses for all subjects. Open symbols – non-significant difference between unimodal and AV distributions ( $P > 0.05$ , t-test for reaction times, F-test for signed localisation errors).

posite hemifield as the V target, the size of the spatial disparity systematically delayed the AV responses, by up to  $\sim 38$  ms for a disparity of 70 deg (Figure 3A). A similar pattern was observed for the localisation errors (Figure 3B): the error variance of the first saccade with respect to the visual target increased as AV-disparity increased, with standard deviation in localisation errors increasing up to  $\sim 16$  deg for the largest spatial disparity.

These data are in nice agreement with previous studies that reported a breakdown of AV-integration when spatial disparity between the A- and V-stimulus increased (Harrington & Peck, 1998; Wallace et al., 2004; Van Wanrooij et al., 2009).

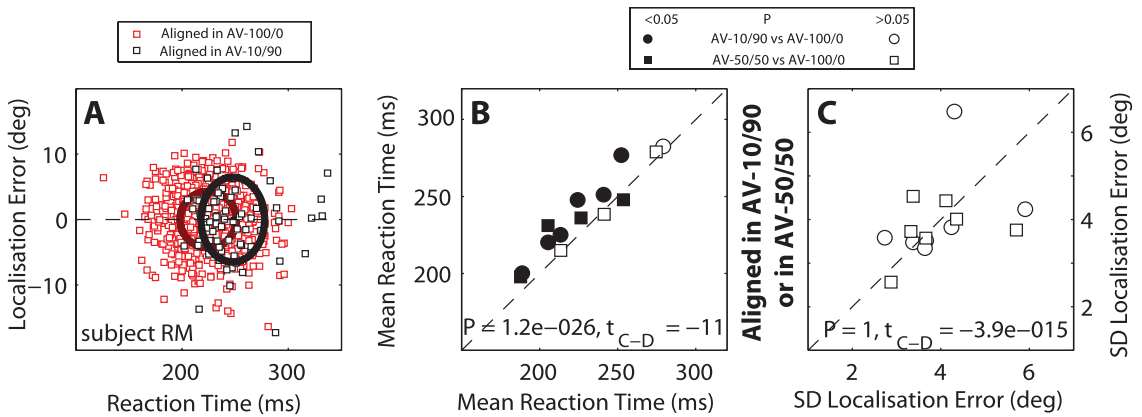


**Figure 3.** Spatial AV disparity influences AV integration in the median plane. **(A)** Average reaction time as a function of spatial disparity. **(B)** Average localisation error standard deviation as a function of spatial disparity. Reaction time and localisation error of aligned responses (spatial disparity 0 deg) was taken as baseline. **Grey lines** – different subjects and experiments; **black bold lines** – average over subjects and experiments; closed circles – significantly (t-test across subjects,  $P < 0.05$ ) different from 0.

*Effect of congruence-disparity distribution.* So far, we have only shown the effects of the current stimulus properties on AV integration. We now address the question whether the ongoing distribution of AV disparities affects AV integration. This was investigated by presenting spatially aligned stimuli, either among other spatially aligned stimuli (AV-100/0 experiment), or among spatially disparate stimuli (AV-10/90, 10% aligned stimuli, Fig. 1F; and AV-50/50 experiment, 50% aligned stimuli, Fig. 1E). We found that spatially aligned AV-stimuli elicited faster in the AV-100/0 experiment than in the AV-10/90 or AV-50/50 experiment, as exemplified for subject RM in Figure 4A. The 2D response distribution of localisation error versus reaction time is faster and more precise in the AV-100/0 experiment (red) than in the AV-10/90 experiment (black;  $K-S=0.32$ ,  $P < < 0.0001$ ). The effect on reaction time is corroborated by the systematic reduction by about 8 ms in the AV-100/0 experiment when compared to the AV-10/90 and 50/50 experiments for all subjects (Figure 4B; for 9 out of 14 subjects/experiments  $P < < 0.001$ , while subject AK reacted faster in the AV-50/50 experiment,  $P=0.02$ ). There was no significant improvement or decrement in localisation precision in the AV-100/0 experiment across all subjects (t-test,  $P > 0.05$ ).

Taken together, these results imply that when subjects might expect a disparate trial, an increase in reaction times is observed, in the absence of an observable change in localisation errors.



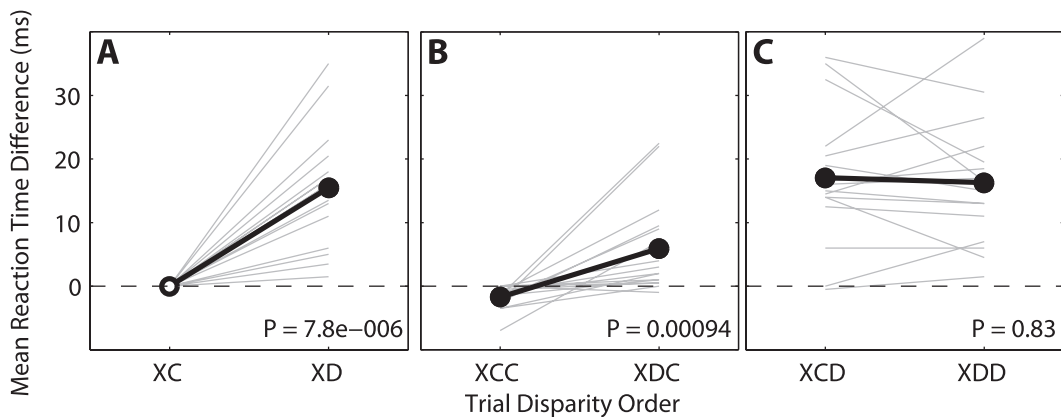


**Figure 4.** Effect of disparity distribution on aligned AV responses. **(A)** Localisation error versus reaction time of responses to aligned AV stimuli in the aligned AV-100/0 (red) and AV-10/90 distractor (black) experiments for subject RM. Ellipses correspond to one standard deviation around the mean. Note the shift toward shorter reaction times, and reduction of the variance in error when all stimuli in an experiment are aligned. **(B)** Average reaction time of AV-aligned responses in the AV-10/90 (circles) and AV-50/50 (squares) experiment versus AV-aligned responses in the AV-100/0 experiment for all subjects. **(C)** Standard deviation in localisation error of AV-aligned responses in AV-10/90 and AV-50/50 experiments versus AV-aligned responses in AV-100/0 experiment for all subjects. Closed markers in panels **B** and **C** represent significant differences for a single subject,  $P < 0.05$ .

*Influence of previous trial disparity on reaction time.* If the prior expectation of subjects is continuously updated, it is expected that the spatial configuration of a previous trial will influence the response to a novel audiovisual event, even if the novel stimuli themselves were presented at different locations. Since in both the AV-50/50 and AV-10/90 experiments trials were randomly interleaved, we had a large number of congruent and disparate trials that were preceded by either a congruent, or a disparate trial (congruent is defined as a spatial disparity less than 15 deg). Current-trial spatial alignment has a large effect on AV integration (Fig. 3), but so does the audiovisual disparity of the previous trial.

Figure 5 shows the effect of the immediate trial history on the head-saccade reaction times when a currently congruent stimulus was preceded by either another congruent trial, or by a disparate trial (cf. Fig. 1A). Fig. 5A shows the two possible current-trial configurations: either congruent (XC), or disparate (XD), regardless of the spatial alignment of the previous trial (indicated by X). Clearly, the average reaction time in the current trial increases for a disparate target configuration (see also Fig. 3). In Fig. 5B we show the average reaction times when the current trial is congruent, whereas the previous trial could be either congruent (XCC), or disparate (XDC). Interestingly, the double-congruent condition yielded significantly faster reaction times (about 8 ms) than trials in which the previous trial was disparate (t-test,  $P = 0.00094$ ). Note that this effect disappears entirely when the current trial is disparate (Fig. 5C). In that case all reaction times are elevated, regardless of the alignment of the previous trial (congruent – XCD; disparate – XDD). No effects were observed for the localisation errors (not shown).





**Figure 5.** Influence of trial history. **(A)** Average reaction times for the current disparate trial (XD) versus the baseline congruent trial (XC) regardless of the alignment of the previous trial (X). **(B)** Reaction times are compared for congruent trials that were preceded by another congruent trial (XCC), or by a disparate trial (XDC). The double-congruent condition yields the fastest reaction times, even significantly faster than the general XC trials. **(C)** The improvement is not obtained when the current trial is disparate (XCD and XDD trial sequences). **Grey lines** – individual subjects and experiments; **black bold line** – average across subjects and experiments. **Closed circles** – significant difference between the two sequences in the subplots,  $P < 0.05$ , t-test.

Hence, the target history concerning spatial alignment vs. spatial disparity influences the reaction time of a current congruent trial in a highly nontrivial way. In the Discussion we present a simple theoretical account for this finding.

## 6.4 Discussion

The present study tested whether human subjects adaptively account for the expected alignment of audiovisual stimuli when programming a rapid orienting response. In conditions where audiovisual stimuli were spatially aligned within the median plane, reaction times were lower, accuracy was higher, and endpoint variability was reduced, when compared to the unisensory evoked response statistics (Fig. 2). When stimuli were spatially disparate, the benefit of audiovisual integration broke down (Fig. 3). We measured the reaction time and accuracy of head orienting responses to simultaneously presented audiovisual stimuli, and explored whether there was a difference when the spatial stimulus statistics in the experiment changed. We hypothesized that if subjects keep track of the audiovisual congruency of prior trials to update their expectation of current stimulus alignment, we should observe an effect of stimulus history on their orienting responses.

Indeed, our main finding was that head saccades were systematically altered by the stimulus statistics, which was evidenced by a decrease in the average reaction times for experiments with only spatially-aligned stimuli, when compared to experiments in which the probability for spatial congruency was reduced (Fig. 4). Interestingly, sequences of trials that contained a larger proportion of congruent stimuli had decreased reaction times when compared to disparate stimulus sequences (Fig. 5). The change in reaction time depended systematically on the particular order of congruent and disparate stimuli of previous trials, regardless

of their actual spatial locations. This suggests that the brain indeed constructs a dynamic expectation regarding the spatial alignment of novel stimuli, which readily adjusts multisensory integration on the basis of recently acquired evidence.

*Audiovisual Integration.* In line with earlier research (Hughes et al., 1994; Nozawa et al., 1994; Frens et al., 1995; Goldring et al., 1996; Colonius & Arndt, 2001; Corneil et al., 2002; Van Wanrooij et al., 2009), we found that simultaneous presentation of spatially aligned stimuli in the median plane also systematically reduced the reaction times and accuracy for head saccades. Thus, the best-of-both-worlds principle (Corneil et al., 2002; Van Wanrooij et al., 2009) also applies to head movements, which are much slower and more variable than gaze saccades, especially in the vertical plane (Goossens & Van Opstal, 1997). This is a strong indication that the effects of audiovisual integration are real, and independent of either the particular pointer used, the stimulus environment (few or many targets, with or without a perturbing noisy background), or response dimension (1D, 2D, horizontal or vertical).

*Effect of stimulus statistics on saccade generation.* Note that we presented stimuli at randomly selected locations within the median plane over a large range (70 deg). In other words, it was highly unlikely that a particular spatial stimulus configuration would be repeated in the next trial. The only attribute that was systematically varied in the experiments was the probability for spatial congruency of the stimuli. Apparently, the mechanisms that subserve multisensory integration are able to extract and use this particular stimulus statistic.

The effect of stimulus statistics on saccade reaction times to visual stimuli has also been studied by Dorris and colleagues (2000) during single-unit recording of saccade-related activity in the monkey SC. In that study, the target could only occupy one of two possible locations (left vs. right), but the saccade reaction times, as well as the SC premotor activity levels, systematically varied with the trial history in much the same way as observed in the current study. For example reactions times were shortest (and premotor activity was highest) after a sequence of identical target locations, while RT increased (activity decreased) when the sequence alternated erratically between left and right.

Their data showed that the history effect occurred for the two particular stimulus locations employed, but they did not test whether the effect would transfer to other locations. Our study indicates that the RT effect on audiovisual integration generalises across all locations when the subject is tested for a large range of targets. Hence, the gaze control system is able to extract the relevant parameter from the trial statistics, which in our case was audiovisual spatial disparity. In line with the results of Dorris et al. (2000), and based on multisensory integration studies in anaesthetised preparations (Meredith & Stein, 1986a; b; Meredith et al., 1987) we conjecture that AV integration would be reflected in the activity of saccade-related cells of the SC. It would therefore be interesting to verify whether and how the rules of multisensory integration that emerge from our study would

be reflected in the sensory and/or preparatory activity epochs of these neurons.

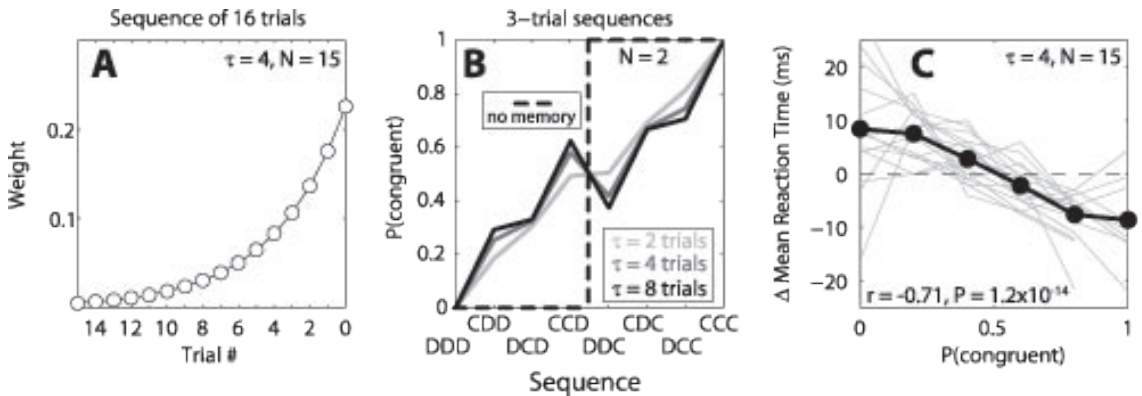
*Model for Reaction Time Modulation by Prior Likelihood Estimation.* The findings shown in Figs. 4 and 5 strongly suggest that audiovisual integration has a dynamic component that depends on the evidence for stimulus congruency as acquired from prior experience. To explain these results within a probabilistic framework we constructed a simple model that estimates the likelihood that audiovisual events may be congruent by weighting prior evidence. We further assume that the likelihood,  $P(\text{congruent})$ , modulates the saccade reaction time: the higher the likelihood, the earlier the response onset. In the model the prior probability of a congruent stimulus configuration is found by continuously updating the weighted averaged probability for stimulus alignment. In this model the occurrence of a disparate stimulus configuration is given weight zero (i.e. a zero post-hoc probability for being aligned). For simplicity, we took the dynamic probability for congruent stimulus configurations to follow an exponential decay (with a time constant  $\tau$ , with a memory  $N$  of up to 50 trials):

$$P(\text{congruent}; n=0) = \sum_{n=0}^N w(n) \times P(\text{congruent} | \text{congruent}, \text{incongruent})$$

$$\text{with } w(n) = w_0 \times \exp\left(-\frac{n}{\tau}\right)$$

$$P(\text{congruent} | \text{congruent}) = 1 \text{ and } P(\text{congruent} | \text{incongruent}) = 0 \tag{2}$$

$w_0$  is a normalization factor, which is 0.2253 for  $\tau=4$  and  $N=15$ . Such a model favours the most recent trials, and mimics a leaky memory (Fig 6A). For example,



**Figure 6.** Explanatory model for dynamic probability estimation of audiovisual congruency. **(A)** Memory weights for a sequence of 15 trials follow an exponential decay with time constant,  $\tau=4$  trials. **(B)** Example  $P(\text{congruent})$ , for 3-trial sequences for time constants of  $\tau=2, 4$  and  $8$  trials. Note that if multisensory integration would have no memory of previous trials, the prediction would follow a step function (**dashed line**), with  $P(\text{congruent})=1$  for all XXC trials and  $P(\text{congruent})=0$  for all XXD trials. **(C)** Average reaction time plotted against  $P(\text{congruent})$  for  $N=50, \tau=4$ , for all subjects and experiments. Reaction times and  $P(\text{congruent})$  of individual trials have been binned on the basis of 0.2-wide windows of  $P(\text{congruent})$ . Note that reaction time is clearly modulated by  $P(\text{congruent})$ . **Grey lines** – individual subjects and experiments; **black bold line** – average across subjects and experiments.

the contribution of an aligned trial that occurred six trials before the current trial is given a weight  $w(6) = 0.0503$ , for  $\tau=4$  and  $N=15$ . In Figure 6B we show the probabilities when the memory extends over  $N=2$  trials, for all eight possible triplet-sequences, for different decay time constants,  $\tau = 2, 4$  and  $8$  trials, respectively. Note that the model predicts that the trial sequence CCD (current trial disparate) has a higher probability for alignment (shorter reaction time) than DDC (current trial congruent) for time constants of  $4$  and  $8$ . Applying the model with  $\tau=4$ , and  $N$  up to  $50$  (depending on the number of previous trials in a block) to our data, we show a nice correspondence between the average change in reaction times (positive: increase, negative: decrease) and the estimated probability for congruent stimuli for the far majority of subjects (binned in  $0.2$ -wide windows, Fig. 6C). A time constant of  $4$  trials yielded the largest  $r^2$ -value ( $r=-0.71$ ).

In summary, our experiments provide clear evidence that the brain makes a dynamic evaluation of the multisensory scene to program a rapid orienting response that may or may not be based on multisensory integration. Such a strategy is particularly useful in unpredictable and complex environments where the statistics of auditory and visual stimuli may continuously vary.

## Acknowledgements

We thank Sharon Smits for help in the earlier phase of this study. We are grateful to Hans Kleijnen, Dick Heeren and Stijn Martens for valuable technical assistance. This research was supported by a Marie Curie Early Stage Training Fellowship of the European Community's Sixth Framework Program (MEST- CT-2004-007825 PB), a VICI grant within the Earth and Life Sciences of NWO (ALW 865.05.003; AJVO, MMVW), and the Radboud University Nijmegen (AJVO).

## Abbreviations

**A**, auditory; **V**, Visual; **AV**, audiovisual; **AV-100/0**, audiovisual experiment containing 100% aligned stimuli; **AV-50/50**, audiovisual experiment containing 50% aligned stimuli; **AV-10/90**, audiovisual experiment containing 10% aligned stimuli; **2D**, two-dimensional; **SC**, Superior Colliculus; **C**, congruent; **D**-disparate.

## 6.5 References

Alais D, Burr D. The ventriloquist effect results from near-optimal bimodal integration. (2004) *Curr Biol* 14:257–62.

Anastasio TJ, Patton PE, Belkacem-Boussaid K. Using Bayes' rule to model multisensory enhancement in the superior colliculus. (2000) *Neural Comput* 12: 1165–87.

Bell AH, Meredith MA, Van Opstal A.J, Munoz DP. Crossmodal integration in the primate superior colliculus underlying the preparation and initiation of saccadic eye movements. (2005) *J Neurophys* 93:3659–73.

Brydges R, Dubrowski A. Collision error avoidance: influence of proportion congruency and sensorimotor memory on open-loop grasp control. *Experimental brain research*. (2009) *Experimentelle Hirnforschung* 198:445–53.

Colonius H, Arndt P. A two-stage model for visual-auditory interaction in saccadic latencies. (2001) *Perception & Psychophysics* 63:126–47.

Corneil BD, Van Wanrooij MM, Munoz DP, Van Opstal AJ. Auditory-visual interactions subserving goal-directed saccades in a complex scene. (2002) *J Neurophys* 88:438–54.

Dorris MC, Pare M, Munoz DP. Immediate neural plasticity shapes motor performance. (2000) *J Neurosci* 20:RC52.

Ernst MO. A Bayesian view on multimodal cue integration. (2005) In Knoblich, G., Grosjean, M., Thornton, I., Shiffrar, M. (eds) *Human body perception from the inside out*. Oxford University Press, New York, pp. 105–31.

Ernst MO. Learning to integrate arbitrary signals from vision and touch. (2007) *J Vision* 7:1–14.

Frens MA, Van Opstal AJ. Visual-auditory interactions modulate saccade-related activity in monkey superior colliculus. (1998) *Brain Res Bull* 46:211–24.

Frens MA, Van Opstal AJ, Van der Willigen RF. Spatial and temporal factors determine auditory-visual interactions in human saccadic eye movements. (1995) *Perception & Psychophysics* 57:802–16.

Goldring JE, Dorris MC, Corneil BD, Ballantyne PA, Munoz DP. Combined eye-head gaze shifts to visual and auditory targets in humans. (1996) *Exp Brain Res* 111:68–78.

Goossens HH, Van Opstal AJ. Human eye-head coordination in two dimensions under different sensorimotor conditions. (1997) *Exp Brain Res* 114: 542–60.

Harrington LK, Peck CK. Spatial disparity affects visual-auditory interactions in human sensorimotor processing. (1998) *Exp Brain Res* 122:247–52.

Hillis JM, Ernst MO, Banks MS, Landy MS. Combining sensory information: mandatory fusion within, but not between, senses. (2002) *Science* 298: 1627–30.

Hughes HC, Nelson MD, Aronchick DM. Spatial characteristics of visual-auditory summation in human saccades. (1998) *Vision Res* 38:3955–63.

Hughes HC, Reuter-Lorenz PA, Nozawa G, Fendrich R. Visual-auditory interactions in sensorimotor processing: saccades versus manual responses. (1994) *J Exp Psych* 20:131–53.

Kording KP, Beierholm U, Ma WJ, Quartz S, Tenenbaum JB, Shams L. Causal inference in multisensory perception. (2007) *PLoS ONE* 2:e943.

Meredith MA, Nemitz JW, Stein BE. Determinants of multisensory integration in superior colliculus neurons. I. Temporal factors. (1987) *J Neurosci* 7:3215–29.

Meredith MA, Stein BE. Spatial factors determine the activity of multisensory neurons in cat superior colliculus. (1986a) *Brain Res* 365:350–4.

Meredith MA, Stein BE. Visual, auditory, and somatosensory convergence on cells in superior colliculus results in multisensory integration. (1986b) *J Neurophys* 56:640–62.

Nozawa G, Reuter-Lorenz PA, Hughes HC. Parallel and serial processes in the human oculomotor system: bimodal integration and express saccades. (1994) *Biol Cybern* 72:19–34.

Press WH, Flannery BP, Teukolsky SA, Vetterling WT. *Numerical Recipes in C: the art of scientific computing*. (1992) Cambridge University Press, Cambridge MA, USA.

Robinson DA. A Method of Measuring Eye Movement Using a Scleral Search Coil in a Magnetic Field. (1963) *IEEE Trans BioMed Eng* 10:137–145.

Sato Y, Toyoizumi T, Aihara K. Bayesian inference explains perception of unity and ventriloquism aftereffect: identification of common sources of audiovisual stimuli. (2007) *Neural Comp* 19:3335–55.

Shore DI, Simic N. Integration of visual and tactile stimuli: top-down influences require time. (2005) *Exp Brain Res* 166:509–17.

Stein BE, Meredith MA. *The Merging of the Senses*. (1993) MIT, Cambridge, MA.

Van Wanrooij MM, Bell AH, Munoz DP, Van Opstal AJ. The effect of spatial-temporal audiovisual disparities on saccades in a complex scene. (2009) *Exp Brain Res* 198:425–37.

Wallace MT, Meredith MA, Stein BE. Multisensory integration in the superior colliculus of the alert cat. (1998) *J Neurophys* 80:1006–10.

Wallace MT, Roberson GE, Hairston WD, Stein BE, Vaughan JW, Schirillo JA. Unifying multisensory signals across time and space. (2004) *Exp Brain Res* 158:252–8.





## **Chapter 7**

# **Summary**

## Chapter 2

Chapter 2 of this thesis described the extension of the DMI method to head-unrestrained paradigms. The results were obtained with a horizontally rotatable dummy platform that simulated the subject's eye and head movements. We showed that the addition of a third perpendicular magnetic field overcomes the limited measurement range of the method (maximally 20 deg from straight-ahead) and at the same time extends the measurement range to a full 360 deg. By using a simple, three-layer feed forward neural network with only a few hidden units, in combination with knowledge about the head orientation in space, the small ring signal could be extracted from the difference coil signal. This was even possible, when the pickup coil and the anti-coil were not exactly aligned, or when small inhomogeneities were present in the fields that induce a complex head position dependent offset. The absolute calibration precision with the neural networks (eight hidden units) was  $\pm 1$  deg over the entire measurement range.

## Chapter 3

After having shown in chapter 2 that in principal the DMI method can be used to record head-unrestrained gaze shifts chapter 3 dealt with the implementation of the method for the use with human subjects. The pickup coil and anti-coil assembly needed to pickup the ring signal and a calibration routine were described. A lightweight frame contained the pickup coil and anti-coil. The distance of ring and pickup coil could be adjusted with wax to yield an optimal signal to noise ratio and to fit the subject's anatomy. In addition, a small coil was wound around a laser pointer that was mounted on the nose bridge. With the coil the subject's head movements could be measured. The laser pointer guided the subject's head movements during the calibration routine. For a successful calibration the subject had first to align the head with a visual target, keep the head pointing at this target and subsequently fixate another target with the eye.

The data were calibrated with artificial neuronal networks. The quality of the DMI calibration was tested by simultaneously recording the other eye with the SSC technique. Correlation coefficients for the horizontal and vertical eye positions, velocity and acceleration of the calibrated DMI and SSC signals were close to one. So that the difference in position between the two methods remained within a few degrees. Due to a slightly worse signal-to-noise ratio for the DMI method the resolution of this method was 0.3 deg compared to 0.2 deg for the SSC technique (0.5 deg for the oculomotor system). The results demonstrated that the DMI method is a valuable alternative to the SSC technique, especially for patient and laboratory animal studies.

## Chapter 4

Chapters 2 and 3 introduced the extensions needed to employ the DMI method in head-unrestrained paradigms. In particular chapter 3 demonstrated the technical feasibility of these extensions with human subjects. The goal of our new DMI method is its implementation for use in electrophysiological recordings with freely behaving laboratory animals. Implants on the animal's skull can be used to rigidly fix the pickup coil assembly in a reproducible way between recording sessions. In this way a calibration database can be collected over a longer period increasing the quality of calibration. Chapter 4 described the application of the head-unrestrained DMI method to rhesus monkeys. The calibration paradigm was simplified compared to the coordinated eye-head paradigm employed with humans. The animal only needed to fixate randomly presented visual targets distributed in the frontal hemisphere. The distribution of targets was found by simulating gaze shifts and the corresponding DMI signals *in silico*. The simulations allowed for the objective quantification of the quality of calibration.

## Chapter 5

Chapter 5 dealt with the orienting behaviour of human subjects toward two simultaneously presented sounds in the vertical plane. Natural acoustic environments are complex containing a mixture of multiple sound sources. The auditory system needs to select behavioural relevant information from this mixture to program a response. In the experiment human listeners rapidly oriented their heads towards either a single sound source (broad-band buzzer, or Gaussian noise), or towards two simultaneously presented sounds (buzzer and noise) at a wide variety of locations in the mid-sagittal plane. In the single-sound condition localisation was accurate. However, in the double-sound condition three response modes were observed: 1) the loudest sound dominated the responses, 2) weighted averaging for small spatial disparities and about equal intensities and 3) bimodal distributions for spatial disparities exceeding about 45 deg. This response behaviour could be explained by the subject's elevation dependent filter function of torso, head and outer ear. This peripheral mechanism is in stark contrast to the central mechanism that accounts for double stimuli responses in the visuomotor system.

## Chapter 6

The last chapter of this thesis described the localization behaviour of human subjects elicited by audiovisual stimuli presented in elevation. The performed experiment expanded earlier studies by analyzing the influence of stimulus statistics on audiovisual integration. The subjects oriented their head to the perceived location of a visual target. Simultaneously an accessory auditory stimulus was presented. Throughout an experimental block the probability of spatial alignment and misalignment of the auditory stimulus with the visual target was varied and its

influence on the orienting response analyzed. The obtained results suggest that the strength of audiovisual integration, measured as a reduction in reaction time with respect to the unimodal responses, is dynamically updated. This is done on the basis of an integration of the probability of an occurrences of congruent or incongruent audiovisual stimuli in previous trials. A simple model based on an exponential weighting of the probability was able to predict the measured reaction times.

## Chapter 8

# Samenvatting

## Hoofdstuk 2

Hoofdstuk 2 van dit proefschrift beschrijft een belangrijke uitbreiding van de zogeheten “Dubbel-Magnetische Inductie” (DMI) methode, die ons in staat stelt om oog bewegingen onder hoofd-vrije situaties nauwkeurig te meten. De beschikking over een dergelijke methode is belangrijk omdat de klassieke oogspoel het grote nadeel heeft dat deze gemakkelijk kan breken, en a.g.v. het verbindingsdraadje snel kan irriteren. Met name voor toepassing in proefdieren is het cruciaal om over een duurzame en precieze meetmethode te kunnen beschikken. Om de theoretische voorspellingen van deze vernieuwde methode te testen, beschrijf ik de resultaten van metingen die zijn verkregen met behulp van een roterbaar dummy platform. Met dit platform konden we de onafhankelijke oog- en hoofdbewegingen van de proefpersoon simuleren over het gehele 360 graden bereik van hoeken in het horizontale vlak. Wij laten zien dat het mogelijk is om met een derde, loodrecht op de overige twee velden staand magnetisch veld, de sterke reductie van het meetbereik van de klassieke DMI-methode (maximaal 20 deg ten opzichte van recht vooruit) uit te breiden naar het volledige 360 graden rotatiebereik. Door vervolgens gebruik te maken van een eenvoudig drie-laags neurale netwerk met slechts enkele “hidden units”, in combinatie met gegevens over de gecalibreerde hoofdoriëntatie in de ruimte, kan het relatief zwakke en sterk niet-lineaire ringsignaal worden gecalibreerd. Deze procedure werkt even goed als oppikspoel en anti-spoel niet precies zijn opgelijnd, en als de ring zich niet precies achter het centrum van de oppikspoel bevindt. Tevens kan het neurale netwerkje eventuele inhomogeniteiten in de magnetische velden goed compenseren. De absolute precisie van de calibratie met het neurale netwerk (met acht “hidden units”) was  $\pm 1$  deg over het gehele meetbereik, bij een resolutie van ongeveer 0.3 graden. De precisie en resolutie van de hoofd-vrije DMI methode zijn daarmee vergelijkbaar met de “gouden standaard” van de klassieke oogspoel methode, en kan op termijn daarvoor een volwaardige vervanger worden in zowel proefdieronderzoek, als bij hoofd-vrije oogbewegingsmetingen aan patiënten.

## Hoofdstuk 3

Hoofdstuk 3 beschrijft een implementatie van de hoofd-vrije DMI methode voor gebruik in proefpersonen. De oppikspoel en de antispoel die beide nodig zijn om het zwakke ringsignaal op te pikken zijn ondergebracht in een eenvoudig brilmontuur. Het hoofdstuk beschrijft tevens een calibratieprocedure, waarmee het mogelijk is om met een relatief gering aantal calibratiedoelen ( $n=108$ ) een goede twee-dimensionale calibratie te verkrijgen. Het montuur was dusdanig gemaakt dat de afstand tussen ring en oppikspoel eenvoudig kon worden aangepast zodat een optimale signaal-ruis verhouding kon worden bereikt. Deze eenvoudige constructie maakte het mogelijk om de bril eenvoudig aan de anatomie van de proefpersoon aan te passen, en het geheel goed-passend en trillingsvrij aan het hoofd te bevestigen. Daarnaast was een klein spoeltje om een kleine laser



pointer aangebracht op de neusbrug van de bril, waarmee de hoofdbewegingen van de proefpersoon onafhankelijk konden worden gemeten. De laserpointer kon worden gebruikt om de hoofdbewegingen van de proefpersoon tijdens de calibratieroutine te leiden. Voor de calibratie moest de proefpersoon eerst het hoofd oplijnen met een visueel doel, en vervolgens - terwijl het hoofd stil werd gehouden - met het oog naar een tweede visueel doel kijken. Hierdoor konden verschillende oog-hoofd combinaties worden verkregen.

De data konden met behulp van neurale netwerken worden gecalibreerd. De kwaliteit van de hoofd-vrije DMI calibratie werd vervolgens getoetst door tegelijkertijd het andere oog met de klassieke oogspoeltechniek te bemeten. De correlatiecoëfficiënten tussen de metingen van de DMI methode en de oogspoelmethode voor de horizontale en verticale oogposities, de oogsnelheden en de oogversnellingen waren dicht bij de maximale waarde van 1.0. Het verschil in gecalibreerde posities tussen de twee methoden bleef binnen enkele graden, wat in de grootteorde valt van normale binoculaire verschillen. Omdat de signaal-ruis verhouding voor de DMI methode iets lager was dan die van de oogspoeltechniek, was de resolutie van de DMI methode 0.3 graden, en die van de oogspoelmethode 0.2 graden. Deze resoluties vallen binnen de fixatienauwkeurigheid (0.5 graden) van het oculomotor systeem. De resultaten laten daarmee zien dat de hoofdvrije DMI methode een goed en bruikbaar alternatief vormt voor de kwetsbare en relatief dure oogspoeltechniek, met name voor toepassing bij patiënten en bij proefdieren. Voor deze laatste categorie zou dan echter wel een eenvoudiger calibratieroutine moeten worden ontwikkeld.

## Hoofdstuk 4

Een belangrijk doel van mijn onderzoek was om de hoofd-vrije DMI methode in electrofysiologische experimenten bij zich gedragende proefdieren te kunnen gaan toepassen. Hoofdstuk 3 liet zien dat de methode goed werkt, maar dat een eenvoudiger calibratieroutine nodig zou zijn om de training van proefdieren niet nodeloos te compliceren. De implantaten op de schedel van het dier kunnen worden gebruikt om het zeer lichte "brillettje" met oppikspoel-antispoel en hoofdbewegingsspoeltjes/laser reproduceerbaar, stabiel en trillingsvrij aan te brengen. Daardoor wordt het mogelijk om over meerdere experimenteersessies een steeds beter wordende calibratie database te verzamelen. Hoofdstuk 4 beschrijft de toepassing van de DMI methode op de hoofd-vrije rhesusaap. We ontwierpen daartoe een sterk vereenvoudigd calibratie paradigma in vergelijking tot de toepassing bij mensen. De aap hoeft hierbij slechts een random verspringend visueel doel nauwkeurig te volgen, zonder dat daarvoor een precieze instructie op de hoofdbeweging noodzakelijk is. Uitgebreide computersimulaties lieten zien hoe een optimale verdeling van visuele doelposities kan worden gekozen, gebruikmakend van de eigenschap dat oog en hoofd niet synchroon en even snel bewegen. Hierdoor kan in elk trial op efficiënte wijze een groot aantal variërende oog-hoofd combinaties worden verkregen. De simulaties stelden ons tevens in staat om de kwaliteit van de cali-

bratie objectief te bepalen, en tot een goede schatting te komen voor het benodigde aantal hidden units in de neurale netwerken, en het benodigde aantal doelsprongen. Deze hoofd-vrije DMI methode wordt daardoor ook geschikt voor toepassing in andere proefdieren (bijv. katten, marmosets), waarbij de enige restrictie is dat het dier getraind kan worden om visueel-uitgelokte oogbewegingen te maken.

## Hoofdstuk 5

Een belangrijke vraag bij geluidslocalisatie is hoe het auditief systeem in staat is om geluid, wat in een natuurlijke omgeving samengesteld kan zijn uit een onbekend aantal verschillende gesuperponeerde bronnen, te scheiden in doel vs. niet-doel. Het auditief systeem dient dus relevante informatie uit dit akoestisch mengsel te selecteren om een doelgerichte localisatierepons te kunnen programmeren. Van de mechanismen die een dergelijk selectieproces mogelijk maken is nog weinig bekend. Hoofdstuk 5 beschrijft het orientatiegedrag van proefpersonen naar simultaan gepresenteerde breedbandige geluiden in het verticale vlak. In onze experimenten dienden menselijke luisteraars een snelle hoofdbeweging te maken naar òf een enkele geluids bron (hetzij breed-bandige zoemer, hetzij breedbandige Gaussianse witte ruis), òf naar twee tegelijkertijd afgespeelde geluidsbronnen (zoemer en ruis). De geluiden werden op verschillende locaties in het mid-sagittale vlak aangeboden, en bij de dubbelstimuli werden de relatieve intensiteiten van zoemer (het doel) en ruis (het niet-doel) systematisch gevarieerd. In het geval van één bron waren de responsies nauwkeurig, en onafhankelijk van de geluidsterkte. In het geval van twee simultaan gepresenteerde bronnen observeerden we drie verschillende responspatronen, die afhankelijk waren van de relatieve intensiteiten en relatieve bronposities: 1) het hardste geluid domineerde de responsies voor alle relatieve bronposities, 2) gewogen middeling werd waargenomen voor kleine ruimtelijke dispariteiten én kleine intensiteitsverschillen, en 3) we verkregen bimodale responsdistributies als de bronnen een bijna gelijke intensiteit hadden, maar een ruimtelijke dispariteit die groter was dan ongeveer 45 graden. Het interessante was dat deze gedragpatronen volledig verklaard kunnen worden met behulp van de elevatie-afhankelijke akoestische filterkarakteristieken van torso, hoofd en oorschelp van de proefpersoon. Dit perifere akoestische mechanisme staat daarmee in groot contrast met conclusies die getrokken zijn in een groot aantal studies aan het visuomotor systeem. In dat geval blijkt namelijk dat vergelijkbare responspatronen moeten worden verklaard met centrale interactiemechanismen in structuren als bijvoorbeeld de oculomotor middenhersenen. Onze studie legt daarmee een fundamenteel verschil bloot tussen de verwerking van visuele en auditieve informatie.

## Hoofdstuk 6

Het laatste hoofdstuk van mijn proefschrift beschrijft het localisatiegedrag van proefpersonen naar audio-visuele stimuli in het verticale vlak (elevatie). Het uitgevoerde experiment vormt een uitbreiding van eerdere studies uit ons laboratorium, waarin proefpersonen dergelijke stimuli dienden te localiseren in een omgeving met vele afleidende stimuli, maar waarin bekend was dat de auditieve en visuele stimulus altijd zouden zijn opgelijnd in de ruimte. In de echte wereld echter, is dergelijke voorkennis niet aanwezig, en het is daarom de vraag hoe het brein kan weten of auditieve en visuele informatie dient te worden geïntegreerd of niet. We vroegen ons daartoe af in hoeverre proefpersonen kennis uit de omgeving opdoen met betrekking tot de spatiële eigenschappen van stimuli, door de invloed van de stimulus statistiek op audio-visuele integratie te onderzoeken. Proefpersonen hadden tot taak om een snelle hoofdbeweging te maken naar de locatie van een visueel doel. Tegelijkertijd met dit doel werd een auditieve stimulus aangeboden, waarover de proefpersoon geen a-priori informatie kreeg. Hij diende deze bron dan ook te negeren (niet-doel). De proefpersoon werd in een gegeven blok met stimuli geconfronteerd met één van drie verschillende audio-visuele “werelden”, waarin de waarschijnlijkheid dat visuele en auditieve stimuli waren opgelijnd werd gevarieerd. De proefpersoon had over deze werelden geen expliciete kennis. In wereld 1 waren de stimuli altijd opgelijnd. In de andere twee werelden gebeurde dit slechts in 16% (wereld 2) of in 50% (wereld 3) van de gevallen. In wereld 2 waren de relatieve locaties tussen auditief en visueel volledig random (uniform verdeeld), maar in wereld 3 waren in de niet-opgelijnde condities de stimuli steeds ver uiteen. Op deze manier was het mogelijk om de invloed van “geschatte opgelijndheid” op het gedrag te bestuderen. De resultaten laten zien dat de sterkte van audio-visuele integratie (gemeten door verandering in de reactietijden) dynamisch wordt bijgewerkt. Dit proces blijkt te zijn gebaseerd op een interne schatting van de waarschijnlijkheid dat stimuli in het volgende trial zullen zijn opgelijnd, of juist niet. Als bijv. in de voorgaande drie trials steeds oplijning was waargenomen, wordt de kans op weer oplijning als hoog geschat, en vindt sterke audio-visuele integratie plaats. Een eenvoudig model dat is gebaseerd op een exponentieel met de tijd verlopende weging van deze waarschijnlijkheden kan de gemeten reactietijden goed verklaren.



## Publications

Van Wanrooij MM, Bremen P, Van Opstal AJ. Acquired prior knowledge modulates audiovisual integration. *Eur J Neurosci*. accepted

Singheiser M, Plachta DT, Brill S, Bremen P, Van der Willigen RF, Wagner H. Target-approaching behavior of barn owls (*Tyto alba*): influence of sound frequency. (2010) *J Comp Physiol A*. Feb Epub ahead of print.

Bremen P, Van Wanrooij MM, Van Opstal AJ. Pinna cues determine orienting response modes to synchronous sounds in elevation. (2010) *J Neurosci*. 30(1):194–204.

Bremen P, Van der Willigen RF, Van Opstal AJ. Applying double magnetic induction to measure two-dimensional head-unrestrained gaze shifts in human subjects. (2007) *J Neurophysiol*. 98(6):3759–69.

Wagner H, Asadollahi A, Bremen P, Endler F, Vonderschen K, Von Campenhausen M. Distribution of interaural time difference in the barn owl's inferior colliculus in the low- and high-frequency ranges. (2007) *J Neurosci*. 27(15):4191–200.

Bremen P, Van der Willigen RF, Van Opstal AJ. Using double-magnetic induction to measure head-unrestrained gaze shifts. I. Theory and validation. (2007) *J Neurosci Methods*. 160(1):75–84.

Bremen P, Poganiatz I, Von Campenhausen M, Wagner H. Sensitivity to interaural time difference and representation of azimuth in central nucleus of inferior colliculus in the barn owl. (2007) *J Comp Physiol A*. 193(1):99–112.

## **Published Abstracts/Conference Presentations**

Bremen P, Hovingh R, Van Opstal AJ. Orienting towards Auditory Double Stimuli in Elevation. 32nd Annual Association for Research in Otolaryngology MidWinter Meeting, Baltimore, MD, February 2009.

Bremen P, Van der Willigen RF, Van Opstal AJ. Head-unrestrained 2D Gaze Shifts Measured With Double Magnetic Induction (DMI): A Direct Comparison with The Scleral Search Coil (SSC). Proceedings of the 17th Annual Meeting of the Society for the Neural Control of Movement (NCM), Sevilla, Spain, March 2007.

Van der Willigen RF, Bremen P, Van Opstal AJ. Head-unrestrained 2D Gaze Shifts Measured With Double Magnetic Induction (DMI): Theory and Validation. Proceedings of the 17th Annual Meeting of the Society for the Neural Control of Movement (NCM), Sevilla, Spain, March 2007.

Bremen P, Singheiser M, Plachta DTT, Van der Willigen RF, Wagner H. Barn owls do not depend on high frequency auditory signals to approach a distant target. Proceedings of the 6th Meeting of the German Neuroscience Society/30th Goettingen Neurobiology Conference [CD\_ROM], supplement to Neuroforum 2005, February; ISSN 0947-0875, 2005.

# Acknowledgements

It's unbelievable how fast time flies by when you're having fun! A number of people deserve credit for the fact that I had such a pleasant time at the Department of Biophysics.

First of all I would like to thank the captain. John, your ability to motivate people is admirable. In my memory there are a lot of incidences when I entered your office making a gloomy face and uttering in a grave voice something like: "John, dit ziet er niet goed uit...". But miraculously after an intense talk and in-depth analysis I left your room with high spirits and the feeling that what I had done so far was not so stupid after all and might actually lead to the solution of the problem at hand. Thank you for being a great supervisor, colleague and co-author. Too sad, however, that you don't share my interest in brain-dead zombi animals and flies.

First-mate, Rob. Without you I would never have ended up in Nijmegen. Your thoughtful advice has helped me to keep a steady course during those stormy four years as a PhD student. It continues to amaze me that you had enough faith in me to come aboard and embark on the voyage of the HMS DMI with me as the helmsman considering the fact that we had quite often what I would call "lively discussions" and what less sturdy characters probably would describe as serious quarrels. I know already that I will miss your refreshing opinions and the fact that you always had an open ear to discuss whatever was on my mind. I owe you a lot. Thanks for everything!

When I had learned everything I could from Rob, Marc (aka Little-John) came on board to teach me a few things about human sound localization, multi-sensory integration and matlabin. As true nerds we also shared some hobbies that are too shocking for the general public to be named here. Marc, I really enjoyed our trip to the ARO 2009 in Baltimore and I for one won't forget our appointment at the 2011 revival of the meeting to discuss future cooperations. See you there!

During the voyage several junior seawomen and seamen in the form of students were picked up. In chronological order: Jerimiah. Your numerical simulations really helped in understanding the complicated DMI signal. And of course, I will never forget our one-and-only recording session with you as a subject! Don't tell your pupils about that or they won't dare to study "Natuurwetenschappen". Next came Robert. You liked the group so much that you not only completed your Bachelor thesis but also your Master thesis with us. Although our initial audio-visual integration project did not yield any significant new results the second project was the starting point for the experiments described in chapter 5. I am sure you will recognize the sounds that were used. Not able to admit defeat and with the expertise of Marc we revived the audio-visual integration experi-



ments. Enter student number three: Sharon, you had to put up with two supervisors who most of the times “slightly” disagreed about how to proceed and who never stopped asking for “just this last plot”. Last but not least, Lieveke hopped on board for a mini-internship lasting only two month. Enough time to expand the auditory double experiments to two dimensions. Lieveke, from the beginning on you amazed me with your ability to quickly and critically analyze the data we collected. Sadly, your time in the lab went by too fast so that we did not have the time to delve into the nasty details of two-dimensional double sound localisation!

I want to thank you all for your commitment and enthusiasm. It’s been a pleasure to work with you all. Good luck for your future endeavors.

What would a cruise be without fellow mates? Therefore, a big thanks for all the junior lunches and the inspiring conversations on the work-floor goes out to Andrew, Artem, Bart, Denise, Gleb, Joke, Joris, Julian, Maaike, Rens, Sigrid, Tom and Yoola. Good luck to you all for your present and future projects!

Obviously a ship needs craftsmen to fix the little leaks that turn up along the way. These skillful artisans came in the form of Stijn, Hans, Dick and occasionally Gunther and Ger. The development of the DMI method would not have been possible without your commitment. I want to thank you, Stijn, for the extra hours you spent on making the little changes we needed until yesterday and of course for posing as a model for the nice picture that made it into paper number two. I hope you enjoyed working on the project as much as I did!

A big thanks is due to the secretarial staff. Judith, Magriet and Irene, thank you for taking care of the bothersome details of ordering stuff and explaining for the hundredth time how to fill in the “aanvraag reis- en verblijfkosten formulier”.

Without the expertise of the CDL personnel the monkey DMI experiments could never have been performed. Jeroen dB, Henk and Annette fed the animals during the weekends and cleaned the cages. Alex, Wilma, Jeroen M and Dr. Eric Schaling assisted in/performed the surgeries. Geertje and Matthieu as veterinarians took care of the well-being and health of the animals. Thank you for your assistance and conversations on the work-floor.

On a more private note I would like to extend my gratitude to my family who supported me unconditionally throughout those years and who had the faith that my choice to study biology (of all fields) couldn’t be that bad. Zu Deutsch: Mama, Papa, vielen Dank, dass ihr mich von klein auf an bedingungslos unterstuetzt habt. Ich weiss euer Vertrauen in mich und meine Entscheidung Biologie (und nicht etwa BWL) zu studieren, sehr zu schaeetzen.

Tamara, thank you for the wonderful years we shared. I will never forget you.

## Curriculum Vitae

I was born on the 29th of November 1979 in Stolberg (Rhld.), Germany. My thirteen years of pre-university education were spent at the Grundschule Breinig and the Ritzefeld Gymnasium Stolberg after which I obtained my Abitur (university-entrance diploma) in 1999. After wasting nine months of my life at the Heinrich Hertz Kaserne, Daun, fulfilling basic military service I started my biology studies at the RWTH-Aachen in 2000. Two years later I obtained my Vordiplom (B.sc.). Doing my studies in the time of the Human Genome Project, initially my interest was mainly focused on molecular biology and its possibilities of creating new enhanced mutant species. During the Vordiplom, however, I became intrigued by the neurosciences, especially by the auditory system when I first heard about the barn owl's astonishing ability to strike prey in total darkness solely relying on its sense of hearing. Therefore, I performed the work on my Diploma thesis in the laboratory of Hermann Wagner working on the binaural system of the barn owl. The degree of "Diplom Biologe" (M.sc.) was obtained mid 2005 and in September 2005 I started my PhD project under the supervision of John Van Opstal at the Department of Biophysics in Nijmegen as a Marie-Curie SensoPrim fellow. The results of the four years as a PhD student in John's lab are described in this thesis. As of October 2009 I am a postdoc in the lab of Philip Joris in Leuven, Belgium, working - this time - on the dark side of the binaural system.