

On the language of Space:
Neurocognitive studies in blind and sighted individuals

Marijn Struiksmā

Reading Committee

Prof. dr. G. R. Semin

Utrecht University, the Netherlands

Prof. dr. T. Vecchi

University of Pavia, Italy

Prof. dr. R. A. Zwaan

Erasmus University Rotterdam, the Netherlands

Prof. dr. N. F. Ramsey

Utrecht University Medical Center, the Netherlands

Dr. R. H. J. van der Lubbe

University of Twente, the Netherlands

PhD thesis, Experimental Psychology, Helmholtz Institute, Utrecht University

ISBN: 978-90-8891-246-7

Cover design: Rianne Poortman

Printed by: Proefschriftmaken.nl | | Printyourthesis.com

Published by: Uitgeverij BOXPress, Oisterwijk

© 2011. All rights reserved. No parts of this work may be reproduced, stored in a retrieval system, or transmitted, in any form or by any means, electronic, mechanical, photocopying, recording, or otherwise, without the prior permission of the author.

On the language of Space: Neurocognitive studies in blind and sighted individuals

De taal van de ruimte:

Neurocognitieve studies bij blinde en ziende individuen

(met een samenvatting in het Nederlands)

Proefschrift

ter verkrijging van de graad van doctor aan de Universiteit Utrecht op gezag van de rector magnificus, prof. dr. G. J. van der Zwaan, ingevolge het besluit van het college voor promoties in het openbaar te verdedigen op woensdag 6 april 2011 des middags te 2.30 uur

door

Marijn Eefje Struiksma

geboren op 23 oktober 1980
te Kingston, Jamaica

Promotor: Prof. dr. A. Postma
Co-promotoren: Dr. M. L. Noordzij
Dr. S. F. W. Neggers

Contents

Chapter 1	7
General introduction	
Chapter 2	17
What is the link between language and spatial images?	
Behavioral and neural findings in blind and sighted individuals	
Simple spatial sentences	41
Chapter 3	43
Spatial language processing in the blind: Evidence for a supramodal representation and cortical reorganization	
Chapter 4	63
The level of preparation and control in reference frame processing	
Chapter 5	93
Reference frame acceptability in haptics differs for the blind and sighted in the horizontal but not in the vertical plane	
Complex spatial descriptions	107
Chapter 6	109
Neural correlates of learning route and survey knowledge from complex spatial descriptions	
Chapter 7	133
Neural correlates of route and survey distance judgments from spatial descriptions of different perspectives	

Grounded cognition	155
Chapter 8	157
Embodied representation of the body contains veridical spatial information	
Chapter 9	175
Spatial iconicity in blind and sighted individuals	
Chapter 10	185
Summary and conclusions	
Nederlandse samenvatting	195
References	207
Publications and conference abstracts	221
List of co-authors and their affiliations	223
Curriculum Vitae	224
Dankwoord / Acknowledgments	225

Chapter 1

General introduction

Molyneux's question

On July 7th 1688 William Molyneux (1656-1698) wrote a letter to John Locke (1632-1704) posing a question about a congenitally blind man¹:

A problem proposed to the author of the Essai Philosophique concernant L'Entendement.

A man, being born blind, and having a globe and a cube, nigh of the same bignes, committed into his hands, and being taught or told, which is called the globe, and which the cube, so as easily to distinguish them by his touch or feeling; Then both being taken from him, and laid on a table, let us suppose his sight restored to him; Whether he could, by his sight, and before he touch them, know which is the globe and which the cube? Or whether he could know by his sight, before he stretched out his hand, whether he could not reach them, tho they were removed 20 or 1000 feet from him?

The question posed by Molyneux on whether a blind man, able to distinguish between a sphere and a cube, with restored vision would be able to distinguish between the objects based on vision, is still relevant today. Molyneux's question referred to whether processing spatial information is innate, or based on experience. A slightly different interpretation of this question could address the nature of conceptual reasoning. Is spatial perception and understanding modality-specific, or rather converted into an encompassing representation that communicates with all our input senses (vision, touch, language, etc.)?

The first part of Molyneux's question refers to the apprehension of shapes and forms. However, shape and form perception also include a spatial dimension. In the original question Molyneux added a second part specifically referring to another spatial dimension: distance. Molyneux's question sets the stage for the present thesis that contains a new theoretical framework and novel experimental work on the nature of the mental representations built up from simple or more complex spatial descriptions. Are these representations propositional, i.e. a set of abstract processing instructions, or do they have analogue, spatially organized qualities? Vital evidence regarding this can be obtained from comparing early blind to sighted. This again bears on Molyneux's question: If spatial mental representations exist are they subject to experience or are they more innate?

The present thesis discusses a series of experiments in which we have studied the nature of spatial mental representations and the link with spatial language and imagery. In Chapter 2 we review the current literature on this topic. The first part (Chapters 3, 4 and 5) deals with simple spatial sentences and reference frame processing. Reference frames determine the point of view for communication and are very important for successful communication. Potentially, reference frames allow for communication between different sensory modalities, as illustrated by Molyneux's question. Spatial language hardly ever occurs in isolation but should also be considered in a larger

¹ This is the original question posed in the letter by Molyneux (Degenaar, 1992). Molyneux's question was first discussed in Locke's *Essay concerning Human Understanding* (book 2, ix, 8) making it publicly available, however, Locke only included the first part.

context such as talking about space and getting from one place to the other. Therefore, the second part of this thesis (Chapters 6 and 7) examined the processing of complex spatial descriptions. The reference frames that clarify how objects are located support different perspectives and possibly different representations, both behaviorally and neurally. This part describes neuroimaging studies that investigate whether the representations built up from these complex descriptions are indeed spatial, that is whether they represent metric information about the locations described. Finally, the third part (Chapters 8 and 9) focuses on how compelling such spatial information is when understanding language. Can spatial information also be accessed through language that is not necessarily spatial at first? This question links spatial representations to the influential idea of grounded cognition that states that language is understood by linking it to perceptual simulations.

Spatial cognition is essential to our interactions with the world

In everyday life we spend a large amount of time navigating through our surroundings. Throughout evolution, we have developed a range of complex vocabularies to efficiently communicate such environmental information to others. Successful wayfinding as well as adequate memory for objects and their locations, for example remembering where food and water can be found, have been essential for our survival. We obtain information about the spatial world from different sensory modalities, such as sight, hearing and touch. Sharing our spatial knowledge by means of spatial language, allows us to combine efforts to understand, remember and find our way in the world.

The classic debate on spatial information representation has contrasted a propositional, amodal semantic view (Fodor, 1975; Pylyshyn, 1981), a visuo-spatial mental imagery view (Kosslyn, 1988; Kosslyn et al., 2003), or a combination of the two (Noordzij & Postma, 2005; Paivio, 1971). In line with the latter two options, Barsalou (1999) has claimed that conceptual representations are grounded in a modality-specific perceptual symbols system. These perceptual symbols are re-enacted through perceptual simulations which are analogue to their modal referent. Translating this grounded view to spatial representations leads to a multimodal spatial representation. A multimodal spatial representation seems superfluous, moreover a certain level of convergence is essential for consistency in spatial locations. Instead of storing the same spatial information in multiple modalities a more parsimonious approach could be a supramodal spatial representation. In a supramodal mental representation a link to modality-specific information is maintained, but essential spatial information is extracted (Barsalou, 1999; Cattaneo et al., 2008). Chapter 2 discusses these possible models for spatial representations and reviews the current literature on the nature of spatial representations and the role of spatial language in generating them.

Processing simple spatial sentences

In essence spatial language is used to communicate the relation between objects. The most straightforward way to do that is using a simple sentence to describe the relation between two objects in a scene. This scene can be readily perceivable, or referring to a future scene, for example during route instructions. The relation between the objects is described using a spatial preposition. Spatial prepositions such as *above*, *in front of*, and *to the left of* are part of a small category of words (± 80 in most languages (Landau & Jackendoff, 1993)).

Classic work on processing simple spatial sentences, such as “*the plus is above the star*”, has been done using a sentence-picture verification paradigm (Clark & Chase, 1972). Participants had to judge whether or not the sentence and the picture described the same situation. The discussion at that time focused on two possible representational formats: a visual image or a set of propositions. Other options have been suggested, such as a dual code where both representations are available in parallel (Paivio, 1971). More recent evidence suggests a strategic model where a propositional code is generated by default and strategically a visual image can be created (Noordzij et al., 2004; Noordzij et al., 2005, 2006a). Based on the literature reviewed in Chapter 2 we suggest an alternative model: a supramodal representation. Spatial information is not modality-specific, in contrast to for example color which is uniquely visual. Rather, spatial information can be obtained from a number of different input modalities, but also through language. The supramodal representation extracts this basic spatial information, but a bi-directional link with the modalities is maintained. This supramodal representation allows us to describe a fictitious spatial scene and externalize characteristics of this scene through multiple output modalities. For example, the following description allows you to draw the scene, but also pick up the kiwi based on its location and texture: “*the girl picked up the kiwi, which was to the left of the orange in the glass bowl*”.

Particularly relevant to the discussion on the nature of spatial images are the results from blind people. Research with early blind people without any visual memories provides a unique opportunity to determine the importance of visual experience in processing spatial information. Although there are subtle to moderate differences between blind and sighted people, there is ample evidence that blind are able to navigate, similar to sighted people (for reviews see Cattaneo et al., 2008; Kaski, 2002; Thinus-Blanc & Gaunet, 1997). A supramodal representation of spatial information predicts a large degree of overlap in neural networks employed. This idea is tested in Chapter 3 where we describe a functional neuroimaging study of blind and sighted participants on processing simple spatial sentences.

When a spatial utterance refers to a directly visible scene verbal and perceptual information have to be compared with each other. Seminal work on the coupling between linguistic and visual information has been done by Carlson and colleagues (Carlson-Radvansky & Irwin, 1994; Carlson-Radvansky & Logan, 1997; Logan, 1994). Although, the research in this part deals with simple spatial sentences, understanding such a spatial sentence, e.g. “the man is to the left of the house”, is not simple at all. Information needs to be converted into a spatial mental representation which allows

comprehension of the situation (Tversky, 1991) and this requires several steps. Carlson-Radvansky and Logan (1997) suggested a model indicating processing steps required in spatial language comprehension (based on theories from e.g. Carlson-Radvansky & Irwin, 1994; Levelt, 1984; Logan & Sadler, 1996).

An important element involved in spatial language comprehension is imposing a reference frame, or point of view. In order for communication to be successful it is essential that the reference frames of the interlocutor and the listener align. Only then can the spatial situation be correctly understood. Reference frames consist of three axes, which parse up space into different directions specifying location, and have several parameters that can be adjusted: e.g. origin, orientation, direction, and distance (Carlson-Radvansky & Logan, 1997; Carlson & Van Deman, 2004). The relevant axis used to match a spatial sentence to a scene differs according to which of the three reference frames is adopted: absolute, relative or intrinsic (Carlson, 1999; Levinson, 1996). The absolute reference frame uses environmental characteristics, such as gravity and cardinal directions, to determine the orientation of the axes. In the relative reference frame the orientation of the axes is based on the viewer, while in the intrinsic reference frame the reference object defines the relevant axis (see Figure 1 for examples).

Cross-linguistic evidence has demonstrated that reference frame preferences can be different across cultures. Dutch and European languages prefer a relative frame of reference, while certain Mayan languages and Australian Aboriginal communities prefer, or might even only have access to, the absolute frame of reference (Levinson, 1996; Pederson et al., 1998). These reference frame preferences are not limited to language use, but are also found across modalities in a wide variety of nonverbal tasks such as inference tasks, recognition memory tasks and unconscious gesture (Levinson, 1996). This striking finding implies that reference frames are an essential part of spatial representations. The fact that reference frame preferences have been found across modalities suggests that these reference frames are not modality-specific. Instead, a supramodal representation could account for the fact that these reference frames are available to different sensory modalities. We explore the flexibility of reference frame use in spatial language understanding as applied to visual (Chapter 4) and haptic (Chapter 5) situations.

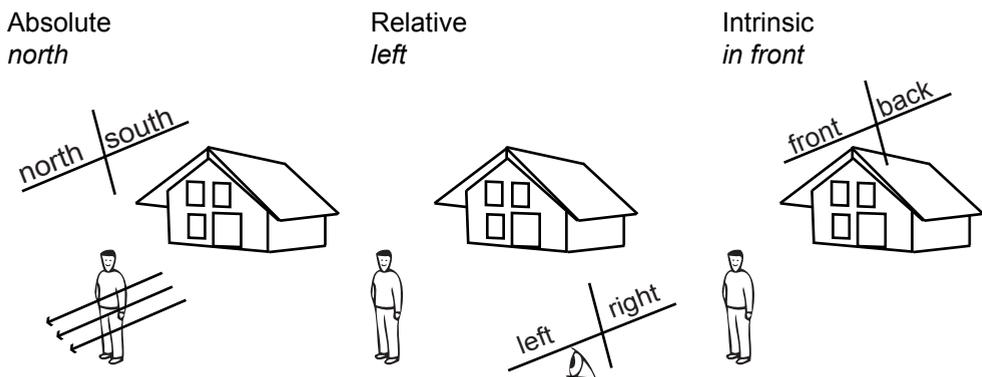


Figure 1 Examples of the intrinsic, relative, and absolute reference frame (adapted from Levinson, 1996).

Perspective in complex spatial descriptions

Of course everyday communication is not restricted to simple sentences. Spatial language can be extremely helpful in wayfinding situations. For example a tourist visiting your home town might approach you and ask you for directions to the museum. The ultimate goal of this action is for you to describe to the tourist how to get to the museum and for the tourist to build up a spatial mental representation of the town and the way to get from the current position to the museum. When the tourist is on foot, without a map, it is likely that you provide directions from a first-person perspective giving a mental tour through town. This mental tour, or route description, describes the relation between landmarks and the observer in a linear fashion using indications such as left and right. This property of linking landmarks to the observer is characteristic for an egocentric reference frame and shows some parallel with the relative reference frame described above. Although there are some parallels between the egocentric and relative reference frames and the route perspective, they are not identical and are related to different levels of processing (Pederson, 2003). Alternatively, when the tourist is viewing a map it is more useful to provide a birds-eye view describing the environment in a hierarchical fashion using cardinal directions, such as north, east, south and west to relate landmarks to other landmarks. This type of description uses a survey perspective and shows similarities with abstract, intrinsic or allocentric reference frames (Hund et al., 2008; Taylor & Tversky, 1992).

In order to navigate successfully the tourist most likely builds up an integrated spatial mental representation of the environment (Taylor & Tversky, 1992; Zwaan & Radvansky, 1998). The different strategies available to provide the directions to the museum suggest that the resultant spatial mental representations could be different. However, as argued before, spatial information can be encoded from different modalities resulting in a unique spatial image. A large body of literature has focused on different aspects of complex spatial descriptions and how they are encoded into memory. The reported results are mixed. There is both evidence for maintaining different aspects of perspectives and convergence into a single representation. For example, Taylor and Tversky (1992) have shown that participants who had read route descriptions performed similar on a number of tasks compared to participants who had read survey descriptions or viewed an actual map. This result suggests that all participants had built up a functional equivalent spatial mental representation after over-learning, i.e. different learning perspectives have yielded highly similar performance on a number of tasks. On the other hand, the encoding of a route description has shown to take longer and produce a higher load on visuo-spatial and central executive memory compared to the encoding of a survey description (Brunyé et al., 2008; Deyzac et al., 2006; Hubona et al., 1998). The resultant spatial representation from a route description has shown to be less flexible and has yielded more difficulty with using a different perspective.

These mixed results reported in the literature indicate that building up a supramodal spatial mental representation is a progressive process. At first, perspective dependent information is available, however, after over-learning a perspective independent supramodal spatial mental representation can be built up. This progressive process has been investigated using neuroimaging (Chapter 6)

in order to determine what exactly supports such a spatial mental representation at a neural level.

Another important issue is whether the type of information derived from the constructed representation could make a difference, both on a behavioral and on a neural level. Do we make different computations when assessing a route distance than when inferring a survey distance? And do participants who have learned a configuration from a route description employ the same neural network as those who have learned a configuration from a survey description? In Chapter 7 the retrieval and usage of route distance information against that of survey distance information has been contrasted in two groups with different original learning perspectives.

Grounded cognition

The perceptual symbols account states that language is understood by simulating the content of a sentence and linking it to perceptual representations (Barsalou et al., 2003; Pecher et al., 2003; Stanfield & Zwaan, 2001; Wu & Barsalou, 2009; Zwaan & Madden, 2005). As such, grounding cognition is modal. Spatial and temporal information claim a special position in such a modal system and these issues have not yet been fully accounted for. Clearly, spatial and temporal processing is not limited to a single modality. Chapter 2 discusses the possibility of a multimodal spatial system. It could be that each modality has its own spatial processing system. However, the fact that the same spatial information can be obtained from multiple different modalities suggests a level of convergence, for example a supramodal representation. Currently, there is no univocal answer to what extent spatial information is modality-specific or modality-general. Based on speculation such a modality-general supramodal representation could be seen as an additional ‘modality’, in this case the spatial modality, with perceptual-like qualities. If so, then a logical question is whether this spatial information is compelling and can also be accessed when understanding language not directly pertaining to space. The current section describes two experiments that test whether language understanding also yields spatial simulations.

A prominent role in the perception of space is played by the body. The abilities of our body, including morphology, motor systems and sensory systems influence how we act in situations and possibly how we understand language (Glenberg, 2010). Chapter 8 describes a study that examines to what extent spatial information is accessed when understanding body part names.

If conceptual knowledge indeed is grounded in the perceptual modality by which it has been obtained, one would expect that the quality of the perceptual system has a direct effect on the knowledge representations of an individual. For instance, the sensory/perceptual loss that elderly persons suffer could lead to subtle changes in their semantic system. Even stronger differences would be expected for people blind from birth. For example, their concept of tomato would be dominated by how a tomato feels or tastes rather than by the hue of its color. Because perceptual symbols are associated with experiential traces one could argue that direct experience modulates the strength of simulations. Vision is the prime modality for spatial information processing. Hence, if spatial representations are thought to be modality-specific then blind individuals might have weaker

spatial organization of their semantic knowledge. Alternatively, no clear differences might exist as other sensory inputs provide ample spatial information. The fact that blind people are able to successfully perform a wide variety of spatial tasks suggests that vision is not a prerequisite and other modalities can provide adequate spatial information sources, eventually leading to a supramodal representation (Struiksma et al., 2009). Chapter 9 examines whether spatial information in conceptual knowledge differs between blind and sighted individuals and to what degree the supramodal representation of spatial information also holds for blind individuals.

Chapter 2

What is the link between language and spatial images? Behavioral and neural findings in blind and sighted individuals

M. E. Struiksma, M. L. Noordzij & A. Postma

Acta Psychologica, 132 (2009), 145–156

Abstract

In order to find objects or places in the world, multiple sources of information, such as visual input, auditory input and asking for directions, can help you. These different sources of information can be converged into a spatial image, which represents configurational characteristics of the world. This paper discusses the findings on the nature of spatial images and the role of spatial language in generating these spatial images in both blind and sighted individuals. Congenitally blind individuals have never experienced visual input, yet are able to perform several tasks traditionally associated with spatial imagery, such as mental scanning, mental pathway completions and mental clock time comparison, though perhaps not always in a similar manner as sighted. Therefore, they offer invaluable insights in the exact nature of spatial images. We will argue that spatial imagery exceeds the input from different input modalities to form an abstract mental representation while maintaining connections with the input modalities. This suggests that the nature of spatial images is supramodal, which can explain functional equivalent results from verbal and perceptual input for spatial situations and subtle to moderate behavioral differences between the blind and sighted.

Introduction

Suppose you want to buy a pineapple. You need to go to the supermarket and you have to find the pineapples. In order to solve this task you may activate several information sources from memory: a visual image containing the shape and form of a pineapple; a sense of the tactile information of the texture one feels when grasping the fruit; olfactory information about its smell; a quasi motor program containing movement and directional information on how to walk towards the fruit stall. Yet, it is possible that you still do not have enough information to find the target location. In that case, an option is to ask for directions. Language can be used to communicate an abstract set of instructions allowing one to find the pineapples. Together all these information sources may be converged into a spatial image: a mental representation of the physical world which includes relations between objects and extensive information on location, orientation and distance (De Vega et al., 2001).

There are two main types of information sources for spatial images. One involves imagery processes which are directly coupled to modality-specific sensory information, for example: visual, olfactory, gustatory and tactile representations. Motor imagery can be thought to be coupled in a top-down fashion to proprioceptive representations and efferent copies of the motor program. For these forms of imagery there is a clear overlap in neural circuitry between perception and imagery (Kosslyn et al., 2001). Secondly, language can also function as a source of information. Whether or not language is modal has been an issue of debate in recent literature (e.g. Barsalou, 1999; Barsalou, 2008; Barsalou et al., 2008; Chomsky, 2000; Jackendoff, 2002; Zwaan & Radvansky, 1998). According to Barsalou (1999) words can only have meaning if they are grounded, or coupled to perceptual concepts. During perception relevant information is extracted as perceptual symbols. This information is gathered from any modality, resulting in multimodal perceptual symbols. These symbols can be combined into a perceptual concept, which can be reenacted through the process of simulation. According to this theory words are linguistic symbols which become associated to simulators for perceptual concepts. On the other hand, we can also consider language as a communication system which provides a set of processing instructions to generate a mental representation of a situation (Barsalou, 1999; Zwaan & Radvansky, 1998) and which does not incorporate meaning directly (Fauconnier, 1999). Such a set of instructions, or propositions, can be considered abstract and amodal, since it does not involve a link to modal information specifically. There is, however, a strong relationship between words and modality-specific information, which is referred to by Barsalou and colleagues (2008) as the linguistic form. The linguistic form could be visual, tactile or auditory and can subsequently be used as pointers to conceptual information (Barsalou et al., 2008). These two theories are important, but the focus of this review is not on discussing how language is grounded. Therefore, we refer to language as words, which are in essence amodal, but can also be related to perceptual modal experience.

The different sources of information that can lead to a spatial image are organized along several levels. The lowest level corresponds to the input, which can either be perceptual bottom-up or top-down retrieved from long-term memory (Cornoldi & De Beni, 2007; Cornoldi et al., 1998), while

the higher levels denote the resulting mental representations. There is a distinction between on the one hand modality based input channels, yielding images coupled to sensory information, and on the other hand language as an amodal input channel, yielding a set of propositions. Together these different sources from both input levels can form a spatial image (Chatterjee, 2001).

The nature of spatial images has been the subject of many studies (e.g. Barsalou, 1999; Carpenter et al., 1999; Lambert et al., 2004; Noordzij et al., 2008; Noordzij et al., 2006a; Reichle et al., 2000). A central question has been to what extent spatial images can be considered amodal, multimodal or supramodal, however, no consensus has yet been reached. Multimodal representations are linked to the perceived modality and are established in the modality-specific brain areas. A supramodal representation exceeds the input from different modalities but maintains modality-specific information (Barsalou, 1999). When discussing neural correlates it is important to consider evidence from multiple sources of information which should yield converging evidence. Behavioral experiments can determine whether direct physical input is analogue to mental representations and can generate similar behavioral results. In addition, neuroimaging research, such as electroencephalogram (EEG) and functional magnetic resonance imaging (fMRI) can offer insight into the neural correlates of the overlap between perception and imagery. This neuroimaging evidence should be treated with some caution since it establishes correlations between task performance and neural activation. Therefore, the use of transcranial magnetic stimulation (TMS) and data from lesion studies can provide converging evidence bearing upon the causal relation between an anatomical region and a specific function. Research with blind people can be linked to the latter two techniques and it provides particularly relevant evidence to test the modality specificity of spatial images. Although congenitally blind people have never processed visual information, they are able to generate spatial images (e.g. Noordzij et al., 2006b, 2007; Vecchi, 2001) on the basis of motor, tactile, auditory, olfactory and linguistic information (for reviews see: Cattaneo et al., 2008; Kaski, 2002; Thinus-Blanc & Gaunet, 1997). This paper reviews studies investigating mental imagery, in particular the nature of spatial images, and the role of spatial language in generating these images in the blind and sighted.

Imagery processes

For visual, haptic, olfactory, gustatory and auditory imagery the corresponding mental images are the contents of working memory, which have a quasi-perceptual and isomorphic nature and are derived either from sensory inputs or from memory traces of these inputs. Spatial imagery is different from sensory imagery in that it is not bound to a single input modality. In order to establish the format of spatial mental imagery, the role of the input from different modalities needs to be determined. The underlying processes that are involved in sensory imagery and spatial imagery could be different. By investigating the neural correlates of different types of imagery processes the underlying mechanisms can be compared.

Most research on imagery has concentrated on visual imagery (see Kaski (2002), for a review).

The general consensus seems to be that visual imagery and visual perception are two distinct processes that partially share a neural network. In an event-related potential (ERP) study Farah and colleagues (1988b) have shown that visual perception and imagery of letters interact at an early stage over the visual cortex, suggesting that the mental images are represented visually. In a positron emission tomography (PET) and repetitive TMS (rTMS) study Kosslyn and colleagues (1999) have demonstrated that mental imagery of patterns of stripes recruits area 17 similar to visual perception (see Kosslyn et al. 2001, for a review). However, imagery does not require sensory processing. Visual imagery and perception are closely related and share similar mechanisms, but with unique characteristics. The idea of partially overlapping networks for visual imagery and perception has been further explored by Ganis and colleagues (2004). In an fMRI study where subjects visualized or saw faint drawings of objects, they also found a partially overlapping network. The overlap between imagery and perception was found predominantly in the frontal and parietal regions and less in the temporal and occipital regions. For an overview of a selection of literature on visual imagery see Table 1.

While less work has been done on tactile, auditory, olfactory, gustatory and motor imagery, partially overlapping networks for perception and imagery have also been reported here (for example Goyal et al., 2006; Jeannerod, 1995; Kobayashi et al., 2004; Kosslyn et al., 2001; Stevenson & Case, 2005). Yoo and colleagues (2003) observed in an fMRI study that the primary and secondary somatosensory areas are active during tactile imagery of gentle brushing of the hand and that there is partial overlap with the activity found during actual tactile stimulation. However, there are also parts of the primary and secondary somatosensory areas that are only active during actual tactile stimulation. Moreover, tactile imagery activates the inferior parietal lobe, which seems to share commonalities across imagery modalities (Uhl et al., 1994; Yoo et al., 2003). Yoo and colleagues (2001) also investigated auditory imagery and its neural correlates. Again they found that imagery and perception have a partially overlapping network. Auditory imagery of a simple monotone can elicit bilateral activation in the primary and secondary auditory areas (Yoo et al., 2001). Although olfactory imagery is difficult to assess, it has been shown to activate similar brain areas compared to actual odor sensation, but once more to a lesser extent (Stevenson & Case, 2005). Stevenson & Case (2005) report converging evidence from behavioral findings, olfactory illusions in patients with Parkinson's and Alzheimer's disease, as well as EEG and fMRI findings.

In conclusion, modal perception and mental imagery seem highly alike. The overlap between perception and imagery agrees with the perceptual symbol systems theory proposed by Barsalou (1999). However, modality-specific inputs are not the only manner through which spatial representations can be achieved. Language provides an alternative elegant way to build a spatial image.

Spatial language

An intriguing question is whether a more indirect input, such as spatial language, goes beyond a mere propositional format and can provide information about the environment needed for

navigation and other spatial actions such as distance comparison. There are two ways in which a spatial sentence can be interpreted in working memory. Either, it is processed by a verbal strategy, or by a pictorial strategy (MacLeod et al., 1978). When spatial language is processed by means of a verbal strategy the propositional information of the sentence is analyzed and maintained in working memory. However, when a pictorial strategy is used the spatial description is parsed and then transformed into a spatial image (Reichle et al., 2000). This dichotomy is similar to the Dual Coding theory proposed by Paivio (1971) which states that information can be represented in verbal and image codes. Noordzij and colleagues (2005) have shown in a sentence-picture and sentence-sentence verification task that subjects use a dual-representational model to process simple spatial sentences. Subjects automatically generate a verbal representation and depending on the expected modality of the second stimulus they generate a pictorial representation as well.

The finding by Noordzij and colleagues (2005) stresses an important issue. In their task the subjects only generated a spatial image when a picture was expected. Perhaps the generation of a spatial image depends on the task and the instructions. In an experiment with more complex descriptions of an environment subjects performed a priming task (Noordzij & Postma, 2005). There were three priming conditions: close in text/close in space, far in text/close in space, and far in text/far in space. The results showed that a spatial priming effect occurred only for metric distances and not for text distances. Subjects responded faster to targets close in space to a prime compared to targets far in space to a prime. This effect was not found for textual distance. This finding indicates that when the memory load increases people adopt a pictorial strategy. The level of automatic generation of a spatial image from spatial language can also be influenced by the specific task instructions. Avraamides and colleagues (2004) demonstrated that in remembering the orientation and distances of multiple targets spontaneous processing results in less variability in the responses for visual perception than for spatial language. However, after forced spatial updating, by letting the subject walk backwards after learning, the results for visual perception and spatial language were similar. These results confirm that visual and verbal inputs can generate spatial images that support functionally equivalent behavior.

Spatial imagery

The general pattern of overlap between sensory processing and corresponding mental imagery cannot easily be transferred to spatial imagery since there is no such thing as direct spatial sensory perception. Instead spatial images are actively constructed from different sources of information (Reisberg, 1997). Moreover, evidence from lesion studies has shown that there is a distinction between visual imagery and spatial imagery. Farah and colleagues (1988a) describe a patient with damage to the temporal-occipital regions, right temporal lobe and right inferior frontal lobe, who is impaired on visual imagery tasks, but performs normal on spatial imagery tasks. Kosslyn and colleagues (1978) performed a series of experiments in which subjects had to scan mental images. In one of their experiments subjects had to study a map with locations and were later asked to men-

tally scan from one location to the other. The results show that subjects took longer to scan longer distances. This indicates that the metric properties of the map are preserved during imagery, and that functional proximity on the image is similar to the physical proximity on the map (Reisberg, 1997).

Notably, the analogy between the physical world and the spatial image is also found when the image is derived from a verbal description (De Vega et al., 2001; Denis et al., 1999; Loomis et al., 2007; Noordzij & Postma, 2005) and the relevant spatial information can be retrieved from both types of input. For example, Denis & Zimmer (1992) reported on a series of experiments to test whether verbal descriptions could be adequately converted into a mental representation. They conducted a spatial priming, distance comparison and mental scanning task. Their findings suggest that verbal descriptions can be used to generate a mental representation similar to visual experience. Additionally, both mental representations also show reliable metric properties. This finding was replicated by Denis (2008) in a distance comparison task based on verbal descriptions. In addition to this task subjects were tested on the Minnesota Paper Form Board, which measures visuo-spatial imagery capabilities. Denis found that high visuo-spatial imagers outperformed low visuo-spatial imagers in the difficult trials. Although the spatial images based on spatial language and perception seem to be highly alike on a behavioral level, the question remains whether they also share a similar neural representation. The functional equivalence of spatial images from verbal and perceptual input could be due to similar, but distinct processes that combine modal representations, or because both inputs feed into an abstract amodal representation.

Modal information can be combined in either a multimodal or a supramodal representation. As stated in the introduction multimodal representations are established in modality-specific brain areas (Barsalou, 1999). In his theory on perceptual symbol systems Barsalou (1999) argues that spatial concepts are also grounded and are simulators of schematic perceptual symbol information and, therefore, multimodal. If spatial images are indeed multimodal then only modality-specific brain areas will be activated when evaluating the image, for example through simulation. There will be no specific brain areas which represent a spatial image, but are not linked to a sensory input modality. An example of a multimodal representation is given in panel A of Figure 1. The mental images of different input sources are combined into a multimodal spatial image. Since the modality-specific input is maintained, this model can explain subtle to moderate behavioral differences on spatial tasks with different input modalities. For example, there was functional equivalence from visual perception and verbal descriptions of a spatial layout. However, without forced spatial updating subjects performed slightly better on distance judgments after visual perception compared to verbal descriptions of a spatial layout (Avraamides et al., 2004). On the other hand, supramodal representations exceed the input from different modalities but maintain modality-specific information (Barsalou, 1999). Therefore, if the imagery mechanism involves a supramodal process we can expect both a supramodal representation in a distinct area, and an overlap with the associated modalities.

For example, when spatial images are generated from spatial language we could expect language

areas to be active as well as a supramodal area. When spatial imagery is a supramodal process the input modality will be reflected in the activation pattern. If the input for spatial images is visual, then a trace of the input should remain when the spatial image is activated. This is portrayed in panel B of Figure 1. The bi-directional arrows between the spatial image and the input represent the trace of overlap with the associated modality. This model could also account for the subtle to moderate behavioral differences as mentioned above, since a link with the input modality is preserved. However, if spatial imagery is a more abstract process and considered amodal, then there should be no reference to the original input modality. Amodal processes are independent from input modality and are not necessarily constructed from multiple input sources. Regardless of the input modality that contributed to the spatial image, the same pattern of activation in a specific area is to be expected. This model has some similarities to the model described by Bryant (1997). He describes a spatial representation system (SRS) framework in which the SRS receives input from different modalities, as well as linguistic information. The relevant spatial information is extracted and combined with general knowledge into a spatial mental model. This final step, the spatial mental model, could be our amodal representation, although we do not specify a SRS as Bryant does. Panel C of Figure 1 demonstrates such an abstract representation. The arrows from the input levels to the spatial image only go in one direction. Once the spatial image is formed, the modality-specific input can no longer be reactivated. In order to account for the behavioral differences between verbal descriptions and visual perception we need to make the assumption that both inputs generate a highly similar spatial image, but the quality differs slightly. As a result behavior is equivalent on a functional level, but might differ at the level of availability of information.

Moreover, different input channels can have different relative contributions to the spatial image, as tentatively displayed in Figure 1. The emphasis on visual input stresses the strong dependence on visual information, which might account for the initially slightly better performance after visual perception in the study by Avraamides and colleagues (2004). Language and motor input are the next most important sources of information that contribute to spatial images (Cattaneo et al., 2008; Knauff & May, 2006). The presence of bi-directional arrows also enables switching between modalities during imagery. For example, in a tactile task sighted subjects might also retrieve visual information, especially if a delay is induced in the task (Zuidhoek et al., 2003). The combination of modalities can strengthen the mental image. Whether this process works with direct input or through mental imagery will not be discussed. Transfer between modalities was also found by Cattaneo & Vecchi (2008). In a 2D mental pathway task subjects had to explore a matrix visually or haptically and were tested on a modality congruent or incongruent matrix. The results show that there were modality-specific effects: a general advantage for visual exploration and similar results for matrices explored in the same modality (irrespective of the testing modality). However, the fact that there was no significant difference between congruent or unimodal and incongruent or cross-modal conditions indicates that there was also a supramodal representation. Further research is needed in order to disentangle the relative contributions of different sources of input and the associated spatial image.

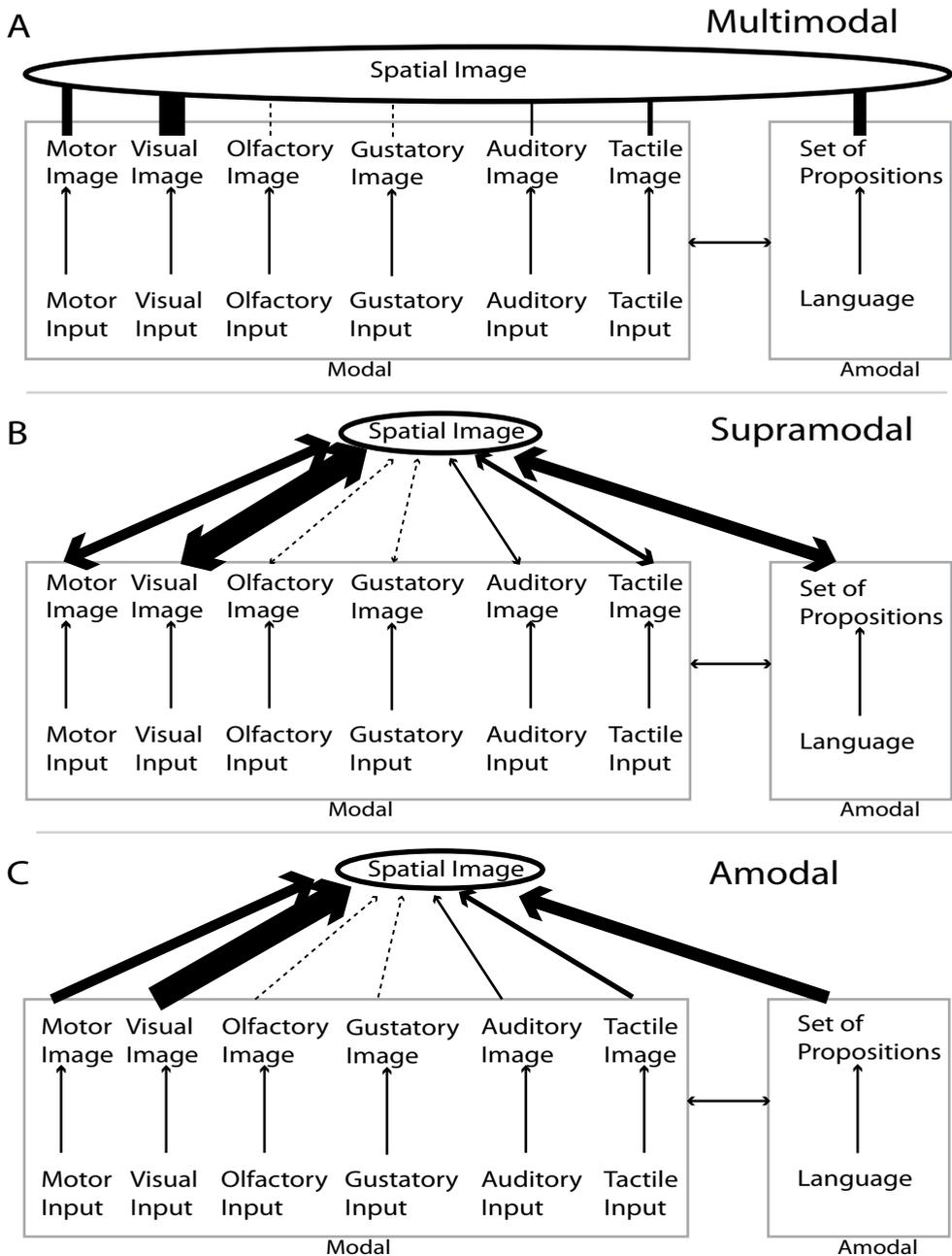


Figure 1 Three different models of how different sources of information can contribute to the generation of a spatial image. The line width represents a schematic weighting of the contribution of the different sources. Panel A describes a multimodal representation established in modality-specific brain areas. Together these form the multimodal representation. Panel B describes a supramodal representation, which exceeds modality-specific input to generate a spatial image, but maintains a bi-directional link with modality-specific input. Panel C is an amodal representation in which a spatial image is extracted from the input and no backward connections remain.

Neural correlates of spatial imagery

The first systematic studies of the neural correlates of visuo-spatial mental imagery go back three decades. The research on visuo-spatial mental imagery started out with determining the different components of imagery. Farah (1989) and Kosslyn and colleagues (1990) revealed several independent components of visuo-spatial mental imagery. These components are image generation, image maintenance, image scanning and image rotation. In an ERP study Farah (1989) found that image generation and image rotation are lateralized. Image generation depends on activation in the left posterior areas while mental rotation depends on activation in the right posterior areas. The parietal lobe has also been associated with spatial updating (see Creem & Proffitt, 2001, for a review) and coding different elements of space, see Table 1 for an overview of relevant papers. For example, Galati and colleagues (2000) reported activation of the right superior parietal lobule (SPL) and right intraparietal sulcus (IPS), common to egocentric and allocentric judgments of horizontal lines in an fMRI task. The IPS was also reported by Mellet and colleagues (2000). In their fMRI experiment this area was associated with spatial mental imagery components in a mental navigation and mental map task. The activation of bilateral IPS was part of a larger fronto-parieto network of activation, similar to the network found for the egocentric judgments in the study by Galati and colleagues (2000). The IPS has also been associated with cross-modal processing of visual and tactile representations. Grefkes and colleagues (2002) conducted an fMRI study with a delayed match-to-sample task and found that the left IPS was activated stronger when object information was transferred between the visual and haptic modality. The posterior parietal cortex, including bilateral IPS and precuneus, was also activated when spatial information on tactile and visual stimuli had to be maintained (Ricciardi et al., 2006). The authors argue that the similar patterns of activation between visual and haptic stimuli was generated by a spatial mental representation. In yet another task Saito and colleagues (2003) found bilateral IPS activation during the integration of shape information from haptic and visual stimuli. The findings in the studies discussed above suggest that the IPS seems involved in a number of spatial tasks for both visual and haptic stimuli.

Another aspect of coding space is the type of spatial relation. There are two types of spatial relations: categorical and coordinate (Kosslyn, 1987; Kosslyn et al., 1989). Categorical spatial relations encode discrete relations while coordinate encode precise metric relations. Trojano and colleagues (2002) tested the mental clock task, based on Paivio (1978), in an fMRI paradigm. They found that the categorical task activated the left SPL and left angular gyrus (AG), while the coordinate task activated the right SPL and right AG. The activation of the AG was replicated in a dot-bar task, which also involved categorical and coordinate judgments (Baciu et al., 1999). As mentioned before, neuroimaging techniques such as EEG and fMRI establish correlations between neuroanatomy and function. The lateralization for categorical and coordinate spatial processing was confirmed by rTMS (Trojano et al., 2006) as well as evidence from stroke patients. Laeng (1994) tested 30 patients with left hemisphere damage and 30 patients with right hemisphere damage on a perceptual categorical/coordinate task. The findings from these patients also strengthen the evidence of

lateralization of categorical and coordinate processing (see Jager & Postma (2003), for a review on a wide variety of research on hemispheric lateralization).

Categorical spatial relations are closely related to spatial language, which is also demonstrated by the left hemispheric involvement. These categorical spatial relations are often used in spatial imagery instructions and spatial language in general. Damasio and colleagues (2001) have shown, in a PET study, that in sighted subjects areas in the left parietal lobe are involved in processing spatial relations. In particular, the left supramarginal gyrus (SMG) was activated while naming spatial relations between objects. In an fMRI study Carpenter and colleagues (1999) asked subjects to read spatial sentences which had to be compared to a picture. This task activated the left posterior temporal gyrus, which is considered a language comprehension area. Furthermore, bilateral parietal areas were associated with spatial processing. These areas included the SPL, the posterior SMG and the AG. The left SMG and left AG are both involved in processing spatial relations and could be the interfaces for converting spatial language into spatial representations. The activity found in the left AG and SMG has been associated with processing spatial relations and locative prepositions in particular (for reviews see Chatterjee, 2008; Kemmerer, 2006).

Tranel and Kemmerer (2004) conducted a lesion study and found that the highest region of lesion overlap in patients, who were impaired on processing locative prepositions, was found in the left frontal operculum and left SMG. Subsequently, it has been shown with fMRI that the left SMG is activated when people process a spatial sentence in a verbal or a visuo-spatial context (Noordzij et al., 2008). In this experiment subjects had to verify a spatial sentence to either another sentence (verbal context), or a picture (visuo-spatial context). Noordzij and colleagues (2008) concluded that the function of the left SMG is to generate an amodal representation of locative prepositions which can be flexibly compared to verbal or visuo-spatial information.

Based on our current definitions of amodal, multimodal and supramodal representations we wish to conclude that the left SMG might be a neural correlate of a supramodal representation instead of an amodal representation. It seems unlikely that this representation is multimodal, since it is found in the same area in both verbal and visuo-spatial context. However, it could be that the left SMG supports a supramodal representation while other areas support the modality-specific activation. The current neuroimaging evidence does not provide enough information to dissociate between amodal and supramodal spatial imagery.

Blindness and spatial imagery

By investigating the process of spatial imagery in individuals who have been blind since they were born it is possible to disentangle whether or not spatial imagery is bound to a modality, with a particular emphasis upon the visual modality. Studying the results from experiments with blind and sighted can provide evidence for the modality specificity of spatial images. Suppose that spatial imagery is a multimodal representation of spatial configurations, then the visual representations will be absent in the blind. This will provide an opportunity to evaluate the development of the

system in the absence of input from the visual channel. However, other modal and amodal representations will still be present. The absence of visual spatial representations in blind might lead to behavioral differences with sighted. If we assume that spatial imagery is a supramodal process it will use both forward and backward connections with the modal areas. In congenitally blind subjects the modal connection to the visual cortex for visual mental images will be absent. As a result the distribution of the input from different input channels in Figure 1 might change, which could result in behavioral differences between the blind and sighted. However, if spatial imagery is amodal, then one would not expect any differences in performance between blind and sighted since the spatial representations need not necessarily be constructed from the same input. The results from behavioral and neuroimaging studies can indicate whether or not spatial imagery is functionally and neuro-anatomically equivalent in blind and sighted subjects.

The distribution of the forward and backward connections in supramodal processing might change in the absence of vision. There might be little or no compensation as evidenced by equal weight-distributions to all input channels, irrespective of the presence of visual input (see Figure 2 panel A for a schematic distribution in sighted subjects and panel B for a comparison with blind subjects). However, another plausible situation might be complete compensation for the visual input channel by higher weight-distributions to the other channels (see Figure 2 panel C). The second option seems plausible considering the plasticity of the visual cortex of blind people. People who are born blind lack any visual stimulation of the visual cortex. Instead of losing the functionality of the occipital cortex, reorganization occurs (e.g. Noppeney, 2007; Noppeney et al., 2005). The visual cortex has become associated with non-visual functions such as tactile processing, Braille reading and other language functions (e.g. Amedi et al., 2004; Burton & McLaren, 2006; Cohen et al., 1997; Ofan & Zohary, 2007; Pascual-Leone et al., 2005). Moreover, the proportion of brain area involved in for example tactile processing increases. Therefore, the relative weights of the non-visual sources of information that can contribute to spatial images can change to compensate for the absence of vision.

Lambert and colleagues (2004) investigated the functionality of the visual cortex in blind and sighted with fMRI. Their subjects were instructed to form mental images from animal names and to listen to abstract words. They suggested that during mental imagery apart from the regular circuit (superior occipital, inferior and superior parietal areas, premotor area and visual association area) blind individuals also activate somatosensory areas and the temporal and fusiform gyrus. When generating mental images from verbal instructions in blind subjects the dorsal pathway is activated, which now relies on haptic sensitivity instead of visual sensitivity. They conclude that the occipital cortex is active in blind individuals for haptic information and in sighted individuals for visual information. Research on plasticity in the blind has not only focused on tactile tasks, but also on auditory tasks (Amedi et al., 2003; Burton et al., 2002b; Noppeney et al., 2003; Röder et al., 2002; Sadato et al., 2002). The results all indicate occipital cortex activity in early blind subjects. Interestingly, in all these experiments the task involved language processing, but in different modalities. This led Burton and colleagues (2002b) to believe that the occipital cortex in blind individuals is involved

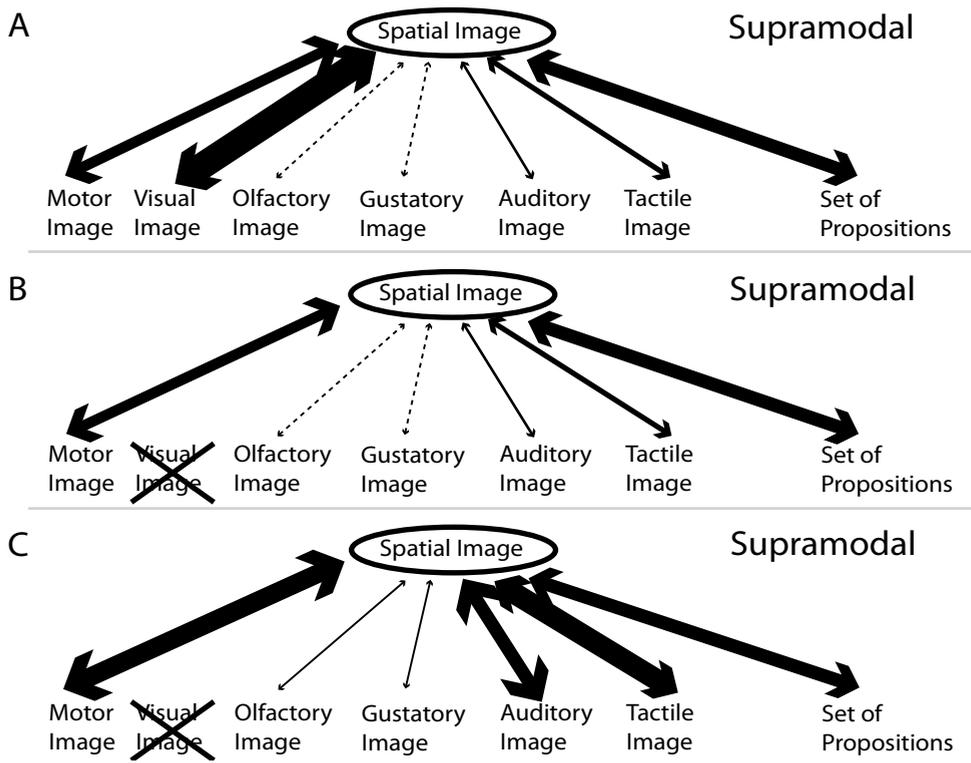


Figure 2 Three schematic models of the weighting of different sources of information, which is represented by the line width. Panel A describes a supramodal representation with the weighting in sighted subjects. Panel B represents a supramodal representation, with a possible scenario for no visual input and no compensation in blind subjects. Panel C is a supramodal representation, with a possible scenario for no visual input and compensation from other sources of information in blind subjects.

in dealing with linguistic aspects. Indeed, their results support the notion that linguistic processes activate the occipital cortex in blind individuals. They postulate that blindness results in new mechanisms in especially the left occipital cortex. These mechanisms reach across multiple visual areas, the degree to which depends on the age of onset of the blindness. Burton and colleagues (2002b) only found evidence of reorganization in blind individuals in the occipital cortex. During a verb generation task both early and late blind and the sighted control group showed similar activity in the language areas: the left inferior frontal gyrus, left dorsolateral prefrontal gyrus and left posterior superior temporal gyrus. Since general language processing seems to be preserved in the general language areas, with additional striate activation in the blind (for a review see Noppeney, 2007), it would be interesting to try and tease apart the involvement of the visual cortex in spatial imagery and its link to spatial language.

Although blind people lack visual information about their environment they are able to navigate. Apparently, the information they obtain from their other senses provides enough knowledge about the environment to find their way. Evidently, and as argued above, space can be represented

by more than just visual information. There seems to be a common compound that is extracted from different sources of information. A reasonable candidate for such a common compound could be a spatial image, as shown in Figure 1, panel B and C.

Thinus-Blanc & Gaunet (1997) give an elegant review of behavioral differences between blind and sighted subjects. They discuss literature on a variety of experiments measuring spatial ability. According to the authors, the inconsistencies in the experiments performed at that time made it hard to draw general conclusions. On the one hand, the trend in most experiments, discussed in their review, seemed that spatial memory is not affected by the age of onset of blindness (e.g. Loomis et al., 1993; Rieser et al., 1986). On the other hand, inferential abilities to generate spatial relationships from information, that is not actually experienced at that moment, did appear to be affected by the age of onset (e.g. Byrne & Salter, 1983; Dodds et al., 1982; Rieser et al., 1986). The latter abilities are essential when you need to make a detour or a shortcut.

Thinus-Blanc & Gaunet (1997) describe a two-level model of spatial processing. In this model they make a distinction between schemata, which are abstract representations from specific maps, and simple rules of encoding, which can be deduced from specific spatial memories. The schemata are independent from the position of a person and are encoded with an exocentered reference frame. Simple encoding rules and specific spatial memories are encoded with an egocentered reference frame. These two different levels of encoding can yield a similar representation and give the same behavioral results on how spatial information is acquired and how this spatial knowledge is used. A logical assumption is that blind people rely more, or perhaps only, on the egocentered organized spatial information. All their spatial information is experienced from a body centered perspective. This assumption is supported by Noordzij and colleagues (2006b), who found that blind subjects showed a preference for route descriptions (with a body-centered perspective), while sighted subjects preferred a survey description (with an environment-centered description). In a task where subjects had to haptically match shapes to the cut-outs in a board and subsequently give a verbal description of the location, blind subjects referred more to other objects on the board, while the blindfolded sighted referred more to the surrounding frame (Postma et al., 2007). This finding is in line with the route and survey preference respectively in the study from Noordzij and colleagues (2006b).

Research on spatial processing abilities of the blind has revealed that blind individuals are able to generate spatial mental maps (e.g. Aleman et al., 2001; Kerr, 1983; Klatzky & Golledge, 1995). They can use descriptions to generate a spatial image in which spatial properties such as distance and location are preserved (see Table 1 for a selection of literature relevant to this discussion). Kerr (1983) showed that metric properties of objects and locations were preserved in blind and sighted, although the blind were slightly slower than the sighted. Blind subjects are also able to generate a mental representation of a clock and compare clock times (Noordzij et al., 2007). All subjects, early blind, late blind and sighted, showed the classical symbolic distance effect (Moyer & Bayer, 1976). This suggests that each group used spatial imagery to solve the task. Nevertheless, the early blind made more errors than the sighted and the late blind seemed to perform slightly better, which

suggests a small advantage for visual experience.

In a recent review Cattaneo and colleagues (2008) discuss papers on the growing evidence that blind can rely on sensory input, such as auditory or haptic, and verbal input to overcome their visual deprivation. Vanlierde and colleagues provide supporting evidence that blind subjects can rely on non-visual sensory input (Vanlierde et al., 2003; Vanlierde & Wanet-Defalque, 2004). They describe similar performance for early blind, late blind and sighted on a pattern symmetry task about mental representations of verbal 2D patterns in a grid. While the accuracy was the same for blind and sighted, both groups used a different strategy. The sighted adopted a visuo-spatial strategy, while the blind adopted a “coordinate XY” strategy. Performance on several tasks appeared highly similar for blind and sighted, however, subtle to moderate differences were also found. For example Knauff & May (2006) showed that blind subjects are not susceptible to the “visual-impedance-effect” in deductive reasoning. Subjects had to make a decision about a situation and when the visual relations in the situation were easily imagined the sighted subjects were hindered. Although blind subjects were not hindered in this particular case, compared to the sighted, they were overall slower and less accurate. The seminal work by De Beni & Cornoldi (1988) has shown that similarities between blind and sighted can also rely on task properties. They conducted a memorization experiment where single nouns, pairs or triplets had to be remembered in connection with a cue. They found that blind subjects were impaired when the memory load increased (with noun pairs and triplets), but when the memory load was small the blind were able to create interactive images. A similar pattern was found by Vecchi and colleagues (2004). Blind subjects and sighted subjects performed on a par when a series of locations had to be remembered on a single matrix. However, when the locations from two matrices had to be remembered, blind subjects performed worse compared to when the locations of the two matrices had to be integrated. This pattern was reversed for the sighted subjects. During visual perception it is common to handle multiple inputs, for haptic perception this is quite impossible.

These results indicate that when the memory load increased by the sequential memorization of two matrices the performance of the blind subjects decreased. Cattaneo and colleagues (2008) suggested that although performance of blind and blindfolded sighted subjects can be highly similar this could be due to the employment of different strategies, non-visual versus visual respectively, as was also suggested by Thinus-Blanc & Gaunet (1997) and Vanlierde & Wanet-Defalque (2004). Cornoldi and colleagues (2009) systematically examined strategy effects in blind and sighted on a 2D mental pathway task. Three different strategies were found in both blind and sighted. Their strategies could be spatial, verbal or mixed. The results showed that when the entire pathway had to be remembered both groups performed the same when using a verbal strategy. However, when using a spatial strategy the blind performed worse than the sighted. This finding is in line with their previous work (De Beni & Cornoldi, 1988) in that this task with a high memory load is more difficult for the blind when using a spatial strategy.

Nevertheless, even in the absence of vision, congenitally blind people are able to generate a proper spatial image and use this information for a wide variety of spatial tasks and navigation.

Given the subtle to moderate behavioral differences between blind and sighted, spatial imagery cannot be amodal, which would have resulted in identical behavioral performance for blind and sighted. Moreover, an amodal representation would not be sensitive to time manipulations. In a haptic parallel-setting task early blind and blindfolded sighted performed similar in the direct setting condition (Postma et al., 2008). When a delay between feeling the reference bar and rotating the test bar was induced performance increased for late blind and blindfolded sighted participants, but did not change for early blind participants. According to the authors a delay induces visual mental imagery, which increases accuracy. The early blind are not able to use visual mental imagery, therefore, their performance remains the same as in the direct condition. Given an amodal representation it would not be possible to retrieve a visual mental representation to improve performance. Alternatively, spatial imagery could be multimodal or supramodal. In his perceptual symbol systems theory Barsalou (1999) initially referred to simulators of spatial relations as multimodal symbols. Moreover, through the process of conceptual combination simulations might easily lead to simulations of related concepts (Wu & Barsalou, 2009). This might also work during mental imagery and explain the possibility to take advantage from visual mental imagery in addition to spatial mental imagery in the haptic parallel-setting task, as mentioned above. According to our model (see panel A of Figure 1) neuroimaging results should identify brain areas which are associated with the input modality. However, this would mean that in the sentence-sentence and sentence-picture verification task by Noordzij and colleagues (2008) the context manipulation should have revealed activation in different areas for each context, verbal and visual. Rather, they found activation in the left SMG that supported both contexts. This would suggest a supramodal representation with the left SMG as a possible neural correlate.

In a response to the perceptual symbol systems theory Freksa and colleagues (1999) suggested that space can have a two-fold role. It can either be a single perceptual symbol representing a specific location. This role would suggest a multimodal perceptual symbol similar to regular perceptual symbols of, for example, objects. Alternatively, space can organize perceptual symbols in providing structural constraints. Such spatial symbols combine a shared structure for different perceptual modalities. The latter role of spatial symbols supports our supramodal representation of spatial images, as shown in panel B of Figure 1. Different perceptual inputs can feed into a shared spatial image, which represents the spatial organization in this specific example. In a reaction to the response by Freksa and colleagues (1999), Barsalou (1999) reasons that spatial symbols may depend on modality-specific, but also modality-general systems. The findings discussed here, for example the functionally equivalent spatial images from verbal descriptions and perceptual input, suggest that these modality-general systems might indeed exist. The possible employment of different strategies for blind and sighted (Cattaneo et al., 2008; Cornoldi et al., 2009; Vanlierde & Wanet-Defalque, 2004) could be a result of the recruitment of different neural networks. However, using completely different neural networks would probably yield large behavioral differences between the blind and sighted. Given the only subtle to moderate behavioral differences between blind and sighted, a certain degree of overlap in the underlying neural networks seems more plausible and

would support the idea that spatial representations are supramodal. Further neuroimaging evidence on research with blind is needed to answer these questions.

Conclusions

Spatial images allow a mental representation preserving configurational layout of the outside world. Importantly, spatial images can be derived from different sources of information: e.g. visual inputs, somatosensory inputs, and spatial language. The main question addressed in this review was to what extent spatial images are intrinsically linked to modality-specific sources or whether they appear at a superordinate level.

For sensory imagery, such as visual imagery, there is a high degree of overlap in the neural networks between perception and imagery. For spatial imagery several sources of information can be used for construction. Perceptual and verbal input can generate functionally equivalent spatial images and the relevant information can be addressed from both types of input. This supports the idea that spatial images are not modality-specific, but seem to appear at a superordinate level. According to the perceptual symbols system from Barsalou (1999) spatial images should be considered multimodal. However, this can not explain the finding by Noordzij and colleagues (2008) that the left SMG is activated regardless of the context (verbal or visuo-spatial). Noordzij and colleagues (2008) have argued that their finding supports an amodal representation, however, according to our definitions this finding could also support a supramodal representation.

Given that spatial images seem to appear at a superordinate level and can produce functionally relevant information, research with blind people has provided insight into the link between input and spatial images. Spatial language has proven to be a very elegant tool to generate spatial images, especially in the blind. Although behavior on numerous spatial tasks has shown a high degree of similarity between blind and sighted subject, there are subtle to moderate differences. This supports a supramodal representation of spatial information, since an amodal representation would result in behaviorally identical results. As such, spatial imagery exceeds different input modalities, while forward and backward connections with the input modalities are preserved and multiple sources of information can be used to construct spatial images. The absent or functionally different connections in the occipital cortex of blind subjects could account for the behavioral differences. Based on our definition of supramodal we would like to suggest that the left SMG is involved in supramodal spatial imagery. The idea of a supramodal organization of the brain is supported by Cattaneo and colleagues (2008) who reported on supramodal organization for tactile perception in two other areas, namely the lateral occipital tactile-visual area and an area in the dorsal visual pathway involved in visual motor perception. Both areas were active not only during visual perception, but also during tactile perception. These findings could be an extension of the perceptual symbol systems proposed by Barsalou (1999). Besides multimodal perceptual representations which reside in modality-specific areas, spatial images can be considered supramodal with neural correlates that respond to different sources of input.

Table 1 A selection of important literature on imagery, blindness and spatial language comprehension.

Author	Modality	Task	Comments	Brain Areas
Yoo et al., 2001	Auditory imagery	fMRI: imagery of a single tone	No perception condition was included.	Medial, inferior and middle frontal, precuneus, superior temporal and anterior cingulate gyrus. Bilateral primary and secondary auditory areas.
Jeannerod, 1995	Motor imagery	Review	Motor imagery and action execution share similar networks.	
Stevenson & Case, 2005	Olfactory imagery	Review	Olfactory imagery exists, but can be hard to measure.	Olfactory imagery can produce similar effects as experiencing an odor.
Carpenter et al., 1999	Spatial imagery	fMRI: spatial sentence-picture verification	There is activation of linguistic and spatial networks during spatial sentence comprehension.	Left posterior temporal gyrus: language comprehension, left and right parietal regions: visuo-spatial processing.
Creem & Proffitt, 2001	Spatial imagery	Review	"where" and "how" systems are structurally and functionally organized in the posterior parietal lobe.	Posterior parietal cortex: spatial transformations.
Galati et al., 2000	Spatial imagery	fMRI: egocentric and allocentric judgment of horizontal lines	The allocentric task activates a subset of the areas activated by the egocentric frames.	Common to both tasks: right SPL and right IPS.
Grefkes et al., 2002	Spatial imagery	fMRI: delayed match-to-sample task	Two objects had to be compared haptically or visually, or crossmodal.	Crossmodal matching of visual and haptic object information: left IPS.
Laeng, 1994	Spatial imagery	Lesion study: categorical and coordinate judgments	30 patients with left hemispheric damage and 30 patients with right hemispheric damage	LH-patients: impaired on categorical judgments. RH-patients: impaired on coordinate judgments.
Mellet et al. 2000	Spatial imagery	fMRI: mental navigation and mental map task based on route and survey perspective	Both tasks activate a parieto-frontal network involved in spatial mental imagery.	Spatial mental imagery: IPS, superior frontal sulcus, middle frontal gyrus and pre-supplementary motor area.
Ricciardi et al., 2006	Spatial imagery	fMRI: haptic and visual spatial one-back matrix task	Locations on two matrices had to be compared.	Haptic and visual maintenance activates a frontoparieto network including the bilateral precuneus and IPS.
Saito et al., 2003	Spatial imagery	fMRI: haptic and visual cross-modal shape matching	2D shape matching.	Shape information is integrated in the posterior IPS.

Author	Modality	Task	Comments	Brain Areas
Trojano et al. 2002	Spatial imagery	fMRI: mental clock task with categorical and coordinate judgments	Categorical and coordinate spatial relations are lateralized.	Categorical: left SPL and AG. Coordinate: right SPL and AG.
Trojano et al. 2006	Spatial imagery	rTMS: mental clock task with categorical and coordinate judgments	rTMS over the posterior parietal cortex.	Categorical: impaired after left parietal stimulation. Coordinate: impaired after right parietal stimulation.
Goyal et al., 2006	Tactile imagery	fMRI: tactile imagery and perception of a static or moving face/non-face	In late blind, but not congenitally blind, extrastriate visual areas show cross-modal activation for visual and tactile processing.	hMT/V5 and FFA is involved in imagery of moving patterns or faces in late blind and sighted. In late blind these areas were also active during tactile perception.
Uhl et al., 1994	Tactile imagery	EEG: tactile imagery of texture	The blind also activate the primary visual cortex. The sighted also reported visual imagery, which could account for their occipital lobe activation.	Parietal sensory association cortex: tactile imagery in the blind and sighted.
Yoo et al., 2003	Tactile imagery	fMRI: Tactile stimulation versus tactile imagery	Specific areas in the somatosensory areas were involved in tactile perception only. Left parietal lobe was only active during imagery.	Contralateral primary and secondary somatosensory areas, precentral gyrus and left parietal lobe.
Kosslyn et al., 2001	Visual and auditory imagery	Review	Mental imagery recruits similar neural networks as actual perception, but imagery does not require sensory processing. This is also true for auditory processing and imagery.	Primary visual cortex active during perception and imagery. Primary auditory cortex active during perception only.
Farah, 1989	Visual imagery	Image generation and rotation	ERP's showed lateralization for different tasks. There is also evidence for a 'what' and 'where' distinction in imagery.	Left posterior: generation. Right posterior: rotation.
Ganis et al., 2004	Visual imagery	fMRI: visualization or perception of faint drawing with judging specific aspects	Visual imagery and visual perception recruit very similar networks. The overlap is largest in the frontal and parietal regions and less in the temporal and occipital regions.	Imagery regions were a subset of visual perception regions.

Author	Modality	Task	Comments	Brain Areas
Kaski, 2002	Visual imagery	Review	Perception and visual imagery are closely related and have common mechanisms with unique characteristics.	Occipital cortex and higher order visual areas in the parietal and temporal lobe are involved in perception and imagery.
Kosslyn et al., 1978	Visual imagery	Mental scanning task	Metric spatial information is preserved during visual mental imagery.	
Kosslyn et al., 1990	Visual imagery	4 tasks: image generation, image maintenance, image scanning and image rotation	These 4 components were present in children and adults. However, scanning and rotation involves a spatial component which was more difficult for younger subjects.	
Kosslyn et al., 1999	Visual imagery	PET and rTMS: visualization of stripe patterns	Early visual cortex is involved in the visualization of stimuli and the activation is linked to information processing.	V1.
Mazard et al., 2005	Visual imagery	fMRI: mental inspection and perception of object and non-object drawings	Imagery resulted in stronger left inferior frontal and inferior temporal gyrus activation than perception of objects. This could reflect verbal retrieval of the object.	The left occipito-temporal-frontal network is stronger activated when imagining objects versus non-objects. Figurative imagery could involve V1.
Aleman et al., 2001	Imagery and blindness	Odd-one-out and spatial imagery task with spatial interference	Blind were able to perform the spatial and pictorial imagery tasks, but made more errors than sighted subjects.	
Cattaneo et al., 2008	Imagery and blindness	Review	Blind and sighted perform similar on spatial tasks, but subtle differences exist which could be due to sequential instead of parallel processing in the blind.	In blind and visually impaired the visual cortices reorganize.
Cornoldi et al., 2009	Imagery and blindness	2D mental pathway task	Blind and sighted perform the same when using a verbal strategy. When using a spatial strategy blind perform poorer than sighted.	

Author	Modality	Task	Comments	Brain Areas
De Beni & Cornoldi, 1988	Imagery and blindness	Memorization of single nouns, pairs or triplets	When memory load was small, blind could generate interactive images. When memory load was high blind subjects were impaired.	
Kerr, 1983	Imagery and blindness	Mental scanning, scanning objects with different sizes, cued-recall test	Blind process spatial images similar to sighted, but require slightly more time. Metric spatial information is preserved.	
Klatzky & Golledge, 1995	Imagery and blindness	A battery of 'tabletop' and spatial locomotor tasks	Performance of early blind, late blind and sighted was highly comparable. However, a linear discriminant analysis could distinguish three groups.	
Knauff & May, 2006	Imagery and blindness	Deductive reasoning	Congenitally blind perform less accurate and slower than sighted subjects. But, sighted subjects suffer from the visual impedance effect, congenitally blind do not.	
Lambert et al., 2004	Imagery and blindness	fMRI: mental imagery from animal names vs. listening to abstract words	Mental imagery of animals activates V1 in blind subjects. The ventral pathway processes figurative features while the dorsal pathway processes dimensional and spatial features. Blind subjects generate mental images in the dorsal pathway and they rely on haptic sensitivity.	V1 active during imagery. Visual for sighted, haptic for blind.
Noordzij et al., 2006b	Imagery and blindness	Route and survey descriptions of an environment: 3 spatial tasks	Early and late blind can build up a spatial mental model from a verbal description, but more efficiently from route descriptions. Visual experience is not essential for the development of spatial representations.	

Author	Modality	Task	Comments	Brain Areas
Noordzij et al., 2007	Imagery and blindness	Visual, spatial, auditory imagery tasks	There were no differences between blind and sighted subjects on the auditory imagery task. However, for the spatial imagery task there was a benefit from visual experience.	
Postma et al., 2007	Imagery and blindness	Haptic matching shapes to a cut-out board and giving verbal labels to the locations	Early and late blind were faster than blindfolded sighted. Blind used more object-references in the verbal labels, while sighted referred more to the surrounding frame.	
Postma et al., 2008	Imagery and blindness	Haptic parallel-setting task	In a direct setting condition early blind, late blind and sighted performed similar. With a delay, late blind and sighted improved due to the use of visual imagery. Early blind did not improve.	
Vanlierde et al., 2003	Imagery and blindness	PET: 2D pattern symmetry recognition task	Performance the same for blind and sighted. Verbal memory control task.	Visuo-spatial imagery activates precuneus, SPL and occipital gyrus in both blind and sighted.
Vanlierde et al., 2004	Imagery and blindness	2D pattern symmetry recognition task	Performance was the same for blind and sighted, but with different strategies. Sighted and late blind use a visuo-spatial strategy. Early blind a coordinate (X,Y) strategy.	
Vecchi, 2001	Imagery and blindness	3D matrix spatial task	Visuo-spatial working memory is involved in spatial mental imagery. However, it is differently organized in blind subjects due to serial processing.	
Vecchi et al., 2004	Imagery and blindness	2D matrix spatial task	Congenitally blind and sighted performed the same on single matrices. Blind were better at integrating two matrices than at remembering both. In sighted the pattern was reversed.	

Author	Modality	Task	Comments	Brain Areas
Baciu et al., 1999	Spatial language comprehension	fMRI: dot-bar categorical/coordinate task	There is hemispheric specialization for spatial relations. During practice the activity in the right angular gyrus decreases, while the activity in the left homologue increases. This could be due to the formation of new categories.	Left AG: categorical spatial relations. Right AG: coordinate relations.
Damasio et al., 2001	Spatial language comprehension	PET: naming actions and spatial relations	The left SMG is perhaps related to retrieving spatial prepositions. The right SMG is associated with the conceptual part of processing spatial relations.	Left SMG: processing spatial relations. Right SMG: naming spatial relations from abstract shapes.
Jäger & Postma, 2003	Spatial language comprehension	Review	Categorical/coordinate spatial relations are lateralized. Both networks involve posterior parietal areas.	Left posterior parietal: categorical spatial relations. Right posterior parietal: coordinate spatial relations and navigation.
Kemmerer, 2006	Spatial language comprehension	Review	The left inferior parietal lobe is an important region as a neural correlate of categorical spatial relations.	Left SMG and AG: spatial relations.
Kosslyn et al., 1989	Spatial language comprehension	4 tasks: judging categorical and coordinate spatial relations	Categorical/coordinate is lateralized. Language processing has categorical properties and is processed in the left hemisphere. Navigation has coordinate properties and is processed in the right hemisphere.	Left hemisphere: categorical. Right hemisphere: coordinate.
Kosslyn, 1987	Spatial language comprehension	Review	There exists a distinction between categorical and coordinate spatial relations.	
Noordzij et al., 2006a	Spatial language comprehension	EEG: spatial sentence-sentence and sentence-picture verification tasks	Spatial language is processed by means of a duo-representational model. People automatically generate a propositional code of a spatial sentence and can additionally generate a visuo-spatial representation.	In a spatial sentence a stronger parieto-occipital negativity, indicative of a visuo-spatial strategy in the visual context.

Author	Modality	Task	Comments	Brain Areas
Noordzij et al., 2008	Spatial language comprehension	fMRI: spatial sentence-sentence and sentence-picture verification	The left SMG is activated in both verbal and visual-spatial context	Left SMG: processing spatial relations.
Reichle et al., 2000	Spatial language comprehension	fMRI: spatial sentence-picture verification, with linguistic and visual-spatial strategies	The linguistic strategy activates language areas and the visual-spatial strategy activates parietal areas. Better verbal skills reduce the activity in the linguistic network, while better imagery skills reduce the activity in the parietal network.	Language areas: linguistic strategy, parietal areas: visual-spatial strategy.
Tranel & Kemmerer, 2004	Spatial language comprehension	Lesion study: production, comprehension and semantic analysis of spatial relations	The left prefrontal region: specialized for processing phonology of locative prepositions. The left inferior parietal region: specialized in processing the meaning of locative prepositions.	Left inferior parietal: meaning of spatial relations.

AG = angular gyrus, FFA = fusiform face area, hMT/V5 = human middle temporal/V5, IPS = intra-parietal sulcus, SMG = supramarginal gyrus, SPL = superior parietal lobule, V1 = primary visual cortex

Simple spatial sentences

VIS
ON
LIND
SPACE
URVE
DESC
RIPTI
ON
AN
GE
PH
MO
DAL
LE
FT
RA
GH
CON
GE
NIT
ALL
ROU
SUP
RA
EYE
RNEC
VAL
REN
M
ENT
AL
REP
RESEN
TATI
ON
POIN
TOFW

Chapter 3

Spatial language processing in the blind: Evidence for a supramodal representation and cortical reorganization

M. E. Struiksma, M. L. Noordzij, S. F. W. Neggers, W. M. Bosker & A. Postma

(submitted)

Abstract

Neuropsychological and imaging studies have shown that the left supramarginal gyrus (SMG) is specifically involved in processing spatial terms (e.g. *above, left of*), which locate places and objects in the world. The current fMRI study focused on the nature and specificity of representing spatial language in the left SMG by combining behavioral and neuronal activation data in blind and sighted individuals. Data from the blind provide an elegant way to test the supramodal representation hypothesis, i.e. abstract codes representing spatial relations yielding no activation differences between blind and sighted. Indeed, the left SMG was activated during spatial language processing in both blind and sighted individuals implying a supramodal representation of spatial and other dimensional relations which does not require visual experience to develop. However, in the absence of vision functional reorganization of the visual cortex is known to take place. An important consideration with respect to our finding is the amount of functional reorganization during language processing in our blind participants. Therefore, the participants also performed a verb generation task. We observed that only in the blind occipital areas were activated during covert language generation. Additionally, in the first task there was functional reorganization observed for processing language with a high linguistic load. As the visual cortex was not specifically active for spatial contents in the first task, and no reorganization was observed in the SMG, the latter finding further supports the notion that the left SMG is the main node for a supramodal representation of verbal spatial relations.

Introduction

Spatial language is used to locate places and objects in the world. In the human brain in particular the left SMG has been found to be crucial for processing spatial language. Neuropsychological patients with lesions to the left supramarginal gyrus (SMG) show remarkable and specific difficulties in producing and understanding spatial terms, in particular locative prepositions such as *above* and *to the left of* (Amorapanth et al., 2010; Kemmerer, 2005; Tranel & Kemmerer, 2004). Furthermore, a recent fMRI study by Noordzij et al. (2008) revealed higher activity in the left SMG for spatial sentences, containing locative prepositions, than for non-spatial sentences (see also Damasio et al., 2001; and Emmorey et al., 2002). This difference in activity was present in both a verbal and a visual-spatial context. The results by Noordzij et al. (2008) could be explained by a supramodal representation of spatial information, implying that activity in this region exceeds information from the stimulus modality yielding similar activation for verbal and visual-spatial contexts (Barsalou, 1999; Bonino et al., 2008; Pietrini et al., 2004; Ricciardi et al., 2006; Struiksma et al., 2009). Furthermore, a supramodal representation maintains a link with the input modality, which can explain behavioral differences between different input modalities. Noordzij et al. showed that participants responded faster when a spatial sentence was followed by a picture than when it was followed by a sentence. An alternative explanation could be that the left SMG is activated by the spatial representation of perceptual or verbal input presented visually. This would be in line with a multimodal representation which is linked to the perceived modality, in this case visual, and established in modality-specific brain areas (Barsalou, 1999; Struiksma et al., 2009).

The aim of the present fMRI study was to further examine the nature of the spatial language representation. By including blind individuals, who never have had any visual experience and are unable to use visual imagery to represent a spatial description, we can dissociate between supramodal and multimodal representations of spatial language. In case of a supramodal representation both blind and sighted individuals would show the same activation independently from the sensory modality that conveys spatial information. Namely, spatial information, available in visual and non-visual modalities, is represented modality-independent. In case of a multimodal representation we would expect to find differences in neural activation for blind and sighted individuals, since both groups would recruit the network most suitable and available for the task. This would mean that the blind group would recruit a tactile/auditory network while the sighted would recruit a visual network. In addition to distinguishing between supramodal and multimodal representations the results from the blind participants also provide information about the functional development of this neural activation. If the left SMG is indeed activated by a supramodal representation of spatial language, i.e. this area is also activated in the blind, this demonstrates that the role of the left SMG is hard-wired and does not require visual experience.

Research with blind and sighted individuals has shown that there are only subtle differences in performance on several spatial tasks (Cattaneo et al., 2008), which suggests that there might be overlap in the neural networks employed and that these functions also develop in the absence of vision.

Indeed, such overlap in neural networks has been found in dorsal and ventral occipito-parietal areas (Bonino et al., 2008; Mahon et al., 2009; Ricciardi et al., 2006; Stilla et al., 2008; Vanlierde et al., 2003). Yet, the literature reports both similar and different findings on the connectivity within these networks in blind and sighted individuals, as well as on the strategies used. In a visual and/or tactile spatial one-back recognition task of 2D and 3D matrices sighted and blind individuals similarly activated a fronto-parietal network comprising bilateral posterior parietal cortex and dorsolateral and inferior prefrontal areas. These results indicate that visual experience is not a prerequisite for the development of spatial working memory (Bonino et al., 2008; Ricciardi et al., 2006). Vanlierde et al. (2003) also found a similar pattern of activation for blind and sighted participants in a spatial imagery task of 2D matrices, involving the precuneus, superior parietal lobule and occipital gyrus, however, participants differed in their strategy. Sighted participants used a visual imagery strategy, while blind participants used an X-Y coordinate strategy (Vanlierde et al., 2003; Vanlierde & Wanet-Defalque, 2004). Whereas Stilla et al. (2008) also observed a similar network for blind and sighted individuals in a tactile microspatial discrimination task, the effective connectivity differed between the blind and sighted. The findings by Stilla et al. (2008) can easily be explained by a supramodal representation of spatial information. Information derived from different modalities or different strategies, contains spatial properties that evoke a supramodal, modality-independent, representation yielding similar results in blind and sighted (Lacey et al., 2007; Struiksmma et al., 2009). At the same time, there may also exist subtle differences, because participants partly maintain the original traces of the input modalities, with sighted still having access to prior visual information, while the blind only have access to nonvisual information.

When comparing blind and sighted individuals it is important to keep in mind that the primary visual cortex of early blind people, who have been blind since they were at most four years old, has received very little or no visual input, and is therefore subject to neuroplastic changes resulting in reorganization (Noppeney, 2007; Noppeney et al., 2005; Pascual-Leone et al., 2005; Shimony et al., 2006). As a consequence, the primary visual cortex of early blind people may get involved in performing non-visual tasks for which sighted people do not show any primary visual cortex activity. For example a wide variety of linguistic tasks have shown occipital cortex activity, e.g. letting blind participants read Braille (Amedi et al., 2004; Burton et al., 2002a; Gizewski et al., 2003; Melzer et al., 2001; Sadato et al., 1996; Sathian, 2005; Uhl et al., 1991). Moreover, when Cohen et al. (1997) used transcranial magnetic stimulation (TMS) on the visual cortex, comprising BA 17, 18, 19, the performance of blind participants decreased on Braille symbol identification. In contrast, sighted participants did not decrease in their performance on embossed Roman letter identification during TMS stimulation to the visual cortex, also comprising BA 17, 18, 19 (see Pascual-Leone et al. (2005) for a review). The primary visual cortex of early blind is not only involved in tactile reading tasks, but also in verbal word association tasks indicating that the functionality of the reorganization goes beyond the analysis of tactile information (Raz et al., 2005). Covertly generating an associated verb to a noun has been shown to elicit primary visual cortex activation in blind participants but not in sighted participants. Supporting the functional role of this activation accuracy was reduced during

repetitive TMS (Amedi et al., 2004; Burton et al., 2002b). Moreover, not only activity related to language processing has been found in the reorganized primary visual cortex, but also activity related to spatial imagery (Vanlierde et al., 2003) and tactile discrimination (Matteau et al., 2010; Ptito & Kupers, 2005). Learning to discriminate the orientation of a letter *T* applied to the tongue with a tongue display unit (TDU) resulted in significant visual cortex activation in blind, but not in sighted individuals. In a follow-up study transcranial magnetic stimulation was used to stimulate the visual cortex before and after training with the TDU (Kupers et al., 2006). Kupers et al. found that blind participants reported tactile sensations on the tongue. Interestingly, these sensations were somatotopically organized.

The foregoing results suggest the possibility that the reorganized visual cortex in the blind is also specifically suitable for processing spatial information. In the context of our spatial language experiment we therefore were interested in the functional relevance of the expected visual cortex activation in the blind. Hamilton et al. (2000) reported on a congenitally blind patient who suffered bilateral occipital strokes. She was a profound Braille reader, but after the stroke she was unable to discriminate tactile information necessary for the complex spatial decoding involved in Braille reading. The involvement of the visual cortex in spatial discrimination was supported by a low-frequency rTMS study by Merabet et al. (2004). They found that rTMS to the visual cortex specifically impaired distance, but not roughness, judgments. The hypothesized additional occipital activation in blind during the spatial language task could follow from language processing in general, comparable to what has been found with other language tasks (such as the classic verb generation task), or demonstrate specific involvement in processing language pertaining to space. Nevertheless, the findings on the spatial language task only allow backward inference about any expected activation in the occipital cortex, i.e. comparing them to previous findings from other tasks. To test to what extent the reorganization is functional and comparable to the established body of literature on reorganization, the same participants also performed a covert verb generation task.

In light of the foregoing, the aim of the present study was twofold: first, to determine whether left SMG activation is associated with supramodal representations which develop in the absence of visual experience; second, to investigate the possible functional reorganization of spatial language processing in the blind. The current task was adapted from Noordzij et al. (2008) to an auditory presentation and included different sentence types in order to determine the specificity of the left SMG.

Methods

Participants

Fourteen early blind and fourteen blindfolded sighted control participants, with no neurological or motor deficits, participated in this experiment. One early blind and one sighted control participant were excluded from the analysis due to task performance at chance level, resulting in two groups of

thirteen participants. There were twelve congenitally blind participants, who have been blind since birth. One participant was born with buphtalmus and lost his eyes in an operation at the age of four. However, before that age he was severely visually impaired and he has no memory of vision. The blind participants and healthy controls were matched for sex, education, age and handedness (for details, and etiology of the blind participants see Table 1). All participants signed an informed consent prior to the experiment, which was approved by the Medical Ethical Board (Medische-toetscommissie (METC-protocolnumber 05/186-E)).

Table 1 Description of early blind participants

<i>Subject number</i>	<i>Occupation</i>	<i>Education level</i>	<i>Sex</i>	<i>Handedness</i>	<i>Age</i>	<i>Etiology*</i>	<i>Age of onset</i>
1	Unemployed	University	f	r	39	ROP	0
2	Translator	Higher Education	f	a	35	ROP	0
3	Computer Programmer	Higher Education	m	r	38	CG	0
4	Sports Masseuse	Vocational Education	f	a	45	LCA	0
5	Office Assistant	Higher Education	f	r	32	CG	0
6	Educator	University	m	l	40	A	0
7	Unemployed	Higher Education	f	l	31	LCA	0
8	Policy Worker	University	f	l	30	LCA	0
9	Student	Higher Education	m	l	19	LCA	0
10	Sales Person	Higher Education	m	a	41	ROP	0
11	Student	Higher Education	m	r	22	ND	0
12	Programmer	Higher Education	m	l	49	B	4
13	Sound Technician	Higher Education	m	r	53	ROP	0

*definitions of etiology: A = Anophthalmia, B= Buphtalmus, CG = Congenital Glaucoma, LCA = Leber's Congenital Amaurosis, ND=Norrie Disease, ROP = Retinopathy of Prematurity. Mean age Early Blind: 36.46 ± 9.80. Mean age Sighted Controls: 37.15 ± 11.16

Design and procedure

Both tasks were presented auditorily through MR-compatible headphones. The task was administered on a PC with Presentation software 9.90 (Neurobehavioral Systems, Albany, CA). Prior to the fMRI experiment participants were instructed about the scanning procedure and both tasks were explained. During the instruction phase participants practiced the two tasks and received feedback.

The spatial language task was based on a sentence verification paradigm. The trials consisted of a sequence of two auditorily presented sentences (see Figure 1A). The experiment included four sentence types with a spatial and non-spatial compound preposition and compound adverb (*left/right of, together with¹, taller/smaller than and older/younger than*). The task was to judge whether or not

1 It is a matter of debate to what extent conjunctions contain a spatial element. Importantly, this type of sentence does not require imposing a reference frame and setting a spatial template in order to allow for an explicit comparison along a spatial dimension (Logan & Compton, 1996). This is precisely the type of contrast we were interested in.

the two sentences described the same situation. All participants were instructed to finish listening to the entire second sentence before responding as quickly and as accurately as possible. Participants responded with the thumb of their right hand on a MR-compatible pneumatic response-box. They pressed the right button when they thought the trial was correct, and the left button when the trial was considered incorrect (see Table 2). There were three types of trials: repetitions, reversals and catch trials (see Table 2). A repetition trial consisted of two sentences in which the order of the two names was identical. In a reversal trial the order of the two names was different for the two sentences. Catch trials contained a third name (i.e. not mentioned in the first sentence) in the second sentence. The participants did not have to pay attention to the different trial types, but only to the different sentence types, which were separated in time by a short silent interval. There were 144 trials presented in blocks (15s) of two trials of one sentence type. Each sentence type consisted of 36 trials, with 16 reversals, 16 repetitions and 4 catch trials. There were four sessions of 18 blocks. The inter-block interval varied between 6s and 9s. The different sentence types were presented pseudo-randomly. No feedback was given.

Table 2 Five different options for sentence 2 in relation to sentence 1.

Sentence Option (e.g. S1 is “ <i>Fien taller than Max</i> ”)	Trial Type	Correct Response
Relation and position of names identical S2: “ <i>Fien taller than Max</i> ”	Repetition	yes
Relation changed and position of names identical S2: “ <i>Fien smaller than Max</i> ”	Repetition	no
Relation and position of names changed S2: “ <i>Max smaller than Fien</i> ”	Reversal	yes
Relation identical and position of names changed S2: “ <i>Max taller than Fien</i> ”	Reversal	no
Introduction of a third name S2: “ <i>Fien taller than Stein</i> ”	Catch	no

The verb generation task consisted of auditorily presented nouns and participants had to covertly generate an associated verb (Amedi et al., 2004; Amedi et al., 2003; Burton et al., 2002b; Raz et al., 2005). Since giving a verbal response during scanning produces movement artifacts the participants were required to only give a mental response. Therefore, no behavioral data was collected.

The verb generation task consisted of three conditions: *word*, *nonword* and *rest*. In the *word* condition participants heard a noun and were instructed to covertly generate an associated verb. The *nonword* condition consisted of trials with passive listening to reversed speech. The sound-spectrum of regular words was reversed with the program CoolEdit 2000 (www.cooledit.com). In this condition the stimuli sounded like words, but had no semantic interpretation, which was used as a control condition for auditory input (Binder et al., 2000).

We used a block-design with 18 blocks of 10 trials of 2.8s, 6 blocks of each condition (*word*, *nonword* and *rest*), presented in pseudorandom order. At the beginning of each block with sound stimuli a beep was presented. A short beep indicated a block with words while a long beep indicated a block with non-words. No feedback was given.

MR data acquisition

Scanning was performed with a 3.0T Philips Achieva scanner using an eight-channel SENSE head-coil to acquire T2*-weighted images with blood oxygenation level dependent (BOLD) contrast. We used the principles of echo shifting with a train of observations (PRESTO) scanning sequence, combined with 2D-SENSE acquisition. This sequence uses three dimensional imaging in combination with a delayed echo read-out after the next RF pulse (Neggers et al., 2008). Also, T2* acquisition was accelerated in 2 directions (2D-SENSE) by skipping lines in K-space. Together this resulted in a four-fold increase in imaging speed in PRESTO. PRESTO-SENSE has been demonstrated to yield fast and reliable activation for 1D-SENSE (Klarhöfer et al., 2003), and is even more sensitive than conventional EPI when using 2D-SENSE (Neggers et al., 2008). An entire volume was acquired in 500.3 ms (TE = 32.4 ms, TR = 21.75 ms, flip angle = 10°, 56×64 acquisition matrix, 32 sagittal slices, isotropic voxels of 4 mm, FOV(ap, fh, rl) = 224×256×128 mm and a SENSE factor of 2 in the AP and 1.8 in the LR direction). Each scanning session was preceded by ten dummy volumes in order to accomplish steady state transversal magnetization.

The spatial language task consisted of four sessions of 800 volumes each. After the final session a reference-scan was acquired, with a flip angle of 25°, but otherwise identical to the PRESTO-SENSE functional MRI images. Due to the increased flip angle this image had slightly more anatomical contrast and was used for coregistration with the anatomical scan. After the spatial language task a T1-weighted anatomical scan was acquired (TE = 4.6 ms, TR = 9.86, flip angle = 8°, 224×224 acquisition matrix, 160 coronal slices, voxel size = 0.875×0.875×1 mm and FOV(ap, fh, rl) = 224×160×168 mm). During the anatomical scan the participants could rest.

The same PRESTO-SENSE sequence was used to acquire T2*-weighted images with blood oxygenation level dependent (BOLD) contrast. The verb generation task consisted of one session of 1024 volumes. After the task another reference scan with slightly more anatomical contrast was collected.

Data analysis

Behavioral data

For each participant individual mean reaction times and performance scores were collected. Group analyses were performed with SPSS (SPSS for Windows, Rel. 14.0.2. 2006. Chicago: SPSS Inc.). Behavioral data were analyzed with a 2 × 2 × 2 × 2 mixed Analysis of Variance (ANOVA). Space (spatial/non-spatial), Category (compound preposition/compound adverb) and Trial Type (repetition or reversal) were the within-subject factors and Group (blind or sighted) was the between-

subject factor. The catch trials were excluded from the analysis: they were included (11%) to keep participants alert and to make sure they paid attention to both names mentioned in the sentence. The results reported below show the effects for correct answers with a significance level of $p \leq 0.05$. When pairwise comparisons were tested the significance level was corrected for multiple comparisons using the Bonferroni method. SPSS multiplies the p -value with the Bonferroni multiplier instead of dividing α by the Bonferroni multiplier. However, the results are equal and we will denote the Bonferroni corrected p -values by p_B .

Since the verb generation task required participants to covertly generate associated verbs there were no behavioral data which could be analyzed.

Functional imaging data

Imaging data was analyzed with SPM5 (Wellcome Department of Imaging Neuroscience, Institute of Neurology, London, www.fil.ion.ucl.ac.uk/spm/software/spm5) and the MarsBaR toolbox for SPM5 (marsbar.sourceforge.net) running under Matlab (R2007b, The MathWorks, Inc., Natick, MA). Preprocessing included coregistration and realignment. The anatomical scan was segmented and spatially normalized with medium regularization (0.01). The spatial normalization parameters from the 'unified segmentation' routine were used to normalize all functional scans (Ashburner & Friston, 2005; Crinion et al., 2007), which were then spatially smoothed with a kernel of 8 mm FWHM.

First level statistics was performed for each participant individually. A high-pass filter with a cutoff period of 128s was applied to remove low frequency fluctuations. The model for the spatial sentence comprehension task contained 62 regressors. The design matrix consisted of eight regressors (Space \times Category \times repetition and reversal trials), three or four nuisance-regressors for the sporadic catch-trials in each session and three additional nuisance-regressors for each session to filter out a very systematic scanner-related oscillation in a very narrow frequency band exactly at 0.5 Hz. Each functional regressor was convolved with a hemodynamic response function.

Using a general linear model the parameter estimates were calculated for all brain voxels. Several effects, mentioned below, were tested by means of linear contrasts between the parameter estimates for different conditions. These contrast images were then passed to a second-level analysis, to model any group effects.

The contrast of interest in the spatial language task were: (i) specific language effects: (a) *left/right* > *together*, (b) *spatial (left/right and taller/smaller)* > *non-spatial (together and older/younger)*. The distinction between spatial and non-spatial might not be as straightforward as assumed. The three sentence types *left/right of*², *taller/smaller than* and *older/younger than* can all be represented on an imaginary axis. *Left/right* can be ordered on a horizontal axis, *taller/smaller* on a vertical axis and *older/younger* on a horizontal time line. The control condition *together with* does not automatically generate

2 Spatial prepositions can be divided into directional prepositions indicating a change in direction, and locative or relational prepositions, describing relations between objects (Coventry & Garrod, 2004). Here we extend the relational aspect of spatial prepositions.

such a line analogy. Therefore, we included a third contrast (c) *relational* > *together* to test this analogy (for the (a), (b) and (c) contrasts the results for repetition and reversal trials were collapsed). In a fourth contrast we tested the difference between (d) *reversal* > *repetition*. The final contrast was: (ii) general language and reorganization effects: *task activation* > *rest*.

In order to determine the differences between the blind and sighted we performed two-sample t-tests at the second level with a contrast between blind and sighted individuals. The commonalities between the two groups were tested by means of a conjunction analysis that tested the conjunction null hypothesis over two orthogonal contrasts (Friston et al., 2005; Nichols et al., 2005). *P*-values were adjusted for the search volume using random field theory and inferences were drawn at the cluster level (details are explained below). The null distribution for the minimum statistic was based on two statistics. This enabled us to infer a conjunction of activation in an area in both blind and sighted groups (Friston et al., 2005).

We report the results of a random effects analysis, with inferences drawn at the cluster level. *P*-values were corrected for multiple comparisons by combining a $p < 0.001$ voxel-level threshold with a cluster extend threshold to obtain a $p < 0.05$ whole-brain corrected significance level (Friston et al., 1996). For the specific language effects a small volume correction was applied (Friston, 1997). This procedure constrained our search space to a spherical region of interest (ROI) in the left SMG (a radius of 10mm around -36, -48, 40 MNI coordinates, based on the coordinates from the peak voxel of a significant cluster of 18 voxels reported by Noordzij et al. (2008)).

The model of the verb generation task consisted of two regressors for the *word* and *nonword* condition and three additional nuisance-regressors to filter out a very systematic scanner-related oscillation in a very narrow frequency band exactly at 0.5 Hz. The two functional regressors were convolved with a hemodynamic response function.

The contrast of interest in the verb generation task was the analysis: *word* > *nonword*. In order to determine the differences between the blind and the sighted we performed a two-sample t-test at the second level with a contrast between blind and sighted individuals. The commonalities between the two groups were tested by means of a conjunction analysis. We report the results of a random effects analysis, with inferences drawn at the cluster level, with similar correction applied as in the spatial language task.

Results

Spatial language task

Behavioral data

The behavioral data indicated that participants performed very accurately on the spatial language task (>90% correct trials). A 2 (Space) × 2 (Category) × 2 (Trial Types) × 2 (Group) mixed ANOVA on the percentage of correct trials, revealed no group differences between the blind and sighted participants ($F(1,24) < 1$). There was a significant main effect of Space. Participants performed

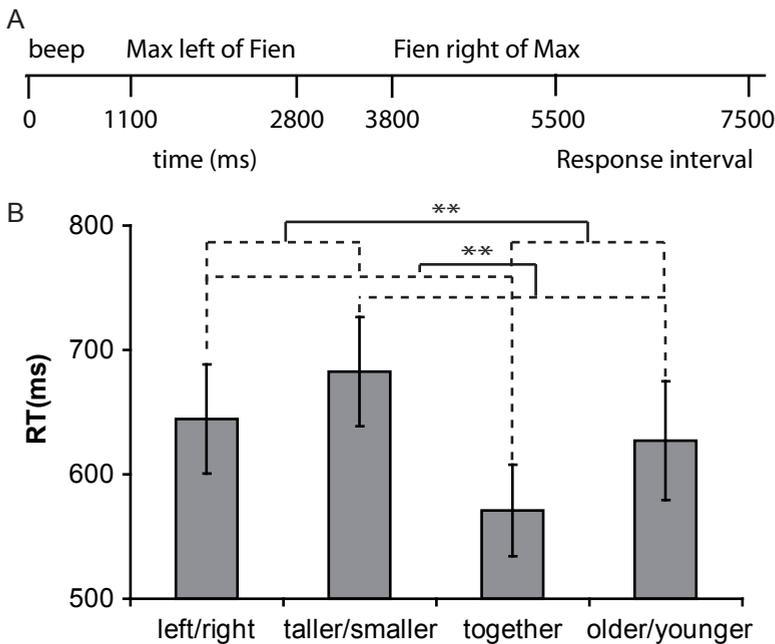


Figure 1 Example of a trial and behavioral results (RTs) from the spatial language task.

(A) An example of a single trial in the spatial language task. Each trial starts with a beep and consists of two spoken sentences and a response interval. In this situation the response was affirmative. (B) Mean reaction time and standard errors in ms on the different sentence types, collapsed for blind and sighted participants. Significant ($p < .05$) effects for Space and Category are marked with **.

slightly better on the non-spatial sentences ($97\% \pm 1.0$ correct), compared to the spatial sentences ($96\% \pm 1.1$ correct, $F(1,24) = 4.38, p = .047$). There was also a significant main effect of Trial Type. Participants made slightly, but significantly, more errors on the reversal sentences ($95\% \pm 1.5$ correct), compared to the repetition sentences ($98\% \pm .7$ correct, $F(1,24) = 7.19, p = .01$).

The behavioral results on the mean response times are shown in Figure 1B. There was a significant main effect of Space ($F(1,24) = 18.24, p < .001$), Category ($F(1,24) = 6.73, p < .016$), and Trial Type ($F(1,24) = 15.95, p = .001$). The interaction between Category and Trial Type was also significant ($F(1,24) = 5.61, p = .026$). Planned pairwise comparisons between repetitions and reversals for both categories showed that participants evaluated repetitions significantly faster than reversals in all conditions (prepositions: $t(11) = -2.88, p_B = .008$, adverbs: $t(11) = -3.51, p_B = .002$), but the difference was larger for adverbs than prepositions.

Although there was no main effect of Group ($F(1,24) < 1$), the interaction between Group and Trial Type was significant ($F(1,24) = 5.85, p = .026$). Further analysis indicated that blind participants were significantly slower on reversal than repetition sentences ($t(11) = -4.53, p_B < .001$), while the sighted participants were equally fast on both sentence types ($t(11) = -1.11, p_B = .276$).

Functional imaging data

The neuroimaging results focus on four main contrasts in the spatial language task. The general task activation was analyzed by means of a whole brain conjunction analysis on the contrast *task activation* > *rest*. The network of significant activation in blind and sighted comprised bilateral parietal areas, bilateral thalamus, right cerebellum and right lingual gyrus (see Table 3 for details and Figure 2 for a visual representation).

Table 3 General task activation compared to rest common to blind and sighted subjects

Region	Cluster Size	Peak voxel Z-score	MNI coordinates (mm)		
			x	y	z
R cerebellum	472	5.97	32	-52	-28
		5.51	20	-55	-20
		5.19	12	-60	-16
L postcentral gyrus	655	5.96	-36	-28	52
L inferior parietal lobule		5.82	-36	-52	52
L supplementary motor area		5.71	0	4	56
R superior temporal gyrus	136	5.64	52	-24	0
		5.15	52	-36	4
		5.05	52	-12	4
L superior temporal gyrus	373	5.42	-56	-16	0
L insula		5.40	-36	16	4
L superior temporal gyrus		5.10	-52	4	-12
R thalamus	192	5.37	12	-16	8
		5.04	12	-28	-4
L thalamus	112	4.98	-12	-16	0
R precentral gyrus		5.16	48	4	44
		5.03	36	-12	60
R postcentral gyrus	102	4.89	28	-4	52
		5.01	44	-28	44
		4.95	36	-32	44
R inferior parietal lobul	16	4.47	36	-55	48
R inferior frontal gyrus		5.16	32	20	4
R lingual gyrus		4.71	4	-76	4
		4.51	12	-72	12

L = left, R = right. T-threshold = 4.20, spatial extent threshold: >10 voxels.

Further analyses were based on a previous fMRI study in sighted individuals (Noordzij, et al., 2008). For the contrast spatial preposition *left/right* versus non-spatial preposition *together* an ROI conjunction analysis was performed which resulted in a significant cluster of activation in both blind and sighted individuals in the left SMG (T-threshold = 2.6, spatial extent threshold: >5 voxels, peak voxel MNI coordinates: -36, -48, 44, Z-score = 2.80, cluster size: 6 voxels, see Figure 3A). We thus replicated the findings by Noordzij et al. (2008) and extended them to blind participants. The present study aimed to investigate the specificity of this finding by adding a spatial and non-spatial adverb to the stimulus set. The ROI conjunction analysis on the *spatial* > *non-spatial* contrast did not yield any significant results.

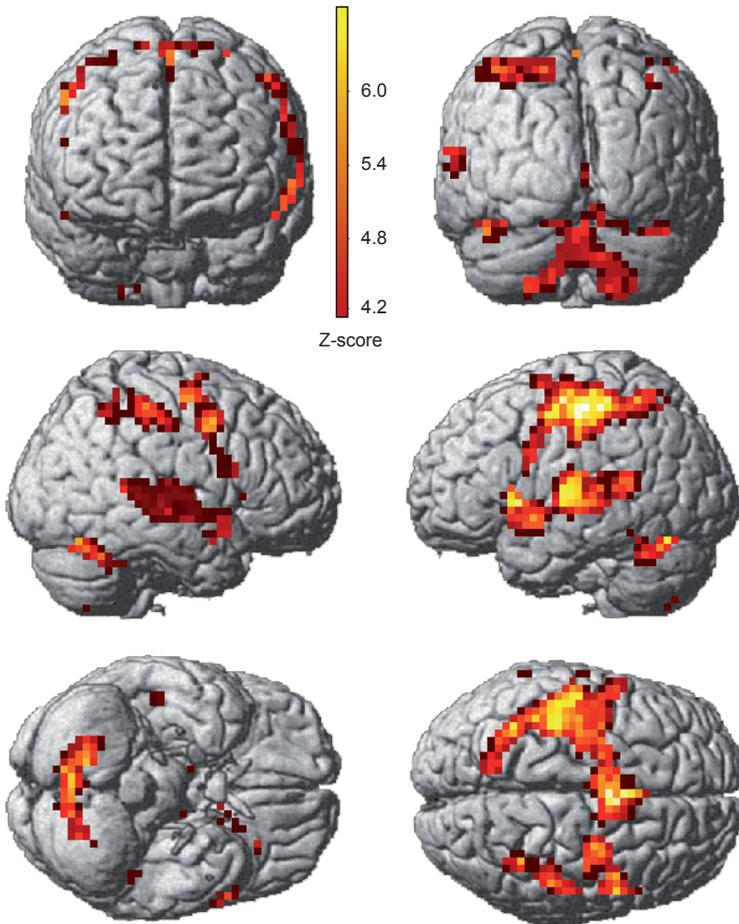


Figure 2 fMRI results from the spatial language task, commonalities between blind and sighted. Network of general task related activation in blind and sighted during the spatial sentence comprehension task. A whole brain conjunction analysis of the contrast *task activation* > *rest*. T-threshold = 4.2, spatial extent threshold: >10 voxels.

However, as stated before, the sentence type *older/younger* can be considered spatial when represented on a horizontal axis. One can imagine comparing ages of people on a timeline, which, as space, can be represented by a canonical axis. We therefore analyzed the contrast between relational sentence types (*left/right*, *taller/smaller* and *older/younger*) and the non-relational sentence type *together*. The conjunction analysis on this *relational* > *together* contrast also revealed a significant cluster in both blind and sighted individuals in the left SMG (T-threshold = 2.6, spatial extent threshold: >5 voxels, peak voxel MNI coordinates: -36, -48, 44, Z-score = 3.20, cluster size: 15 voxels, see Figure 3B).

The regression parameters of the left SMG for all four sentence types (Figure 3C), in both the blind and sighted participants, were higher for sentences in which an evaluation about a relation

(*left/right*, *taller/smaller* or *older/younger*) was required than for sentences in which only the correspondence of the two names had to be verified (a sentence with *together*).

The behavioral results showed that the responses of blind participants were significantly slower for reversals than repetitions, while the sighted participants responded equally fast to both. However, there was no significant difference in activation in the left SMG between the reversals and repetitions for both blind and sighted individuals, nor for the blind in particular.

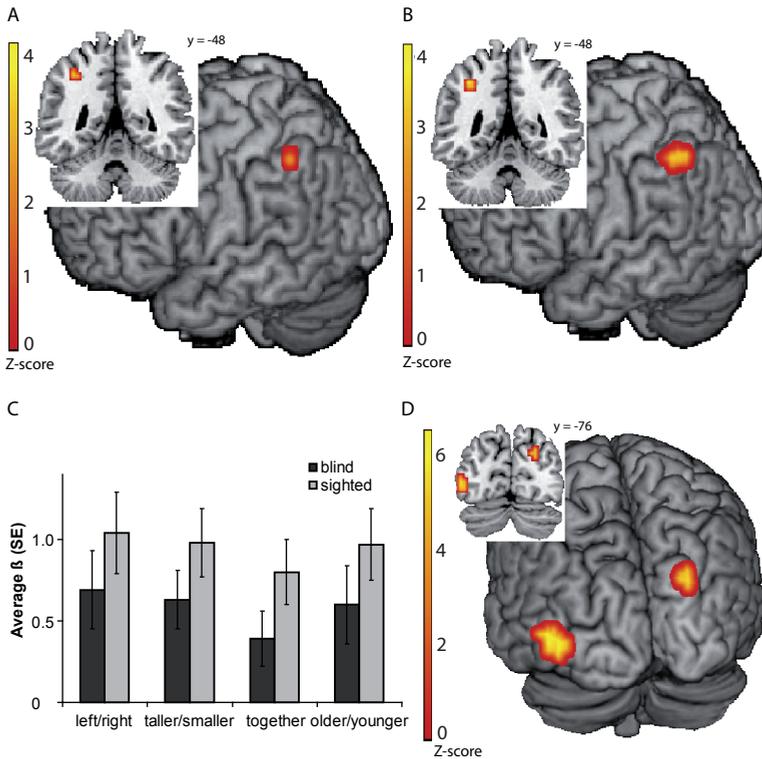


Figure 3 fMRI results from the spatial language task, commonalities and differences between blind and sighted. (A) ROI conjunction analysis of the contrast *left/right* > *together* showing a significant cluster of activation in the blind and sighted individuals. T-threshold = 2.6, spatial extent threshold: >5 voxels, peak voxel MNI coordinates: -36, -48, 44, Z-score = 2.80, cluster size: 6 voxels. (B) ROI conjunction analysis of the contrast *relational(left/right, taller/smaller and older/younger)* > *together* showing a significant cluster of activation in the blind and sighted individuals. T-threshold = 2.6, spatial extent threshold: >5 voxels, peak voxel MNI coordinates: -36, -48, 44, Z-score = 3.20, cluster size: 15 voxels. (C) The average parameter estimates (β) for the functional ROI of the left SMG from the contrast *relational* > *together*. The dark gray bars indicate the parameter estimates for the blind and the light gray for the sighted individuals. The error bars denote the standard error of mean. (D) Difference between blind and sighted individuals for the contrast *task activation* > *rest* in the spatial language task. T-threshold = 4.2, spatial extent threshold: >10 voxels. A cluster of 23 voxels: middle occipital gyrus (peak voxel MNI coordinates: -48, -80, 4, Z-score = 4.66). A cluster of 10 voxels: cuneus (peak voxel MNI coordinates: 24, -80, 32, Z-score = 4.17).

Apart from commonalities the differences between blind and sighted were also analyzed to determine the level of reorganization. For the whole-brain analysis on the general contrast *task activation* > *rest* we found a significantly higher activation for the blind compared to the sighted individuals in the left middle occipital gyrus and right cuneus (see Figure 3D for details). There were no significantly greater activations in the sighted, compared to the blind. Furthermore, there were no general task activation differences between blind and sighted, tested in both directions, in the left SMG.

Verb generation task

Functional imaging data

The verb generation task was included in order to determine the level of reorganization for general language processing. A conjunction analysis between the blind and sighted participants on the contrast *word* > *nonword* revealed those areas that are specific to generating verbs. In blind and sighted participants covert language generation activated bilateral language areas, including the inferior frontal gyrus and middle temporal gyrus also known as Broca's and Wernicke's area respectively (see Table 4 for details). Further significant activation was found bilaterally in the cerebellum and the supplementary motor area as well as the left precentral gyrus. These latter areas are involved in the covert generation of the words (Ackermann et al., 1998; Herholz et al., 1996).

The differences between blind and sighted participants were found mainly in the occipital cortex (see Figure 4A for details). The left cuneus and bilateral middle occipital gyrus showed significantly higher activation for blind compared to sighted participants. In the right hemisphere the activation of the middle occipital gyrus extended slightly into the middle temporal gyrus.

Reorganization

The *word* > *nonword* contrast in the verb generation task revealed reorganized areas in the blind, similar to what has been previously found (Amedi et al., 2003; Burton et al., 2002b). The functional result from the verb generation task was used to define an ROI for further analysis on the reorganization within the spatial language task. For the general contrast *task activation* > *rest* we found a significantly higher activation for the blind compared to the sighted individuals in the left middle occipital gyrus (T-threshold = 3.5, spatial extent threshold: >5 voxels, peak voxel MNI coordinates: -48, -76, 4, Z-score = 4.52, cluster size: 31 voxels) and left cuneus (T-threshold = 3.5, spatial extent threshold: >5 voxels, peak voxel MNI coordinates: -16, -84, 28, Z-score = 3.92, cluster size: 7 voxels, see Figure 4B for details). The reorganized areas did not show a modulation for the *left/right* > *together* contrast, neither for the *spatial* > *non-spatial* contrast, nor for the *relational* > *together* contrast. Interestingly, the contrast *reversal* > *repetition* did show significant activation within the reorganized areas in the blind. The activation was found in the left middle occipital gyrus (T-threshold = 3.5, spatial extent threshold: >5 voxels, peak voxel MNI coordinates: -48, -76, 4, Z-score = 4.55, cluster size: 14 voxels) and left lingual gyrus (peak voxel MNI coordinates: -32, -72, -12, Z-score = 4.38, cluster size: 51 voxels, see Figure 4C). This finding suggests that the reorganized areas are involved

in processing information with a higher linguistic load, which has been reported before.

For example, Amedi et al. (2003) showed that superior verbal memory is correlated with occipital cortex activation, while Van der Lubbe et al. (2010) have shown that superior duration discrimination abilities were related to enhanced occipital negativity in the blind, during an electroencephalogram (EEG) experiment. We performed a correlation analysis on the spatial language data in order to test this explanation. For each blind participant the number of significantly activated voxels, within the verb generation ROI, on the contrast *task activation > rest* was counted. There was a significant negative correlation between the number of voxels in the reorganized area and the difference in reaction time between reversals and repetitions ($r(11) = -.54, p = .03$, see Figure 4D). This means that blind participants who were better at the reversal trials, i.e. responded equally fast to reversals and repetitions, were most likely to have a higher level of reorganization.

Table 4 General language processing effects common to blind and sighted subjects

Region	Cluster Size	Peak voxel Z-score	MNI coordinates (mm)		
			x	y	z
Supplementary motor area	127	5.40	0	0	60
		4.69	0	16	44
		4.53	12	0	68
R Cerebellum	34	4.84	32	-64	-24
		4.67	40	-60	-28
L inferior frontal gyrus	123	4.83	-48	16	-4
		4.68	-20	12	4
		4.52	-52	16	4
L precentral gyrus	21	4.82	-48	-4	48
R inferior frontal gyrus	48	4.65	36	24	4
		4.51	48	16	-4
L middle temporal gyrus	33	4.46	-52	-36	0
		4.42	-56	-24	0
L superior temporal gyrus		4.08	-60	-52	8
L cerebellum	36	4.31	-40	-56	-24
		4.24	-36	-48	-24
		4.08	-32	-60	-24
R cerebellum	22	4.28	12	-68	-16
		3.92	8	-60	-8

L = left, R = right. T-threshold = 4.20, spatial extent threshold: >10 voxels.

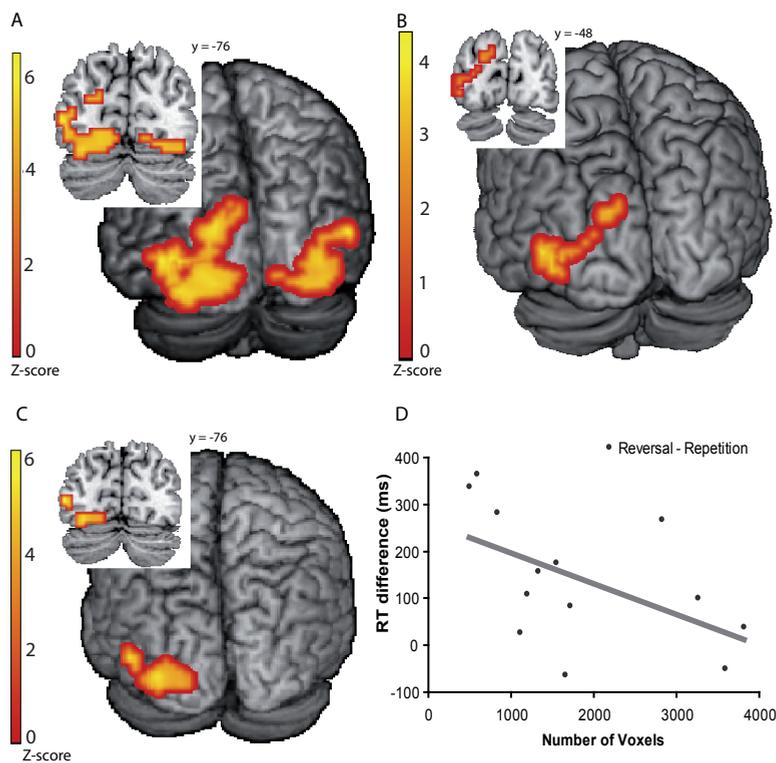


Figure 4 fMRI results from the verb generation task and reorganization analysis.

(A) Difference between blind and sighted individuals for the contrast word > nonword in the verb generation task. T-threshold = 4.2, spatial extent threshold: >10 voxels. Left cluster of 195 voxels: middle occipital gyrus (peak voxel MNI coordinates: -48, -76, 4, Z-score = 4.86); cuneus (peak voxel MNI coordinates: -48, -76, 4, Z-score = 4.86). Right cluster of 94 voxels: middle temporal gyrus (peak voxel MNI coordinates: 56, -56, 4, Z-score = 4.83); middle occipital gyrus (peak voxel MNI coordinates: 40, -60, -8, Z-score = 4.34). (B) Difference between blind and sighted individuals for the contrast task activation > rest in the spatial language task within the verb generation ROI. T-threshold = 3.5, spatial extent threshold: >5 voxels. A cluster of 31 voxels: middle occipital gyrus (peak voxel MNI coordinates: -48, -76, 4, Z-score = 4.52). A cluster of 7 voxels: cuneus (peak voxel MNI coordinates: -16, -84, 28, Z-score = 3.92). (C) Difference between blind and sighted individuals for the contrast reversal > repetition in the spatial language task within the verb generation ROI. T-threshold = 3.5, spatial extent threshold: >5 voxels. A cluster of 14 voxels: the left middle occipital gyrus (peak voxel MNI coordinates: -48, -76, 4, Z-score = 4.55, cluster size: 14 voxels) A cluster of 51 voxels: left lingual gyrus (peak voxel MNI coordinates: -32, -72, -12, Z-score = 4.38, cluster size: 51 voxels). (D) Correlation between the number of activated voxels in the reorganized occipital lobe in the blind and the difference in RT (ms) between reversals and repetitions ($r(11) = -.54$, $p = .03$).

Discussion

The aim of this study was to determine whether spatial language is represented in a supramodal representation in the left SMG that does not require visual experience to develop its functionality. Sustained activation in the left SMG in blind and sighted participants during an auditory spatial language task supported a supramodal representation of spatial language. In addition, the verb generation task further established the reorganized cortical areas in the blind. Activation in reorganized visual cortical areas in the blind was not associated with specific spatial processing. However, this activation did have functional relevance because it increased with higher linguistic load.

Besides differences in the occipital lobe most of the activated areas were similar in the blind and sighted individuals. The large language network that was revealed in the verb generation task indicated that reorganization in the blind was limited to the occipital areas. This finding was supported by the large network found in the spatial language task for both blind and sighted as well. The overlapping result in the left SMG during the spatial language task also supports this notion. The present study used an auditory presentation and extends the finding by Noordzij et al. (2008) on the contrast *left/right > together*.

Noordzij et al. (2008) used a visual paradigm with verbal and visual-spatial contexts. In their study the left SMG was activated regardless of the nature of the visual stimulus. The first stimulus was always a sentence while the second stimulus could be another sentence or a picture. Noordzij et al. analyzed activation related to the second stimulus separately and found sustained left SMG activation for both types of second stimuli. In contrast, the behavioral results revealed that participants responded significantly faster to pictures compared to sentences indicating that sighted participants used a visual-spatial representation to compare a spatial sentence to a picture (Noordzij et al., 2008; Noordzij et al., 2006a). If the activation found in the left SMG had been due to linguistic processing of spatial sentences, then no sustained activation for the second stimulus, in particular the picture, should have been found. Instead the maintained activation found in the left SMG for both verbal and visual-spatial context might support a supramodal representation of spatial information.

Interestingly, the auditory paradigm in the current study generated left SMG activation similar to the visual paradigm used by Noordzij and colleagues. Moreover, while there were subtle behavioral differences between the blind and sighted participants, the blind also activated the left SMG and the behavioral differences were not correlated to left SMG activation. Together these results imply that the activation in the left SMG does not depend on visual experience, nor on the nature of the stimulus or the sensory input channel being used, but rather suggests a superordinate level of processing with maintained traces to the input modality and thus is associated with the supramodal representation of spatial terms. Importantly, the activation of the left SMG in the blind participants, who have never received visual input, demonstrates that visual experience is not a prerequisite for developing a supramodal representation of spatial information. This finding is in line with recent findings on object and spatial processing in cortical areas that were previously thought to be involved in encoding visual information only (Bonino et al., 2008; Mahon et al., 2009; Pietrini et al.,

2004; Ricciardi et al., 2006; Ricciardi et al., 2008). For example, the motion-sensitive middle temporal cortex (hMT+) not only responds to optic flow, but also to tactile flow, auditory movement or movement *per se* (Matteau et al., 2010; Poirier et al., 2006; Ricciardi et al., 2007). More recently, a supramodal sensory representation has also been found for the mirror system. This supramodal mirror system develops without access to visual experience and allows blind individuals to interact effectively with the world around them (Ricciardi et al., 2009).

In addition to replicating the *left/right > together* contrast from Noordzij et al. (2008) the *relational > together* contrast also showed significant activation in the left SMG. Space and time are closely related, as described by the Metaphoric Mapping Theory (Boroditsky, 2000; Haspelmath, 1997; Heine et al., 1991; Kemmerer, 2005; Santiago et al., 2007), which states that spatial schemas and temporal schemas share a relational, basic spatial structure. This allows organization of temporal concepts. Santiago et al. (2007) have shown that there exists a space-time conceptual metaphor along a mental time line: left-past and right-future. Accordingly, the temporal sentence type *older/younger than*, used in the current study, could also have been analyzed within a quasi-spatial schema along a mental time line. As a result there were three relational sentence types which ordered information along an axis. We suggest that the left SMG might be involved in the discrimination of information ordered along an axis, instead of being selectively involved in processing spatial prepositions.

The current study tested brain activation in two very different tasks. One might wonder why it is relevant to include the verb generation task. As mentioned in the introduction we were interested in the functional reorganization of the occipital cortex especially during the spatial language task. A recent review by Kriegeskorte et al. (2009) explained the possible problems of using the same data for selection and selective analysis, i.e. using the difference between blind and sighted on the *task activation > rest* contrast in the spatial language task to determine the ROI for testing the *relational > together* contrast could yield distortions and invalid statistics. The verb generation task, however, is inherently independent of the spatial language task. As such it was useful as a localizer for language related areas in the blind and provided an independent ROI which allowed for further analysis of occipital cortex activation in the spatial language task.

There is a large body of literature on the application of the covert verb generation task and it has proven to elicit robust language related activity as well as reveal occipital areas that have been subjected to reorganization in the blind (Amedi et al., 2004; Amedi et al., 2003; Bookheimer, 2002; Burton et al., 2002b; Cuenod et al., 1995; Gernsbacher & Kaschak, 2003; Raz et al., 2005). The results in the verb generation task verified that the blind activate classical language areas, similar to the sighted participants. Importantly, the blind also showed additional language related activity in the occipital cortex for two very different tasks: the verb generation and the spatial sentence comprehension task. This constitutes further evidence of reorganization of the occipital cortex and adds to the established body of literature on reorganization. There is an increasing amount of evidence that the reorganization of the occipital cortex is functionally relevant and involved in the processing of language processing, Braille reading, spatial imagery and tactile discrimination (Cohen et al., 1997; Kupers et al., 2006; Ofan & Zohary, 2007; Pascual-Leone et al., 2005; Röder et al., 2002; Uhl

et al., 1991; Van der Lubbe et al., 2010; Vanlierde et al., 2003).

Based on the somatotopic reorganization found by Kupers et al. (2006) we hypothesized that the reorganized occipital cortex of the blind might also be specifically suitable for processing language pertaining to space. Our results do not support this idea. The reorganized areas did not show a modulation for the spatial or relational conditions. Interestingly, we did find a significant difference between the blind and sighted individuals in the occipital cortex for the contrast *reversal* > *repetition*. The left middle occipital gyrus showed an increased activation for reversals in the blind. This might be related to the behavioral difference that blind participants responded slower to reversal trials than repetition trials, while this difference was absent in the sighted individuals. The reversal trials have a higher linguistic complexity since the relation between the persons was changed, but the situation described could still be the same (e.g. “*Max left of Wies*” and “*Wies right of Max*”). The activation in the left middle occipital gyrus could therefore be associated with processing linguistically more complex information in the blind. A negative correlation between the level of reorganization in the blind and their performance on linguistically more complex trials suggests this functional role. Participants with larger activity in the reorganized visual cortex responded more equally to reversal and repetition trials.

A possible explanation for the fact that the sighted participants responded equally fast to both reversal and repetition trials is that they used a visual-spatial strategy to solve the task. In contrast, the blind participants might rely more on a verbal strategy, which is more sensitive to linguistic complexity (Noordzij et al., 2005, 2006a). It should be noted here that no differences in activation between reversal and repetition trials were found within the left SMG. Even with a possible difference in strategy between blind and sighted, both groups showed significant and comparable activation in the left SMG. This further strengthens the idea that a supramodal representation may underlie spatial language processing. Moreover, the concept of a supramodal representation does not exclude possible differences in strategy, such as previously found by Vanlierde et al. (2003, 2004) and described in the introduction. Rather it focuses on the common underlying types of information, such as spatial information elements.

In conclusion, the present study offers further insights in the patterns of brain activation underlying spatial language processing. We found that during language processing in general the blind and sighted individuals activate a highly similar network. In addition the blind participants also showed activation in the reorganized occipital cortex. This reorganization activity was also found in the spatial language task, and was modulated by linguistic complexity. There thus seems to be a functional relevance of the reorganized areas; they support processing of linguistically more complex trials. However, within these regions there is no further distinction for a particular semantic category, in this case spatial relations. Importantly, the left SMG appears particularly involved in parsing relational terms ordered along a single dimension. The finding that blind participants also activate the left SMG when processing relational terms is very interesting since it implies that the role of the SMG is hard-wired. Regardless of visual experience, the left SMG supports the supramodal representation of spatial and other dimensional relations in language.

Chapter 4

The level of preparation and control in reference frame processing

M. E. Struiksmā, M. L. Noordzij & A. Postma

(in preparation)

Abstract

When giving verbal spatial directions the visual scene might not be readily available. The current study describes three sentence-picture and picture-sentence verification experiments that investigated to what extent we can prepare mapping linguistic and perceptual components and whether it is subject to control. The first experiment showed that spatial indexing, reference frame activation and spatial template construction occurred after having only seen a picture, but not after having only read a sentence. However, when participants were cued to use a specific reference frame in experiment two they were able to activate these frames already after having read a sentence. The third experiment, in which the expectancy of a specific cue was manipulated, tested whether the non-cued reference frame was still available, or whether participants would adjust to a more efficient strategy, in which only a single reference frame was activated. The results suggested that after reading a sentence the absolute/relative reference frames were default, while the availability of the intrinsic reference frame was strategic. On the other hand, after seeing a picture there was clear evidence of a strategy focusing on a specific reference frame as indicated by significantly lower response times to the unexpected reference frame. The current set of experiments provides evidence that participants adjust their strategy and can prepare several processing steps, in particular reference frame activation, based on the type of information available. This supports a supramodal spatial representation of spatial information from different modalities.

General introduction

When communicating about objects one can discuss multiple different properties. One can talk about their form, color, or function. In addition one can describe the relation between the object and other objects, in particular the spatial relations. During communication about spatial relations it is important to use the same point of view as the addressee, also known as reference frame. For example, when you want to convey the message “the bird is above the tree” this could point to different locations depending on the orientation of the tree, standing up or lying down.

Seminal work on understanding spatial sentences has been done by Carlson and colleagues (Carlson-Radvansky & Irwin, 1994; Carlson-Radvansky & Logan, 1997; Logan, 1994). When a spatial sentence describes a particular scene verbal and perceptual information have to be compared. This information is converted into a spatial mental representation which allows comprehension of the situation (Tversky, 1991) and requires several steps. Carlson-Radvansky and Logan (1997) suggested a model indicating processing steps required in spatial language comprehension, (based on theories from e.g. Carlson-Radvansky & Irwin, 1994; Levelt, 1984; Logan & Sadler, 1996) According to this model when a spatial sentence and a scene co-occur several processes are triggered by one’s aim to understand the situation. First, spatial indexing links an object from the perceptual representation to a symbol in the conceptual representation, based on the verbal message, and identifies the reference object (“tree” in the example above) (Step 1). Second, multiple reference frames are imposed coordinating the mapping between language and perception (Step 2). Third, spatial templates are constructed, which denote good, acceptable and bad regions of space associated with the spatial term and relevant reference frame (Step 3). Fourth, a reference frame is selected by inhibiting the other reference frames (Step 4). Fifth, the spatial templates are weighted according to the selected reference frame and combined into a composite template (Step 5). Finally, a search is performed on the composite template to make an acceptability judgment which is made on the regions of acceptability as defined by the reference frames (Carlson-Radvansky & Logan, 1997) (Step 6).

Reference frames consist of three axes, which parse up space into different directions specifying location, and have several parameters that can be adjusted: e.g. origin, orientation, direction, and distance (Carlson-Radvansky & Logan, 1997; Carlson & Van Deman, 2004). The relevant axis used to match a spatial sentence to a scene differs according to which of the three reference frames is adopted: absolute, relative or intrinsic (Carlson, 1999; Levinson, 1996). The absolute (A) reference frame uses environmental characteristics, such as gravity and cardinal directions, to determine the orientation of the axes. In the relative (R) reference frame the orientation of the axes is based on the viewer, while in the intrinsic (I) reference frame the reference object defines the relevant axis. In the canonical situation these three reference frames are aligned (ARI), for instance when the tree, in the example above, is upright and the viewer is upright. However, when the tree is lying down, the intrinsic (I) reference frame is misaligned with the absolute and relative (AR) reference frame and when the viewer is lying down the relative (R) frame is misaligned with the absolute and intrinsic

(AI). These specific situations can provide insight in the availability and use of the different reference frames.

The model proposed by Carlson-Radvansky and Logan (1997) was based on visual search situations where the spatial sentence and visual scene were presented simultaneously. Thus the hypothesized set of processing steps may apply specifically when there is direct comparison between the verbal description and visual scene. However, in real life such a situation might not always occur, for example when giving directions the visual scene is not readily available and spatial language is presented in isolation or, when explaining the set-up of your new room to another person you first observe the room and later use spatial language to describe what you have seen. This raises the question whether we are forced to wait with mapping linguistic and perceptual components until we can directly perceive the visual scene, while keeping the sentence in short-term memory, or whether we can prepare some of the processing steps anticipating the visual input. Alternatively, one might wonder whether at the moment of observing a visual scene you are aware of the possible uncertainty in sentence options and whether you prepare for these different sentence options. Questions closely related to the possibility of preparation are what exactly we prepare for, which processing steps can be prepared and whether this preparation is automatic, or subject to strategies.

In line with the foregoing, the aim of the current study was to determine the level of preparation in spatial sentence processing. In particular we focused on reference frame processing, the critical step in which differences between situations occur when perspective could be ambiguous, for example when a reference object is rotated. Regardless of perspective, spatial indexing, multiple reference frame activation and spatial template construction (steps one, two and three) occur similarly in these situations. However, the selection of a reference frame differs between situations through dissociation of the relevant axes.

Most research on reference frames has used a verification paradigm where the sentence and picture were presented simultaneously. Early work on sentence-picture verification paradigms investigating linguistic complexity has used simultaneous presentation with the instruction to first read the sentence and then examine the picture (Clark & Chase, 1972). Later work on linguistic complexity has used a serial presentation starting with a sentence, followed by a picture (MacLeod et al., 1978). The serial sentence-picture verification paradigm has also been used to examine spatial attention and the influence of context on spatial sentence verification (Logan, 1994; Noordzij et al., 2005). However, serial presentation has not received much attention in reference frame research. By disentangling the presentation of verbal and pictorial input the response patterns can provide information on the type of information required for processing. Equal response patterns for serial sentence-picture (SP) and picture-sentence (PS) verification suggests that both verbal *and* pictorial information is required in order to initiate the processing steps described by Carlson-Radvansky and Logan (1997). Different response patterns for SP and PS could indicate that participants can already prepare (some) processing steps on the basis of one type of information. In the present study we discuss three experiments which shed light on the level of preparation and control in spatial sentence processing. In Experiment 1 we examine which information is required to initiate

the processing steps necessary for comprehension. In Experiment 2 we subsequently test whether participants can control the onset of reference frame activation. Finally, in Experiment 3 we investigate how flexible the control is that participants can have on reference frame activation.

Experiment 1

In order to determine whether reference frame activation is susceptible to preparation Experiment 1 combines sentence-picture and picture-sentence verification. Critically, preparation in the form of processing steps, as mentioned above, require time. Given that participants are able to prepare, limiting preparation time should result in an increase in response times, simply because preparation might not have been completed. Therefore, in the current experiment the delay between the sentence and the picture is manipulated to obtain information about the preparation process, especially the time point at which reference frames can be activated. If reference frames are activated during the delay, then changing the length of the delay should interfere with the response times. Previous research has used a number of different procedures. In the current experiment the delay is 0ms, 500ms or 1000ms. The 0ms is adopted from research by Carlson and colleagues (e.g. Carlson-Radvansky & Logan, 1997; Regier & Carlson, 2001) who usually present a sentence which is replaced by a picture. The picture then remains on the screen until a response is given. The 500ms is adopted from a series of experiments conducted by Logan (1994). Finally, the 1000ms is adopted from Taylor et al. (1999) who presented a picture for 2000ms followed by a delay of 1000ms and then the sentence.

Since the experimental setup involved participants sitting on a chair, the absolute and relative frame are always aligned. In order to create a conflict between reference frames the reference object is rotated, dissociating the intrinsic (I) from the absolute/relative (AR) frames. In the critical trials the participants gave a yes-response to either the canonical orientation in which all frames were aligned (ARI) or to a non-canonical orientation in which the AR or I frame provided the correct response.

There are two possible scenarios: first, participants cannot prepare any of the processing steps resulting in no effect of inducing a delay. As such, reference frames can only be activated and selected after the second stimulus, when both verbal *and* perceptual information are available. The model by Carlson-Radvansky and Logan predicts this scenario where verbal information is used to guide visual attention (see also Logan, 1995, 1999). Second, spatial information can be represented in a supramodal fashion in a spatial image (Struiksma et al., 2009). Such a supramodal representation can be built up from different modalities and extracts the relevant spatial information. This could indicate that from both verbal and visual information spatial information can be stored as a spatial image which is later used for comparison to the second stimulus. As such a certain level of preparation could occur during the delay resulting in interference between delay and response times, yielding shorter response times for longer delays. In order to determine whether participants need one type of information in order to prepare for the second type of information or whether

they need both types of information to yield a response the analysis focuses on the three-way interaction between stimulus order, delay and reference frame.

In sum, the first experiment was designed to determine whether reference frame activation is limited to situations where verbal *and* perceptual information are present, or whether there are processing steps which are subject to preparation anticipating the next piece of information.

Yes response	No response
Ball left of car	Ball right of car
Ball above car	Ball below car
Ball smaller than car	Ball taller than car



Figure 1 Example of sentence types and a stimulus picture.

This picture provided yes-response in combination with the sentences in the left column. And no-responses in combination with the sentences in the right column

Methods

Participants

In the first experiment we tested 24 right-handed participants (12 male), mean age 21.1 ± 0.5 years old. All participants were native Dutch speaking students from Utrecht University and were paid €9,- for participation. Three participants did not understand the purpose of the task and responded only according to a relative and intrinsic frame respectively. These participants were excluded from the analysis. All participants gave informed consent and were naïve with respect to the purpose of the experiment.

Stimulus material

Two sequentially presented stimuli comprised a trial: a sentence followed by a picture or a picture followed by a sentence. There were 16 different pictures of a car and a ball (4 orientations of the car \times 4 positions of the ball, for an example see Figure 1). Participants always saw the back of a car. This way the intrinsic left and right of the car were aligned with the participants' left and right. There were 6 possible sentences (see Figure 1). For every delay each picture was presented twice with each spatial sentence (3 delay \times 16 pictures \times 2 = 96 trials per sentence). In total there were 480¹ trials per order (SP or PS). In 45% of the trials participants were required to give a yes-response.

¹ We also included nonspatial sentences (taller/smaller: 98 trials) to encourage participants to verify the identity of the objects. These trials turned out to be subject to a short-cut strategy, namely the ball was always smaller than the car, resulting in ceiling level performance. These trials will not be discussed further.

Design and procedure

Participants were tested for approximately one and a half hour. The experiment consisted of two types of verification tasks: SP and PS. Each order (SP or PS) was tested with three blocks with different delays (Delay=0, 500 or 1000ms). In total there were six different blocks which were presented to each subject in a counterbalanced order.

The experiment started with an instruction about the possible answers. Participants saw an example of a picture, such as in Figure 1, and were explained that “*above*” could be interpreted from you own point of view, but also from the cars’ point of view. Their task was to judge as quickly and accurately as possible whether the sentence and the picture described the same situation according to any, or both, of the possible interpretations. They were given three practice trials for each order (SP and PS) in which the delay decreased from 1000 to 500 to 0ms. Subsequently they started with the six different blocks. Prior to the start of each block participants were informed about the order of the stimuli and the delay.

A trial consisted of a fixation cross (500ms), the first stimulus (1000ms either a sentence or a picture), a varied delay (0, 500 or 1000ms) and a second stimulus (2000ms during which the response was given). The response was given on the keyboard. Participants pressed the left button (f) for “yes” and the right button (j) for “no”. After the practice trials the participants received feedback from the experimenter. During the six blocks no feedback was given. After each block participants could take a short break and continue by pressing the left button.

Data analysis

Overall performance was analyzed, after which the yes responses were analyzed in detail, because they provided affirmative information on the use of different reference frames. The response time analysis was performed on correct yes responses that were within the 2000ms response interval. Behavioral data on performance and response times was analyzed with a $2 \times 2 \times 3 \times 3$ repeated measures Analysis of Variance (ANOVA). The within subject variables were: Order (SP, PS), Preposition (*left/right, above/below*), Reference Frame (AR, I, ARI) and Delay (0, 500, 1000ms). The reported results include effects with a significance level of $p < .05$. Further analysis by means of pairwise comparisons used a significance level corrected for multiple comparisons with the Bonferroni method. SPSS multiplies the p -value with the Bonferroni multiplier instead of dividing α by the Bonferroni multiplier. The results are, however, equal and this method corrects for multiple comparisons. The Bonferroni corrected p -values will be denoted by p_B .

Results

Performance

Overall performance was high for both SP (93.1% \pm 4.3 correct) and PS (92.3% \pm 5.4 correct) and did not differ between the two orders ($t(21) = 1.01, p = .324$). Moreover, the task was to judge whether the sentence and the picture described the same situation according to any of the possible

interpretations. This led to three specific sentence-picture combinations for which a single frame indicated a correct answer. In both conditions participants were able to perform the task. They responded by using the intrinsic frame in 47% of the cases, while the ideal subject would have reached 50%.

The 2 (Order) \times 2 (Preposition) \times 3 (Reference Frame) \times 3 (Delay) repeated measures ANOVA on the percentage correct yes responses revealed a main effect of Preposition ($F(1,20) = 21.96, p < .001$) and Reference Frame ($F(2,40) = 19.82, p < .001$). Participants performed better on *above/below* judgments ($93.0\% \pm 1.1$) compared to *left/right* judgments ($89.2\% \pm 1.6$). Pairwise comparisons showed that participants were more accurate in the ARI condition, when both frames were aligned ($95.5\% \pm .8$), and slightly less accurate in the AR condition, when the absolute/relative frame provided a yes-response ($91.5\% \pm 1.7, t(20) = -3.42, p_B = .009$). Performance in the I condition, when the intrinsic frame provided a yes-response, was less accurate compared to the AR condition ($86.2\% \pm 2.0, t(20) = 3.31, p_B = .009$). The interaction between Order \times Reference Frame \times Delay was not significant ($F(4,80) = 1.44, p = .230$).

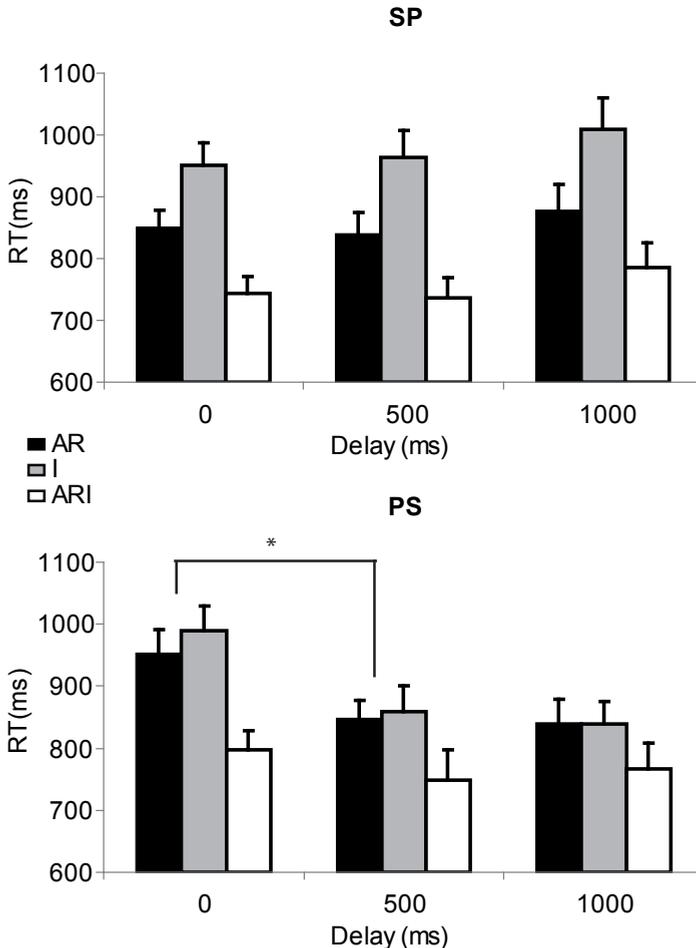


Figure 2 Mean response times (ms) for different delays for SP and PS ($N = 21, \pm$ SEM)

Reaction time

A separate $2 \times 2 \times 3 \times 3$ repeated measures ANOVA on the mean reaction times revealed a main effect of Preposition ($F(1,20) = 68.69, p < .001$) and Reference Frame ($F(2,40) = 105.15, p < .001$). Participants responded slightly faster to *above/below* judgments ($812\text{ms} \pm 28$) than to *left/right* judgments ($898\text{ms} \pm 34$). Pairwise comparisons showed that participants responded fastest in the ARI condition, when both frames were aligned ($763\text{ms} \pm 29$), a little slower in the AR condition, when the absolute/relative frame provided a yes-response ($867\text{ms} \pm 30, t(20) = -11.75, p_B < .001$), and slowest in the I condition, when the intrinsic frame provided a yes-response ($935\text{ms} \pm 35, t(20) = -5.41, p_B < .001$).

The two-way interaction between Order \times Reference Frame ($F(2,40) = 17.44, p < .001$) was significant, as was the three-way interaction of interest between Order \times Reference Frame \times Delay ($F(4,80) = 4.12, p = .004$). The two-way interaction followed the pattern of the main effect of Reference Frame for SP (ARI ($755\text{ms} \pm 34$) $<$ AR ($854\text{ms} \pm 32, t(20) = -9.29, p_B < .001$) $<$ I ($975\text{ms} \pm 38, t(20) = -9.26, p_B < .001$)). For PS ARI judgments ($771\text{ms} \pm 36$) were faster than AR judgments ($879\text{ms} \pm 31, t(20) = -8.27, p_B < .001$) which were as fast as I judgments ($895\text{ms} \pm 34, t(20) = -9.2, p_B = 1.000$). The three-way interaction further explained the two-way interaction. In SP for all three reference frames there were no significant differences between the three delays (all $ts < 1.60$, see also Figure 2). In PS for ARI there were no significant differences between the three delays (all $ts < 1.40$). However, for the AR and I reference frame we found the pattern that responses were significantly slower in the 0ms delay compared to the 500 (AR: $t(20) = 2.80, p_B = .040$, I: $t(20) = 3.00, p_B = .021$) and 1000ms delays (AR: $t(20) = 3.00, p_B = .021$, I: $t(20) = 5.16, p_B < .001$).

The three-way interaction between Order \times Preposition \times Reference Frame ($F(2,40) = 10.38, p < .001$) also followed the interaction between Order \times Reference Frame, as explained above, in addition to the main effect of Preposition where *above/below* judgments were slightly faster than *left/right* judgments. The two-way interaction between Order \times Preposition was significant ($F(1,20) = 23.02, p < .001$) as well as the three-way interaction between Order \times Preposition \times Delay ($F(2,40) = 6.35, p = .004$). This interaction showed a similar pattern for Order and Delay as the three-way interaction explained above. In addition, it followed the main effect of Preposition in that participants responded faster to *above/below* judgments than to *left/right* judgments.

Discussion

The main question was whether processing linguistic and perceptual components in order to comprehend spatial sentences can only be instantiated when all information is available, or whether participants can prepare, for example, by activating reference frames while awaiting visual input. In SP there was no effect of manipulating the delay, which confirms the model of Carlson-Radvansky and Irwin (1994) and implies that after reading a sentence participants did not prepare anticipating the picture. Only when both verbal and perceptual information was available processes were instantiated that allowed for the comprehension of the situation. In PS on the other hand, there was an effect of delay. Responses were significantly slower in the 0ms delay condition than in the 500 or

1000ms delay condition. This suggests that participants were able to prepare some processing steps during the delay, and that this preparation was possible within a 500ms delay.

Since the visual information was presented there are a few candidates for processing steps that can be prepared. The first step described in the introduction is spatial indexing. Although the exact verbal description was not yet presented, participants could have made an educated guess and could have indexed the reference object and located object. After the objects had been indexed, multiple reference frames could have been activated together with their associated spatial templates (steps two and three). However, no selection of a specific reference frame could have been made and no composite template could yet have been constructed (steps four and five) without the information presented in the sentence. This order of processing steps, as described in the model, was based on the information available from both verbal and perceptual representations. In PS the verbal information was not yet available, which might have altered the order of processing steps, or even altered the possible steps in itself. We could speculate that participants might have been able to generate prototypical sentences based on the multiple reference frames and spatial templates that could have been activated.

The assumption that participants could have generated prototypical sentences as response options can be an explanation for the fact that in PS ARI judgments were fastest, but there was no difference between AR or I judgments. In the ARI judgments all frames were aligned and there was no competition which could have slowed down responses. However, in the AR and I judgments the spatial template belonging to each reference frame has one region of perfect fit, which dictates only a single spatial term in both conditions, therefore, selection of a specific reference frame is similar for AR and I judgments and the difference disappears.

Our results on SP extend previous results by Carlson-Radvansky and Irwin (1994) in that ARI judgments were fastest. In the third experiment of Carlson-Radvansky and Irwin participants were also instructed to answer 'yes' if the sentence and picture matched according to any reference frame. Without this manipulation participants responded faster to AR than I judgments, however, with this instruction the difference between AR and I on response times and performance disappeared. Our experiments, with 'accept all' instruction, showed that participants were faster and more accurate on AR judgments than I judgments. A critical difference between the Carlson-Radvansky and Irwin paradigm and our paradigm is the simultaneous versus serial presentation of the sentence and picture. Moreover, the Carlson-Radvansky and Irwin paradigm tested only the spatial term *above*, while the current experiment tested four different spatial terms. The *above/below* judgments benefited from the asymmetric factor gravity (Franklin & Tversky, 1990). The *left/right* judgments on the other hand are symmetrical, and more difficult as has been shown by slower response times and more errors both in the literature (Carlson & Van Deman, 2008; Franklin & Tversky, 1990; Levelt, 1984; Levinson, 1996; Logan, 1995) and in the current experiment. The inclusion of different spatial terms also decreased the level of predictability in the current study. This could have been the reason that baseline differences between I judgments and AR judgments remained in the current experiment.

Manipulating the delay resulted in changes in response times in PS, but not in SP. This could be interpreted as reference frame activation not being task-driven, but rather an automatic process instantiated upon presentation of the picture. Apparently, there were (at least) three processing steps relevant in this situation: multiple reference frames activation, spatial templates construction and selection of a single reference frame (Carlson-Radvansky & Irwin, 1994; Carpenter et al., 1999). When the picture was the second stimulus, activation and selection of a reference frame was only instantiated at the end of the trial. However, when the picture was the first stimulus, reference frame activation and spatial template construction could have occurred in the delay, but a selection could only have been made after the sentence was presented. There is an interaction between the order of the stimuli and activation and selection of reference frames. As mentioned above, in PS the order of processing steps might even have been altered, which implies a certain level of voluntary control. In Experiment 2 we focus on this interaction and try to identify more precisely when frames are activated and whether this activation is subject to voluntary control.

Experiment 2

In this experiment a cue is introduced, instructing the subject which reference frames is required to solve the task. If susceptible to voluntary control, then in SP introducing a cue should influence the activation of reference frames, because participants need to wait until the picture is presented to start activating the frames without the cue. However, in PS participants already start activating reference frames during the delay, therefore, introducing a cue can provide information on the time-course of activating reference frames and can enhance selection. Given the fact that Experiment 1 showed that participants are able to activate reference frames in PS within the 500ms delay, in Experiment 2 we use the same 500ms delay. Furthermore, to obtain information on the flexibility of reference frame activation a cue is introduced at three different time points: early, middle and late during the delay. The interaction of interest is between Order, Reference Frame and Cue, because we might expect different patterns for SP and PS.

Methods

Participants

24 right-handed participants (11 male), mean age 21.1 ± 0.5 years old. All participants were native Dutch speaking students from Utrecht University and were paid €10,- for participation. Two participants were excluded from the analyses, one participant due to a failure of the program to save the responses in the first part of the experiment and one participant failed to correctly perform the control trials. All participants gave informed consent and were naïve with respect to the purpose of the experiment.

Stimulus material

The same stimulus material as in experiment 1 was used. Two sequentially presented stimuli comprised a trial: a sentence followed by a picture (SP) or a picture followed by a sentence (PS). The pictures and sentences were the same as used in Experiment 1. In this experiment we presented a cue between the first and second stimulus. The cue consisted of a small picture from the car (intrinsic cue) or the man (absolute/relative cue). For each sentence there were three pictures that provided a unique intrinsic response and three pictures that provided a unique absolute/relative response, see Figure 3. For each sentence these six unique pictures were combined with all possible cues (no cue/100/250/400ms \times AR/I = 8 cues). This set was repeated (6 pictures \times 8 cues \times 2 = 96 trials). In addition for each sentence the remaining 10 pictures were assigned randomly to the eight cues (10 trials). The control condition taller/smaller was combined with the 16 different pictures and the no-cue condition (16 trials). In total there were 456 trials (106 trials \times 4 spatial sentences + 16 trials \times 2 control sentences) for each order (SP and PS). In 57% of the trials a yes-response was required.

Similar to Experiment 1 participants were first explained what the possible interpretations could be. Subsequently they were instructed about the different options for the cue. The cue could be a picture of a car, implying an Intrinsic interpretation, or a picture of a man, implying a Relative interpretation. Their task was to judge as quickly and accurately as possible whether the sentence and the picture described the same situation according to the cued interpretation. When there was no cue, participants were instructed to use both interpretations. They were given 16 practice trials for each order (SP and PS) in which the four cue conditions were present. Subsequently they started with the two different sessions. Prior to the start of each session participants were informed about the order of the stimuli. Half way through each session there was a break.

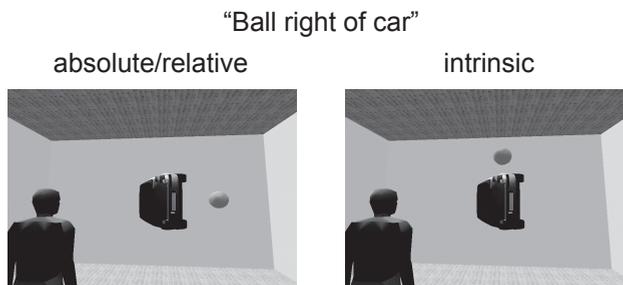


Figure 3 Two unique pictures for the sentence “Ball right of car”.

Design and procedure

Participants were tested for approximately one and a half hour. The experiment consisted of two types of verification tasks: SP and PS. Each order (SP or PS) was tested with a fixed delay of 500ms and four different cue presentations (no-cue, 100, 250 or 400ms). The different cue presentations were presented randomly. The presentation of the two orders, SP or PS, was counterbalanced over participants.

A trial consisted of a fixation cross (500ms), the first stimulus (1000ms either a sentence or a picture), during a 500ms delay a varied cue-condition was presented for 50ms (no-cue, 100, 250 or 400ms) and a second stimulus (the response interval was increased to 4000ms, since participants also had to process the cue). The response was given on the keyboard. Half of the participants pressed the left button (z) for “yes” and the right button (/) for “no”. The other half of the participants used the reverse response coding. After the practice trials the participants received feedback from the experimenter. During the two sessions no feedback was given. After each session participants took a short break.

Data analysis

Overall performance was analyzed, after which the yes responses were analyzed in detail, because they provided affirmative information on the use of different reference frames. The response time analysis was performed on correct yes responses that were within the 4000ms response interval. Behavioral data on performance and response times was analyzed with a $2 \times 2 \times 2 \times 4$ repeated measures ANOVA. The within subject variables were: Order (SP, PS), Preposition (*left/right, above/below*), Reference Frame (AR, I) and Cue (-400, -250, 100ms, no-cue). The reported results include effects with a significance level of $p < .05$. Further analysis by means of pairwise comparisons used a significance level corrected for multiple comparisons with the Bonferroni method. The Bonferroni corrected p -values will be denoted by p_B .

Results

Performance

Overall performance was high and did not differ between SP ($92.3\% \pm 4.0$) and PS ($93.6\% \pm 3.4$, $t(21) = -1.92$, $p = .068$). The 2 (Order) \times 2 (Preposition) \times 2 (Reference Frame) \times 4 (Cue) repeated measures ANOVA on correct yes responses revealed main effects of Preposition, Reference Frame and Cue. Performance on *above/below* judgments ($94.9\% \pm .6$) was better than on *left/right* judgments ($92.5\% \pm .7$, $F(1,21) = 15.26$, $p = .001$). Participants also performed better on I judgments ($94.9\% \pm .5$) compared to AR judgments ($92.6 \pm .8$, $F(1,21) = 8.64$, $p = .007$). The effect of cue type was also significant ($F(3,63) = 7.83$, $p < .001$) and further analyzed by means of pairwise comparisons. When no cue was presented ($90.8\% \pm 1.3$) performance was less accurate than with an early cue at -400ms ($95.7\% \pm .6$, $t(21) = -3.27$, $p_B = .022$) and in comparison with an intermediate cue at -250ms ($94.3\% \pm .6$, $t(21) = -2.83$, $p_B = .049$). There was a trend that indicated that performance without a cue was also less accurate than with a late cue at -100ms ($90.8\% \pm 1.3$, $t(21) = -2.67$, $p_B = .078$).

The two-way interaction between Reference Frame \times Cue was significant ($F(3,63) = 3.65$, $p = .017$) as well as the three-way interaction between Order \times Reference Frame \times Cue ($F(3,63) = 5.72$, $p = .002$). Pairwise comparisons for the three-way interaction showed that in SP during the late cue at -100ms condition performance was better for I judgments than for AR judgments (see Table 1 for details). For the other cue conditions there was no difference between I judgments and

AR judgments. The same pattern was found during PS for the late cue condition and no cue condition (see Table 1 for details). For the early and intermediate cue condition there was no difference between I judgments and AR judgments.

Table 1 Performance for different rrders and reference frames for each cueing condition (N = 22)

Cue	AR judgments (%)	I judgments (%)	<i>t</i> -value	p_B -value
SP				
-400	96.0 ± 1.1	96.0 ± 1.0	0.00	1.000
-250	93.8 ± 1.4	93.4 ± 1.0	0.22	.835
-100	89.6 ± 1.9	96.8 ± 0.8	-3.79	.001 *
none	89.3 ± 1.9	89.1 ± 2.5	0.21	.847
PS				
-400	94.7 ± 0.9	96.2 ± 1.1	-1.15	.247
-250	94.5 ± 1.1	95.5 ± 1.0	-0.75	.448
-100	93.6 ± 0.9	96.2 ± 0.7	-2.25	.040 *
none	89.0 ± 2.3	95.6 ± 1.0	-2.87	.010 *

* $p_B < .05$

Reaction time

Mean correct reaction times were analyzed with a repeated measures ANOVA and revealed a main effect of Preposition, Reference Frame and Cue. Mean reaction times were significantly faster for *above/below* judgments (975.4ms ± 45.3) compared to *left/right* judgments (1098.1ms ± 57.7, $F(1,21) = 24.27$, $p < .001$). Overall, I judgments (1020.9ms ± 52.1) were given faster than AR judgments (1052.7ms ± 49.6, $F(1,21) = 4.69$, $p = .042$). Furthermore, there were significant differences between the four different types of cues ($F(3,63) = 47.16$, $p < .001$). Pairwise comparisons revealed that early cues facilitated the response, but there was no difference between a late cue at -100ms and no cue at all (-400ms cue (935.4ms ± 53.7) < -250ms cue (1004.1ms ± 49.8, $t(21) = -5.08$, $p_B < .001$) < -100ms cue (1105.5ms ± 52.9, $t(21) = -8.33$, $p_B < .001$) = no cue (1102.0ms ± 49.1, $t(21) < 1.00$, $p_B = 1.000$)).

Table 2 Pairwise comparisons on mean reaction times for the Order × Cue and Order × Reference Frame interactions (N = 22)

Contrast	SP		PS	
	<i>t</i> -value	p_B -value	<i>t</i> -value	p_B -value
<i>Order × Cue</i>				
Early – Intermediate cue	-5.09	.000 *	-3.08	.034 *
Intermediate – Late cue	-7.18	.000 *	-5.39	.000 *
Late – No cue	-0.73	1.000	0.92	1.000
<i>Order × Reference Frame</i>				
AR – I	-2.10	.048 *	4.83	.000 *

* $p_B < .05$

The three-way interaction of interest between Order \times Reference Frame \times Cue was not significant ($F(3,63) = 1.61, p = .190$). However, the two-way interactions between Order \times Cue and Order \times Reference Frame were significant ($F(3,63) = 6.85, p < .001$ and $F(1,21) = 21.51, p < .001$ respectively). Pairwise comparisons revealed that the main effect of Cue also drove the interaction indicating that early cueing facilitated responses, see Table 2. The second significant two-way interaction explained the pattern between AR and I responses for SP and PS. Pairwise comparisons revealed that in SP participants always responded faster to AR judgments ($1031.7\text{ms} \pm 55.0$) than to I judgments ($1077.7\text{ms} \pm 51.7$, see Table 2 for details). Interestingly in PS this pattern was reversed, participants now responded faster to I judgments ($964.0\text{ms} \pm 58.1$) compared to AR judgments ($1073.6\text{ms} \pm 47.7$, see Table 2 for details). However, the difference between AR and I judgments seemed more pronounced in PS than SP. In order to test this suggestion post-hoc pairwise comparisons were performed. Indeed the difference between AR and I judgments in the early cueing condition

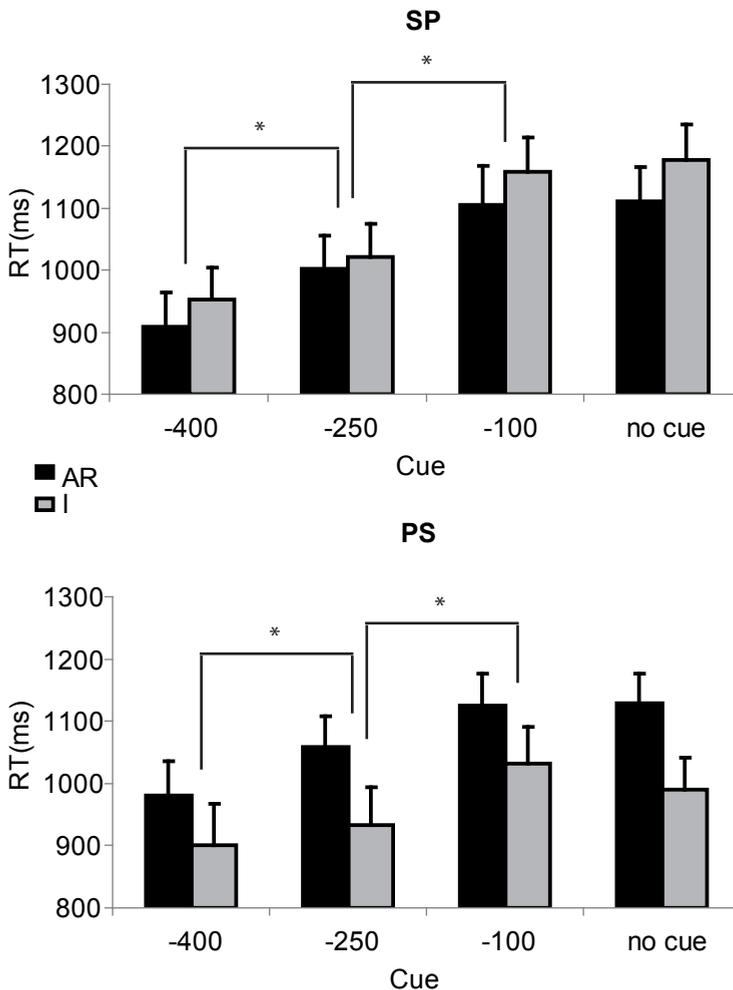


Figure 4 Mean response times (ms) for the effect of cueing different reference frames for SP and PS ($N = 22, \pm \text{SEM}$)

during SP diminished ($t(21) = -1.99, p = .060$) while it remained during PS ($t(21) = 2.67, p = .014$). Although there are different patterns for AR and I judgments during SP and PS, the effect of cueing did not differ for both orders, see Figure 4, as demonstrated by the non-significant three-way interaction.

Finally the two-way interaction between Preposition \times Cue was significant ($F(3,63) = 3.86, p = .013$) as well as the three-way interaction between Order \times Preposition \times Reference Frame ($F(1,21) = 5.01, p = .036$). However, these are ordinal interactions which all followed the pattern of the main effect of Preposition in that responses were faster on *above/below* in comparison with *left/right* (all $p_{Bs} < .010$), although the difference can be more pronounced in some cells yielding the interaction. However, for the current experiment the effect of Preposition is most relevant.

Discussion

The aim of Experiment 2 was to shed light on the level of voluntary control on reference frame activation. The results from Experiment 1 indicated that participants were able to prepare several processing steps after seeing a picture. Which implies that presentation of an early cue may facilitate reference frame activation and perhaps even selection. In contrast, when participants first read a sentence they waited until the picture was presented to start reference frame processes, which suggests that in the SP condition participants might not be susceptible to cueing. Alternatively, the results from Experiment 1 also showed that the order of processing steps from the Carlson-Radvansky and Logan model (1997) do not have to follow a fixed order. Perhaps when attention is focused on a specific reference frame in SP participants might be able to activate that reference frame prior to the presentation of the picture.

The results from Experiment 2 revealed that early cueing facilitated responses in both SP and PS. In SP cueing might have enhanced the activation of reference frames, which suggests that this processing step could be subject to voluntary control. Overall responses were still faster on AR judgments, compared to I judgments, however, the difference seemed to become smaller with early cueing which indicates flexibility in activating reference frames.

In PS, cueing facilitated responses, as was hypothesized. There was now a striking difference between the two reference frames. Participants were much faster on I judgments, compared to AR judgments. Based on their self-reports we suspect that people might have adopted a strategy focusing on intrinsic reference frame processing. This has also been suggested by Taylor et al. (1999) who conducted an EEG experiment with a picture-sentence paradigm. They first presented the picture for 2000ms, followed by a fixation cross for 1000ms and then the sentence was presented word for word, each word lasting 300ms. The final word was either the intrinsic object (e.g. 'car') or referring to the observer ('you'). Taylor et al. found, contrary to their hypothesis, a larger N400 elicited by the AR judgments, indicating a semantic conflict. Their explanation was that participants might adopt an intrinsic strategy, which was supported by their P3 effect. The intrinsic judgments showed an earlier P3 indicating an easier decision and suggesting that the absolute/relative judgment was determined from memory. The strategy also indicates that people might be expecting to read 'car',

consequently, an unexpected *you* then creates a semantic violation. The present findings support the intrinsic strategy hypothesis. Nevertheless, when a cue dictated the use of the absolute/relative frame, participants could have switched to an alternative response. The question that remains is how strong the influence of strategy is. How flexible is the control participants have over the availability of different reference frames? What happens to the alternative reference frames when a cue dictates use of a specific reference frame will be tested in Experiment 3.

Experiment 3

The main question in Experiment 3 is to what extent one can prepare for a specific reference frame and to what extent the other reference frames are still activated. Or in other words what are the strategic and automatic components during spatial sentence verification. Carlson-Radvansky and Irwin (1994) argued that multiple reference frames are activated and subsequently a single frame is selected even when participants are instructed to use only a single reference frame. For example, participants with the instruction to accept all AR placements were slower on rejecting an object in a perfect I placement than an object in a bad placement according to all reference frames. This competition suggests automatic activation of the alternative reference frame. On the other hand, they also found that participants respond faster to the absolute/relative frame than to the intrinsic frame (Carlson-Radvansky & Irwin, 1994), which can be explained as baseline preferences. Perhaps the absolute/relative frame is the dominant or default frame for evaluating the spatial term *above*.

In the literature there is no consensus on the dominant reference frame. For example, Taylor and colleagues found that participants rarely use the relative frame for *front*, *left* and *right*, and even with explicit instruction participants favored the intrinsic frame (Taylor et al., 2001; Taylor et al., 1999). Others state that the relative frame is most economical since it is in line with how we perceive the world (Levelt, 1984; Linde & Labov, 1975). We would argue that these different outcomes are a consequence of the fact that preferences are dynamic and determined by the spatial terms used and the order of the stimuli. The results from Taylor and colleagues where participants preferred the intrinsic frame over the relative frame could have been due to the choice of the spatial term *in front*. The term *in front* is associated with the side of objects facing the viewer, especially when movement is involved (Levelt, 1984). This bodily asymmetry can also benefit *above* and *below* relations, which are also consistent with gravity, perhaps resulting in a more natural use of the relative frame. On the other hand, *left* and *right* do not benefit from either bodily asymmetries or gravity (Logan, 1995; Miller & Johnson-Laird, 1976). Therefore, Experiment 3 combines SP and PS and used four different spatial terms, similar to Experiments 1 and 2 (*above*, *below*, *left*, *right*). Moreover, in order to determine whether reference frame activation is strategic rather than automatic the expectancy of the cue is manipulated. This paradigm was adopted from Noordzij et al. (2005). Prior to each block of trials participants were informed on the proportion of AR and I cues. For example in one condition the probability of AR cues was 80% and 20% for the I cue. If reference frame activation is automatic, then participants will be equally fast on unexpected trials and expected trials. However, if participants adopt a strategy to take advantage from the prior information and focus on the

predominant reference frame then there will be a switch cost and responses will be slower in the unexpected trials.

The current experiment uses *above*, *below*, *left* and *right*. Based on previous findings from Carlson and colleagues on *above* and the results from Experiment 1 the absolute/relative frame seems default in SP. As such, we predict that an unexpected AR judgment will be faster than an unexpected I judgment. In PS, however, there is no clear default reference frame. It seems that participants are better able to use different strategies, which predicts no differences between unexpected AR judgments and unexpected I judgments.

We include three different expectancy conditions where the percentage AR and I cues is varied (80-20, 50-50 and 20-80% respectively). The 80% AR judgment and 80% I judgments conditions create situations in which participants can adopt a strategy focusing on the dominant cue. In the 50-50% condition there is no advantage for either one of the frames, since the prevalence is equal. This condition is comparable to the early cue condition in Experiment 2. Any baseline differences between AR and I judgment will become apparent in this condition. The differences between AR and I judgments for different orders of stimulus presentation and cue expectancy can be tested with the three-way interaction between Order, Reference Frame and Expectancy. Furthermore, in order to test whether there are similar preferences for a default frame for the different spatial terms we will also look at the three-way interaction between Preposition, Reference Frame and Expectancy.

Methods

Participants

17 participants participated (13 right-handed participants, 11 male), mean age 24.76 ± 2.3 years old. All participants were native Dutch speaking students from Utrecht University and were paid €18,- or received course credits for participation. All participants gave informed consent and were naïve with respect to the purpose of the experiment.

Task design

Two sequentially presented stimuli comprised a trial: a sentence followed by a picture (SP) or a picture followed by a sentence (PS). The pictures, cues and sentences were the same as used in Experiment 2. In this experiment we presented a cue for 50ms between the first and second stimulus at 100ms. In this experiment we used the same unique pictures for each spatial sentence as in Experiment 2. We focused on the spatial sentences and added four control trials for each sentence in which both reference frames were aligned, but a different located object was used. Moreover, in this experiment the expectancy of the cues was manipulated. For each order (SP or PS) there were three blocks in which the expectancy of the cue was varied. The three blocks consisted of 20% AR cues – 80% I cues, 50% AR – 50% I and 80% AR – 20% I. In total there were 400 trials ((24 trials \times 4 spatial sentences + 16 trials \times 4 control sentences) \times 10) for each block and six blocks in total (SP/PS \times 20-80/50-50/80-20). In 50% of the trials a yes-response was required.

Procedure

Participants were tested for two sessions of approximately one and a half hour. The experiment consisted of two types of verification tasks: SP and PS. Each order (SP or PS) was tested with a fixed delay of 500ms and a cue presented at 100ms within the delay. The expectancy of the cue was manipulated, which yielded six blocks (SP/PS \times 20-80/50-50/80-20). Each session consisted of a single order with the three expectancy manipulations. Order presentation and expectancy were counterbalanced over participants.

Similar to experiments 1 and 2 participants were first explained what the possible interpretations could be. Subsequently they were instructed about the different options for the cue. Their task was to judge as quickly and accurately as possible whether the sentence and the picture described the same situation according to the cued interpretation. They were given 16 practice trials for each order (SP and PS) in which the cue expectancy was 50%AR – 50%I. Subsequently they started with the three different blocks for a fixed order. Prior to the start of each block participants were informed about the expectancy of the cue. There were three breaks during each block where participants could continue by pressing a button.

A trial consisted of a fixation cross (500ms), the first stimulus (1000ms either a sentence or a picture), during a 500ms delay a cue was presented for 50ms at 100ms, and a second stimulus (4000ms during which the response was given). The response was given on the keyboard. Participants pressed the left button (f) for “yes” and the right button (j) for “no”. After the practice trials the participants received feedback from the experimenter. During the six blocks no feedback was given. After each block participants took a short break and the second session measuring the other order was tested during a separate event.

Data analysis

Overall performance was analyzed, after which the yes responses were analyzed in detail, because they provided affirmative information on the use of different reference frames. The response time analysis was performed on correct yes responses, which meant that participants had responded within the 4000ms response interval. There were four yes response options, two were dictated by AR cues and two were dictated by I cues. For each cue type there were two alternatives: responses belonging to a canonical oriented reference object and responses belonging to a rotated reference object (see Figure 5).

Behavioral data was analyzed with a $2 \times 2 \times 4 \times 3$ repeated measures ANOVA. The within subject variables were: Order (SP, PS), Preposition (*left/right, above/below*), Reference Frame (AR, AR-canonical, I, I-canonical, see Figure 5) and Cue-expectancy (20-80/50-50/80-20). There were four different reference frames. In the AR condition there was a cue indicating to use the AR reference frame. The correct response was yes only according to the relative reference frame, see Figure 5 row 1. In AR-canonical the cue also indicated the use of a AR reference frame, however the correct response was yes according to both the AR and the I reference frame, see Figure 5 row 2. In the I condition the correct yes response was according to the intrinsic reference frame, see

Figure 5 row 3. While in I-canonical the cue indicated the use of the I reference frame, but the yes response was correct according to both the AR and the I reference frame, see Figure 5 row 4. The reported results include effects with a significance level of $p < .05$. Further analysis by means of pairwise comparisons used a significance level corrected for multiple comparisons with the Bonferroni method. The Bonferroni corrected p -values will be labeled p_B .

Sentence	Cue	Picture	Reference Frame
"ball right of car"	AR		AR
"ball right of car"	AR		AR-canonical
"ball right of car"	I		I
"ball right of car"	I		I-canonical

Figure 5 Examples of response options for yes responses, AR: absolute/relative, I: intrinsic.

Results

Performance

Overall performance was very high for SP ($93.8\% \pm 1.0$) and PS (93.9 ± 1.1) and did not differ between the two orders ($t(16) = -.29, p = .815$). The rest of the analysis was performed on correct yes responses. A 2 (Order) \times 2 (Preposition) \times 4 (Reference Frame) \times 3 (Expectancy) repeated measures ANOVA revealed main effects of Preposition, Reference Frame and Expectancy. Similar to Experiment 1 and 2 performance on *above/below* ($93.6\% \pm 1.2$) was better than performance on *left/right* ($91.6\% \pm 1.2, F(1,16) = 11.37, p = .004$). The main effect of Reference Frame ($F(3,48) = 39.23, p < .001$) was further analyzed by means of pairwise comparisons and was driven by the difference in performance between single frame answers and ARI conditions. Performance on the single frame answers did not differ between AR judgments ($89.7\% \pm 1.6$) and I judgments ($86.9\% \pm 1.9, t(16) = 3.00, p_B = .066$). Nor did performance for AR-canonical ($97.0\% \pm .9$) differ from I-canonical ($96.8\% \pm .7, t(16) = .25, p_B = 1.000$). However, performance on ARI responses was significantly better than performance on single frame responses (AR-canonical - AR: $t(16) = 6.08, p_B < .001$ and I-canonical - I: $t(16) = 6.60, p_B < .001$). There was also a main effect of Expectancy ($F(2,32) = 5.97, p = .006$). Performance was better in the 50% Relative cueing condition ($93.5\% \pm 1.6$) compared to the 80% Relative cueing condition ($91.0 \pm 1.6, t(16) = 3.13, p_B = .013$) but did not differ from performance in the 20% Relative cueing condition ($93.4 \pm 1.0, t(16) < 1, p_B = 1.000$).

The two-way interaction between Preposition \times Reference Frame was significant ($F(3,48) = 3.76, p = .017$), but followed the pattern of the main effect of Preposition in that performance was better for *above/below* than for *left/right* ($p_B < .02$). This was true for all reference frames,

except for the AR-canonical frame where performance was equal for the different prepositions ($t(16) = -.34, p_B = .787$). Another two-way interaction between Reference Frame \times Expectancy was significant ($F(6,96) = 20.39, p < .001$). As was the three-way interaction between Order \times Reference Frame \times Expectancy ($F(6,96) = 3.15, p = .007$). The three-way interaction was further analyzed by means of pairwise comparisons.

Table 3 Pairwise comparisons on performance for SP and PS for the Order \times Reference Frame \times Expectancy interaction ($N = 17$)

Contrast	SP		PS		
	<i>t</i> -value	<i>p_B</i> -value	<i>t</i> -value	<i>p_B</i> -value	
<i>20 % Relative cues</i>					
AR – I	-1.96	.369	-2.53	.134	
AR-canonical – I-canonical	-0.44	1.000	0.55	1.000	
AR-canonical – AR	4.86	.001	* 4.48	.002	*
I-canonical – I	4.64	.002	* 3.19	.043	*
<i>50 % Relative cues</i>					
AR – I	0.57	1.000	-0.69	1.000	
AR-canonical – I-canonical	0.64	1.000	-0.30	1.000	
AR-canonical – AR	4.81	.008	* 4.93	.001	*
I-canonical – I	4.30	.003	* 4.69	.001	*
<i>80 % Relative cues</i>					
AR – I	5.35	.000	* 5.66	.000	*
AR-canonical – I-canonical	-0.89	1.000	1.42	1.000	
AR-canonical – AR	2.32	.019	* 5.00	.001	*
I-canonical – I	7.15	.000	* 4.95	.001	*

* $p_B < .05$

The details of the pairwise comparisons for SP are given in Table 3. For SP there was no significant difference in performance between unexpected AR ($85.8\% \pm 2.3$) and I judgments ($90.3\% \pm 1.2$, see also Figure 6) when 20% of the cues was Relative. There were also no differences between AR-canonical ($96.0\% \pm 1.3$) and I-canonical ($96.8\% \pm 1.1$). But performance on the canonical responses were significantly better than the single frame responses. The same pattern was found in the condition with equal Relative and Intrinsic cues. Performance on AR ($89.0\% \pm 2.6$) and I ($87.7\% \pm 2.5$) were not different from each other, but performance was better for AR-canonical ($97.1\% \pm 1.1$) and I-canonical ($96.3\% \pm 1.6$), which did not differ from each other. In the condition where 80% of the cues was Relative, participants performed significantly better on expected AR responses ($91.5\% \pm 2.0$) compared to unexpected I responses ($82.4\% \pm 2.4$). And again performance was better, but did not differ from each other, for AR-canonical ($95.9\% \pm 1.1$) and I-canonical responses ($97.7\% \pm .9$).

The details of the pairwise comparisons for PS are given in Table 3. The pattern observed in PS was similar as observed in SP. In the 20% Relative cueing condition there was a trend indicating

that performance on unexpected AR judgments ($88.2\% \pm 2.1$) was worse than performance on expected I judgments ($93.1\% \pm 1.9$, see also Figure 6). In the ARI conditions there was no difference between AR-canonical ($98.5\% \pm .9$) and I-canonical ($98.2\% \pm .8$), but performance on these condition was significantly better than on the single reference frame conditions. In the 50% Relative expectancy cueing condition there were no significant performance differences between AR ($90.6\% \pm 1.5$) and I ($91.7\% \pm 1.9$) nor between AR-canonical ($97.5\% \pm .6$) and I-canonical ($97.8\% \pm .8$), but performance on the latter was significantly better than performance on the first two. In the 80% Relative cueing condition performance on expected AR responses ($92.9\% \pm 1.7$) was significantly better than performance on unexpected I responses ($76.5\% \pm 3.8$). There was no significant performance difference between AR-canonical ($96.9\% \pm 1.6$) and I-canonical responses ($95.2\% \pm 1.7$), but the responses were more accurate than the single reference frame responses.

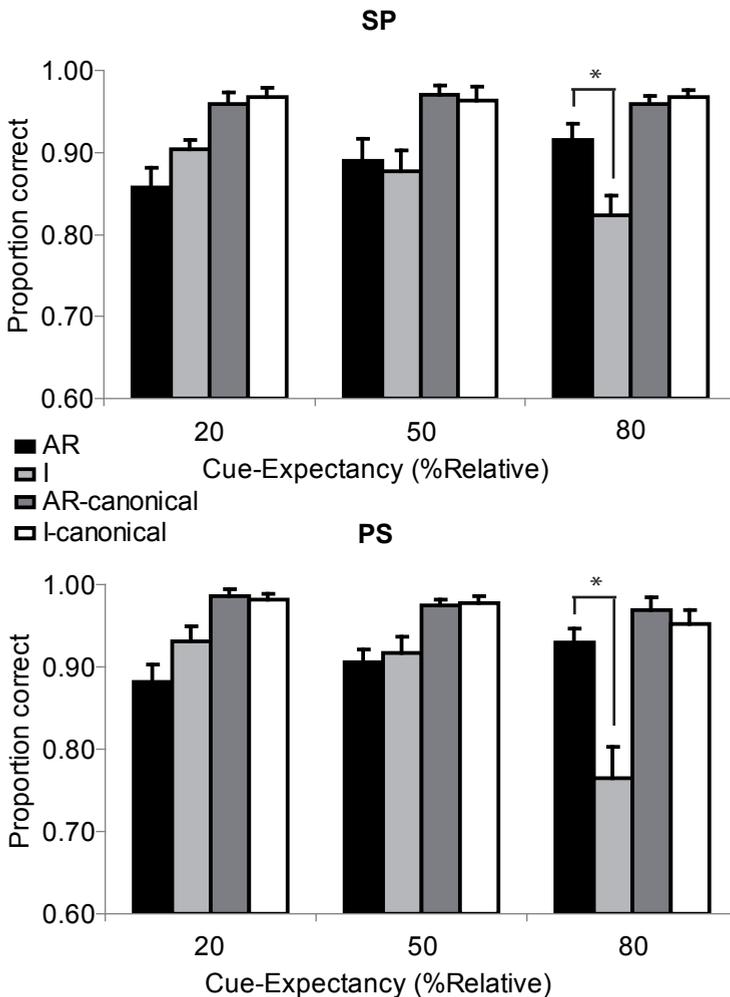


Figure 6 Mean performance for SP and PS for different reference frames and cue expectancies ($N = 17$, \pm SEM)

Reaction time

The behavioral analysis of the mean correct reaction times revealed main effects of Preposition, and Reference Frame. Consistent with the findings in Experiment 1 and 2 participants were again faster on *above/below* judgments (744ms \pm 49) compared to *left/right* judgments (838ms \pm 57, $F(1,16) = 23.27, p < .001$). The main effect of Reference frame ($F(3,48) = 45.33, p < .001$) was driven by the difference between single frame answers and the ARI conditions. Pairwise comparisons revealed that overall there were no significant differences between AR and I responses for the single frame (AR: 848ms \pm 57, I: 902ms \pm 59, $t(16) = -2.57, p_B = .132$) and ARI conditions (AR-canonical: 719ms \pm 52, I-canonical: 695ms \pm 48, $t(16) = 1.41, p_B = 1.000$), but participants responded faster in ARI condition in comparison with single frame conditions (AR-canonical - AR: $t(16) = -8.06, p_B < .001$ and I-canonical - I: $t(16) = -9.90, p_B < .001$).

The two-way interaction between Preposition \times Reference Frame ($F(3,48) = 12.32, p < .001$) and the three-way interaction between Preposition \times Reference Frame \times Expectancy ($F(6,96) = 4.81, p < .001$) were significant and followed the main effect of Preposition. For almost all cells responses were faster for *above/below* than *left/right* ($p_B < .036$). Only in the 20% Relative cueing condition for AR judgments there was a trend in the same direction (*above/below*: 848ms \pm 50, *left/right*: 915ms \pm 65, $t(16) = -2.09, p_B = .056$). This pattern where *above/below* responses were faster than *left/right* responses did not differ for the different reference frames.

Table 4 Pairwise comparisons on reaction times for SP and PS for the Order \times Reference Frame \times Expectancy interaction (N =17)

Contrast	SP		PS			
	<i>t</i> -value	<i>p_B</i> -value	<i>t</i> -value	<i>p_B</i> -value		
<i>20 % Relative cues</i>						
AR – I	-1.35	1.000	3.97	.000	*	
AR-canonical – I-canonical	1.77	.595	2.05	.081		
AR-canonical – AR	-3.19	.031	*	-5.29	.000	*
I-canonical – I	-12.41	.000	*	-5.24	.001	*
<i>50 % Relative cues</i>						
AR – I	-2.74	.095	1.64	.751		
AR-canonical – I-canonical	2.52	.143	1.83	.518		
AR-canonical – AR	-4.17	.003	*	-7.27	.000	*
I-canonical – I	-10.44	.000	*	-5.42	.000	*
<i>80 % Relative cues</i>						
AR – I	-4.50	.002	*	-4.11	.005	*
AR-canonical – I-canonical	-1.50	.876	-1.76	.587		
AR-canonical – AR	-4.83	.001	*	-7.21	.000	*
I-canonical – I	-6.97	.000	*	-5.67	.000	*

* $p_B < .05$

The two-way interaction between Order \times Reference Frame was significant ($F(3.48) = 6.95$, $p = .001$) as well as the two-way interaction between Reference Frame \times Expectancy ($F(6.96) = 14.27$, $p < .001$). These two-way interactions were further specified by the final three-way interaction between Order \times Reference Frame \times Expectancy ($F(6.96) = 5.84$, $p < .001$), which was the interaction of main interest. The details of the pairwise comparisons for SP are given in Table 4. For SP we found a similar pattern as for the performance, where higher performance was associated with faster responses. When 20% of the cues was Relative responses were equally fast for unexpected AR ($826\text{ms} \pm 73$) and expected I judgments ($877\text{ms} \pm 61$, see also Figure 7). Responses were also equally fast for AR-canonical ($728\text{ms} \pm 70$) and I-canonical ($666\text{ms} \pm 56$). But the canonical responses were significantly faster than the single frame responses. The same pattern was found in the 50% Relative and 50% Intrinsic cueing condition. AR ($824\text{ms} \pm 75$) and I ($917\text{ms} \pm 69$) were

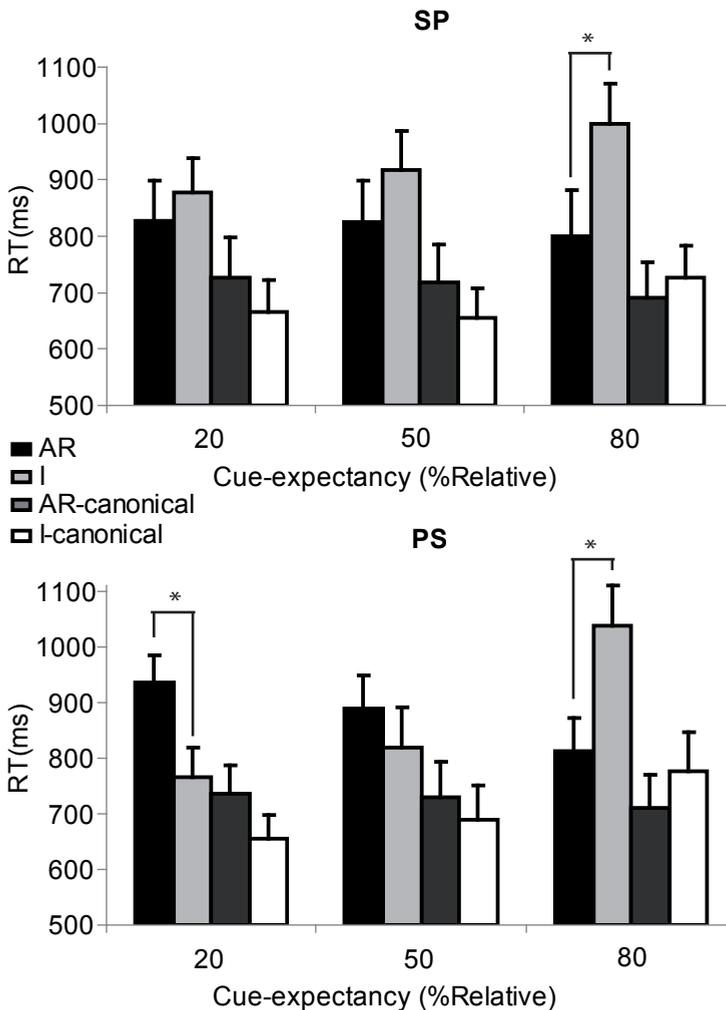


Figure 7 Mean response times (ms) for the effect of expectancy for different reference frames for SP and PS ($N = 17$, \pm SEM)

equally fast, but slower than AR-canonical ($719\text{ms} \pm 67$) and I-canonical ($656\text{ms} \pm 53$), which did not differ from each other. A different relationship was found in the condition where 80% of the cues was Relative. Now participants were significantly faster on expected AR responses ($801\text{ms} \pm 81$) compared to unexpected I responses ($998\text{ms} \pm 73$), as was hypothesized.

The details of the pairwise comparisons for PS are given in Table 4. Similar to the performance results in PS participants adopted a different strategy, depending on the Cue-expectancy. In the 20% Relative cueing condition we now found that unexpected AR judgments ($936\text{ms} \pm 49$) were significantly slower than expected I judgments ($765\text{ms} \pm 54$, see also Figure 7). In the canonical conditions there was no difference between AR-canonical ($735\text{ms} \pm 52$) and I-canonical ($656\text{ms} \pm 41$), but they were significantly faster than the single reference frame responses. In the 50% Relative expectancy cueing condition there responses were equally fast for AR ($890\text{ms} \pm 59$) and I ($818\text{ms} \pm 73$) and for AR-canonical ($730\text{ms} \pm 62$) and I-canonical ($688\text{ms} \pm 63$), but the latter were significantly faster than the first two. In the 80% Relative cueing condition expected AR responses ($812\text{ms} \pm 60$) were significantly faster than unexpected I responses ($1038\text{ms} \pm 73$). There was no significant difference between AR-canonical ($711\text{ms} \pm 59$) and I-canonical responses ($776\text{ms} \pm 70$), but the responses were faster than the single reference frame responses.

Discussion

In Experiment 3 we set out to examine whether reference frame activation is strategic or automatic. In SP for the 50% Relative cueing condition we found no difference between AR and I judgments, which implies that there were no baseline switch costs for changing reference frames. In PS we also found this pattern, in contrast to Experiment 2. However, since there were different blocks with different prevalence of cues there was no longer an advantage of strongly focusing on the I judgments in the 50% condition and participants arguably have adjusted their strategy.

In SP we found clear evidence from performance and response times that the relative reference frame was the default frame and was automatically activated. Performance on AR judgments was better and was equally fast regardless of the cueing condition, which could be interpreted as participants always activating the AR frame during SP. On the other hand, the I judgments did vary with the different cueing conditions, responses were significantly slower and less accurate when unexpected compared to when expected, which indicates that the activation of the I frame was strategic, rather than automatic.

For PS there was a different pattern. When pictures were presented first, strategies seem to play an important role in all cueing conditions. The strategic activation of reference frames became particularly clear in the 20% and 80% cueing conditions. The predominant reference frame should always have yielded the fastest response while switching to the other reference frame should have produced a cost. This cost also became clear from the performance results in the 80% cueing condition, where participants were less accurate on the unexpected I judgments. This suggests that participants exert a certain level of voluntary control and could have strategically activated either type of reference frame in PS.

The pattern of results in SP could have been instantiated by two different mechanisms. First the activation of the intrinsic reference frame is strategic while the absolute/relative reference frames are always activated by default. Second, in line with the Multiple Frame Hypothesis (Carlson-Radvansky & Irwin, 1994) all frames are initially activated, but the weighting of the frames is strategic for intrinsic. After reference frames are activated a selection is made and the spatial templates of all reference frames are weighted according to the selection. In this possibility the absolute/relative frames would always receive a relatively high weight while the weight of the intrinsic frame varies with the expectancy of the cue. Similarly, for PS the strategic results could have been due to these two different possibilities. The current paradigm cannot distinguish between these two possibilities and further research is needed.

General discussion

The aim of the present study was to determine the level of preparation and control in spatial sentence processing. The elegant model by Carlson-Radvansky and Logan (1997) proposes six processing steps that are required when using spatial language. However, the model is based solely on a situation where the linguistic and perceptual information are both available. The central aim of the current study was to examine what happens when visual and linguistic information is asynchronous. Are we forced to wait with mapping linguistic and perceptual components until we can directly perceive the visual scene, while keeping the sentence in short-term memory, or can we prepare some of the processing steps anticipating the visual input.

The first experiment was designed to show which type of information was necessary in order to initiate the processing steps described in the model by Carlson-Radvansky and Logan. The results from Experiment 1 may suggest that participants did not prepare for seeing a picture after having heard a sentence, since there was no effect of manipulating the delay. The visual uncertainty that followed the presentation of the sentence did not initiate processes to reduce this uncertainty. However, after having seen a picture shortening the delay interfered with the response times. This could infer that participants were already preparing while awaiting the sentence to reduce the verbal uncertainty. A shorter preparation time resulted in slower responses. Although the model by Carlson-Radvansky and Logan was based on reading the sentence and linking that information to a perceptual representation we can argue that participants can prepare the first three processing steps after having seen a picture. Since the experiment consisted only of sentences involving a car and a ball they were able to spatially index the objects and determine the reference object and located object. Moreover, multiple reference frames could have been imposed and their associated spatial templates could have been constructed without the exact verbal input. The next step in the model is the selection of a reference frame. Since the sentence was not yet presented, this step could not have been taken and as such the composite template, which is a combination of the weighted spatial templates, could also not have been constructed. However, we could speculate that participants were able to generate prototypical sentences that would provide a correct answer. Since participants

were instructed to accept situations according to both absolute/relative and intrinsic reference frames they could make an educated guess about the response options when seeing a picture. A car in its canonical orientation would yield only one prototypical sentence which would yield a yes response. However, when the car was rotated there would be two alternative response options which would yield a yes response.

The first three steps in the Carlson-Radvansky and Logan model seemed to apply irrespective of the order of the stimuli. The difference between the two orders of stimulus presentation arose at the level of reference frame selection and the other processing steps could have been different for SP and PS. This interaction could be interpreted as a level of control and flexibility, which was further investigated in Experiments 2 and 3.

In the second experiment we forced participants to use a specific reference frame. The research question was whether participants were able to prepare the first three steps when they are forced to. If cueing did not have an effect then the processing steps are part of a rigid process and cannot be controlled. Alternatively, benefits from cueing indicates flexibility and a certain amount of voluntary control over the processing steps and further supports the interaction between information type available and the set of processing steps that can be initiated. Moreover, by manipulating the time point at which the cue for a specific reference frame was given we could get an indication about the level of voluntary control that participants could have had on reference frame activation. Clearly, the results from Experiment 2 indicate that cueing worked in both SP and PS. Cues presented early during the delay facilitate responses even stronger than cues presented in the middle of the delay.

Interestingly, in SP the difference in response times between AR and I judgments seemed to be diminished during the early cue condition. This effect could have been the result of two different processes. Either the cue facilitated I judgments, compared to AR judgments, or the cue hindered AR judgments. Although the second experiment involved processing a cue as well as processing the verbal and perceptual information the results on the performance could give an indication about which process occurred. In Experiment 1, during SP, performance was worse for I judgments ($86.3\% \pm 1.9$) compared to AR judgments ($92.1\% \pm 1.7$). When a cue was introduced in Experiment 2 overall performance for I judgments improved and specifically for the early cueing condition in SP performance was the same for I judgments ($96.0\% \pm 1.1$) and AR judgments ($96.0\% \pm 1.0$). These findings suggested that especially the intrinsic reference frame benefited from cueing, boosting both performance and reaction time.

According to the Multiple Frame Hypothesis (Carlson-Radvansky & Logan, 1997) all reference frames are initially activated and the irrelevant frames have to be inhibited subsequently, ultimately selecting the reference frame(s) to determine the response (Carlson-Radvansky & Jiang, 1998; Carlson & Van Deman, 2008). The Multiple Frame Hypothesis is true for SP, but what happens in PS? Moreover, another interesting question that remained was what happened to the alternative reference frames when a cue dictated the use of a specific reference frame. Or in other words, what are the strategic and automatic components during spatial sentence verification?

Experiment 3 again supported the importance of the context, since different patterns were observed for SP and PS. However, there were two possible mechanisms that can explain the results. The strategic choices could influence the activation of reference frames, or the selection process affecting the weighting of the reference frames and the composite template. When participants first read a sentence, regardless of the mechanism at hand, the results imply that the absolute/relative reference frames were always available, regardless of the expectancy of the cues. There were no switch costs for unexpected AR judgments. However, the intrinsic reference frame seems subject to strategy. Only when the majority of the cues was intrinsic participants benefited from this prior knowledge. On unexpected I judgments, participants were significantly slower, which suggests that this reference frame was not yet available. The current study cannot distinguish between the strategic activation or selection model. In order to distinguish between these two models reaction time and performance might not be the best measure. Online electro-encephalogram (EEG) measurements might provide a better dependent variable to help dissociate between these two options (see also Carlson et al., 2002).

In PS we found strong evidence for strategic availability of reference frames. In both conditions where one of the cue types was prevalent the unexpected cue yielded significantly slower responses. This finding could indicate that a specific reference frame might be preselected resulting in a single response option, but this is highly susceptible to strategies. The fact that we found different patterns for SP and PS stresses the different roles that the context plays. In a communicative situation where for example you are giving a route description it seems more natural to use the absolute/relative reference frame. By using this perspective the receiver can later link this information to his or her own position and match the instruction with the environment. This idea is supported by research on different perspectives in giving route descriptions on the basis of a visual map (Hund et al., 2008). Participants used a route perspective, describing the route according to an egocentric or relative reference frame linking objects to the viewer, when the receiver was thought to drive through the city. In a visual situation, for example describing an apartment, participants take the listener on a mental tour, focusing on the relative reference frame (Linde & Labov, 1975). On the other hand, when the receiver was also in front of a map in the study by Hund et al., participants were more inclined to use a survey perspective, describing the relations between objects in an allocentric manner focusing on the intrinsic reference frame, not including the relation between the viewer and the objects. The flexibility might be induced by the presence or absence of visual uncertainty. When objects are in a non-canonical orientation the possibility of different reference frames might be stressed, which might explain the higher level of flexibility that participants showed in PS in Experiment 3.

In conclusion, the current set of experiments have shown that people in a communicative setting can prepare several processing steps, in particular spatial indexing, reference frame activation and spatial template construction, based on the type of information available. The fact that preparation can take place during verbal and visual input provides supporting evidence for a supramodal spatial representation (Struiksma et al., 2009). Both types of input can initiate preparatory steps, in

particular reference frame activation which parses up space. Moreover, participants are able to exert a certain level of voluntary control on reference frame availability based on the context. Given an expected situation participants might adopt their strategy and become more efficient by reducing the number of alternatives. This strategy is adopted mainly during perceptual situations and to a limited extent during verbal situations. During verbal processing the absolute/relative reference frames seem default, but when a high level of perceptual uncertainty is expected participants can change their strategy to focus more on the orientation of the reference object in order to determine the spatial relation with the located object. Nevertheless, participants can easily switch back to their default reference frames.

Chapter 5

Reference frame acceptability in haptics differs for the blind and sighted in the horizontal but not in the vertical plane

M. E. Struiksma, M. L. Noordzij & A. Postma

(under revision)

Abstract

The current study investigated which reference frames are preferred when matching spatial language to the haptic domain. Sighted, low-vision and blind participants were tested on a haptic sentence-verification task where participants had to judge the relation between a ball and a shoe after haptic exploration. Results from the spatial relation “above”, in the vertical plane, showed that multiple reference frames were available after haptic inspection of a configuration. Moreover, the pattern of results was similar for all three groups and resembled patterns found for the sighted on visual sentence-verification tasks. In contrast, when judging the spatial relation “in front”, in the horizontal plane, the blind showed a markedly different response pattern. The sighted and low-vision participants did not show a clear preference for either the absolute/relative or the intrinsic reference frame when these frames were dissociated. The blind, on the other hand, showed a clear preference for the intrinsic reference frame. In the absence of a dominant cue, such as gravity in the vertical plane, the blind might emphasize the functional relationship between the objects due to enhanced experience with haptic exploration of objects.

Introduction

When communicating about the spatial relations between objects it is important to use the same point of view. For example while having tea you might be looking for the sugar and you ask your companion. In response you might receive the instruction: “the sugar is in front of the teapot”. The exact location of the sugar could be different according to your point of view, the point of view of your companion, or the point of view of the teapot. In order to make this communication successful it is essential that you and your companion use the same point of view, also known as reference frame.

Seminal work on understanding spatial descriptions has been done by Carlson and colleagues (Carlson-Radvansky & Irwin, 1994; Carlson-Radvansky & Logan, 1997; Carlson, 1999; Logan, 1994). When a spatial sentence describes a particular scene verbal and perceptual information have to be compared. This information is converted into a spatial mental representation which allows comprehension of the situation (Tversky, 1991) and requires several steps. The reference object (the teapot in the example above) needs to be identified in order to impose the origin of a reference frame, yielding a set of endpoints defining the searching space for the located object (the sugar in the example above, Carlson, 1999).

Reference frames consist of three axes, which parse up space into different directions specifying location, and have several parameters that can be adjusted: e.g. origin, orientation, direction, and distance (Carlson-Radvansky & Logan, 1997; Carlson & Van Deman, 2004). The orientation and direction parameters can be determined by different sources of information, for example: environmental characteristics, such as gravity and cardinal directions; the listener, or observer; or the reference object, yielding the absolute (A), relative (R) or intrinsic (I) reference frames respectively (Carlson, 1999; Levinson, 1996). In a canonical situation these three reference frames are aligned (ARI), for instance when the teapot, in the example above, is facing you and your companion is sitting next to you at the same side of the table. However, when the teapot is rotated and no longer facing you, the intrinsic (I) reference frame is misaligned with the absolute and relative (AR) reference frame.

Reference frames form a core aspect of spatial mental representations, which are of key importance to human cognition. Representing spatial location is essential for perception in order to localize objects and track them over time, eventually allowing efficient actions on stable and moving objects. Research on reference frames has focused mainly on comparing sentences to visual scenes (Carlson-Radvansky & Jiang, 1998; Carlson, 1999; Carlson & Van Deman, 2008). Furthermore, in a visual mental rotation task the environmental and retinal reference frames could be dissociated by tilting the head of the subject (McMullen & Jolicoeur, 1990). Evidence from neglect patients showed a double dissociation between body-centered and stimulus-centered reference frames on visual figure discriminative cancellation tasks (Ota et al., 2001).

Surprisingly little work has been done on how spatial language is matched to perception in the nonvisual modalities. How would a particular spatial sentence match with a concurrent haptic

inspection of a configuration? It has been suggested that the initial inputs can be either modal or linguistic, but both types of information eventually converge into a supramodal representation (Röder et al., 2007; Struiksma et al., 2009). The supramodal representation contains the basic spatial information, such as reference frames, however, a connection with the input modalities is maintained. A spatial representation can be accessed from perceptual or verbal input, yielding functional equivalence (Denis & Zimmer, 1992; Loomis et al., 2007; for a review see Struiksma et al., 2009), nevertheless, since modal connections are maintained subtle differences are possible. The functional equivalence might suggest that spatial information, for example reference frames, is also available through haptic input. The aim of the current study was to investigate which reference frames are preferred when matching spatial language to the haptic domain. A group of individuals that is of particular interest here are the blind. Blind people no longer have, or never have had, experience with the visual domain. Instead they rely on hearing and touch to obtain information about the world. In addition, low-vision people still have some access to visual input, but might not be able to use it to its full extent and thus turn to rely on the other sensory modalities. The behavior and preferences of this group could give insight in the importance of the quality of the visual input for reference frame use based on haptic input. The current study tested sighted, low-vision and blind participants on a haptic sentence-verification task where participants had to judge the relation between a ball and a shoe after haptic exploration.

Previous research on reference frames in the blind has suggested that congenitally blind people rely on egocentric encoding since proprioceptive, vestibular, touch and movement information is available with a body-centered reference (Millar, 1994; Röder et al., 2007) and more distal information is reduced and more difficult to process (for reviews see Cattaneo et al., 2008; Thinus-Blanc & Gaunet, 1997). The distinction between egocentric and allocentric reference frames refers to a body-centered coordinate system and an environment-centered coordinate system respectively. However, the term egocentric can also be associated with the relative reference frame explained above (Levinson, 1996). As such, we might expect that blind participants in our haptic sentence-verification task are more inclined to use the relative reference frames than sighted participants. Alternatively, since haptic object handling and search is quite common in blind individuals, a reverse pattern might be observed instead. In line with this, Postma et al (2007) found that blind individuals used more object related spatial language (i.e. intrinsic reference frame) terms than sighted persons whereas the latter produced slightly more board references (i.e. absolute reference frame) when describing the layout of a previously inspected object array.

Preference differences between the reference frames can be observed especially in situations where the reference frames are dissociated. We explored this for two different spatial relations. According to the spatial framework model from Franklin and Tversky (1990) the vertical axis with spatial terms *above/below*, aligned with gravity, is the predominant axis followed by *in front/back*, which has a perceptual and functional asymmetry, and subsequently *left/right*. Previous research in the visual domain has primarily used *above* in the vertical plane (Carlson, 1999; Carlson & Logan, 2001; Logan, 1995; Logan & Compton, 1996), therefore, we first tested this spatial relation. The

critical trials are those in which the shoe is rotated 90° with respect to its canonical orientation. In this case a “yes” response implies the use of a different reference frame than for a “no” response. For example, in Figure 1, depending on the location of the ball a “yes” response to the question “*Is the ball above the shoe?*” indicates the use of the absolute/relative or intrinsic reference frame. In order to stress the different reference frames we also tested *in front*, which is in the horizontal plane. This latter relation is more ambiguous and can be influenced by motion. For example, in a stationary situation *in front* refers to the object closest to an observer, however, when moving *in front* refers to the object farthest away (Alloway et al., 2006). The increased ambiguity might stress the intrinsic reference frame, making it a more plausible response option. This expectation is based on research with visual paradigms. We might speculate that blind people will not show such a switch of interpretation since they rely more on haptic input, even during movement. Hence, *in front* might be more stable and refer to the object closest to the experiencer. Therefore, we expected that blind participants rely stronger on the absolute/relative reference frame, similar to the *above* situation, and that the sighted and low-vision participants are more inclined to use the intrinsic reference frame, compared to the *above* situation.

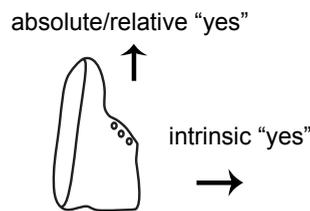


Figure 1 Example of positions that correspond to an absolute/relative and intrinsic “yes” response.

Methods

Participants

Participants were all visitors of a Dutch annual event for low-vision and blind individuals, from two consecutive years. The first group of participants consisted of 343 visitors: 154 sighted controls (SC), 146 (VI) and 43 blind (BL) participants, for details, see Table 1. A Kruskal Wallis test revealed a significant effect of group on age : $\chi^2(2, N = 343) = 13.18, p = .001$. A post-hoc test using Mann-Whitney Pairwise comparisons with Bonferroni correction ($\alpha_B = \alpha / 3 = .017$) showed that the low-vision participants were significantly older than the sighted ($U = 8738.00, N_1 = 146, N_2 = 154, p = .001$, two-tailed) and the blind ($U = 2214.00, N_1 = 146, N_2 = 43, p = .003$, two-tailed). The blind and sighted did not differ from each other in age ($U = 2987.50, N_1 = 43, N_2 = 154, p = .328$, two-tailed). There was also a significant effect of education level: $\chi^2(2, N = 343) = 13.18, p = .001$. Pairwise comparisons revealed that the sighted controls had a higher education level than the low-vision participants ($U = 8689.00, N_1 = 154, N_2 = 146, p < .001$, two-tailed), but there was no difference with the blind group ($U = 2811.00, N_1 = 154, N_2 = 43, p = .105$, two-tailed) or between the low-vision and the blind participants ($U = 2911.00, N_1 = 146, N_2 = 43, p = .445$, two-tailed).

The number of males was lower in the sighted group compared to the other two ($\chi^2(2, N = 343) = 8.288, p = .016$). The second group of participants consisted of 403 visitors: 170 sighted controls (SC), 181 low-vision (LV) and 52 blind (BL) participants, for details, see Table 1. Again there was a significant difference in age ($\chi^2(2, N = 403) = 21.50, p < .001$). The low-vision participants were significantly older than the sighted controls ($U = 12766.00, N_1 = 181, N_2 = 170, p = .006$, two-tailed), who were significantly older than the blind ($U = 3293.50, N_1 = 170, N_2 = 52, p = .005$, two-tailed). Similar to the first group there was a significant difference in education level ($\chi^2(2, N = 403) = 23.10, p < .001$). The sighted controls had a significantly higher education level than the low-vision participants ($U = 11153.50, N_1 = 170, N_2 = 181, p < .001$, two-tailed), but there was no significant difference with the blind group ($U = 4060.00, N_1 = 170, N_2 = 52, p = .343$, two-tailed). Also similar to the first group the number of males was lower in the sighted group compared to the other two ($\chi^2(2, N = 403) = 7.188, p = .027$).

Table 1 Descriptives for all groups for both experiments

Group	N	Male (%)	Age (year \pm SD)	Education (code \pm SD)
<i>Group 1</i>				
SC	154	36.36	47.70 \pm 13.23	5.77 \pm 0.79
LV	146	51.37	54.14 \pm 17.32	5.37 \pm 1.00
BL	43	53.49	45.12 \pm 15.38	5.53 \pm 0.96
<i>Group 2</i>				
SC	170	35.29	50.76 \pm 14.92	5.71 \pm 0.87
LV	181	45.86	55.30 \pm 16.49	5.21 \pm 1.05
BL	52	53.85	44.12 \pm 15.51	5.50 \pm 1.06

Stimulus material

The stimulus set consisted of a small-size baby shoe and a small ball. The stimuli were placed in boxes, which were covered with fabric and placed on one side on a table. All boxes contained identical shoes and balls. In the centre of the fabric an incision was made such that participants could place one hand inside the box to feel the stimuli. For the *above* judgments, the shoe and ball were attached to the far end vertical plane. For the *in front* judgments, the shoe and ball were attached to the bottom of the box. There were different conditions based on the orientation of the shoe and the position of the ball. When the shoe was in the canonical orientation all reference frames aligned, however when the shoe was rotated the absolute/relative and intrinsic frames could be dissociated, see Table 2 for an overview.

Table 2 Example of stimuli and stimulus codes

Stimulus “above”	Stimulus “in front”	Orientation	Reference frame describing “yes”	Code
		Canonical	absolute/relative/intrinsic	cARI
		Canonical	none	cNo
		Canonical	absolute/relative	cAR
		Canonical	intrinsic	cl
		Noncanonical	absolute/relative	ncAR
		Noncanonical	intrinsic	ncI
		Noncanonical	none	ncNo

Design and procedure

Visitors to the stand were asked to place one hand inside the box. They were told that each box contained a small shoe and a ball. Their instruction was to feel both stimuli and to judge whether the sentence “*The ball is X the shoe*” was true, X could be *above* or *in front of* (the Dutch translation is: “*Is de bal boven/voor de schoen?*”). Participants responded by saying: “yes” or “no”¹. For each subject the responses were noted. There were five boxes in the *above* version and seven boxes in the *in front* version. The order of the boxes was randomized and changed every 10 participants to avoid order effects. After the participants had finished the trials they were asked to give a few descriptives: gender, age, information about their visual capabilities and education.

¹ Participants all were visitors of an annual event for low-vision and blind individuals. In a laboratory setting it would have been beneficial to allow participants to rate the description on a graded scale. Since many people participated at the event we have chosen to use only two response alternatives to simplify instructions.

Data analysis

The responses were recoded “no” = 0, or “yes” = 1. The same codes were used for both experiments. Since the number of participants was unequal for the three groups and the level of measurement was nominal nonparametric statistics were appropriate. The main effects of proportion of “yes” responses between different conditions were computed by the McNemar test using a binomial distribution. The between-group effects per condition were analyzed by means of an exact significance test for Pearson’s Chi-Square. The exact significance test was selected because there were cells that had an expected count less than 5, which violated the assumption for the regular Chi-Square test. The group effects were further analyzed by means of pairwise comparisons with the exact significance test for Pearson’s Chi-Square. The significance level was corrected for multiple comparisons with the Bonferroni method.

Results

Above

The proportion “yes” responses for all three groups on the five different conditions in the vertical plane are given in Figure 2. In the cARI condition when all reference frames aligned and the ball was placed above the shoe almost all participants responded “yes”. When the shoe was in the same orientation, but the ball was placed at the heel of the shoe (cNo) most participants responded “no”. When the shoe was rotated 90° to the right in a noncanonical orientation and the ball was placed next to the sole (ncNo) almost all participants responded “no”. The conditions of interest were those where the shoe was rotated and the ball was placed above the shoe according to the subject’s point of view (ncAR), or according to the shoe’s point of view (ncI). In order to verify whether participants had access to multiple reference frames the McNemar test was used to compare cARI with ncAR and ncI with ncNo. In the ncAR condition participants mostly responded with “yes”, although significantly lower as in the cARI condition ($N = 343$, exact $p < .001$). In the ncI condition, participants were more inclined to respond with “no”, although a significant amount of participants responded with “yes”, as indicated by the significant difference from the ncNo condition ($N = 343$, exact $p < .001$). See Figure 2 for an illustration of these effects.

The conditions of interest were further analyzed for group differences by means of an exact significance test for Pearson’s Chi-Square. In the ncAR condition there were no group differences ($\chi^2(2, N = 343) = .590$, exact $p = .742$), however in the ncI condition there were significant group differences ($\chi^2(2, N = 343) = 7.047$, exact $p = .029$). The group differences were further tested with pairwise comparisons, the significance level was corrected for multiple comparisons with the Bonferroni method ($\alpha_b = \alpha / 3 = .017$). The pairwise comparisons revealed a significant difference between sighted and visually impaired. Low-vision participants were more inclined to respond with “no” ($\chi^2(1, N = 300) = 6.412$, exact $p = .013$). There was no significant difference between sighted and blind participants ($\chi^2(1, N = 197) = .001$, exact $p = 1.000$) and low-vision and blind participants

($\chi^2(1, N = 189) = 3.421$, exact $p = .098$). Together these results show a highly similar pattern for all three groups and replicate previous results from visual paradigms.

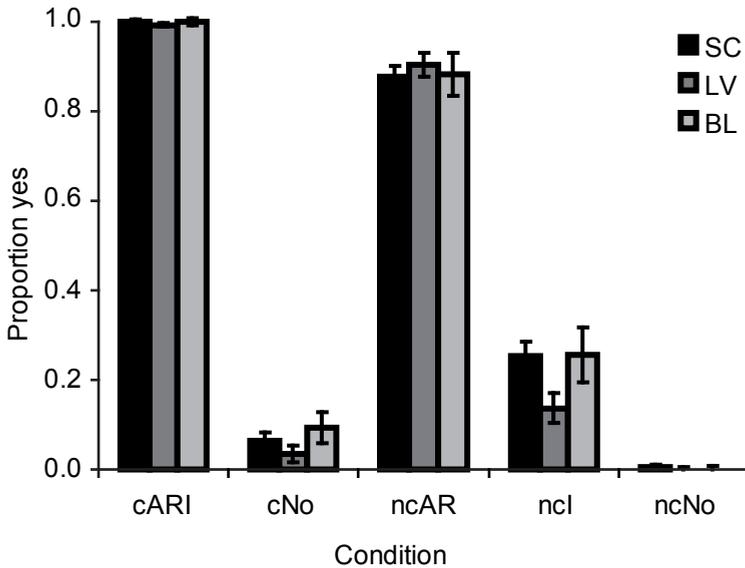


Figure 2 Results for *above* for sighted controls (SC), low-vision (LV) and blind (BL) participants (\pm SEM)

In front

The proportion of “yes” responses for all three groups on the seven different conditions in the horizontal plane are given in Figure 3. In the canonical condition (cARI) when the ball was placed in front of the shoe and a “yes” response was required, almost all participants responded with “yes”. In the canonical situation where a “no” response was required (cNo) most participants also responded with “no”. The same pattern is true for the noncanonical “no” condition (ncNo). The differences arise when the reference frames are in conflict. In order to verify whether participants had access to multiple reference frames the McNemar test was used to compare cARI with ncAR and cAR and ncNo with ncl and cI. In the canonical and noncanonical absolute/relative conditions (cAR and ncAR) the proportion of “yes” responses was significantly lower compared to the cARI condition in which all reference frames were aligned ($N = 403$, exact $ps < .001$). In the canonical and noncanonical intrinsic conditions (cI and ncl), the proportion of “yes” responses was significantly higher compared to the ncNo condition ($N = 403$, exact $ps < .001$). See Figure 3 for an illustration of these effects.

The conditions of interest were further analyzed by means of an exact significance test for Pearson’s Chi-Square. The group differences were further tested with pairwise comparisons, the significance level was corrected for multiple comparisons with the Bonferroni method ($\alpha_B = \alpha / 3 = .017$). When the absolute/relative frames dictated a “yes” response, in both canonical and non-canonical orientations of the shoe, there were group differences (see Table 3 for details, row cAR and ncAR). Approximately half of the sighted and low-vision participants gave a “yes” response,

whereas the blind participants were significantly more inclined to respond with “no”, see Figure 3 and Table 3 for details. In the canonical situation where the intrinsic frame dictated the “yes” response (cI) the proportion of “yes” responses in the blind was significantly higher compared to the sighted and low-vision participants (see Table 3 row cI). In the noncanonical intrinsic situation (ncI) the proportion of “yes” responses was significantly higher for the blind and sighted, compared to the low-vision participants. In the low-vision group approximately half of the participants responded with “yes”, similar to the canonical absolute/relative situation (see Figure 3 and Table 3 row ncI for details). The results for the blind clearly differentiate from the sighted and visually impaired. Blind participants seem to favor the use of the intrinsic reference frame, while sighted and low-vision participants show no clear preference for the intrinsic or the absolute/relative reference frames.

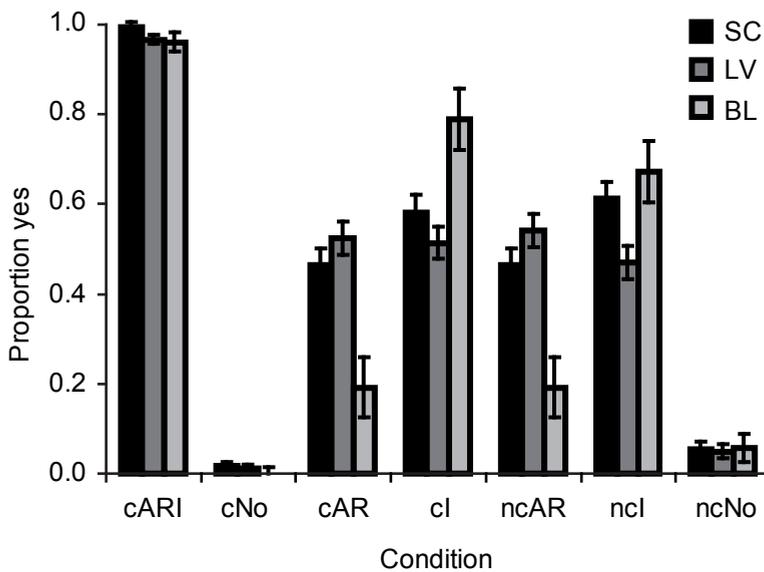


Figure 3 Results for *in front* for sighted controls (SC), low-vision (LV) and blind (BL) participants (\pm SEM)

Table 3 Statistics for the results of *in front*.

	Main effect		SC vs. LV		SC vs. BL		LV vs. BL	
	χ^2	exact p						
	(2, N = 403)		(1, N = 351)		(1, N = 222)		(1, N = 233)	
cAR	18.084	<.001	1.269	.286	12.302	.001	18.045	<.001
ncAR	19.798	<.001	2.065	.166	12.302	.001	19.800	<.001
cI	12.515	.002	1.662	.200	7.262	.008	12.470	<.001
ncI	10.499	.005	7.128	.010	.639	.513	6.695	.012

Columns represent the main effect and pairwise comparisons between different groups (sighted controls (SC), low-vision (LV) and blind (BL)). Rows represent different conditions: cAR = canonical absolute/relative, ncAR = noncanonical absolute/relative, cI = canonical intrinsic, ncI = noncanonical intrinsic.

Above vs. in front

The studies on “above” and “in front” both tested the noncanonical conditions of interest where either the absolute/relative or the intrinsic reference frames dictated the “yes” response (ncAR and ncI respectively). This enabled us to compare both groups on each condition with an exact significance test for Pearson’s Chi-Square. All groups showed a significant decrease in the proportion of absolute/relative responses and consequently an increase in intrinsic responses for “in front” in comparison with “above” (see Table 4 for details on the statistics). This pattern was expected for the sighted and low-vision participants. Surprisingly, the blind participants also showed an increase of intrinsic responses, as was also demonstrated by the results from “in front” alone.

Table 4 Statistics for the results of comparing *above* to *in front*

	SC		LV		BL	
	χ^2 (1, N = 324)	exact p	χ^2 (1, N = 327)	exact p	χ^2 (1, N = 95)	exact p
ncAR	61.140	<.001	50.946	<.001	45.012	<.001
ncI	42.123	<.001	41.016	<.001	16.408	<.001

For the sighted controls (SC), low-vision (LV) and blind (BL) participants. Rows represent different conditions: ncAR = noncanonical absolute/relative, ncI = noncanonical intrinsic.

Discussion

The main goal of the current study was to explore which reference frames are preferred when matching a spatial description to a haptic inspection of a configuration. Specific interest focused on whether people with no or limited visual input show different preferences than sighted controls. Notably, the results from two experiments clearly showed that multiple reference frames are available in these circumstances. Moreover, low-vision and blind participants demonstrated a similar pattern as the sighted in the vertical plane.

The results on *above*, in the vertical plane, are in line with previous results found in the visual domain with sighted participants (Carlson-Radvansky & Irwin, 1993, 1994; Carlson, 1999). In a rating study Carlson-Radvansky & Irwin (1993, 1994) found a clear preference of the absolute/relative reference frames, but there was a significant difference between the canonical condition in which all reference frames were aligned and the noncanonical condition in which the absolute/relative reference frames dictated a “yes” response. This indicated that participants did not resolve to an absolute/relative strategy, using only those reference frames while ignoring the intrinsic reference frame. The induced conflict between the reference frames in the noncanonical absolute/relative reference frame condition resulted in a significant drop in acceptance rate which indicated that the intrinsic reference frame was first available and subsequently discarded. Further proof refuting the absolute/relative strategy came from the difference between the noncanonical intrinsic reference frame condition and the noncanonical No condition, where none of the reference frames dictated a “yes” response. If participants had used such a strategy then there should have been no difference

in accepting the noncanonical intrinsic and noncanonical No. Instead Carlson-Radvansky & Irwin (1993, 1994) found that participants accepted noncanonical intrinsic significantly more often than noncanonical No. The present results from our haptic sentence-verification task also demonstrated these effects. The fact that the difference between the noncanonical No and noncanonical intrinsic condition was significant indicates that the intrinsic reference frame was available, although not the preferred reference frame.

Interestingly, the low-vision and even the blind participants showed the same behavior as the sighted participants, i.e. a clear preference for the absolute/relative reference frame, but also the availability of the intrinsic reference frame. Only in the noncanonical intrinsic condition a difference between the three groups emerged. In this situation the low-vision participants were more inclined to respond with “no” indicating an even stronger preference for the absolute/relative reference frame compared to the sighted and the blind. The idea that blind subjects were more inclined to use the relative reference frame is not supported by these results. Alternatively, it is also not the case that sighted subjects are more inclined to use the relative reference frame. Instead the behavior of blind and sighted is highly similar. The low-vision participants showed a slightly stronger preference for the relative reference frame in the noncanonical intrinsic condition. However, this difference was only marginally significant when compared to the sighted participants and did not differ from the blind participants. Therefore, this minor effect in the low-vision group should not be overemphasized.

With the second spatial relation *in front*, in the horizontal plane, apparent differences between the groups emerged. The low-vision and the sighted participants generally showed the same behavior, while the response pattern from the blind was clearly different. The responses on the control conditions in which all reference frames aligned and indicated either a “yes” or a “no” response were the same for all three groups. This suggests that participants understood the task they were doing. A closer analysis of the conditions of interest in which there was a conflict between the absolute/relative and intrinsic reference frames revealed that for the sighted the clear preference for absolute/relative had diminished in comparison to the findings in the vertical plane. On average approximately 50% of the participants responded according to the absolute/relative reference frames and approximately 50% according to the intrinsic reference frame, while this was approximately 90% and 10% respectively in the vertical plane. As mentioned before, the spatial relation *in front* is more ambiguous as demonstrated by the influence of movement on the interpretation of the relation (Alloway et al., 2006; Miller & Johnson-Laird, 1976). As predicted, this was also demonstrated by the current results where half of the participants have used the intrinsic reference frame. A similar pattern was also found for the visually impaired, implying that the availability of vision has a strong influence on judging the spatial relation.

In contrast, the results from the blind participants showed a markedly different pattern. We hypothesized that for the blind haptic input during movement would provide a more stable interpretation for *in front* according to the absolute/relative reference frames. This would have been reflected by a stronger preference for the absolute/relative reference frames in the blind. Contrary to our

expectations blind participants seemed to prefer the intrinsic reference frame. Instead of using their body-centered coordinate system which would have yielded a bias in favor of the absolute/relative reference frames, blind participants focused on the reference object, in this case the shoe. Evidently the orientation of the shoe was of greater relevance to the blind than to the sighted and visually impaired. We conjecture that the increased experience with haptic object handling in the blind is crucial here. Touch is their primary source of input for information about the world around them and determining the orientation of objects as such is extremely helpful. Directly feeling whether you pick up a shoe by its nose or its heel when you want to put it on offers a great benefit. In line with this idea, various researchers have pointed out that situational context and functional relation between objects also strongly determine which reference frame is most dominant (Carlson-Radvansky et al., 1999; Carlson-Radvansky & Radvansky, 1996; Carlson & Kenny, 2006; Coventry & Garrod, 2004). The functional relation focuses attention on one of the objects, enabling the intrinsic reference frame. On the other hand, the relative reference frame is associated with the relation between the object and the viewer (Carlson-Radvansky & Radvansky, 1996; Miller & Johnson-Laird, 1976). The results from Carlson-Radvansky & Radvansky (1996) corroborate this distinction and showed that the choice of reference frame used to describe a scene was influenced by the presence of a functional relation. For example, in a picture with a mail carrier and a mailbox participants preferred the intrinsic reference frame when the mail carrier was facing the mailbox, as if he was posting a letter. When the mail carrier was facing the opposite direction the functional relation disappeared and participants preferred the relative reference frame.

In the case of our task where participants had to judge whether or not the “*ball was in front of the shoe*” there is also a clear functional relationship. When the ball is positioned in line with the front of the shoe one could imagine kicking the ball away. According to this functional relationship we would then expect participants to prefer the intrinsic reference frame in the canonical condition where all frames overlap, the canonical intrinsic condition and the noncanonical intrinsic condition. In accordance with judging the functional relationship, using the intrinsic reference frame in the other conditions predicts “no” responses for the canonical absolute/relative and noncanonical absolute/relative conditions. This is exactly what we have found for the blind participants for judgments in the horizontal plane. Not only the orientation of the shoe was more important to the blind, but especially the functional relation between the ball and the shoe predicts the pattern of results.

The functional explanation of the results for the horizontal plane does not only apply to the blind, but could equally well hold for the sighted and visually impaired. There is clear evidence of more ambiguity when the spatial relation *in front* needs to be verified, as demonstrated by the larger proportion of intrinsic responses in all groups. We have discussed two possible mechanisms that might explain this increase of intrinsic response: movement induced ambiguity and the functional relation between objects. With the current experiment it is not possible to disentangle the influence of these two mechanisms. However, we speculate that the results from the sighted and low-vision participants are mainly due to the movement induced ambiguity, since in all conditions of interest

the proportion of intrinsic responses was approximately around 50% and no differentiation according to the functional relation was made. On the other hand, the results from the blind follow a pattern that fits more with the functional explanation. They show a clear preference for the intrinsic reference frame with a “yes” response when the ball was in front of the nose of the shoe and a “no” response in the other conditions. In order to fully understand the contribution of the functional relationship between objects further research is needed where the functional relationship is manipulated. If we are correct about the idea that the blind were relying stronger on the functional relationship, compared to the sighted, then in the absence of a functional relationship we would predict that the blind prefer the absolute/relative reference frames. On the other hand, in the sighted and low-vision group we would expect similar results as found for *in front*.

In conclusion, the present study has shown that multiple reference frames are available in sighted, low-vision and blind participants during haptic spatial sentence-verification. These results provide supporting evidence for a supramodal representation of spatial information where reference frames are available for comparison between spatial language and visual, but also haptic situations. Moreover, for judging *above*, in the vertical plane, gravity seems to play a dominant role as demonstrated by the relatively few occasions of using the intrinsic reference frame when it was dissociated from the absolute/relative reference frame. This was true for all three groups, including the blind participants. On the other hand, for the spatial relation *in front*, in the horizontal plane, a different pattern was found. For the sighted and low-vision group due to the absence of a strong cue such as gravity the intrinsic sides of the reference object became an important cue. This became clear from the increased number of intrinsic reference frame responses in the conditions where the reference frames were dissociated. Most notable, the blind participants showed an even stronger intrinsic reference frame preference in this situation. The emphasis on the functional relation between the ball and the shoe may explain this outcome. Evidently, blind participants rely stronger on the details of the reference and located object, which might be due to their enhanced experience with haptic exploration of objects.

Complex spatial descriptions

VIS
ON
LIND
SPACE
URVE
DESC
RIPTI
ON
AN
GE
PH
MO
DAL
LE
FT
RA
GH
CON
GE
NIT
ALL
ROU
SUP
RA
EYE
RNEC
VAL
REN
M
ENT
AL
RE
RESEN
TATI
ON
POIN
TOFW

Chapter 6

Neural correlates of learning route and survey knowledge from complex spatial descriptions

M. E. Struiksma, L. J. Talsma, M. L. Noordzij, S. F. W. Neggers & A. Postma

(in preparation)

Abstract

Behavioral results have shown that descriptions from route or survey perspectives are represented in spatial mental representations. A critical question is whether and to which extent they are equivalent. Therefore, the current study investigated the neural correlates of encoding complex verbal spatial descriptions from different perspectives. We observed a general overlapping network of activation for the route and the survey group associated with the processing of spatial descriptions. A parametric modulation revealed an increase in activation in the caudate nucleus due to the increasing familiarity with the description in both groups as the number of repetitions increased. In addition, distinct hippocampal activation for the survey and route group showed that both groups were able to process allocentric components of the environment. Route and survey descriptions are encoded in a *hierarchical* fashion where allocentric components can also be inferred from egocentric components resulting in a functional equivalent mental representation for both encoding perspectives. Importantly, the current neurocognitive results from verbal complex descriptions overlap to a great extent with previous results from visual paradigms. These findings support the idea that spatial information is represented in a supramodal manner regardless of the input modality.

Introduction

In everyday life we spend a large amount of time navigating through our surroundings. Successful wayfinding as well as adequate memory for objects and their locations, such as the positioning of food and water, are essential to our survival. Throughout evolution, we have developed a range of complex vocabularies to efficiently communicate such environmental information to others. Sharing spatial knowledge by means of spatial language allows us to combine our efforts to find our way in the world.

When understanding a spatial description, information needs to be converted into a spatial mental representation which supports comprehension of the situation (Tversky, 1991). These spatial mental representations can be stored in memory and retrieved again for later use in navigation. Denis and Zimmer (1992) have shown that there is substantial overlap between spatial mental representations built up from visual experience and those derived from spatial descriptions. A large body of research has focused on the visual nature of these spatial mental representations and the different components of visuo-spatial mental imagery (e.g. Farah, 1989; Kosslyn et al., 1990).

A behavioral study by Hund et al. (2008) examined the descriptions participants provided to give an imaginary recipient directions about the lay-out of a city centre. Such a configuration of an area that is either too large or too complex to be overlooked completely from one viewpoint is termed a large-scale environment (Hubona et al., 1998). Hund et al. found that in describing such a large-scale environment participants adequately adapted their descriptions to the need of the recipient. When participants believed the recipient to be truly traveling through the city by car, they were likely to provide directions that included more landmarks and used indications such as *left* and *right*, also referred to as a route description (Taylor & Tversky, 1992). However, if participants believed the recipient to be looking at a map, they would provide descriptions with more cardinal directions, such as *north*, *east*, *south* and *west*, also referred to as a survey description (Taylor & Tversky, 1992). A description from a route perspective relates all spatial information to the recipient using an egocentric frame of reference. In contrast, a verbal description from a survey perspective relates the positions of objects relative to each other, applying an allocentric frame of reference (Taylor & Tversky, 1992).

An interesting issue that has been addressed by Taylor and Tversky (1992) also deals with the nature of the representation. In a series of experiments they contrasted the idea that spatial mental representations are based on visuo-spatial imagery versus the idea that they are like structural descriptions that represent the gist or verbatim record of the text. In order to distinguish between these types of representations they constructed descriptions from different perspectives. Participants had to answer verbatim and inference questions, from the learned and alternative perspective, about the environments. If the construction of a spatial mental representations relied on a structural description then perspective information would be incorporated in the representations yielding different results for different perspectives. In contrast, the results indicated no differences between perspectives, suggesting that spatial information was extracted regardless of perspective.

In line with Denis and Zimmer (1992), Taylor and Tversky (1992) also showed that participants who had studied a map performed similar to participants who had read descriptions of the environment.

Clearly various behavioral studies have shown that subjects are successful in encoding information into a spatial mental model both from a route and a survey description (Brunyé et al., 2008; Brunyé & Taylor, 2008a; Brunyé & Taylor, 2008b; Giudice et al., 2007; Noordzij & Postma, 2005). There exists functional equivalence, i.e. different learning perspectives yield highly similar performance on a number of tasks, after over-learning. These mental representations reflect the spatial properties of the configuration similar to when the representation would be generated from visuo-spatial information (Denis & Zimmer, 1992; Giudice et al., 2007) and include categorical as well as metric information (Denis, 2008; Noordzij & Postma, 2005; Noordzij et al., 2006b; Péruch et al., 2006). Spatial information from different input modalities can be converged into a spatial mental representation. Spatial information is extracted from modal information to form a supramodal representation while maintaining a link with the input modality (Struiksma et al., 2009). This supramodal representation also incorporates verbal information.

Noordzij and Postma. (2005) demonstrated that the spatial mental representation that participants had built up from spatial descriptions also entailed metric spatial information. Although participants who had learned the route description also constructed a spatial mental representation with analogue metric detail, those who had learned the survey description had built up a representation with more fine-grained spatial detail. This finding suggests that, depending on the task, there are qualitative differences between the perspectives used, but both result in a spatial mental representation that is functionally equivalent. This study was repeated with early blind participants (Noordzij et al., 2006b). Interestingly, also the blind participants built up a spatial mental model with analogue spatial detail. However, the blind participants showed a slight preference for the route descriptions, whereas the sighted preferred the survey description.

The foregoing suggests that perspective during learning may not lead to complete functional equivalence. Specifically, the encoding process differs between a route and a survey description on a number of grounds. The encoding of a route description has shown to take longer and produce a higher working load on visuo-spatial and central executive memory than the encoding of a survey description (Brunyé & Taylor, 2008a; Brunyé & Taylor, 2008b; Deyzac et al., 2006). In addition, in the early stages of processing the nature of the resultant spatial mental model seems to differ (Hubona et al., 1998), where representations from a route perspective are less flexible resulting in more difficulty with using a different perspective compared to a survey perspective (Brunyé & Taylor, 2008a).

Possibly, these behavioral differences found building up a mental representation from a route and a survey description are the result of the manner in which spatial information is conveyed in these descriptions. In a route description a series of egocentric moments is presented from which also knowledge about the larger allocentric configuration can be inferred. In a survey description, on the other hand, relations between objects in the environment are presented within the

description and an allocentric representation can be constructed directly but also exclusively. Being presented in a more abstract manner, it requires less working memory resources and takes less time to encode allocentric information from a survey description (Brunyé & Taylor, 2008b). As a consequence, perspective flexibility allowing novel route computation is achieved already after a smaller number of presentations.

Notably however, only allocentric and no egocentric information is presented in a survey description and consequently no egocentric representation can be constructed. This is in contrast to a route description, from which both the egocentric subject-based components and inferred allocentric environment-based information can be stored in memory. Based on the behavioral results described above it can therefore be hypothesized that from a route description both egocentric (directly) and allocentric (indirectly) components are encoded into memory, while from a survey description only allocentric (directly) components can be encoded. We refer to this hypotheses as the *hierarchical* model. Despite differences in the encoding process, it has been shown that after sufficient learning spatial mental models built from route and survey descriptions result in similar performance for subjects on a number of spatial memory tasks (Brunyé & Taylor, 2008a; Taylor & Tversky, 1992), in particular on the allocentric bird-flight distance comparison task (Noordzij & Postma, 2005; Noordzij et al., 2006b).

The presumed functional equivalence thus far has only been measured after the spatial mental representation was fully constructed. The mixed results explained above indicate that building up a supramodal spatial mental representation is a progressive process. At first, perspective dependent information is available, however, after over learning a perspective independent supramodal spatial mental representation can be built up. An intriguing question regarding this issue is what exactly supports such a spatial mental representation at a neural level? The progressive nature suggests that the representation should evolve as people learn the configuration. A very elegant technique to investigate this issue is functional magnetic resonance imaging (fMRI). Examining the neural network involved in encoding spatial descriptions might shed light on the underlying processes involved in building up a spatial mental representation. The neural network involved might evolve over time mirroring the behavioral results that show distinct differences in performance after a single exposure compared to over-learning. The current study aimed to investigate the neural correlates of encoding spatial descriptions from different perspectives. And especially focused on the parametric modulation of number of exposures. This modulation potentially reveals the network that shows increasing activation as the number of exposures increases.

At present, a wide range of neuroimaging studies have investigated different aspects of human spatial memory. Many of these studies have indicated that spatial memory is not likely to be represented by a unitary system but functionally distributed across the cortex, with different aspects of a spatial scene processed in parallel systems in a large overlapping frontal-parietal network, but also with distinct neural correlates (Aguirre & D'Esposito, 1997; Burgess, 2008; Doeller et al., 2008; Hartley et al., 2003; Janzen & Weststeijn, 2007; Zaehle et al., 2007). One distinction that has been made is between egocentric and allocentric components being processed and stored into different

areas in the brain (Burgess, 2006, 2008; Maguire et al., 1998a). Based on this distinction we may propose an alternative model, namely the *distinct representations* model. This model suggests that egocentric and allocentric components are distinct and stored in different neural networks. As such, people will be able to encode only egocentric components from a route perspective, while people will be able to encode only allocentric components from a survey perspective.

The wide range of neuroimaging studies that have investigated spatial memory have shown that a large fronto-parietal network is active during spatial tasks (Mellet et al., 2000; Shelton & Gabrieli, 2002; Zaehle et al., 2007). In addition to overlapping networks, differences between coding egocentric and allocentric components have also been reported. It is well known that the hippocampus is involved in allocentric spatial memory (for a review see Burgess et al., 2002). Activity in the hippocampus reflected navigational accuracy (Maguire et al., 1998a), survey knowledge (Latini-Corazzini et al., 2010; Mellet et al., 2000; Neggers et al., 2006; Wolbers & Buchel, 2005), information about boundary locations (Doeller et al., 2008; Iaria et al., 2003), and flexible wayfinding (Hartley et al., 2003). The structure that is often associated with egocentric spatial memory is the caudate nucleus, e.g. following a well-known route (Hartley et al., 2003; Janzen & Weststeijn, 2007), egocentric response strategies, i.e. remembering turns (Iaria et al., 2003), landmarks (Doeller et al., 2008), and navigational speed (Maguire et al., 1998a). A third structure of interest is the parahippocampal gyrus, which had been linked to processing landmarks. Landmarks are especially relevant in route-perspectives and egocentric representations, however, they are also relevant to allocentric representations. Indeed, the literature reports the parahippocampal gyrus for egocentric (Aguirre & D'Esposito, 1997; Epstein et al., 2007; Janzen & Weststeijn, 2007; Latini-Corazzini et al., 2010; Mellet et al., 2000) and allocentric representations (Aguirre & D'Esposito, 1997; Galati et al., 2010; Latini-Corazzini et al., 2010).

The majority of spatial memory studies have used visual stimulus presentation, such as virtual reality environments. However, the implied supramodal nature of spatial cognition (Struiksma et al., 2009) suggests that verbal presentation should provide a functionally equivalent representation in addition to similar neural correlates with respect to the spatial representation. This supramodal representation also implies that the verbal nature of the task would recruit language related areas, while visual tasks would be associated with activity in visual areas. Zaehle et al. (2007) described an fMRI study where they distinguished between verbal descriptions of egocentric and allocentric spatial relations between a few objects. They indeed found a common fronto-parietal network and distinct hippocampal activation associated with the allocentric coding of space, similar to visual paradigms.

To our knowledge, the current study is the first neuroimaging study to use verbal descriptions (administered auditorily) to describe a large-scale complex environment. Moreover, the two models explained above (progressive and distinct representations models) explain how different sources of information are encoded and possibly stored in the brain. Although it is very tempting to speak about the neural representation of spatial memory, for example egocentric components, it is not very likely that a single neural structure contains the cognitive map of the learned environment. It

is more likely that an entire network is involved in the encoding of spatial information. The current study examined the encoding of processing spatial descriptions while measuring brain activity. Participants listened to six repetitions providing information describing our large-scale complex environment. This enabled us to examine which structures become more or less involved over time. The current parametrically modulated fMRI design provided a unique opportunity to investigate the nature and neural substrates of encoding spatial mental representations into memory.

Methods

Participants

40 participants took part in this experiment (20 male and 20 female, age $M = 24$, $SD = 2.9$). The data of eleven participants (7 male and 4 female, age $M = 23$, $SD = 3.3$) had to be replaced due to scanner problems (three participants), or performance below chance level during the behavioral tests (eight participants). All participants were native Dutch speakers and right-handed as assessed with the Edinburgh handedness questionnaire (Oldfield, 1971). They had normal or corrected to normal vision, no hearing problems and no history of neurological disorders. Participants signed an informed consent as approved by the Medical Ethical Board (Medisch-etische toetscommissie (METC-protocolnumber 05/186-E). Males and females were distributed equally over the two experimental conditions. Apart from this, participants were randomly assigned to the experimental groups.

Materials and design

All descriptions were digitally recorded and edited with the free open source software Audacity 1.2.6 (audacity.sourceforge.net). They were all read by the same voice at the same speed, resulting in three recordings of approximately the same length for the encoding phase. High frequencies were filtered out and the recordings were normalized at 3 dB.

In the encoding task participants studied two descriptions: an experimental and a control description. In the experimental description the spatial configuration of a zoo environment was described. In the control description participants learned a non-spatial story about a suitcase filled with colored clothes. Participants received the experimental zoo-description either in a survey or a route perspective. The control description was the same for both groups.

Experimental descriptions

Both the environment and the descriptions in this experiment were adaptations from the materials used by Noordzij and Postma (2005). Moderate changes were necessary to make the stimuli suitable for use in the fMRI design. The most prominent changes were simplification of the zoo environment and shortening of the descriptions. A fictitious map of a zoo was designed that was square in

shape and measured 20 by 20cm. The zoo was divided into four quadrants. In each part two animals were located, thus a total of eight animals (all three-syllable words) were positioned in the zoo. A road was added that connected the centers of each quadrant, with an extra path leading to the entrance of each animal cage. See Figure 3 for the precise configuration of the zoo-environment.

Two descriptions of this environment were made: one in a survey perspective and one in a route perspective. See appendix A for an English translation of the texts. Because of the different perspectives, the texts differed on a number of points. The survey-description described animal locations independent of the listener. The description was written in the third person and animal locations were either related to each other, or their position was defined within the quadrant. In both cases, cardinal directions (such as *north*, *east*, *south* and *west*) were applied to indicate animal locations. In contrast, in the route description new animal locations were introduced relative to the suggested position of the listener within the zoo. The participant was addressed in the second person and locative propositions (such as *in front* and *to the right*) were used to indicate animal positions.

The descriptions also differed in their overall organization. The survey description started by introducing all four quadrants after which the different animals were pointed out per quadrant (i.e. a hierarchical organization). In contrast, the route description identified the first quadrant, pointed out the animals located within that quadrant and then continued to mention the second quadrant, point out the animals, etc. Information about the overall layout was revealed in a step-wise fashion (i.e. a linear organization).

A number of factors were held constant in both descriptions. To preserve a coherent description, which has shown to result in easier processing and better recall (Denis & Cocude, 1992), both perspectives introduced animals a sequential manner. The descriptions started with the animal closest to the entrance and continued in a clock-wise direction.

Furthermore, the total number of words for both descriptions was approximately the same, i.e. 203 for the survey description and 232 for the route description. In either description all animals and locations were mentioned only once.

Control description

A control description was developed to match the experimental descriptions on several aspects but lacking any spatial or temporal references. The control description describes a grandmother and her granddaughter investigating a suitcase filled with colored clothes. In the description a total of eight pieces of clothing in specific colors (all one-syllable words) were mentioned. Color and clothes combinations were assigned randomly. All objects were mentioned only once and the total number of words of the description was approximately the same as for the experimental descriptions, namely 178. See Appendix for an English translation of the control description.

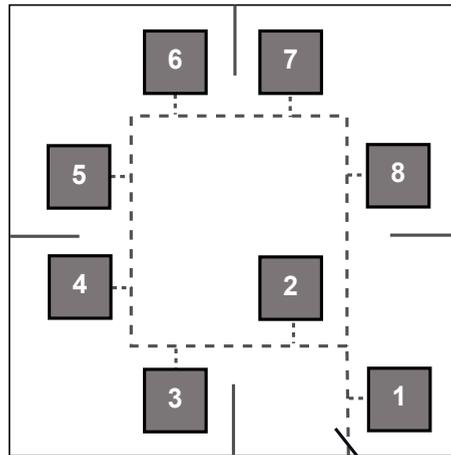


Figure 1 The configuration of the zoo environment.

Presentation of the descriptions

In order to use the auditory descriptions in an fMRI block design, each description was divided into two parts that were presented in an interleaved sequence. At the start, the first half of the zoo description was presented, followed by the first half of the control description. Subsequently, the second half of the zoo description was presented, followed by the second half of the control description. Prior to each part of the description participants heard a beep characteristic for that specific description. By editing the pauses between the sentences all description parts had a duration of 40 seconds.

Procedure

Auditory stimuli were presented using an MR compatible electrodynamic headphone (CONFON HP-SI 01; MR Confon GmbH, Baumgart et al., 1998) with an Aureon 5.1 USB soundcard. The sound level was set at 0dB, however the participants wore ear plugs, which resulted in a comfortable sound level in combination with the scanner noise. An MR compatible pneumatic response box with four buttons was used to record responses. All tasks were administered on a PC with Presentation software 12.0 (Neurobehavioral Systems, Albany, CA).

The experiment consisted of four tasks: one encoding task and three retrieval tasks (discussed in chapter 7 of this thesis). The encoding task included a description of the zoo-environment and a control description about pieces of clothing in a suitcase.

At the start of the experiment participants received information about the procedure of the experiment. After given time to ask questions, they signed the informed consent. Before entering the scanner participants received a first set of instructions. They viewed a drawing of the outer edge of the zoo-environment, on which the experimenter pointed out the indicators that would be used in the description (either survey or route perspective). Except for this first instruction, procedure and instruction was identical for both experimental groups. For participants in the survey group the

north, east, south and west side of the zoo were indicated. For participants in the route group the experimenter clarified the left and the right side of the participant. A pilot study had shown that with this instruction participants were able to correctly draw the animals in their relative locations after studying the descriptions.

Before entering the scanner participants heard the zoo-description and control description twice. Once as a whole and once interleaved with each other. This was done to familiarize the participants with the acoustic characteristics of the descriptions, such as the voice of the narrator and the pace of reading. In addition, participants received instruction about the control task and the response coding together with four practice trials, in order to limit their total time in the scanner.

After the instruction and practice session participants entered the scanner and the experiment started. Participants were instructed to close their eyes during scanning. For the encoding task participants listened to the interleaved sequence of the descriptions six more times (approximately 17 minutes). A behavioral pilot study had shown that participants were successful in making an accurate drawing of the experimental environment after six presentations of this combination sequence. Participants reported to feel confident of the descriptions after four presentations on average. In total participants listened to eight exposures of both descriptions to ensure that they had enough exposure to build up a mental representation. Subsequently, an anatomical scan was recorded (approximately 9 minutes), during which the participant was allowed to rest but told not to fall asleep. This brief rest period was followed by the three retrieval tasks. Scanning time for each task was approximately 11 minutes and prior to each scanning session participants were again instructed to close their eyes.

After finishing the third task participants were taken out of the scanner. As a last assignment they were asked to draw the environment on a scale model in which the correct animals needed to be matched to their relative positions. They were given no time constraint but were told to make the drawing as accurate as possible. Furthermore, they were asked to list all clothes with the correct colors. Only participants who could accurately, i.e. correct locations and recall of all animals, draw the environment and list the clothes-color combinations were included in the analysis.

MR data acquisition

Functional imaging was performed on a Philips Achieva 3T scanner with an eight-channel SENSE head coil, using the blood oxygen level-dependent (BOLD) sensitive, navigated 3D PRESTO pulse sequence (Neggens et al., 2008). An entire volume was acquired in 608 ms (TE = 22.5 ms, flip angle = 10°, 56×64 acquisition matrix, 40 sagittal slices, isotropic voxels of 4 mm, FOV(ap, fh, rl) = 224×256×160mm and a SENSE factor of 2 on the AP and 1.8 in the LR direction). Each scanning session was preceded by ten dummy volumes in order to accomplish steady state transversal magnetization.

The encoding task consisted of 1772 dynamics. After scanning of the encoding task was finished a reference scan was made with a flip angle of 27°, but otherwise identical to the PRESTO-SENSE functional MRI images. Due to the increased flip angle, this image had slightly more anatomical

contrast and was used for coregistration with the anatomical scan. A T1-weighted anatomical scan was acquired for each participant (TE = 4.6ms, TR = 9.86ms, flip angle = 8°, 224×224 acquisition matrix, 160 coronal slices, voxel size = 0.875x0.875x1mm and FOV(ap, fh, rl) = 224x160x168mm).

Data analyses

Behavioral data

The drawings of the participants were judged on the number of correctly recalled animals and their relative locations.

Functional imaging data

Imaging data was analyzed with SPM5 (Wellcome Department of Imaging Neuroscience, Institute of Neurology, London, www.fil.ion.ucl.ac.uk/spm/software/spm5) and the MarsBaR toolbox for SPM5 (marsbar.sourceforge.net) running under Matlab (R2007b, The MathWorks, Inc., Natick, MA). Preprocessing included coregistration and realignment. The anatomical scan was segmented and spatially normalized with medium bias regulation (0.001). The spatial normalization parameters from the ‘unified segmentation’ routine were used to normalize all functional scans (Ashburner & Friston, 2005; Crinion et al., 2007), which were then spatially smoothed with a kernel of 8 mm FWHM.

First level statistics were performed for each participant. A high-pass filter with a cut-off period of 128s was applied to remove low frequency fluctuations of the scanner. The encoding task was designed as an on-off block design, in which the presentation of the experimental description corresponded to the on- and the control condition to the off-situation. The model for the encoding task contained 28 regressors. The design matrix consisted of five regressors of interest (experimental condition, parametric modulation of paragraph, parametric modulation of repetition, rest and beep modeled as (parametrically modulated) delta functions convolved with a hemodynamic response function (see Figure 2 for the first two regressors)). Several nuisance-regressors were added: three to filter out a very systematic scanner-related oscillation in a very narrow frequency band exactly at 0.5 Hz, and 20 regressors to model out physiological blood pulsatility rhythms induced by heart beat and respiration based on the RETROICOR algorithm (Glover et al., 2000). The latter method effectively reduces the substantial physiological noise often observed in BOLD fMRI time series data, enhancing sensitivity for detecting signal changes.

Using a general linear model the parameter estimates were calculated for all brain voxels. Several effects, mentioned below, were tested by means of linear contrasts between the parameter estimates for different conditions. These contrast images were then passed to a second-level analysis, to model any group effects.

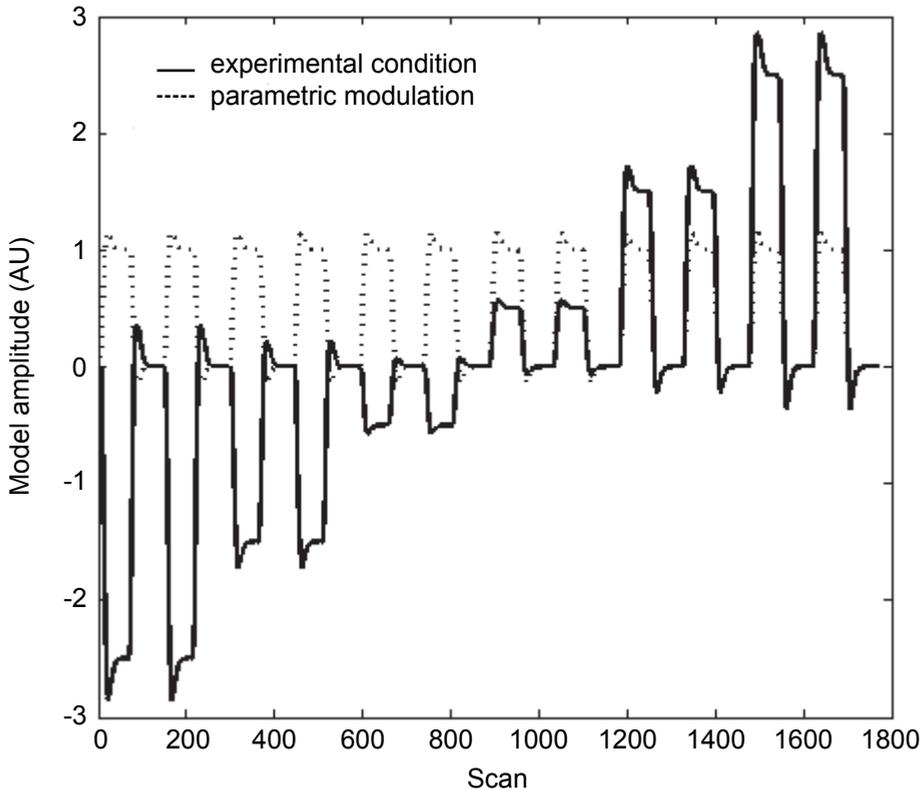


Figure 2 Illustration of the two contrasts of interest in the model of the encoding task.

The contrasts of interest in the encoding task were *experimental condition* and the *parametric modulation repetition*. For each of the two experimental groups, one-sample t-tests were used to test differences in the encoding of a spatial versus a non-spatial description, as well as monitor any time-modulations in this process

We report the results of a random effects analysis with inferences drawn at the voxel level. *P*-values were family wise error-corrected for multiple comparisons using Gaussian random fields theory (Worsley et al., 1996). Temporal autocorrelation in the fMRI model was modeled using autoregression modeling of the first order (Henson, 2003). For the parametric modulation of repetition a small volume correction was applied (Friston, 1997). This procedure constrained our search space to a region of interest (ROI). This ROI was designed using the aal-template of the WFU Pickatlas toolbox, as developed by the functional MRI Laboratory at the Wake Forest University School of Medicine (www.fmri.wfubmc.edu/cms/software). Based on the results of previous neuroimaging studies, the ROI included the hippocampus, parahippocampal gyrus and caudate nucleus.

Results

For each of the two experimental groups, one-sample t-tests were used to test differences in brain activation and deactivation in the encoding of a spatial versus a non-spatial description. The route group had listened to the route description of the environment, which was contrasted with the control description eliminating general auditory processing areas. The general contrast of *encoding condition* revealed a large fronto-temporo-parietal network (see Figure 3 and Table 1 for details), including the bilateral middle frontal gyrus, predominantly left superior temporal gyrus and a large bilateral superior parietal activation including the superior parietal lobule. Also a substantial part of the cerebellum was activated. Deactivation occurred in the superior occipital gyrus and stronger on the left side than on the right side. This is probably caused by the auditory nature of the task.

When we look at the three target areas (hippocampus, parahippocampal gyrus and caudate nucleus) in more detail by means of an ROI analysis we observe bilateral activation in the parahippocampal gyrus, some bilateral activation of the caudate nucleus and one activated voxel in the right posterior part of the hippocampus for the general contrast of encoding route information in comparison with the control description. The activation of the parahippocampal gyrus could be associated with the processing of animal names as landmarks. The activation in the caudate nucleus was expected for the route encoding. When we look at the deactivation within this ROI we find a significantly deactivated voxel in the anterior part of the right hippocampus.

The survey group had listened to the survey description of the environment, which was contrasted with the control description. This revealed an activation and deactivation pattern associated with processing spatial information. Again a large fronto-temporo-parietal network was found (see Figure 3 and Table 2 for details). However, for the survey group there was stronger activation in the frontal regions, encompassing the bilateral frontal eye fields, the right inferior parietal lobe and bilateral thalamic activation. For the survey group no cerebellar activation was found. The only area that showed significant deactivation was found in bilateral posterior cingulate.

Closer inspection of the ROI revealed right parahippocampal activation and bilateral, but more pronounced right-side, activation of the caudate nucleus. The observed parahippocampal activation suggests that even in the survey descriptions the animal cages functioned as landmarks. Interestingly, we also found caudate nucleus activation, usually associated with processing egocentric components from a route description. Within the ROI we found deactivation of two voxels in the right hippocampus.

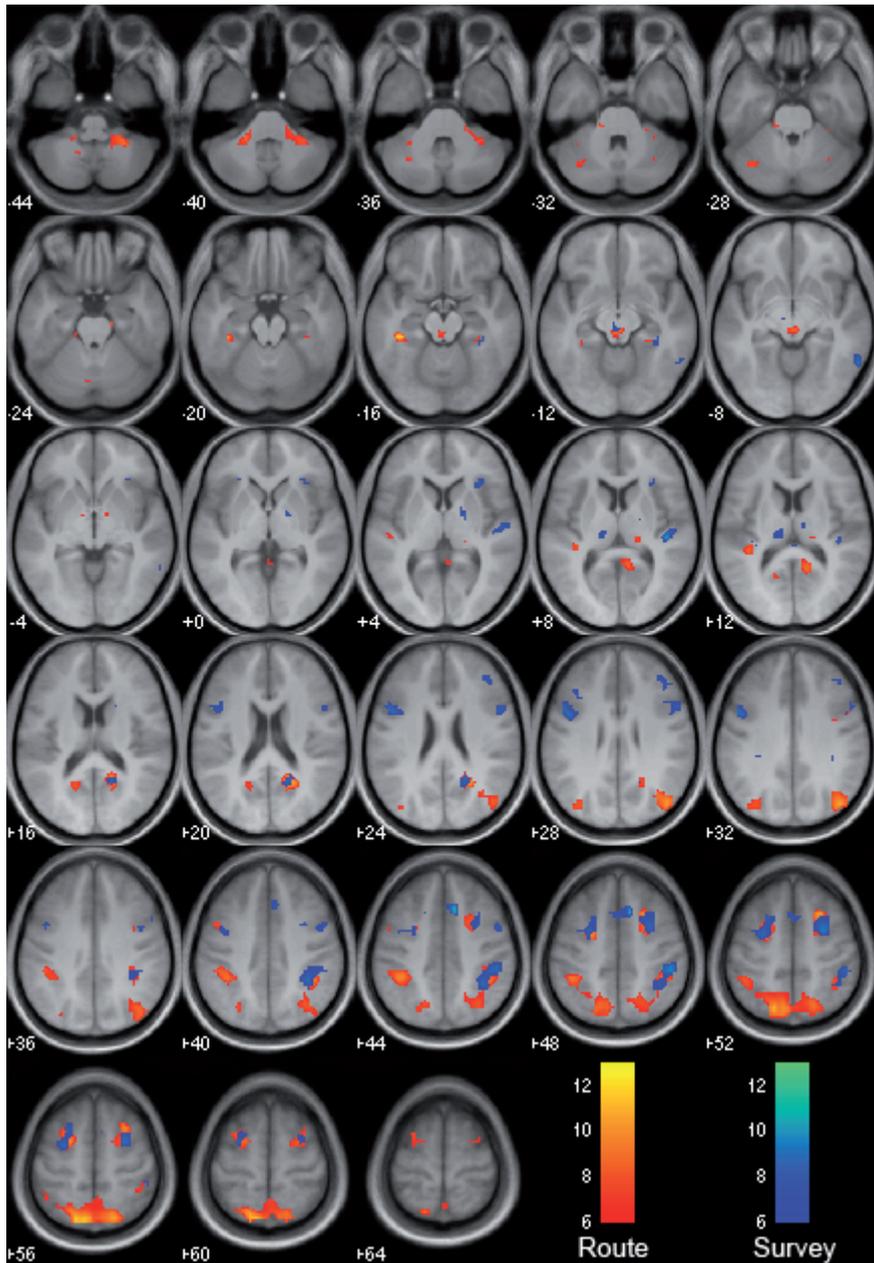


Figure 3 General activation of encoding a spatial description compared to a control description. The areas in red represent the activation found in the route group. The areas in blue represent the activation found in the survey group.

Table 1 Task activation of encoding a route description compared to a control description

Region	L/R	Cluster Size	Peak voxel Z-score	MNI coordinates (mm)		
				x	y	z
<i>Route description activation whole brain</i>						
Middle frontal gyrus	L	82	6.17	-24	0	52
Middle frontal gyrus	R	92	5.98	24	16	52
Precentral gyrus	L	12	5.36	-44	4	40
Precentral gyrus	R	12	5.08	52	8	32
Superior temporal gyrus	L	12	5.55	-40	-36	12
Inferior temporal gyrus	L	15	6.50	-40	-32	-16
Fusiform gyrus	L		4.75	-28	-40	-12
Fusiform gyrus	R	10	4.95	32	-36	-12
Superior parietal lobule	L	638	6.23	-16	-68	56
Precuneus	L		6.21	-8	-72	56
Cuneus	R		6.13	20	-56	20
Cuneus	L	16	5.33	-16	-60	16
Thalamus	R	7	5.11	16	-24	8
Lentiform nucleus	L	4	4.96	-12	-4	0
Lentiform nucleus	R	2	5.54	12	-4	-4
Midbrain	L/R	20	5.77	0	-28	-12
Midbrain	R	4	5.23	16	-20	-24
Cerebellum	L	17	5.58	-20	-44	-40
Cerebellum	R	61	5.51	28	-48	-44
<i>Route description deactivation whole brain</i>						
Superior occipital gyrus	L	22	5.60	-12	-96	24
Superior occipital gyrus	R	4	4.97	20	-92	24
Inferior parietal lobule	L	1	4.94	-52	-64	44
Inferior temporal gyrus	L	4	4.83	-44	0	-32
Superior occipital gyrus	L	1	4.69	-8	-100	4
Posterior cingulate	L	1	4.63	-8	-52	32
<i>Route description activation ROI</i>						
Parahippocampal gyrus	R	21	4.95	32	-36	-12
Parahippocampal gyrus	L	11	4.54	-28	-40	-8
Caudate nucleus	R	2	3.96	16	-8	16
Caudate nucleus	L	2	3.89	-12	-8	16
Hippocampus	R	1	3.82	20	-32	8
<i>Route description deactivation ROI</i>						
Hippocampus	R	1	3.73	20	-8	-12

Whole brain: $p < .05$, FWE-corrected.

ROI (hippocampus, parahippocampal gyrus and caudate nucleus): $p < .05$, FWE-corrected.

Table 2 Task activation of encoding a survey description compared to a control description

Region	L/R	Cluster Size	Peak voxel Z-score	MNI coordinates (mm)		
				x	y	z
<i>Survey description activation whole brain</i>						
Inferior frontal gyrus	L	139	5.79	-44	4	28
Middle frontal gyrus	L		5.52	-28	-4	52
Middle frontal gyrus	R	123	5.84	28	0	52
Inferior frontal gyrus	R		5.44	52	8	24
Precentral gyrus	R		5.23	28	4	40
Supplementary motor area	L	44	4.99	0	8	52
Supplementary motor area	R		6.19	12	16	48
Caudate nucleus	R	1	4.55	20	12	16
Superior temporal gyrus	L	2	4.70	-36	-28	12
Superior temporal gyrus	R	27	6.04	44	-24	8
Inferior temporal gyrus	R	15	5.89	60	-56	-8
Postcentral gyrus	L	3	4.68	-32	-32	32
Inferior parietal lobe	R	104	5.73	48	-36	48
Postcentral gyrus	R		4.54	56	-24	44
Insula	L	3	4.93	-24	28	0
Insula	R	17	5.33	32	24	4
Lentiform nucleus	R	11	5.40	20	-4	4
Fusiform gyrus	R	9	5.40	36	-36	-16
Precuneus	R	24	5.54	20	-52	24
Posterior cingulate	R	2	4.64	4	-32	12
Midbrain	L	1	5.25	-8	-16	-8
Midbrain	L/R	7	5.15	0	-24	-12
Thalamus	L	14	5.19	-12	-20	8
Thalamus	R	6	4.91	8	-12	12
<i>Survey description deactivation whole brain</i>						
Posterior cingulate	L/R	21	5.33	0	-48	28
<i>Survey description activation ROI</i>						
Parahippocampal gyrus	R	9	5.40	36	-36	-16
Caudate nucleus	R	10	4.55	20	12	16
Caudate nucleus	L	3	4.33	-16	4	16
<i>Survey description deactivation ROI</i>						
Hippocampus	R	2	4.05	28	-16	-16

Whole brain: $p < .05$, FWE-corrected.

ROI (hippocampus, parahippocampal gyrus and caudate nucleus: $p < .05$, FWE-corrected.

The current design enabled us to investigate not only those areas that are associated with the complex task of processing spatial descriptions and learning about the environment, but also to specifically investigate those areas that are specifically involved in the learning process and where activation gradually changes over time. The parametric modulation of the spatial description versus control description revealed those areas that change when the descriptions are repeated multiple times. For the route group a whole brain analysis was performed, which revealed significant activation that increased over time in the left caudate nucleus, the right superior frontal gyrus and the right insula (for details see Table 2). There were no areas found that decreased significantly over time.

The most important analysis was the ROI analysis of the repetition contrast. This revealed a significant increase in bilateral hippocampus and bilateral caudate nucleus (for details, see Figure 4 and Table 3). The activation in those areas became stronger with more repetitions of the route description. Interestingly, the hippocampus became active while processing the route description. Similar to the whole brain analysis, also no deactivation was found for the ROI analysis.

The survey group showed significant increasing activation in the left putamen, bilateral anterior cingulate, left insula, right inferior frontal operculum, right middle temporal gyrus and right supramarginal gyrus (for details, see Table 4). Within the ROI of interest we observed significantly increasing activation in the left hippocampus, as was expected for processing survey information, and surprisingly in the left caudate nucleus (see Figure 4 and Table 4 for details). Neither for the whole brain analysis, nor for the ROI analysis was there any significant deactivation.

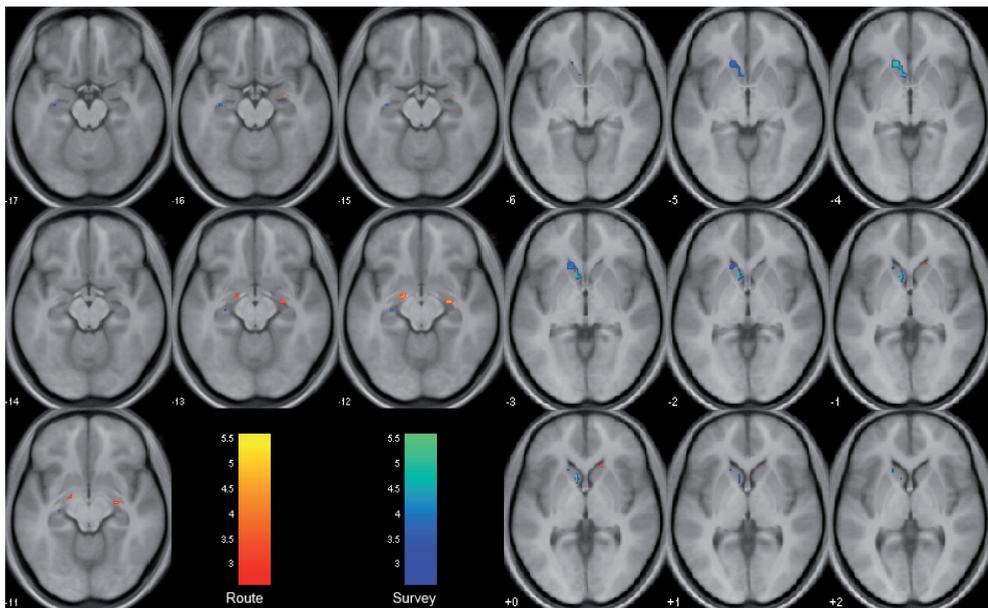


Figure 4 Activity increasing with the number of presentations for encoding spatial descriptions compared to a control descriptions. The areas in red represent the activation found in the route group. The areas in blue represent the activation found in the survey group.

Table 3 Task activation of repeating a spatial description compared to a control description

Region	L/R	Cluster Size	Peak voxel Z-score	MNI coordinates (mm)		
				x	y	z
<i>Route repetition activation whole brain</i>						
Superior frontal gyrus	R	5	5.09	24	44	0
Caudate nucleus	L	2	4.77	-24	16	20
Insula	R	4	4.72	36	-20	24
<i>Route repetition activation whole brain</i>						
none						
<i>Route repetition activation ROI</i>						
Hippocampus	L	3	3.71	-16	-4	-12
Hippocampus	R	3	4.07	28	-12	-12
Caudate nucleus	L	1	3.88	-16	12	20
Caudate nucleus	R	2	3.82	16	24	0
<i>Route repetition deactivation ROI</i>						
none						

Whole brain: $p < .05$, FWE-corrected.

ROI (hippocampus, parahippocampal gyrus and caudate nucleus): $p < .05$, FWE-corrected.

Table 4 Task activation of repeating a spatial description compared to a control description

Region	L/R	Cluster Size	Peak voxel Z-score	MNI coordinates (mm)		
				x	y	z
<i>Survey repetition activation whole brain</i>						
Putamen	L	5	5.04	-16	4	-8
Putamen	L	2	4.82	-28	12	-8
Anterior cingulate	L	15	4.93	-8	36	8
Anterior cingulate	R	9	5.03	12	48	16
Superior medial frontal gyrus	R		4.97	8	56	16
Superior medial frontal gyrus	R	1	4.58	8	44	40
Inferior frontal operculum	R	5	4.92	40	-4	-24
Middle temporal gyrus	R	2	4.90	44	-56	20
Insula	L	1	4.59	-40	-8	16
Supramarginal gyrus	R	2	4.57	56	-48	28
<i>Survey repetition deactivation whole brain</i>						
none						
<i>Survey repetition activation ROI</i>						
Hippocampus	L	1	3.89	-32	-16	-16
Hippocampus	L	1	3.82	-28	-20	-12
Caudate nucleus	L	15	4.24	-12	24	-4
<i>Survey repetition deactivation ROI</i>						
none						

Whole brain: $p < .05$, FWE-corrected.

ROI (hippocampus, parahippocampal gyrus and caudate nucleus): $p < .05$, FWE-corrected.

Discussion

The aim of this study was to investigate the neural correlates of encoding spatial descriptions from different perspectives. We designed a complex zoo environment and created descriptions from a route and a survey perspective and contrasted these spatial descriptions with a control description about a suitcase with colored clothes. The control description was used to control for verbal input modality effects and general processing of repeated information. The contrast between the spatial description and the control description revealed those areas that are involved in processing spatial information and generating a spatial mental representation about the environment.

After the scanning sessions participants were asked to draw the zoo environment in the outline that they had viewed prior to the experiment. They were also asked to name all color-clothes combinations. All participants remembered all the items and their colors from the control descriptions. They were also able to draw all the animals in their correct relative locations. This is an indication that participants indeed learned the environment and built up a spatial mental representation of the zoo that could be accessed later on (see chapter 7 this thesis). There were no qualitative differences between the route and the survey group.

The analyses of the neural correlates of spatial description learning showed a large overlap between the group who listened to a route description and the group who had learned from a survey description. The current design did not allow a direct contrast between the two encoding groups. The use of an on-off block design implies that there is no sufficient amount of rest which functions as a baseline condition. Rather, the control description served as the baseline. Since we have two different groups there might be baseline differences which could cause artificial differences if a direct comparison was made. Instead, we had to turn to a qualitative analysis by looking at the overlays and visually inspecting the overlap and differences. The overlap existed mainly in the middle frontal gyrus and the right inferior parietal lobule. The large overlapping network in combination with the ability to correctly draw the environment, regardless of encoding perspective, suggests that the spatial mental representation that participants had built up is functionally equivalent for the route and survey group. This finding supports the *hierarchical* model which predicts perspective encoding differences during the initial learning phase, but after sufficient learning a functional equivalent spatial mental representation should arise.

In addition to these overlapping areas the route group also showed a large activation in the cerebellum. Activation in the cerebellum is often associated with processing movement. This has been found previously in wayfinding studies as well. For example, Maguire et al. (1998) found cerebellum activation when their movement tasks were compared with their static scenes task. Hartley et al. (2003) also found cerebellum activation when their wayfinding and route-following tasks were compared to a control trail-following task. This suggests that the activation found in the current experiment might be related to the imaginary route that participants could have created during the encoding phase. Participants could have imagined themselves walking through the zoo and remembering the route and the animals they encountered on their way.

The activation in the frontal eye fields (FEF) was more pronounced in the survey group than in the route group. During visual paradigms the FEF are associated with eye movements, of simulated eye movements (Colby & Goldberg, 1999). In the current experiment only auditory stimuli were used and participants were instructed to keep their eyes closed during scanning. The absence of visual cortex activation implies that participants followed this instruction, therefore, it is very unlikely that the current FEF activation is due to real eye-movements. Instead, activation of the FEF has also been found during wayfinding studies. For example, Schinazi and Epstein (2010) reported FEF activation which was significantly larger for decision points while remembering a route, compared to non-decision points. This finding was very similar to another study performed in their laboratory where the FEF were activated for navigationally intensive conditions, compared to a familiarity task that was less navigationally intensive (Epstein et al., 2007). Wallentin et al. (2008a) also report FEF activation during a spatial task. Participants were asked to verify a spatial sentence using either and egocentric reference frame (e.g. “Was she in front of you?”) or an allocentric reference frame (e.g. “Was she in front of him?”). Participants viewed identical pictures prior to the target sentence and FEF activation was found only during the linguistic verification phase. They found stronger activation of bilateral FEF for allocentric recall compared to egocentric recall and baseline. This difference was also strongly correlated with performance: good performers had a larger difference in FEF response compared to poor performers. This suggests that the FEF could have been involved in recalling spatial information and shifting from an egocentric reference frame, that was used in the picture, to an allocentric reference frame, possibly shifting spatial attention. We may speculate that the FEF activation found in the current experiment for the group that had listened to the survey description might be due to the shifting of attention to the different quadrants in the zoo environment. On the other hand, we cannot rule out the option that the survey group used imaginary eye-movements during the encoding phase. Further research is needed to determine the exact nature of the FEF activation.

The most interesting results follow from the three regions of interest (ROI) which might show perspective specific activation. The general comparison between a spatial description and the control condition revealed significant activation within the ROI. For the route group we found substantial bilateral parahippocampal gyrus activation. Previous research has linked this structure to processing landmark information (Epstein et al., 2007; Janzen & Weststeijn, 2007; Latini-Corazzini et al., 2010). Interestingly, we also found right parahippocampal gyrus activation in the survey group. Recently, Latini-Corazzini et al. (2010) showed that the parahippocampal gyrus was activated during route and survey tasks and supports perception of scenes. The parahippocampal gyrus is also referred to as the parahippocampal place area, since it is believed to be critical for place recognition (Epstein et al., 2007). The current study is the first to report parahippocampal gyrus activation after nonvisual verbal input. The activation in the parahippocampal gyrus is not limited to visual input only, but could also be activated by landmarks presented in verbal descriptions. This finding is in line with a supramodal representation of space (Struiksma et al., 2009) where spatial information can be represented in an abstract way independent from input modality.

In addition to parahippocampal gyrus activation we also found bilateral caudate nucleus activation in both the route and the survey group. The parametric modulation of repeating the descriptions revealed an increasing activation in the caudate nucleus, bilaterally for the route group and left-sided for the survey group. Previous reports on the function of this structure have associated it with processing egocentric components of spatial information (Hartley et al., 2003; Iaria et al., 2003; Janzen & Weststeijn, 2007; Latini-Corazzini et al., 2010; Maguire et al., 1998a). In line with Iaria et al. (2003) remembering turns in the route group could activate the caudate nucleus. Surprisingly, also the survey group showed a significant increase in the caudate nucleus after the number of repetitions of the descriptions grew. One of the clear differences between the route and the survey description was the serial presentation in the route description relating the animal locations to the recipient versus the hierarchical presentation in the survey description relating animal locations to the environment. The route description would describe the recipient moving through the environment taking turns as the route progresses. The survey description, on the other hand, does not include these egocentric components. Another mechanism supported by the caudate nucleus is response learning (Doeller et al., 2008; Poldrack et al., 2001). A route can be built up from linking landmarks or locations (Janzen & Weststeijn, 2007). This could be a form of stimulus-response learning, as has also been shown by Janzen and Weststeijn (2007). Additionally, Iaria et al. (2003) demonstrated that the caudate nucleus activation in a visual place-learning task increased with practice and suggests a role for it in performing familiar behavior. The present parametric modulation represent an increase in familiarity. Since we also observed caudate nucleus activity in the survey group this could reflect an increasing familiarity rather than an increasing knowledge of egocentric components.

The third region of interest was the hippocampus, which has been associated with allocentric components of spatial information, in particular navigational accuracy (Maguire et al., 1998a), survey knowledge (Latini-Corazzini et al., 2010; Mellet et al., 2000; Wolbers & Buchel, 2005), information about boundary locations (Doeller et al., 2008; Iaria et al., 2003), and flexible wayfinding (Hartley et al., 2003). For the survey group we found left posterior hippocampal activation that increased with the number of repetitions, suggesting that allocentric information was used to build up a spatial mental map of the zoo environment. The route group also showed a significant increase in hippocampal activation, but this activity was bilateral and more anterior compared to the survey group. This finding suggests that the route group was able to build up a spatial mental map with allocentric information, as was also demonstrated by the good performance on the drawing task. The differences between activation foci for the route and the survey group imply that the representations are not identical, i.e. perspective invariant, but rather functionally equivalent.

Taken together, the overlapping network of activation found in the route and the survey group for the general processing of spatial descriptions supports the *hierarchical* model. The parametric modulation revealed distinct areas that showed an increase in activation as the number of repetitions grew for the route group and the survey group. This shows that the encoding of route and survey information is distinct at first, but after over-learning develops into a functional equivalent spatial mental representation. Moreover, this study is the first to use verbal descriptions of

large-scale complex environments.

Importantly, the large fronto-parietal network found in addition to navigational specific activation in the hippocampus, parahippocampal gyrus and caudate nucleus has been previously reported for visual experiments. The fact that a highly similar neural network is involved in visual and verbal wayfinding situations strengthens the idea that spatial information is represented supramodally (Struiksma et al., 2009). Visual and verbal spatial information are essentially different: visual information is parallel and very detailed while verbal information is serial and mainly categorical. Despite these inherent differences the behavioral and neuroimaging studies, discussed above and including the current study, demonstrate that people are able to use both types of information to a highly similar extent. The essential spatial information can be extracted from different sources of input and is represented in brain areas that are modality independent. Indeed, we found a large amount of overlap between our results and previously reported results that have been obtained with visual paradigms, supporting a supramodal representation of spatial information.

In conclusion, the current results from verbal complex descriptions support the idea that spatial information is represented in a supramodal manner where spatial information is extracted regardless of the input modality. Processing route and survey descriptions activated a large fronto-temporo-parietal network. The strongest overlap existed in the middle frontal gyrus and the right inferior parietal lobule. For both groups we observed activation in the parahippocampal gyrus potentially linked to landmark terms presented in the verbal descriptions. In the survey group we found left posterior hippocampal activation, while the route group showed bilateral anterior hippocampal activation. Together these findings suggest that route and survey descriptions are encoded in a *hierarchical* fashion where allocentric components can also be inferred from egocentric components resulting in a functional equivalent mental representation for both encoding perspectives.

Acknowledgments

We would like to thank Huib Versnel for allowing us to use his MR compatible audio system. This system has greatly improved the audibility of the experiment.

Appendix

Examples of the descriptions translated from Dutch.

Route description example

You are looking at a straight wall. On the right hand side of the wall you see the main entrance of the zoo. You walk through the entrance. Inside the zoo, you notice that all animal residences are square and of the same size. A strip of grass runs around each residence with a path leading you from the main road to its entrance. You stand on the main road in the first section of the zoo. On your right you see the chamois. You take a quarter turn to your left and walk a little further. On your right hand side you now see the hyenas. You walk straight ahead and enter the second section of the zoo.

In the second section of the zoo to your left you see the white bears. You walk a little further and take a quarter turn to your right. On your left hand side are the kangaroos. You walk straight ahead and enter the third section of the zoo. On your left hand side you see the giraffes. You continue a little further again and take a quarter turn to your right. On your left hand side are the elephants. You walk straight ahead and enter the fourth section of the zoo. On your left hand side you see the chimpanzees. You pass them and take a quarter turn to your right again. On your left hand side the arctic foxes appear. In front of you is the first section again with straight ahead the main entrance.

Survey description example

The zoo has the shape of a square and is divided into four sections. The first section is located in the south-east corner of the zoo, the second in the south-west corner, the third in the north-west corner and the fourth part is located in the north-east corner of the zoo. A main road runs through the zoo forming a square by connecting the centers of each of the sections with an extra piece of road that links it to the main entrance. In each section of the zoo two animals are located. All animal residences are square and are of the same size. A strip of grass runs around each residence with a path leading you from the main road to its entrance. The entrance of the zoo is located in the first section, in the middle of its southern wall.

North-east of the entrance, in the first section of the zoo, the residence of the chamois is located. In the north-west corner of this section are the hyenas. South-west of these, in the second section, is the residence of the white bears. In the north-west corner of this section, are the kangaroos. To the north of this, in the third quadrant, the giraffes are located. In the north-east corner of this section the elephants reside. To the east of this, in the fourth section of the zoo, is the home of the chimpanzees. In the south-east corner of this section is the residence of the arctic foxes.

Control description example

Liselot plans to visit her grandmother. Grandmother has recently cleaned out her attic and discovered a suitcase with coloured clothes. Together they have a look if any of them are still suitable for grandmother to wear. Liselot takes a red coat out of the suitcase. She decides that it is still in good shape and hands it to grandmother. Grandmother picks up a pair of yellow trousers. Liselot thinks that they look too old and she throws them away. They also find a brown tanktop in the suitcase. Liselot believes it should find its way to the bin immediately.

Grandmother puts her hand in the suitcase and pulls out a white skirt. She looks at Liselot with hopeful eyes. Liselot nods her head to sign that she can keep it. She picks up a blue sweater. It still looks brand new and grandmother adds it to her pile. Liselot takes a green dress out of the suitcase. She hands it to grandmother. Grandmother pulls out a purple cardigan. Liselot tosses it into the bin. She gets out a black t-shirt. Grandmother shakes her head. The suitcase is empty. And grandmother is glad they have cleaned it out together.

Chapter 7

Neural correlates of route and survey distance judgments from spatial descriptions of different perspectives

M. E. Struiksma, L. J. Talsma, M. L. Noordzij, S. F. W. Neggers & A. Postma

(in preparation)

Abstract

When giving directions one can provide them from a first-person perspective, in which landmarks and their positions are related to the listener's egocentric reference frame. This type of description can be referred to as a route description. In a survey description one provides a birds-eye view with cardinal directions, such as north, east, south and west relating landmarks to other landmarks within an allocentric reference frame. The current study examined whether egocentric route distance judgments activated similar or distinct neural networks compared to allocentric survey distance judgments. Moreover it was also assessed whether the original learning perspective played an additional role. The behavioral results showed high similarity between the route and the survey group, who had learned the configuration of a zoo through a route and survey description respectively. Furthermore, a similar global pattern was found for the route and survey distance judgments tasks. These findings support the idea that over-learning yields a supramodal representation which is perspective independent. The neuroimaging results revealed activation of bilateral precuneus, linked to processing spatial information, and bilateral middle occipital gyrus and left parahippocampal gyrus, involved in processing landmarks, again comparable for the two judgments tasks. The current behavioral and fMRI results further support the idea that after over-learning a complex spatial description people build up a perspective independent spatial mental representation.

Introduction

When giving directions to a tourist you can provide them from a first-person perspective that includes landmarks and indications such as *left* and *right*, also referred to as a route perspective. In this way positions of landmarks are explained to the listener with an egocentric reference frame. However, when the tourist is viewing a map it might be more useful to provide a birds-eye view with more cardinal directions, such as *north*, *east*, *south* and *west* relating landmarks to other landmarks with an allocentric reference frame, also referred to as a survey perspective (Hund et al., 2008; Taylor & Tversky, 1992). The different characteristics of these two types of descriptions suggest that memories of these descriptions might rely upon distinct mental representations.

Indeed, several studies have found evidence for preservation of route and survey perspective dependent information in the mental representations resulting from these initial descriptions (Brunyé & Taylor, 2008a; Brunyé & Taylor, 2008b; Noordzij & Postma, 2005; Shelton & McNamara, 2004). For example, Shelton and McNamara (2004) showed that perspective switching from study to test produced a cost in scene recognition performance. Moreover, when the perspective remained the same there was an advantage when the test view matched the learned orientation. Brunyé and Taylor (2008a) observed that after a single study cycle participants had more difficulty verifying inference statements, about non-described spatial descriptions, when there was a perspective switch. This effect remained even after three study cycles suggesting that perspective characteristics were maintained. Furthermore, participants who had a single study cycle of a survey description performed significantly better on the inference statements and map drawing, compared to the participants who had a single exposure to the route description. Even after over-learning Noordzij and Postma (2005) found a relative advantage for survey descriptions over route descriptions. Participants who had learned the survey descriptions displayed evidence of a more fine-grained localization of the objects.

Whereas the studies discussed above indicate some perspective dependent differences, it is unclear whether these are really qualitative or actually quantitative. In line with the latter, other studies on route and survey descriptions have supported the idea that individuals develop a functional equivalent, i.e. perspective independent, spatial mental representation of the environment allowing inferencing about spatial relations even from the perspective not initially learned (Brunyé & Taylor, 2008a; McNamara, 1986; Taylor & Tversky, 1992). Taylor and Tversky (1992) showed that the relations between landmarks were captured in the same spatial mental representations for readers of route and survey descriptions and observers of maps. When participants were asked to answer inference questions there were no differences between the initially learned and the new perspective, suggesting that participants had used a spatial mental representation rather than a textual representation. Although Noordzij and Postma (2005) found a relative advantage for survey descriptions, a symbolic distance effect was also obtained after a route description, implying that participants had built up a spatial mental representation with analogue spatial detail from both types of verbal descriptions.

The mixed pattern of results does not provide conclusive evidence for the question whether or not spatial mental representations are perspective dependent. The conclusion appears to depend on task properties and number of exposures. After limited exposure evidence of perspective dependent spatial mental representations has been found whereas over-learning has yielded a perspective independent representation. In Chapter 6 (this thesis) we have shown that a large overlapping fronto-temporo-parietal network is activated during learning of a route or a survey description. In addition, hippocampal activation was found for the route and survey perspective associated with the progressive process of building up a spatial representation.

Another important issue is whether the type of information derived from the constructed representation could make a difference, both on a behavioral and on a neural level. Do we make different computations when assessing a route distance than when inferring a survey distance? The current study therefore contrasted retrieval and usage of route distance information against that of survey distance information. Moreover, we assessed whether the original learning perspective played an additional role. Do participants who have learned a configuration from a route description employ the same neural network as those who have learned a configuration from a survey description?

The present study examined the neural networks involved in performing a route task, drawing on egocentric information, and a survey task, drawing on allocentric information, and discusses the retrieval part of the experiment described in Chapter 6. There were two groups who learned the complex configuration of a zoo environment either from a route or a survey perspective. Subsequently, they performed the spatial tasks. Both tasks were distance comparison tasks, however, the instructions varied. In the established allocentric survey task participants were asked to compare bird-flight distances (Denis, 2008; Denis & Zimmer, 1992; Noordzij & Postma, 2005; Noordzij et al., 2006b), while in the egocentric route task participants were asked to compare route distances.

As discussed in Chapter 6 previous neuroimaging research on route and survey perspectives has also shown the dichotomy between perspective dependent and independent representations. Based on these results we can describe two possible models. First, the *distinct representations* model which assumes that egocentric and allocentric components are distinct and stored in different neural networks and yield perspective dependent representations closely connected to route and survey perspectives respectively. As such, people would be able to encode only egocentric components from a route perspective, whereas they can encode allocentric components from a survey perspective, which are stored into different areas in the brain (Burgess, 2006, 2008; Maguire et al., 1998a). This would predict that the route perspective learning group would use only an egocentric representation to perform either task, resulting in a similar network for both spatial tasks. However, it would also predict a behavioral advantage for the perspective congruent route task. Accordingly, the survey perspective learning group would activate a distinct allocentric representation for both tasks and show a behavioral advantage for the survey task.

Second, we may argue that from a route description both egocentric (directly) and allocentric (indirectly) components are encoded into memory, while from a survey description only allocentric

(directly) components can be encoded summarized in the *hierarchical* model (Shelton & Gabrieli, 2002). As a result we might expect similar performance for different encoding perspectives on a number of spatial memory tasks (Brunyé & Taylor, 2008a; Taylor & Tversky, 1992), in particular on the allocentric bird-flight distance comparison task (Noordzij & Postma, 2005; Noordzij et al., 2006b). Possibly, on egocentric spatial tasks differences between encoding perspectives might emerge. While allocentric information is available to both encoding groups, only the route encoding group has access to egocentric information. Therefore, we might hypothesize similar performance on allocentric tasks, but not on egocentric tasks. In retrieval, depending on the task at hand, either egocentric or allocentric information in memory could be selected for use (Burgess, 2008).

A final alternative model is based on a perfect learning situation. Although differences between perspectives might exist after few exposures, after over-learning a completely perspective independent representation is built up. All the essential spatial relations between objects in the zoo configuration have been extracted from the learning perspectives and are combined in a supramodal representation. A supramodal representation is a representation of purely spatial information about a configuration which can be linked to modality specific information, such as the representation of the specific animal in the cage (Struiksma et al., 2009). The *perfect learning* model assumes that the process of over-learning extracts relevant spatial information which is represented independent from perspective. As such, we might expect no differences between learning perspectives or spatial tasks, since they will all draw upon the perspective independent supramodal representation.

Methods

Participants

40 participants took part in this experiment (20 male and 20 female, age $M = 24$, $SD = 2.9$). The data of eleven participants (7 male and 4 female, age $M = 23$, $SD = 3.3$) had to be replaced due to scanner problems (three participants), or performance below chance level during the behavioral tests (eight participants). All participants were native Dutch speakers and right-handed as assessed with the Edinburgh handedness questionnaire (Oldfield, 1971). They had normal or corrected to normal vision, no hearing problems and no history of neurological disorders. Participants signed an informed consent as approved by the Medical Ethical Board (Medisch-etische toetscommissie (METC-protocolnumber 05/186-E). Males and females were distributed equally over the two experimental conditions. Besides this explicit division participants were randomly assigned to the experimental groups. Task-sequences were counterbalanced within the groups.

Materials and design

The encoding phase is described in detail in Chapter 6. Half of the participants learned the route description, the other half learned the survey descriptions. Subsequently, all participants performed all three behavioral tasks.

Retrieval tasks

In the traditional bird-flight distance comparison task, participants are asked to adopt a survey perspective and compare pairs of bird-flight distances between objects in a previously learned environment. Denis and Zimmer (1992), continuing with an idea by Moyer and Bayer (1976), designed this task to test whether participants formed a spatial mental model from a verbal description that included such spatial information. The larger the distance difference, the faster and better participants were in making a comparison judgment. This phenomenon is termed the distance comparison effect. Because in the bird-flight distance comparison task participants were explicitly instructed to adopt a survey perspective and object to object distances are compared, the task is a good measure to test participants' allocentric knowledge of the environment.

To test route knowledge, an adaptation to the traditional distance comparison task was made, in which participants were asked to adopt a route perspective (street view) and make similar judgments about route lengths. In addition to the zoo description participants also learned a control description and performed a control task (see Chapter 6 for more details on the control description) The control task had a similar design as the distance comparison tasks but was non-spatial in nature and tested the knowledge about the clothes and corresponding colors.

Survey task: bird-flight distance comparison

Participants were presented with two pairs of spoken animal names. The first object in both pairs always was the same, e.g. "Chimpanzee – Arctic Fox / Chimpanzee – Elephant". Each word had a duration of approximately 1300ms. Within each pair both spoken words were presented with a 300ms interval. A 1000ms interval was administered between both pairs.

Participants were asked to imagine a map of the environment and mentally focus on the bird-flight distance between each pair. The task was to judge whether the second distance was shorter (left button) or longer (right button) than the first distance.

A list of 36 trials was made consisting of two pairs of animal names each. The trials were divided into three categories of 12 trials each, based on the ratio between the distance difference from the center of the animal cage on the drawing of the environment. The ratios were defined through dividing the smallest distance by the largest distance. Categories were then: small difference (ratios between 0.77 – 1.00), medium difference (ratios between 0.65 – 0.76), and large difference (ratios between 0.30 – 0.64). In half of the trials the longer distance was placed in the first position and in the other half in the second position. This was counterbalanced within each category.

A pseudorandom fixed sequence was developed in which trials of the three categories were equally distributed over the experiment and were interleaved with 12 items of reverse speech (duration 3.5s). The reverse-speech items served as filler items and were also included in the route and control task to allow between task analyses on the imaging data. In the trial sequence there were never two successive trials with the same first animal.

Route task: route distance comparison

In the route task participants were asked which route distance between a pair of animals in the zoo was the longest. Participants were instructed that they could only follow the road in a clockwise direction.

The same trial pairs were used as in the survey task, which inevitably resulted in slightly different category boundaries. They were divided over three categories based on the ratio between the route distance difference from the center of the animal cage: small difference (ratios between 0.77 – 1.00: 9 trials), medium difference (ratios between 0.55 – 0.76: 12 trials), and large difference (ratios between 0.30 – 0.54: 15 trials). The trials were placed in a different sequence than in the survey task in order to equally distribute the different categories. The same 12 filler trials were used as in the survey task, but placed in a different order. In this sequence there were never two successive trials with the same first animal.

Control task

Participants were presented with two combinations of clothing-color combinations, e.g. “jacket – red / sweater – yellow”. Items had a duration of approximately 800ms. Within each pair both spoken words were presented with a 300ms interval and a 1000ms interval was administered between both pairs.

Participants were instructed to press the left button if both combinations were correct or if both combinations were incorrect. They were instructed to press the right button if one of the combinations was correct and the other one was not. A list of 36 pairs of clothing-color combinations was made with nine items per condition: both correct, both incorrect, first correct second incorrect, first incorrect second correct.

The 36 trials together with the 12 filler trials were pseudorandomly placed in a presentation sequence, such that the different response codes and conditions were evenly distributed over the experiment.

Procedure

The detailed procedure of the learning phase is described in Chapter 6. After the participants had learned the environment in the scanner and the anatomical scan had been made the three tasks were presented. The order of administration of the three tasks was either route, control, survey task, or survey, control, route task. The order of task presentation was counterbalanced over participants in each learning perspective group. Before the task started participants received instructions and a set of practice trials. Scanning time for each task was approximately 11 minutes and prior to each scanning session participants were instructed to close their eyes.

After finishing the third task participants were taken out of the scanner. As a last assignment they were asked to draw the environment on a scale model in which the correct animals needed to be matched to their relative positions. They were given no time constraint but were told to make

the drawing as accurate as possible. Furthermore, they were asked to list all clothes with the correct colors. Only participants who could accurately, i.e. correct locations and recall of all animals, draw the environment and list the clothes-color combinations were included in the analysis.

MR data acquisition

Functional imaging was performed on a Philips Achieva 3T scanner with an eight-channel SENSE head coil, using the blood oxygen level-dependent (BOLD) sensitive, navigated 3D PRESTO pulse sequence (Neggers et al., 2008). An entire volume was acquired in 608 ms (TE = 22.5 ms, flip angle = 10°, 56×64 acquisition matrix, 40 sagittal slices, isotropic voxels of 4 mm, FOV(ap,fh,rl) = 224x256x160mm and a SENSE factor of 2 on the AP and 1.8 in the LR direction). Each scanning session was preceded by ten dummy volumes in order to accomplish steady state transversal magnetization.

The retrieval tasks consisted of 1114 dynamics each. After scanning of each retrieval task was finished a reference scan was made with a flip angle of 27°, but otherwise identical to the PRESTO-SENSE functional MRI images. Due to the increased flip angle, this image had slightly more anatomical contrast and was used for coregistration with the anatomical scan. A T1-weighted anatomical scan was acquired for each participant (TE = 4.6ms, TR = 9.86ms, flip angle = 8°, 224×224 acquisition matrix, 160 coronal slices, voxel size = 0.875x0.875x1mm and FOV(ap,fh,rl)= 224x160x168mm).

Data analyses

Behavioral data

Analyses were done for the survey and route task. The data from the control task were only analyzed for overall proportion correct responses to verify that there was no difference between the two groups. Data of all practice trials were discarded. Individual proportion correct responses and mean response times to correct responses on the survey and route task were collected. Trials in which the response time was longer than 4000ms were excluded.

For the survey and route task the behavioral measures were analyzed using a $2 \times 2 \times 3$ mixed ANOVA with Encoding Perspective (Route, Survey) as a between-subjects factor and Task (route, survey) and Distance Difference (small, medium, large) as within-subject factors. Further analysis by means of pairwise comparisons used a significance level corrected for multiple comparisons with the Bonferroni method. SPSS multiplies the p -value with the Bonferroni multiplier instead of dividing α by the Bonferroni multiplier. The results are, however, equal and this method corrects for multiple comparisons. We will denote the Bonferroni corrected p -values by p_B .

In addition to the metric distance difference we also analyzed the data with text distance differences. Instead of a spatial representation participants could have merely remembered the actual text yielding only a surface representation. If so, then we would expect an influence of text distance rather than metric distance. This possibility was analyzed with another $2 \times 2 \times 3$ mixed ANOVA,

but the Distance Difference categories were computed using the ratio between the number of words in the text between the pairs of animal names. This resulted for the survey task in the following categories: Small text differences (0.76 – 1.00), Medium text differences (0.24 – 0.75), and Large text differences (0.00 – 0.23). For the route task this resulted in the following categories: Small text differences (0.69 – 1.00), Medium text differences (0.33 – 0.68), and Large text differences (0.00 – 0.32).

Functional imaging data

Imaging data was analyzed with SPM5 (Wellcome Department of Imaging Neuroscience, Institute of Neurology, London, www.fil.ion.ucl.ac.uk/spm/software/spm5) and the MarsBar toolbox for SPM5 (marsbar.sourceforge.net) running under Matlab (R2007b, The MathWorks, Inc., Natick, MA). Preprocessing included coregistration and realignment. The anatomical scan was segmented and spatially normalized with medium bias regulation (0.001). The spatial normalization parameters from the ‘unified segmentation’ routine were used to normalize all functional scans (Ashburner & Friston, 2005; Crinion et al., 2007), which were then spatially smoothed with a kernel of 8 mm FWHM.

First level statistics were performed for each participant. A high-pass filter with a cut-off period of 128s was applied to remove low frequency fluctuations of the scanner. The model for each task contained 26 regressors. The design matrix consisted of three regressors (experimental condition, filler items and beep) modeled as delta functions convolved with a hemodynamic response function. Several nuisance-regressors were added: three to filter out a very systematic scanner-related oscillation in a very narrow frequency band exactly at 0.5 Hz, and 20 regressors to model out physiological blood pulsatility rhythms induced by heart beat and respiration based on the RETROICOR algorithm (Glover et al., 2000). The latter method effectively reduces the substantial physiological noise often observed in BOLD fMRI time series data, enhancing sensitivity for detecting signal changes.

Using a general linear model the parameter estimates were calculated for all brain voxels. Several effects, mentioned below, were tested by means of linear contrasts between the parameter estimates for different conditions. These contrast images were then passed to a second-level analysis, to model any group effects.

The contrast of interest in the retrieval tasks was *experimental condition* > *filler items*. Because the filler items were equal in all three tasks contrasting the experimental conditions (route, survey and control trials) with the filler items the three tasks could subsequently be compared directly. We use a 3×2 factorial design. In this design, *task* (survey, route or control task) is a within-group factor and *experimental group* (route- or survey-description) is a between-group factor.

The commonalities between the two groups were tested by means of a conjunction analysis that tested the conjunction null hypothesis over two orthogonal contrasts (Friston et al., 2005; Nichols et al., 2005). The contrasts used for the conjunction analysis were one of the experimental conditions > control condition, for example Route group route task > Route group control task. This procedure reduced general task related activity present in both experimental and the control task.

The null distribution for the minimum statistic was based on two statistics. This enabled us to infer a conjunction of activation in an area in both groups (Friston et al., 2005). In order to determine the differences between the Route and Survey group we examined the contrast between the groups for both tasks. The contrast between the tasks was also analyzed for each group.

We report the results of a random effects analysis with inferences drawn at the voxel level. *P*-values were family wise error-corrected for multiple comparisons using Gaussian random fields theory (Worsley et al., 1996). Temporal autocorrelation in the fMRI model was modeled using autoregression modeling of the first order (Henson, 2003).

Results

Behavioral data

Accuracy

The overall proportion correct responses on the control task was high (Survey group $M = .92$, $SE = .015$, Route group $M = .88$, $SE = .023$) and did not differ between the two groups ($t(38) = -1.25$, $p < .220$).

For the survey and route tasks the effect of task order was not significant, therefore the results are collapsed over task order. A 2(Encoding Perspective) \times 2(Task) \times 3(Distance Difference) mixed ANOVA on the proportion correct trials for the metric distance differences revealed no differences between Survey and Route encoding perspectives ($F(1,38) = .11$, $p = .742$). There was a significant main effect of Task ($F(1,38) = 7.58$, $p = .009$). Accuracy on the route task was higher ($M = .76$, $SE = .015$) than accuracy on the survey task ($M = .71$, $SE = .018$). The main effect of Distance Difference was significant ($F(2,76) = 32.31$, $p < .001$). Pairwise comparisons revealed a symbolic distance effect. The accuracy on the Small distance differences was significantly lower ($M = .65$, $SE = .018$) compared to Medium distance differences ($M = .74$, $SE = .017$, $t(38) = -3.83$, $p_B = .001$). And the accuracy on Medium distance differences was significantly lower than accuracy on Large distance differences ($M = .81$, $SE = .016$, $t(38) = -3.83$, $p_B = .001$).

In addition to the main effect of Distance Difference the interaction with Task was also significant ($F(2,76) = 39.54$, $p < .001$). The pairwise comparisons revealed a slightly different pattern for the survey task and the route task. For the survey task there was no significant difference between Small ($M = .71$, $SE = .026$) and Medium distance differences ($M = .63$, $SE = .029$, $t(38) = 2.18$, $p_B = .102$). In line with the symbolic distance effect the accuracy on Large distance differences was significantly higher ($M = .78$, $SE = .020$, $t(38) = -4.75$, $p_B < .001$, see Figure 1). For the route task accuracy on Small distance differences ($M = .59$, $SE = .022$) was lower compared to Medium distance differences ($M = .85$, $SE = .018$, $t(38) = -10.40$, $p_B < .001$), which is in line with the symbolic distance effect. However, accuracy on Large distance differences did not differ from Medium distance differences ($M = .84$, $SE = .019$, $t(38) = .74$, $p_B = 1.000$, see Figure 1).

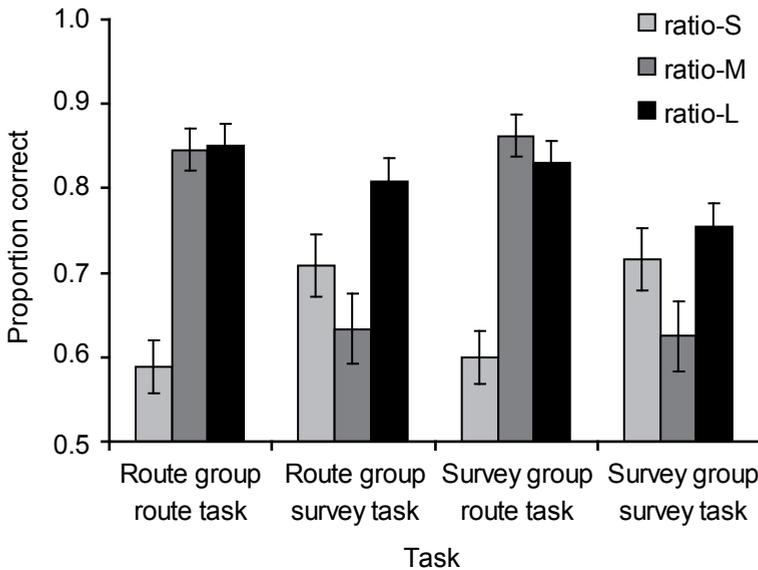


Figure 1 Proportion correct trials as a function of task and group in metric distance differences.

The data was also analyzed using text distance as the dependent variable again with a 2 (Encoding Perspective) \times 2 (Task) \times 3 (Distance Difference) mixed ANOVA. There was a significant main effect of Task ($F(1,38) = 14.66, p < .001$). Similar to the metric analysis accuracy on the route task was higher ($M = .78, SE = .015$) than accuracy on the survey task ($M = .71, SE = .018$). The main effect of Distance Difference was significant ($F(2,76) = 23.86, p < .001$). Pairwise comparisons revealed a symbolic distance effect. The accuracy on the Small distance differences was significantly lower ($M = .68, SE = .017$) compared to Medium distance differences ($M = .79, SE = .016, t(38) = -6.17, p_B < .001$). And the accuracy on Medium distance differences was significantly lower than accuracy on Large distance differences ($M = .76, SE = .015, t(38) = 1.86, p_B = .217$).

The interaction between Distance Difference and Task was also significant ($F(2,76) = 21.59, p < .001$). Pairwise comparisons revealed different patterns for the route and survey task. In the survey task accuracy on Small distance differences was significantly lower ($M = .69, SE = .027$) compared to Medium distance differences ($M = .77, SE = .019, t(38) = -3.16, p_B = .009$). Accuracy on Medium distance differences was significantly higher than Large distance differences ($M = .67, SE = .021, t(38) = 4.525, p_B < .001$, see Figure 2). The pattern in the survey task did not match the symbolic distance effect. For the route task accuracy on Small distance differences ($M = .67, SE = .022$) was lower compared to Medium distance differences ($M = .82, SE = .020, t(38) = -5.54, p_B < .001$), which is in line with the symbolic distance effect. However, accuracy on Large distance differences did not differ from Medium distance differences ($M = .86, SE = .018, t(38) = -2.09, p_B = .130$, see Figure 1), but the trend is in the right direction. Thus, the symbolic distance effect in the route and survey task can be explained by metric distance differences. Importantly, the symbolic distance effect in the route task, but not in the survey task, can also be explained by text distance differences. There were no significant Encoding Perspective differences for the accuracy data.

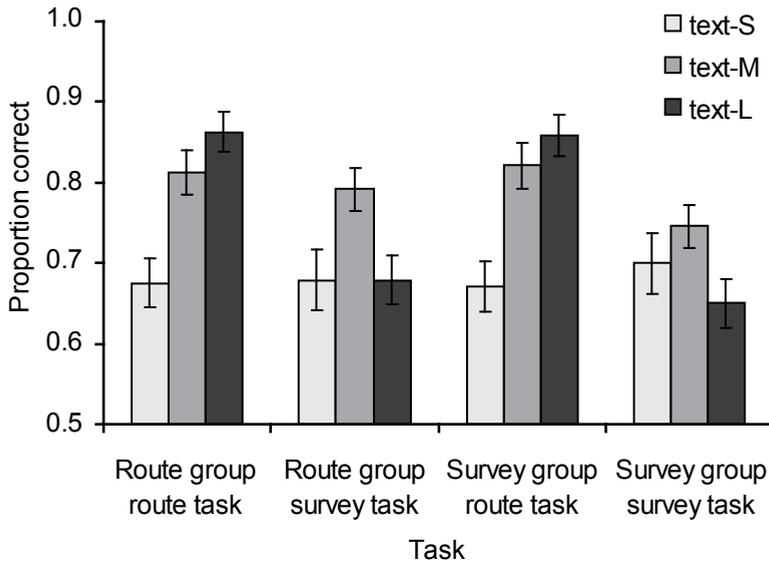


Figure 2 Proportion correct trials as a function of task and group in text distance differences.

Response times

The behavioral results from the mean response times are shown in Figure 3 for the route and survey task. Again we found a significant main effect of Task ($F(1,38) = 17.72, p < .001$). Response times on the survey task were slower ($M = 2657\text{ms}, SE = 67$) than response times on the route task ($M = 2412\text{ms}, SE = 68$). This result in combination with the fact that accuracy was higher in the route task suggests that there was no overall speed-accuracy trade-off.

The main effect of Distance Difference was also significant ($F(2,76) = 14.42, p < .001$). Responses to Small distance differences were slower ($M = 2638\text{ms}, SE = 69$) but did not differ significantly from responses to Medium distance differences ($M = 2570\text{ms}, SE = 67, t(38) = 1.38, p_B = .527$). However, responses to Large distance differences were significantly faster ($M = 2398\text{ms}, SE = 67, t(38) = 3.90, p_B = .001$).

There was a significant interaction between Distance Difference and Task ($F(2,76) = 10.09, p < .001$). Similar to the accuracy results the response time patterns differed slightly for the survey and route task. For the survey task the response times on the Small distance differences ($M = 2751\text{ms}, SE = 81$) did not differ from those on the Medium distance differences ($M = 2798\text{ms}, SE = 90, t(38) = -.62, p_B = 1.000$). In line with the symbolic distance effect response times on the Large distance differences were significantly faster ($M = 2423\text{ms}, SE = 75, t(38) = 5.00, p_B < .001$, see Figure 3). For the route task the response times on the Small distance differences were significantly higher ($M = 2525\text{ms}, SE = 86$) than the Medium distance differences ($M = 2340\text{ms}, SE = 61, t(38) = 2.94, p_B = .017$), which is in line with the symbolic distance effect. On the other hand, the response times on the Large distance difference did not significantly differ from the Medium ($M = 2372\text{ms}, SE = 75, t(38) = -.75, p_B = 1.00$). See Figure 3 for an illustration of these patterns. There were no significant Encoding Perspective differences for the response time data.

The analysis using text distance as the dependent variable revealed a significant main effect of Task ($F(1,38) = 21.41, p < .001$). Similar to the metric analysis response times on the route task were faster ($M = 2392\text{ms}, SE = 66$) than on the survey task ($M = 2651\text{ms}, SE = 70$, see Figure 4). The main effect of Distance Difference was marginally significant ($F(2,76) = 2.78, p = .069$). However, pairwise comparisons revealed no significant differences between the distance differences (all t s < 2.17). The interaction between Distance Difference and Task was also not significant ($F(2,76) = .08, p = .920$). Thus, the symbolic distance effect in the response times on both tasks can be explained by metric distance differences, but not by text distance differences.

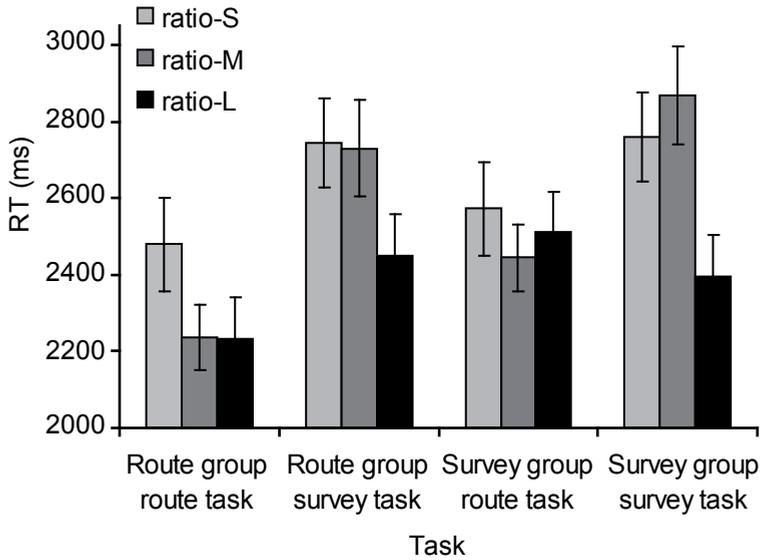


Figure 3 Mean response times as a function of task and group in metric distance differences.

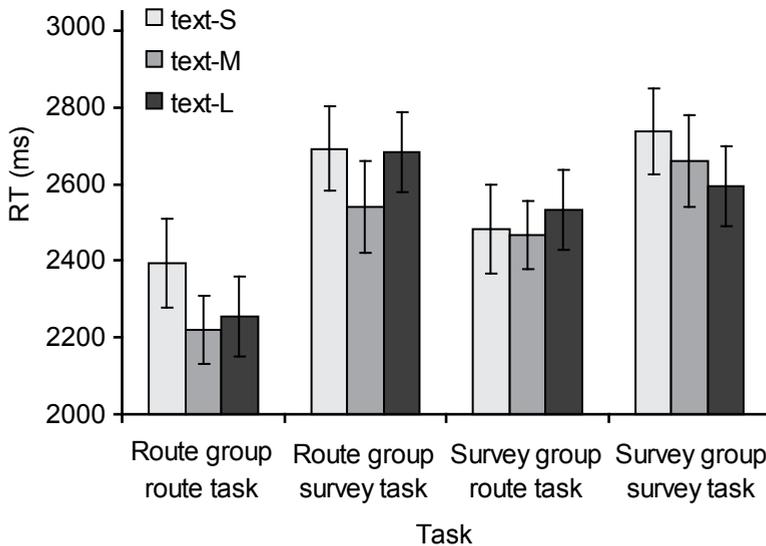


Figure 4 Mean response times as a function of task and group in text distance differences.

fMRI data

We will first discuss the commonalities between the different groups. On a group level we examined the conjunction of the route task and the survey task. The contrasts used in the conjunction analysis for the Route group were *route-control task*, and *survey-control task*. By contrasting the experimental tasks with the control task prior to the conjunction analysis the resulting activation pattern was specific for performing a spatial distance comparison task in the Route group. This conjunction analysis revealed bilateral activation in the middle occipital gyrus (see Figure 5A and Table 1 for details). There was also bilateral precuneus activation, however the activation was more pronounced in the right hemisphere. Finally, both groups showed significant activation in the left parahippocampal gyrus.

The same conjunction analysis was also performed for the Survey group. This analysis revealed areas that were significantly activated in both the route and the survey task (see Figure 5B and Table 1). The activation pattern was very similar to the pattern found in the Route group. Again bilateral middle occipital gyrus activation was found, as well as bilateral precuneus and left parahippocampal gyrus activation. In addition there was a small cluster of activation in the left middle frontal gyrus.

Table 1 Task activation of the conjunction of route-control task and survey-control task for the Route and the Survey group.

Region	L/R	Cluster Size	Peak voxel Z-score	MNI coordinates (mm)		
				x	y	z
<i>Route group</i>						
Middle occipital gyrus	L	13	6.09	-36	-80	32
Middle occipital gyrus	R	40	6.99	44	-72	28
Precuneus	L	11	5.19	-12	-64	20
Precuneus	R	46	6.57	16	-56	20
Precuneus	R	40	5.95	8	-64	56
Parahippocampal gyrus	L	9	5.87	-28	-44	-8
<i>Survey group</i>						
Middle occipital gyrus	L	45	9.76	-40	-76	28
Middle occipital gyrus	R	73	7.72	40	-76	32
Precuneus	L/R	108	7.18	16	-56	20
Precuneus	L	88	5.84	-4	-60	56
Parahippocampal gyrus	L	5	5.30	-28	-44	-8
Middle frontal gyrus	L	3	4.77	-24	0	48

Whole brain: $p < .05$, FWE-corrected.

Another interesting analysis is the conjunction on task level. Since both groups performed the same tasks we can also analyze the areas activated for each task that were in common for both encoding perspective groups. The conjunction analysis for the route task was based on the *route-control task* contrast for the Route group, and *route-control task* contrast for the Survey group (see Figure 6A and Table 2 for details). This analysis revealed significant activation for the route-task compared to the

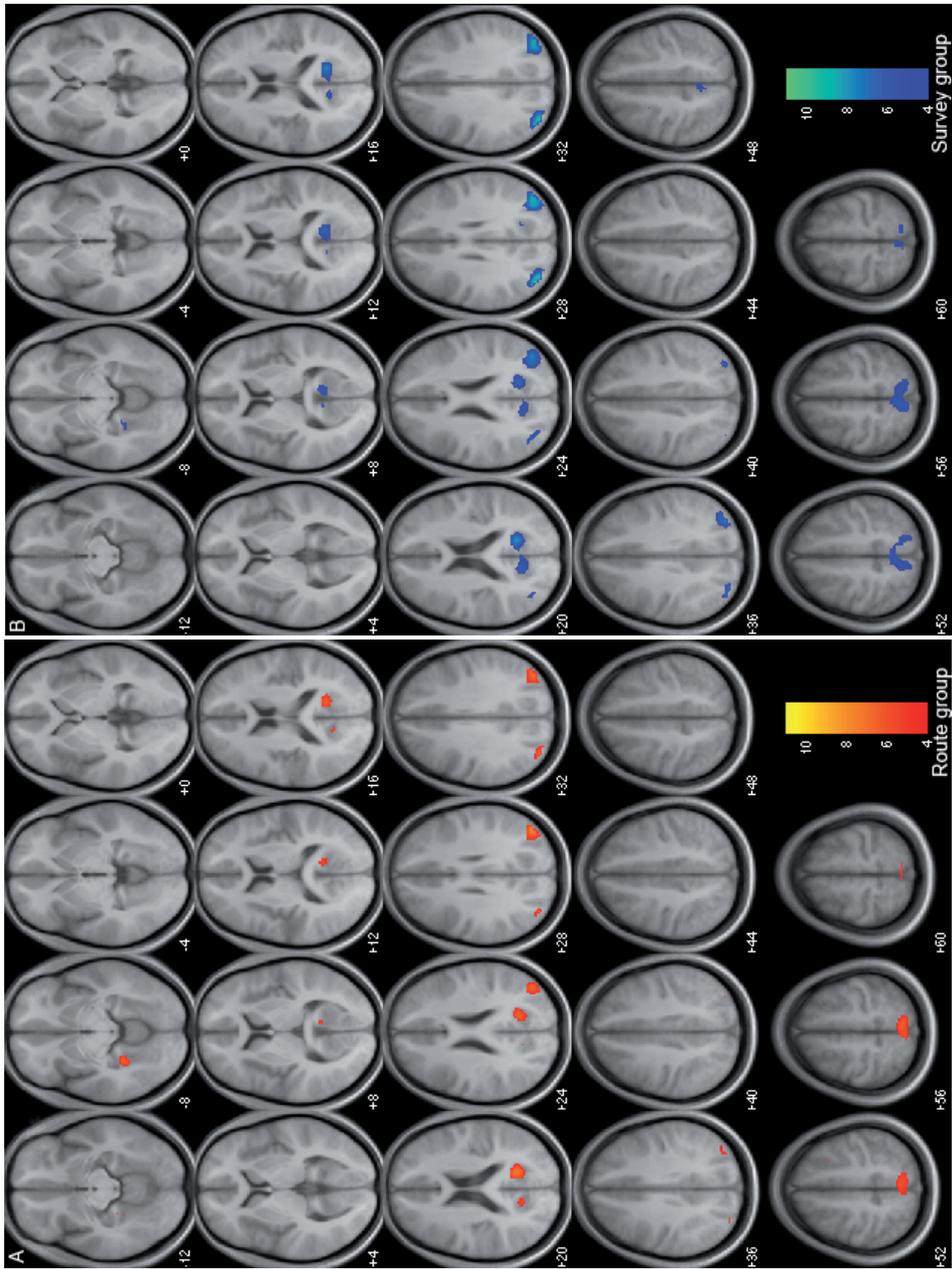


Figure 5 Conjunction of route-control task and survey-control task. A) conjunction results for the Route group. B) conjunction results for the Survey group.

control task for both the Route and the Survey group in bilateral middle occipital gyrus and precuneus. In addition there was significant activation in the left parahippocampal gyrus.

A similar analysis was performed for the survey task in which the *survey-control task* contrast for the Router group, and *survey-control task* contrast for the Survey group were tested (see Figure 6B and Table 2 for details). Again a very similar network of activation was found as for the route task. This conjunction analysis also revealed bilateral middle occipital gyrus and precuneus activation in both encoding perspective groups. Furthermore, there was activation in the left parahippocampal gyrus and the left middle frontal gyrus.

Table 2 Task activation of the conjunction of the Route and the Survey group for route-control task and survey-control task

Region	L/R	Cluster Size	Peak voxel Z-score	MNI coordinates (mm)		
				x	y	z
<i>route-control task</i>						
Middle occipital gyrus	L	14	6.36	-36	-80	32
Middle occipital gyrus	R	40	6.99	44	-72	28
Precuneus	L	16	5.72	-12	-60	20
Precuneus	L/R	35	5.54	0	-64	56
Precuneus	R	62	6.98	16	-60	20
Parahippocampal gyrus	L	8	5.87	-28	-44	-8
<i>survey-control task</i>						
Middle occipital gyrus	L	15	6.34	-36	-80	28
Middle occipital gyrus	R	50	7.82	44	-72	28
Precuneus	L	9	5.19	-12	-64	20
Precuneus	R	44	6.57	16	-56	20
Precuneus	R	49	5.45	12	-64	56
Middle frontal gyrus	L	20	5.48	-24	-4	52
Parahippocampal gyrus	L	4	5.30	-28	-44	-8

Whole brain: $p < .05$, FWE-corrected.

Since the patterns of activation found in the conjunction analyses on group and task level seemed highly similar an overall conjunction analysis was conducted. In this analysis a conjunction of *route-control task*, *survey-control task* contrasts for the Route group, and *route-control task*, *survey-control task* contrasts for the Survey group was performed. The overall conjunction revealed that bilateral middle occipital gyrus, precuneus and the left parahippocampal gyrus were activated in both spatial tasks in both the Route and the Survey group (see Figure 7 and Table 3 for details).

In order to assess the differences between tasks for the Route and the Survey contrasts between the route and survey task were computed. There were no significant differences between the two spatial tasks in the Route group and neither in the Survey group. When the differences between groups were assessed for each spatial task there were also no significant differences.

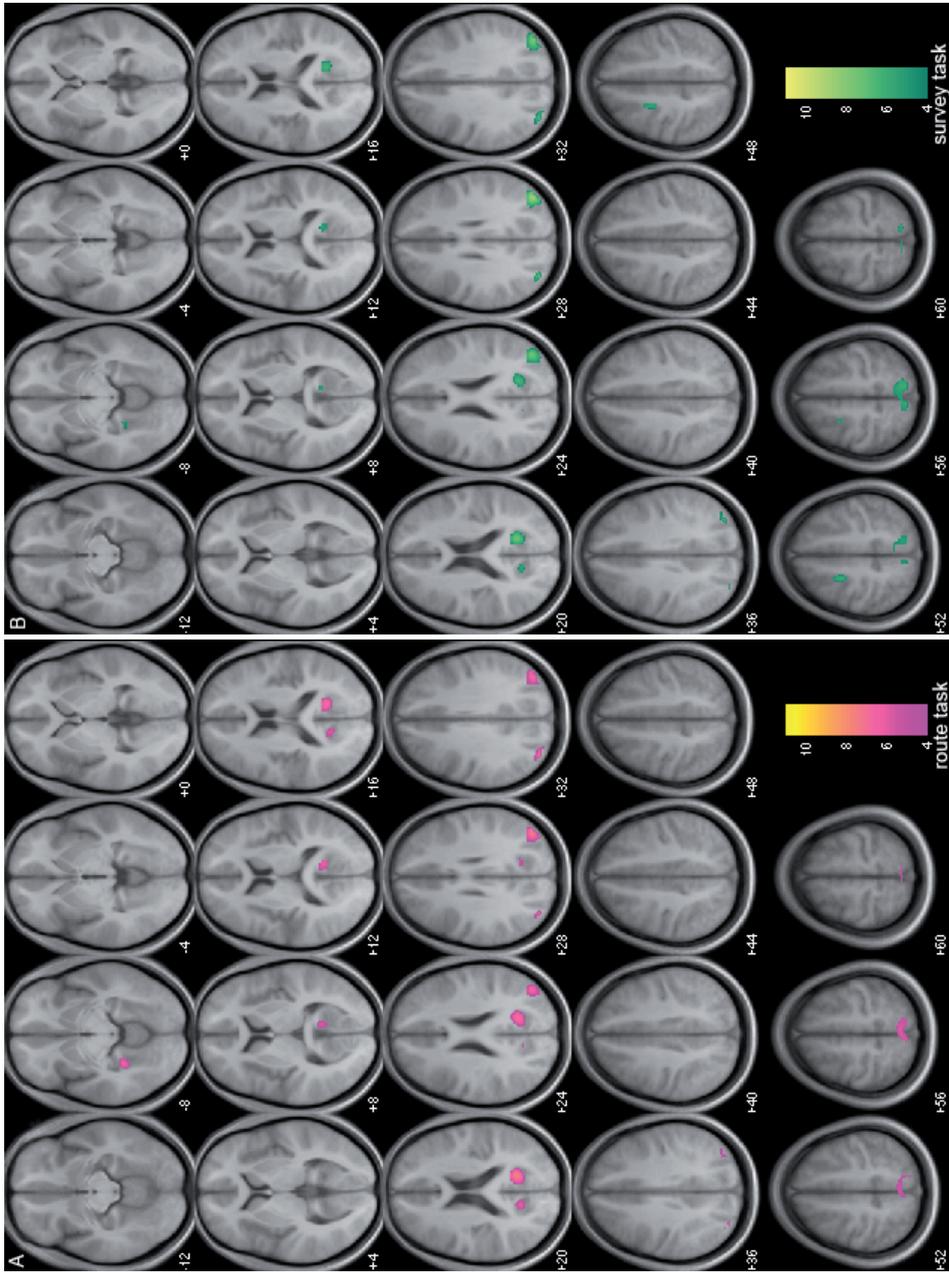


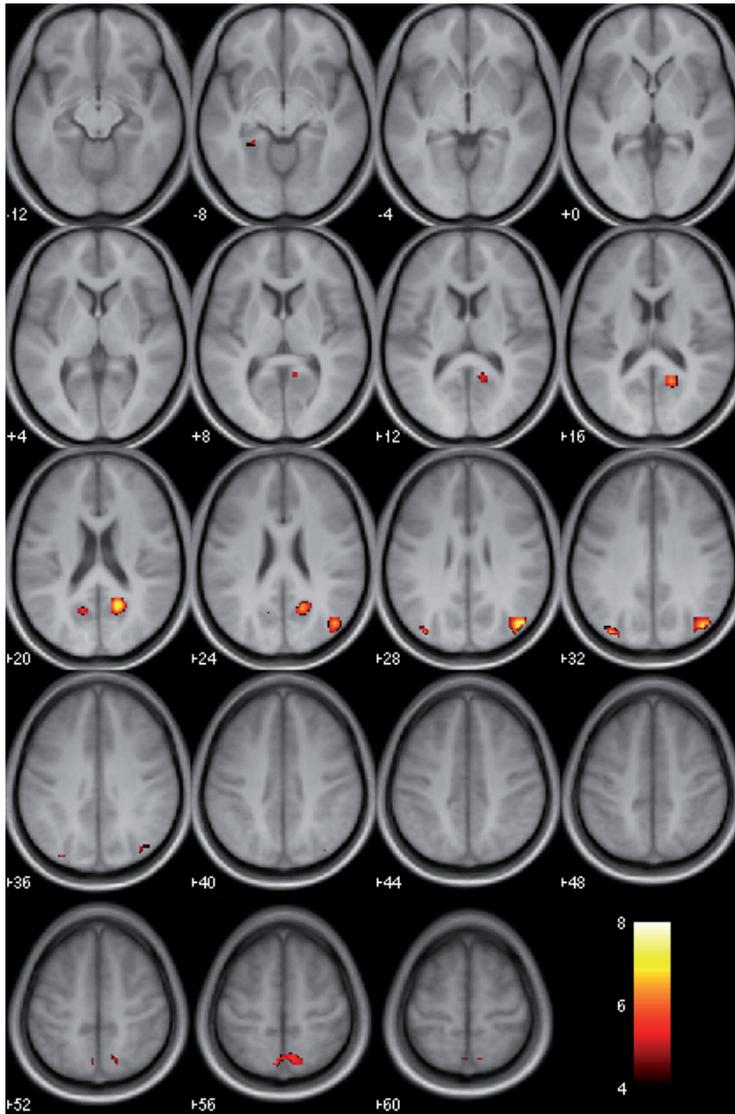
Figure 6 Conjunction of the Route and Survey group on an experimental task vs. the control task.

A) conjunction results for the route task. B) conjunction results for the survey task.

Table 3 Task activation of the overall conjunction of the Route and the Survey group on the route-control task and survey-control task

Region	L/R	Cluster Size	Peak voxel Z-score	MNI coordinates (mm)		
				x	y	z
Middle occipital gyrus	L	13	6.09	-36	-80	32
Middle occipital gyrus	R	40	6.99	44	-72	28
Precuneus	L/R	26	5.35	-4	-68	56
Precuneus	L	9	5.19	-12	-64	20
Precuneus	R	44	6.57	16	-56	20
Parahippocampal gyrus	L	4	5.30	-28	-44	-8

Whole brain: $p < .05$, FWE-corrected.

**Figure 7** Conjunction result for the Route and Survey group on the route and survey task vs. control task.

Discussion

The current study aimed to elucidate the neural networks activated by a route task, drawing on egocentric information, and a survey task, drawing on allocentric information, in two groups with different initial learning perspectives. Two groups learned a complex configuration of a zoo environment from a route or a survey description. Subsequently, both groups performed a route, survey and control task. After the scanning session was finished participants drew the configuration of the zoo in the outline they had viewed at the beginning of the experiment. These map drawings indicated that participants were able to correctly remember all the animals in the zoo and their relative locations. This was true for the survey group as well as the route group.

The behavioral results confirmed that participants had built up a spatial mental representation of the configuration with analogue spatial information. Both performance measures, proportion correct trials and response times, indicated the presence of a symbolic distance effect. However, this symbolic distance effect was not strictly linear, as would have been expected (Denis & Zimmer, 1992), but also contained a significant quadratic component (revealed by the within-subject contrasts). This quadratic component was also present in the slightly different patterns for the route task and the survey task. The route task performance on both measures was worse for the Small distance differences (i.e. lower proportion correct and slower response times) and better for the Medium and Large distance differences. In the survey task performance on both measures was worse for the Small and Medium distance differences and better for the Large distance differences.

When interpreting the behavioral results, it is important to consider the role of the initial textual inputs. Although previous evidence has suggested that listeners will automatically generate a mental representation of a complex environment it is not unlikely that they will also remember (parts of) the literal description. We have therefore tested whether an analysis based on text distance differences would also yield a symbolic distance effect. Only in the accuracy data we observed a partial influence of a textual representation. The symbolic distance effect on the accuracy data in the route task can be explained by metric distance differences, but also by text distance differences. This was true for the Route and the Survey group. Although the survey description was different from the route description in that it had a hierarchical build up, there still was some overlap in the order of presenting the different quadrants. Thus, text distance might have been helpful to some extent even for the survey group on the route task. On the other hand, on the survey task text distance did not support performance and did not result in a distance comparison effect. A symbolic distance effect was only found for the metric distance analysis, suggesting that participants had used metric information to solve this task.

Interestingly, there were no differences between the two different learning perspective groups on the mean accuracy and response times data. This suggests that the route and the survey group used a functionally equivalent spatial mental representation which could have been sensitive to the type of spatial knowledge that needed to be accessed rather than to the type of knowledge that had been available during the learning phase. Although we found subtle differences in response patterns

between the route and the survey task it was not very strong. The general pattern of improving performance with increasing distance difference was found in both tasks, however the non-linear trend varied slightly between the tasks. This could have been due to the fact that different types of spatial knowledge were addressed, or could have been due to other task-related differences such as specific categorization boundaries. Regardless of this, the behavioral results showed high similarity between the route and the survey group and a similar global pattern for the route and survey task. These findings support the idea that over-learning yields a perspective independent spatial mental representation and fits with a supramodal representation.

So far, the behavioral evidence has demonstrated that the constructed mental representations from route and survey perspectives contain metric spatial detail. The next step was to determine whether participants who had learned a configuration from a route description employed the same neural network as those who had learned a configuration from a survey description. The neuro-imaging results revealed activation of similar neural networks for both groups on both tasks. The conjunction analyses on group and task level resulted in overlapping networks. This was confirmed by the overall conjunction analysis. The areas that were active in all conditions were bilateral middle occipital gyrus, precuneus and the left parahippocampal gyrus. These areas of activation have been reported previously in related research on visual paradigms.

The middle occipital gyrus has been associated with processing landmark information. Committeri et al. (2004) reported bilateral middle occipital gyrus activation during their landmark-centered condition where participants judged which target object was closer to the centre of the main landmark. The left middle occipital gyrus showed an increased response to navigationally relevant landmarks in the study by Janzen and Weststeijn (2007). These findings were obtained from virtual reality paradigms. However, middle occipital gyrus activation has also been found after actual navigation. For example, Rosenbaum et al. (2004) tested inhabitants of the city of Toronto on a number of spatial tasks related to the city. They found increased left middle occipital gyrus response to landmarks compared to baseline. Ghaem et al. (1997) also reported left middle occipital gyrus activation which was specifically activated during mental navigation of route segments in a previously navigated environment. Finally, mental scanning of a representation built up from a verbal description also activated the middle occipital gyrus extending into the angular gyrus (Mellet et al., 2002). The current results also involved the processing of landmarks. Participants were asked to judge distances between pairs of landmarks (animal names referring to locations in the zoo), which seems to be related to activity in the middle occipital gyrus.

The results discussed above demonstrate that verbal information about landmarks can activate the middle occipital gyrus similar to visual and navigational inputs. We might argue that the activity related to landmarks found in the middle occipital gyrus is due to processing visual landmarks or visual imagery of these landmarks. However, Mellet et al. (2002) contrasted mental scanning of a representation built up from a verbal description with actual map studying. The middle occipital gyrus was significantly activated in the text condition and deactivated in the map studying condition. This suggests that the activation found may be linked to the spatial relevance of the landmarks,

rather than the visual representation of the objects. In the current experiment only verbal information was provided. Since no visual condition was included we cannot rule out that the participants had used visual imagery of the landmarks which could have resulted in middle occipital gyrus activation. However, if subjects had resorted to a visual imagery strategy we should have observed a larger visual processing network, including the primary visual cortex (Kaski, 2002; Kosslyn et al., 1999; Mazard et al., 2005). On the contrary, we did not observe additional visual imagery activation indicative of a visual imagery strategy. On the other hand, the wide variety of modalities that have been found to generate activation in the middle occipital gyrus could support the idea that spatial information is represented supramodally (Struiksma et al., 2009) and that this activation is related to the spatial relevance of the landmarks rather than the visual characteristics.

Another area that has been associated with processing landmarks is the parahippocampal gyrus (Aguirre & D'Esposito, 1997; Epstein et al., 2007; Galati et al., 2010; Janzen & Weststeijn, 2007; Latini-Corazzini et al., 2010; Mellet et al., 2000). Aguirre et al. (1996) have reported bilateral parahippocampal gyrus activation linked to processing topographic information. Maguire et al. (1996) also found bilateral parahippocampal gyrus activation using real-world stimuli. In a star-maze virtual reality experiment Maguire et al. (1998b) found only right parahippocampal gyrus activity. The authors suggested that these differences in laterality might have reflected different strategies. A common distinction is that the left hemisphere is more involved in verbal tasks while the right hemisphere is more involved in spatial tasks (cf. Jager & Postma, 2003; Kemmerer & Tranel, 2000). In this study activity in the parahippocampal gyrus was on the left. Given the spatial nature of the current task relying on a spatial representation, right parahippocampal gyrus activity would have been more obvious. Arguably, the verbal nature of the task resulted in left lateralized parahippocampal activation, while the star-maze by Maguire et al. (1998b) required a pure spatial strategy and the large-scale environments allowed both verbal and spatial strategies. We might speculate that parahippocampal gyrus activation is modality dependent.

Lastly, we observed bilateral precuneus activation. Activity in this area has also been reported in earlier spatial studies. The precuneus was activated during mental navigation, which could have been due to visuo-spatial imagery or inspection of the mental image (Aguirre et al., 1996; Burgess et al., 2002; Ghaem et al., 1997; Maguire et al., 1998b). However, Mellet et al. (2000) reported precuneus activation linked to mental navigation, but not during a mental map task. They suggested that the precuneus activity could have been due to stronger mental imagery in the mental navigation task which required mental exploration of actual navigation, possibly with higher visual details. On the other hand, the activation could have been due to processing specific detailed spatial information, rather than visual imagery. This has also been suggested by Wallentin et al. (2008b). They compared a purely linguistic spatial task to a visual spatial task and found overlapping activation in the precuneus. In the linguistic experiment the spatial task was contrasted with a relational task and a non-spatial task. All the tasks could have elicited visual imagery easily (e.g. the spatial task: "*was she turned towards him?*", the relational task: "*was she older than him?*", and the non-spatial task: "*was she blushing?*"). Because a visual imagery strategy could have been available in all three tasks

the contrast between the spatial task and the other tasks revealed specific spatial processing. This contrast revealed precuneus activation. As such, the precuneus could be associated with processing spatial information and this activity is not strictly visual or linguistic, but rather could be supramodal (Struiksma et al., 2009).

Taken together, the behavioural and functional imaging result reveal marked similarities between different types of tasks (egocentric and allocentric) and different learning conditions (route and survey descriptions). This suggests that after over-learning participants have built up a perspective independent spatial mental representation. The areas that were activated during these spatial tasks included the bilateral precuneus, linked to processing spatial information, and bilateral middle occipital gyrus and left parahippocampal gyrus, involved in processing landmarks. We did not observe distinct neural activation for the egocentric or allocentric components, nor was there a difference in performance between the groups on different tasks. Therefore, the current data do not support the *distinct representations* and *hierarchical* model. Although the behavioural data support a *perfect learning* model we cannot simply conclude this from the neuroimaging data. A very plausible alternative explanation is that the tasks are too similar and may rely heavily on imagery processes, which could be linked to the middle occipital gyrus and precuneus activity. A possible alternative for the egocentric route task could be an imaginary pointing task. Such a pointing task is solved using an egocentric representation (for a review see Burgess, 2006; Wang & Spelke, 2002) which could form a stronger contrast to a bird-flight distance comparison task, using an allocentric representation.

In conclusion, after over-learning of perspective dependent information, behavioural performance on spatial tasks that required different spatial judgments was very similar supporting a supramodal spatial representation. In addition the corresponding neural activation was very similar for both spatial tasks and both initial learning perspectives.

Grounded cognition



Chapter 8

Embodied representation of the body contains veridical spatial information

M. E. Struiksma, M. L. Noordzij & A. Postma

Quarterly Journal of Experimental Psychology, (in press)

This is a preprint of an article whose final and definitive form will be published in 'The Quarterly Journal of Experimental Psychology' © 2011 Copyright The Experimental Psychology Society reprinted by permission of Taylor & Francis Ltd, on behalf of The Experimental Psychology Society The Quarterly Journal of Experimental Psychology is available online at:
<http://www.informaworld.com/1747-0218>

Abstract

In two experiments the extent to which mental body representations contain spatial information was examined. Participants were asked to compare distances between various body parts. Similar to what happens when people compare distances on a real visual stimulus, they were faster as the distance differences between body parts became larger (Experiment 1), and this effect could not (only) be explained by the crossing of major bodily categories (*umbilicus* to *knee* vs. *knee* to *ankle*, Experiment 2). In addition, participants also performed simple animate/in-animate verification on a set of nouns. The nouns describing animate items were names of body parts. A spatial priming effect was found: Verification was faster for body part items preceded by body parts in close spatial proximity. This suggests automatic activation of spatial body information. Taken together, results from the distance comparison task and the property verification task showed that mental body representations contain both categorical and more metric spatial information. These findings are further discussed in terms of recent embodied cognition theories.

General introduction

When asked to estimate whether a new pair of jeans will fit we need to judge size and distance information of our own body. An interesting question is how well we are able to do this purely based on memory (i.e. without looking at or touching our own body). This task draws on a mental representation of the body, without perceptual input, referred to as either a somatopresentation (Longo et al., 2009) or a body image (e.g. Dijkerman & De Haan, 2007; Head & Holmes, 1912). It includes lexical-semantic knowledge about body parts and structural knowledge about the topology of the body (De Vignemont, 2009; Longo et al., 2009; Schwoebel & Coslett, 2005). Although the topological map contains visuo-spatial relationships between body parts there remains a close link to language (Longo et al., 2009; Semenza & Goodglass, 1985). Most likely the visuo-spatial relationships are stored in categorical relations such as “on top of” and “right side connected” (Laeng et al., 2003). It remains to be seen though whether a spatial organization with veridical distance information is maintained. The present study aimed to explore this possibility as well as to what extent this organization is automatically accessed.

Recently, Smeets et al. (2009) examined, as one of the first, whether distance information is actually preserved in our mental body representation. They used a so called distance comparison task (Moyer & Bayer, 1976) and compared females with high and low body shape concern. Participants were required to make a forced choice about which distance between body pairs was the longest (e.g. *hip-hip* vs. *ear-ear*). Smeets et al. found evidence for a symbolic distance effect, i.e. reaction time decreases, and accuracy increases, with increasing (physical) distance difference. For example, participants responded faster and more accurately in case of a *hip-hip* vs. *ear-ear* comparison than when contrasting *hip-hip* against *shoulder-shoulder*. This symbolic distance effect strongly suggests that participants used spatial information to solve the task. It remains to be seen whether this spatial information is really metric. It should be mentioned that Denis and Zimmer (1992) in their original paper on the distance comparison task argued that their participants were able to convert verbal descriptions of environments into mental representations that contained reliable metric properties similar (isomorphic) to the representations derived from visual experience with an actual map.

A point of criticism on the study by Smeets et al. (2009) regards their stimulus set. They included a range of horizontal distance pairs that were emotionally sensitive to females with high body shape concern, such as hip and thigh. They contrasted these items to emotionally insensitive items, such as eye and ankle. Smeets et al. found that emotionally sensitive pairs, where at least one distance was emotionally sensitive, were judged faster than insensitive pairs, especially in the shorter distance difference spectrum. As such, the symbolic distance effect was less pronounced for emotionally sensitive pairs compared to emotionally insensitive pairs. This effect was similar in high and low body shape concern groups. The authors provide an explanation for this odd finding: Perhaps female participants visualized the emotionally sensitive body parts more often, which speeds up access to these representations. Since no differences between high and low body shape concern participants were found an alternative explanation could be that spatial knowledge about

body parts is not only represented in a spatial manner but also in a semantic factual way (e.g. your conviction that “your hips are huge” makes you always judge the hips as the largest distance). As such, the symbolic distance effect obtained by Smeets et al. (2009) does not provide conclusive evidence for the presence of a spatial representation of the body. Therefore, in a first experiment we designed a distance comparison task without focussing on emotionally sensitive body parts but instead focussing on three different scanning orientations in order to disentangle a spatial representation and a semantic representation: a) horizontal distances b) vertical distances c) mixed horizontal and vertical distances. It might be argued that for comparisons along the horizontal or vertical orientation we cannot rule out the possibility that they are solved on the basis of semantic factual knowledge rather than by mental scanning of spatial distance properties. That is, we might simply know which widths or heights between body parts are the largest without needing to scan a particular mental distance (e.g. you know that your hips are wider than your ears). For a mixed orientation trial, comparing horizontal and vertical distances, we assume it is highly unlikely that the ranking of differently oriented distances is stored as a semantic fact. We rarely have to contrast a horizontal body distance (e.g. the width of your hips) to a vertical one (e.g. the length of your arm) in daily life. Consequently, the only way to achieve accurate performance in the mixed condition would be to really mentally scan the distance from one body part to the other. Thus the strongest evidence for a spatially organized mental body representation would follow from a symbolic distance effect in the mixed condition. Furthermore, if participants use semantic knowledge to solve the horizontal and vertical conditions, but need to perform an additional computation to solve the mixed condition, this would result in generally faster responses for the horizontal and vertical conditions, which do not require any computations, compared to the mixed condition.

The body distance comparison task requires active deliberation on body size information. However, theories on embodied cognition and perceptual simulations argue that language comprehension involves linguistic processing as well as automatic simulation of referents such as multimodal percepts, beliefs and actions (Barsalou, 1999; Barsalou et al., 2008; Zwaan, 1999). In line with these theories recent neuroimaging studies have demonstrated motor resonance when processing action related verbs, e.g. reading the word *kick* activated the leg-area in the motor cortex (for a review see Willems & Hagoort, 2007). Similarly, evaluating body parts could result in automatic activation of a mental body representation. Therefore, we added a second task (a property verification task) assessing the possibility of automatic activation of the underlying mental body representation. In the property verification task no conscious decision is required on any form of spatial information (Denis & Zimmer, 1992; Noordzij & Postma, 2005). Participants simply have to judge whether or not an item falls in a certain semantic class. In this case, the criterion property was animate versus in-animate, and all the animate items were body parts. By comparing the inter-trial relations, however, the presence of a so called spatial priming effect can be established: Faster responses to target items (e.g. *nose*) when preceded by another target item close in space (e.g. *chin*) than when preceded by a target item far in space (e.g. *knee*). Importantly, these effects occur unconsciously and are less sensitive to strategies (McNamara, 1986) than e.g. image scanning tasks (Kosslyn et al., 1978).

Property verification requires participants to use deeper processing which could result in situated simulations with access to spatial information, instead of relying on superficial word-level representations as in lexical decision tasks (Barsalou et al., 2008; Glaser, 1992). We expected that if the mental body representation is immediately accessed, a spatial priming effect should occur signified by faster reaction times for target body parts when primed by another body part near in space than when primed by a body part far in space. However, if the spatial information is unavailable or inaccessible participants can only rely on semantic information resulting in similar reaction times for near and far primed target body parts.

We describe the results from two experiments. The first experiment consisted of a distance comparison task and a property verification task to examine the spatial organization of the mental body representation. However, this organization could involve strong metric information or depend on more coarse spatial categories. In the second experiment the paradigm was adapted to test whether the crossing of major bodily categories could better explain the distance comparison and spatial priming effects.

Experiment 1

In the first experiment participants performed a distance comparison task and a property verification task. Both tasks provide information about the organization of body parts. The distance comparison task explicitly instructs participants to use their mental representation about their own body. The property verification task, on the other hand, measures the mental body representation in an implicit manner, through possible priming effects. Participants started with the property verification task to ensure that they had not consciously activated their body image. Together these tasks may shed light on the hypothesized spatial organization of the mental body representation. In order for participants to focus on their own mental body representation they were blindfolded and the stimuli were presented auditorily.

Methods

Participants

A total of 28 Utrecht University students participated for course credits or monetary compensation. Two participants had misunderstood the instructions in one of the tasks as demonstrated by performance far below chance-level and were excluded from the analysis which resulted in a group of 26 participants (13 male; age $M = 21.69$, $SD = 1.72$). Prior to the experiment all participants signed an informed consent, which was approved by the local ethics committee.

Design

All participants were blindfolded during the auditory experiment. Participants were led into the experiment room by the experiment leader and were seated in front of a table with a computer with two external loudspeakers.

Participants started with the property verification task. The task was to judge whether an item was animate or inanimate. Half of the participants pressed the left button on a response box for animate items and the other half pressed the right button for animate items. Participants were instructed to respond as quickly as possible without compromising accuracy. After the instruction participants received 12 practice trials, where animate items were animal parts (e.g. *wing*). Participants received feedback about their performance. During the task the animate items were body parts.

Subsequently, participants performed the distance comparison task in which they were asked to compare two distances between pairs of body parts and to judge which pair contained the largest distance. When the first pair was the largest distance participants had to press the left key, on the response box, when the second pair was the largest distance they had to press the right key. Again, participants were instructed to respond as quickly as possible without compromising accuracy. Prior to the distance comparison task participants were instructed to imagine their own body as if they were watching a picture of themselves, standing upright with their arms hanging down and their feet together. This instruction was repeated three times during the experiment. After the instruction participants received 10 practice trials with feedback about their performance. During the actual task no feedback was provided.

Stimulus material and procedure

Although the participants first performed the property verification task we will start with discussing the distance comparison task because of theoretical relevance following the line of arguments from the introduction.

Distance comparison task The task consisted of 92 trials divided over three orientations: Horizontal, Vertical and Mix (see Table 1, for examples). Nested within the orientations was the distance difference: Small, Medium and Large distance differences, based on the ratios between the two body distance pairs within a trial. The ratios were computed using a standard body and by dividing the smallest by the largest distance and sorted into Small, Medium and Large distance differences (see Table 1). In a pilot experiment 25 participants were photographed standing in an upright position in accordance with the instruction. For each participant the individual ratios were computed and compared against the ratios from the standard body. The pilot experiment showed that in general individual ratios provided the same categories as the standard body ratios, therefore the standard body ratios are reported in Table 1.

There were four sessions in which trials were presented randomly. The sessions were separated by a reminder of the imagery instruction. A single trial started at 0ms with a short beep (100ms), at

1100ms the first word of pair one was presented (700ms), followed by a short interval (300ms), at 2100ms the second word of pair one was presented (700ms), followed by a short interval (800ms), at 3600ms the first word of pair two was presented (700ms), followed by a short interval (300ms), at 4600ms the second word of pair two was presented (700ms), followed by a response interval. Participants had 4000ms to respond. The next trial started 500ms after the response, or after the response interval when no response was given. The words were recorded with Audacity (audacity.sourceforge.net) and presented auditorily through external loudspeakers with Presentation software 11.0 (Neurobehavioral Systems, Albany, CA).

Table 1 Examples of trials with different orientations and different distance differences and the ratio bins determining the different categories for each orientation.

<i>Orientation</i>	<i>Distance difference</i>	<i>Example trial</i>		<i>Range</i>
Horizontal	Small	hip-hip	elbow-elbow	0.69-1.00
	Medium	shoulder-shoulder	elbow-elbow	0.50-0.69
	Large	shoulder-shoulder	ear-ear	0.30-0.50
Vertical	Small	nose-navel	hip-knee	0.62-1.00
	Medium	hip-knee	forehead-chin	0.37-0.62
	Large	nose-navel	forehead-chin	0.25-0.37
Mix	Small	elbow-wrist	knee-knee	0.76-1.00
	Medium	elbow-wrist	ankle-ankle	0.57-0.76
	Large	ankle-ankle	nose-navel	0.14-0.57

Property verification task For the auditory property verification task a list of 72 objects (body parts or non-living objects) was created. Each object required a response. The trials were presented in a fixed order which determined the prime-target relation. There were two types of prime-target relations: Near (shoulder-ear) and Far (wrist-ankle). In total there were eight prime-target pairs for each spatial distance. The words were presented with the same setup as used in the distance comparison task. A single trial consisted of the presentation of a word, followed by a response interval of 4000ms maximum. The next trial started 1000ms after the response, or at the end of the response interval in the case of no response.

Data analysis

For the distance comparison task we analyzed two behavioral measures: performance measured by the proportion correct trials and mean response times to correct verifications. Behavioral data were analyzed using SPSS (SPSS for Windows, Rel. 16.0.2 2008. Chicago: SPSS Inc.) with a 3×3 repeated measures Analysis of Variance (ANOVA). The within-participant factors were Orientation (Horizontal, Vertical or Mix) and Distance (Small, Medium or Large difference). Further analysis by means of pairwise comparisons used a significance level corrected for multiple comparisons with the Bonferroni method. SPSS multiplies the p -value with the Bonferroni multiplier instead of

dividing α by the Bonferroni multiplier. The results are, however, equal and this method corrects for multiple comparisons. We will denote the Bonferroni corrected p -values by p_B .

For the spatial property verification task we also analyzed two behavioral measures: performance measured by the proportion correct trials and mean response times to correct targets. Behavioral data were analyzed with a one-way repeated measures ANOVA. The within-participant factor was Distance (Near: target is close to the prime, Far: target is far from the prime). The results are reported using the same procedure as in the distance comparison task.

Results

Distance comparison task

Performance The overall proportion of correct trials was high ($M = .816$, $SE = .012$), but differed slightly for the three orientations. The 3 (Orientation) \times 3 (Distance) repeated measures ANOVA on the proportion correct answers revealed a main effect of Orientation ($F(2,50) = 11.22$, $p < .001$). Pairwise comparisons showed that performance was slightly lower on Horizontal trials ($M = .783$, $SE = .018$) compared to Vertical trials ($M = .863$, $SE = .011$, $t(25) = -4.16$, $p_B = .001$). Performance on Vertical trials was slightly higher compared to Mix trials ($M = .801$, $SE = .016$, $t(25) = 4.07$, $p_B = .001$). There was also a main effect of Distance ($F(2,50) = 86.06$, $p < .001$). Pairwise comparisons revealed that performance was higher for Large distance differences ($M = .918$, $SE = .013$) compared to Medium distance differences ($M = .846$, $SE = .015$, $t(25) = 4.00$, $p_B = .001$) and higher for Medium distance differences compared to Small distance differences ($M = .684$, $SE = .018$, $t(25) = 10.19$, $p_B < .001$). Furthermore, the interaction between Orientation and Distance was significant ($F(4,100) = 5.13$, $p = .001$) and was also further analyzed by means of pairwise comparisons between distance differences for each Orientation. These showed that performance increased with distance difference (see Figure 1). For all orientations participants performed significantly better on Medium compared to Small distance differences (all $p_B < .010$). For Mix trials performance was significantly better on the Large distance differences compared to Medium ($p_B = .002$), which can be summarized as a symbolic distance effect. For Horizontal and Vertical trials the difference between Medium and Large distance differences was not significant ($p_B = .321$ and $p_B = .165$ respectively), but pointed in the predicted direction, see Figure 1.

Response times Another 3 (Orientation) \times 3 (Distance) repeated measures ANOVA was performed to analyze the mean response times to correct verifications. Again a main effect of Orientation ($F(2,50) = 5.84$, $p = .005$) was found. Participants responded faster on Horizontal trials ($M = 1389.54\text{ms}$, $SE = 71.09$) compared to Vertical trials ($M = 1523.69\text{ms}$, $SE = 66.84$, $t(25) = -2.58$, $p_B = .048$) and compared to Mix trials ($M = 1505.08\text{ms}$, $SE = 68.78$, $t(25) = -3.85$, $p_B = .025$). The faster responses and lower accuracy on the Horizontal trials might suggest a speed-accuracy trade-off. However, closer inspection of the distance differences revealed that the fastest responses were given to the Large distance differences, while the accuracy was highest in this condition. This pattern argues against a speed-accuracy trade-off. The main effect of Distance was also significant

($F(2,50) = 31.22, p < .001$). Pairwise comparisons showed that participants responded faster to Large distance differences ($M = 1314.21\text{ms}, SE = 61.32$) compared to Medium ($M = 1477.34\text{ms}, SE = 65.29, t(25) = -4.17, p_B = .001$) and faster to Medium distance differences compared to Small distance differences ($M = 1626.77\text{ms}, SE = 77.38, t(25) = -3.50, p_B = .005$). The interaction between Orientation and Distance was not significant ($F(4,100) = 1.62, p = .176$). All results are summarized in Figure 2.

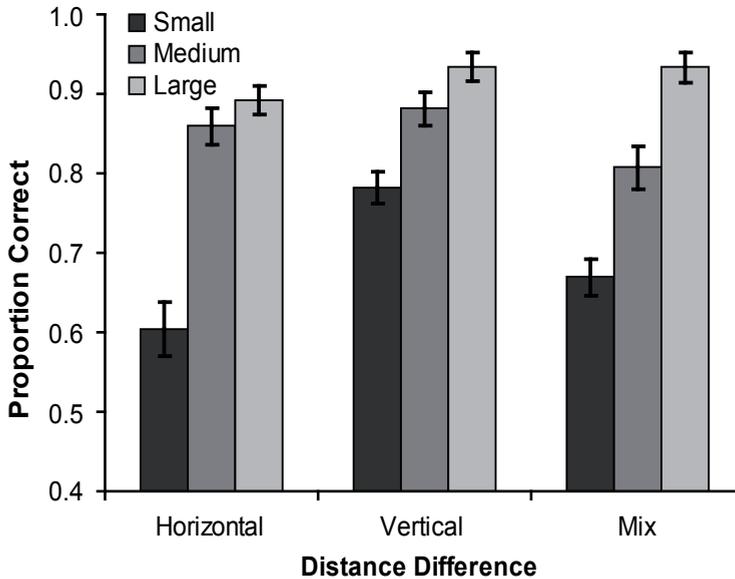


Figure 1 Mean proportion correct responses on the distance comparison task, split up for three orientations. Chance level is 0.5. Error bars indicate standard error of the mean.

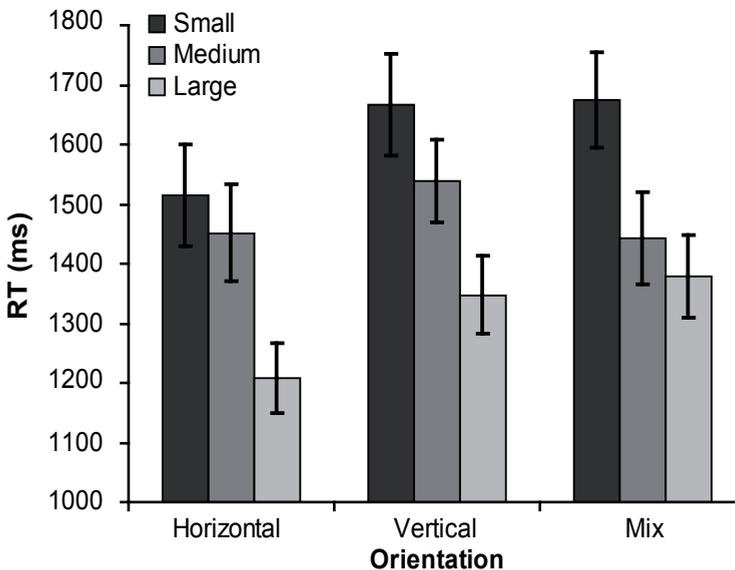


Figure 2 Mean reaction times on the distance comparison task, split up for three orientations. Error bars indicate standard error of the mean.

Property verification task

Performance and response times The overall proportion of correct trials was very high ($M = .98$, $SE = .01$). The one-way repeated measures ANOVA on the proportion correct trials did not reveal an effect of Distance ($F(1,25) = .14$, $p = .713$).

The analysis of the mean response times to correct body parts revealed an effect of Distance ($F(1,25) = 7.71$, $p = .010$). Participants responded significantly faster to Near targets ($M = 772.37\text{ms}$, $SE = 25.60$) than to Far targets ($M = 835.03\text{ms}$, $SE = 38.91$).

Discussion

The purpose of this study was to determine whether the mental representation of the body is spatially organized. In the body distance comparison task we found a symbolic distance effect where responses were faster and more accurate with increasing distance difference. This effect indicates that participants found it easier to compare large distance differences between pairs of body parts, and is in line with previously obtained results when participants compared actual metric distances on a visual map (Denis & Zimmer, 1992). The fact that the symbolic distance effect emerged especially in the Mix condition, where participants had to compare a Horizontal and Vertical distance, suggests that participants could no longer rely on semantic knowledge to solve the task, since different postures might alter the ranking of distance differences. Instead they used spatial information and needed to actively manipulate two differently oriented distances. The active manipulation strategy seems to have been reflected on the Vertical orientations as demonstrated by the presence of the symbolic distance effect in this condition as well. There was also a symbolic distance effect in the Horizontal orientation. Additionally, responses in the Horizontal orientation were faster than in the other two orientations. As we hypothesized, this could be due to a semantic strategy. Certainly the presence of the symbolic distance effect rules out a complete semantic strategy, but like the results reported by Smeets et al. (2009) there might have been some semantic influence. Overall, the presence of the symbolic distance effect in all three conditions suggests that our body knowledge is spatially organized.

In order to assess the possibility of automatic activation of the underlying mental body representation we included the property verification task for which we found that participants responded significantly faster to targets close in space than to targets far in space. This spatial priming effect provides converging evidence that participants use a mental representation of the body with preserved spatial information. Moreover, even without being aware participants activate spatial information about the body in line with theories on embodied cognition.

The results of both the distance comparison task and property verification task suggest that body information is spatially organized. However, they do not inform us whether a strong, exact metric is employed or rather coarser categorical spatial information is used¹. In parallel to the levels of taxonomy based on the principles of categorization by Rosch et al. (1976), Andersen suggested a partonomy: a division of the body into four categories: head, trunk, arm and leg (Andersen,

¹ We would like to thank an anonymous reviewer for this suggestion.

1978). Partonomy represents a hierarchical categorical ordering of body parts. For example, the basic concept *knee* has *leg* as its superordinate concept, *patella* as its subordinate concept and *ankle* as its co-ordinate concept. When we apply this framework to our property verification task it could be that the near distances were all within a superordinate category, while the far distances were across superordinate categories. Notice that we assume that a categorical organization is not merely semantic, but rather is spatial at a coarse grained level, i.e. body parts within the same category are spatially equivalent though relative distance differences emerge when crossing one, or more categories. The use of categorical spatial information could also have been present in the distance comparison task. Experiment 1 was not designed to distinguish between these two spatial alternatives. In order to further examine the categorical spatial possibility we therefore performed a second experiment in which we included two factors: Distance (Near, Far) and Category (Within, Across).

Experiment 2

In the first experiment we have shown that information about the spatial organization of the body is automatically available in the property verification task and deliberately accessed in the distance comparison task. However, this organization could involve either strong metric information or depend on more coarse spatial categories, or even a mix of the two. Adding a factor Category could elucidate the underlying characteristics of the spatial organization of the body.

For the distance comparison task new pairs of distances were constructed such that the factor Category could be added. We only used the vertical condition since it offered the most suitable combinations of body parts. The factor Distance again had three levels: Small, Medium and Large distance differences. These three levels were realized for pairs within a category (Within) and pairs across a category (Across). We used the partonomy by Andersen (1978) with four body categories: head, trunk, arm and leg. In Experiment 1 we have argued that it is unlikely that participants used semantic information in the mixed condition. Indeed the results support this idea. However, in the vertical condition participants could still rely on another strategy than a metric one. In this adapted experiment we manipulated the influence of categorical information. We may conjecture four possible response patterns. If participants indeed based their response on categorical information, then we would expect to find no symbolic distance effect, and an advantage for Within trials compared to Across trials. On the other hand, the use of metric information predicts a symbolic distance effect for both Within and Across trials. Alternatively, both metric and categorical information could be available yielding a symbolic distance effect but easier, i.e. faster and more accurate, for Within trials than for Across trials as demonstrated by a significant interaction. A pure semantic strategy would yield no symbolic distance effect and no categorical effect, which seems unlikely given the results of Experiment 1.

Experiment 2 also included the property verification task with the manipulation of the factor Category. For both Near and Far trials the combination of body parts was chosen either from within a body category or from across body categories. The same rationale as for the distance com-

parison task holds here yielding four predictions: categorical, metric, a combination of categorical and metric or semantic information. Furthermore, to make sure that participants did not activate their body image we also included animal part names, which were also animate.

Methods

Participants

A total of 28 Utrecht University students participated for course credits or monetary compensation. None of the participants had participated in Experiment 1. One participant misunderstood the instructions in one of the tasks, resulting in performance below chance level, and was excluded from the analysis which resulted in a group of 27 participants (11 male; age $M = 23.56$, $SD = 2.69$). Prior to the experiment all participants signed an informed consent, which was approved by the local ethics committee.

Design

The distance comparison task from Experiment 1 was adapted. The task only consisted of vertical trials in which Distance and Category were manipulated. The procedure of administering the tasks was identical to Experiment 1.

Stimulus material and procedure

Distance comparison task Participants received the same instruction as in Experiment 1. The task consisted of 48 trials divided over two categories: Within (pairs were within the same body category) and Across (pairs were from different body categories). Nested within the categories was the distance difference: Small, Medium and Large, based on the ratios between the two body distance pairs within a trial. The boundaries for the ratio bins that determined the three distance differences were the same for Within and Across trials. There were three sessions in which trials were presented randomly, the sessions were separated by a reminder of the imagery instruction. The presentation of the trials was identical to Experiment 1.

Property verification task For the auditory property verification task a list of 138 objects (body parts, animal parts or non-living objects) was created. There were two categories: Within (prime-target pairs were within the same body category) and Across (prime-target pairs were from different body categories). Nested within the categories were two distances: Near and Far. In total there were six prime-target pairs for each condition. In addition to the body parts there were 24 animal parts that also required a yes-response. The number of yes- and no-responses was equal. The trial presentation was identical to Experiment 1.

Data analysis

For the distance comparison task we analyzed two behavioral measures: performance measured by the proportion correct trials and mean response times to correct verifications. Behavioral data were analyzed with a 2×3 repeated measures ANOVA. The within-participant factors were Category (Within or Across) and Distance (Small, Medium or Large difference). The results are reported in the same way as in Experiment 1.

For the spatial property verification task we analyzed performance measured by the proportion correct trials and mean response times to correct body part targets. Behavioral data were analyzed with a 2×2 repeated measures ANOVA. The within-participant factors were Category (Within or Across) and Distance (Near or Far). The results are reported in the same way as in Experiment 1.

Results

Distance comparison task

Performance The 2×3 repeated measures ANOVA for the proportion correct trials revealed significant main effects of Category ($F(1,26) = 14.63, p = .001$) and Distance ($F(2,52) = 43.76, p < .001$), as well as an interaction ($F(2,52) = 8.94, p < .001$, see Figure 3). Performance Within categories ($M = .775, SE = .013$) was better than Across categories ($M = .698, SE = .022$). Pairwise comparisons for the various Distance differences revealed that performance increased with distance difference. Performance was better for Large distance differences ($M = .866, SE = .022$) compared to Medium distance differences ($M = .759, SE = .024, t(25) = 4.28, p_B = .001$), which in turn was better than Small distance differences ($M = .584, SE = .023, t(25) = 5.30, p_B < .001$).

In order to assess the presence of a symbolic distance effect pairwise comparisons between the three distance differences within each category were performed. The pairwise comparisons revealed that in Within trials performance was equal for Large and Medium distance differences ($t(25) = 1.58, p_B = .375$, see Figure 3), but better compared to Small distance differences ($t(25) = 6.95, p_B < .001$). In the Across trials performance was better for Large compared to Medium distance differences ($t(25) = 5.28, p_B < .001$), and equal compared to Small distance differences ($t(25) = 1.26, p_B = .659$).

Response times Another 2×3 repeated measures ANOVA on the mean response times for correct verifications revealed only a significant main effect for Distance difference ($F(2,52) = 18.74, p < .001$, all other $F_s < 1.00$). Pairwise comparisons showed that responses were slower for Small distance differences ($M = 1943.77\text{ms}, SE = 80.67$) compared to Medium distance differences ($M = 1805.44\text{ms}, SE = 71.22, t(25) = 2.94, p_B = .021$, see Figure 4). Furthermore, Medium distance differences were responded to slower than Large distance differences ($M = 1676.09\text{ms}, SE = 61.07, t(25) = 3.72, p_B = .003$).

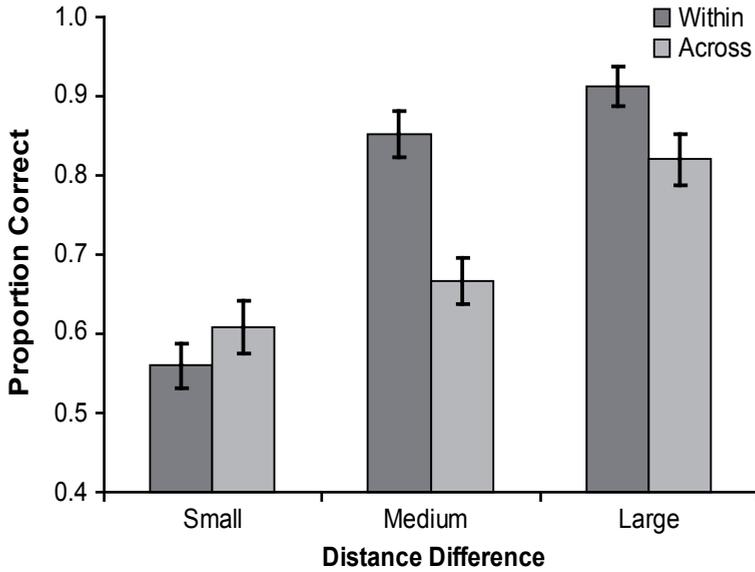


Figure 3 Mean proportion correct responses on the distance comparison task with factors Category and Distance displayed for three distance differences (Small, Medium and Large). Chance level is 0.5. Error bars indicate standard error of the mean.

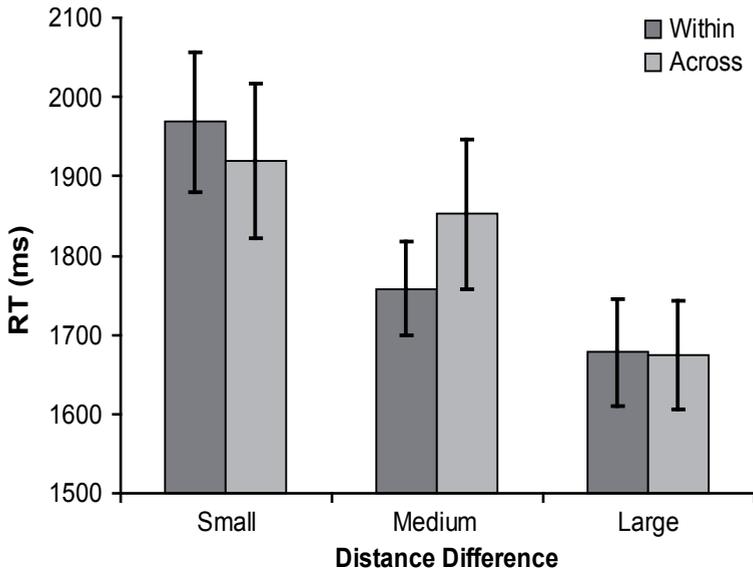


Figure 4 Mean reaction times on the distance comparison task with factors Category and Metric displayed for three distance differences (Small, Medium and Large). Error bars indicate standard error of the mean.

Property verification task

Performance The overall proportion of correct trials was again very high ($M = .97$, $SE = .008$). The 2×2 repeated measures ANOVA on the proportion correct body part targets did not reveal any main effects ($F_s \leq 1.00$), however the interaction between Category and Distance was significant ($F(1,26) = 6.57$, $p = .017$). Pairwise comparisons revealed that for the Near pairs the Within trials were judged better ($M = .99$, $SE = .006$) than the Across trials ($M = .97$, $SE = .013$, $t(25) = 2.08$, $p_B = .043$). Performance on the Far pairs did not differ (Within: $M = .96$, $SE = .017$; Across: $M = .98$, $SE = .015$; $t(25) = -1.19$, $p_B = .265$).

Response times The 2×2 repeated measures ANOVA on the mean response times for correct body part targets showed significant main effects for Category ($F(1,26) = 5.42$, $p = .028$) and Distance ($F(1,26) = 11.05$, $p = .003$), as well as an interaction ($F(1,26) = 8.00$, $p = .009$). Pairwise comparisons for the interaction showed that for the Near trials response times were equal for the Within trials and the Across trials ($t(25) = .53$, $p_B = .598$, see Figure 5). However, for the Far trials responses were faster for Within trials compared to Across trials ($t(25) = -3.33$, $p_B = .003$, see Figure 5).

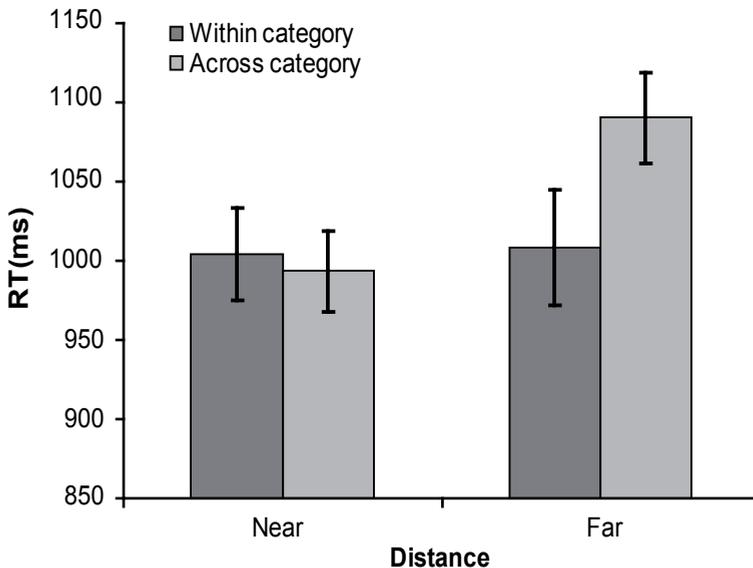


Figure 5 Mean reaction time on the property verification task with factors Category and Metric displayed for two distances (Near and Far). Error bars indicate standard error of the mean.

Discussion

The aim of Experiment 2 was to disentangle the underlying characteristics of the spatial organization of the body. A factor Category was composed including distance relations which crossed major body categories and distance relations which resided within a single body category. The vertical orientation from the distance comparison task of Experiment 1 was adapted in order to comprise both within category trials and across category trials. We hypothesized that if participants used

categorical spatial information responses would be faster for within category trials than for across category trials. Moreover, the symbolic distance effect would then be absent. On the contrary, the results showed the presence of a symbolic distance effect in the response times for both within and across category trials. This effect was supported by the symbolic distance effect in the proportion correct trials although the symbolic distance effect was more pronounced for within than across category trials. The emergence of these effects, especially in the response times, strongly suggests that participants indeed used spatial metric information to solve the task and did not resort to semantic strategies, although a mild influence of categorical information was observed in the proportion correct trials.

The distance comparison task is an explicit task in which participants were instructed to use their own body image during the task. Following this instruction participants were able to use spatial metric information to perform well on the task. However, in the property verification task, which preceded the distance comparison task, participants were unaware of the fact that the underlying body image might be activated. In Experiment 2 this was more strongly controlled for by including animal part names that also required a yes-response. This implicit task showed that besides metric spatial influences categorical spatial influences might emerge when processing body part names. The results from the property verification task of Experiment 2 depict a combination of our metric and categorical hypothesized response patterns. Since performance was at ceiling level this interaction should not be overemphasized. However, the results from the response times show an interesting pattern. The response times on the near distances followed the metric hypothesis in that there were no differences between within and across category trials. On the other hand, the far distances followed the categorical/topological pattern with slower response times to across category trials compared to within category trials. Or to put it differently, for the within category trials we did not find a spatial priming effect, but for the across category trials we did find a spatial priming effect. This mixed pattern of results indicates that both a categorical/topological and a more metric representation might play a role though they are not fully independent.

General discussion

The present set of experiments was designed to elucidate whether the mental representation of the body contains spatial information. More specifically, the spatial organization could be fully metrically, fully categorically/topologically or a mix of both. In two experiments we addressed these questions. In both experiments we found a symbolic distance effect in the distance comparison task indicating that the spatial information available about our body is isomorphic (Denis & Zimmer, 1992). This effect emerged in both experiments in the response times, demonstrated by an inverse linear relationship between response time and distance difference. The performances in Experiment 1 demonstrated a linear relationship between proportion correct trials and distance difference and this general pattern was also present in Experiment 2. However, in Experiment 2 there was also a moderate influence of categorical information which resulted in a more pronounced

symbolic distance effect for within category trials compared to across category trials. The distance comparison results indicate that metric distance was the predominant source of information when participants were explicitly instructed to use this information and the categorical/topological manipulation only affected the proportion correct trials. Moreover, the spatial organization of our body appears quite compelling and may be accessed automatically and obligatorily as indicated by the priming results of Experiment 1 and 2. Reactions were faster to body parts when preceded by body parts in the vicinity. Experiment 2 addressed the nature of the spatial organization in more detail. The response times in the property verification task gave a mixed pattern of results with both signs of a metric organization and of a categorical one.

Hierarchical theories claim that spatial mental representations contain hierarchies of categorical information together with more detailed information (Huttenlocher et al., 1991; Maki, 1981; McNamara, 1986; Stevens & Coupe, 1978). Besides Euclidean distance between items, mental representations can also contain different clusters of items grouped into global spatial categories. Research on object location memory has shown that categories can be determined by a variety of sources, such as geographical boundaries (Friedman, 2009; Maki, 1981), functional clusters (Stevens & Coupe, 1978), artificial boundaries (McNamara, 1986), vertical and horizontal axes (Huttenlocher et al., 1991; Postma et al., 2006) color clusters (Hommel et al., 2000), or spatial proximity (De Lillo, 2004). With regard to the body ‘geographical’ boundaries or spatial proximity could be important to determine the body categories. Interestingly, the present results indicate that hierarchical coding in combination with detailed information might also characterize the mental body representation, which contains metric spatial information, as well as categorical spatial information based on the partonomy division (basic concept: *knee*, superordinate concept: *leg*).

Which type of information we can access depends on the task at hand. The two tasks used in our experiments had different sensitivity to the categorical spatial information. In the implicit property verification task both types of spatial information appear to affect response times. In the distance comparison task participants were explicitly instructed to use metric distance information to solve the task. Although categorical information influenced the proportion correct trials there was not influence found in the response times. Given the explicit instruction participants were able to focus mainly on the metric distance information as demonstrated by the symbolic distance effect in the response times.

The present results are in line with embodied cognition theories of knowledge representations (Barsalou, 1999, 2008; Zwaan, 1999; Zwaan & Radvansky, 1998). According to these theories the meaning of words is obtained through situated simulations, re-enactments of modal states captured during perception (Barsalou, 2008). Hearing body part names starts simulations of the mental body representation where metric and categorical spatial information seem to be available in parallel. The extent of simulation varies with the depth of processing, first superficial processing activates linguistic information subsequently deeper processing activates associated simulations (Barsalou et al., 2008). We would like to speculate that the semantic decision in the property verification tasks evoked relatively deep semantic processing resulting in situated simulations with access to

categorical and metric spatial information. In order to test this idea future research might use a more superficial task, for example judging whether the letter “a” was present in the word. If our assumption is correct then we should observe a lower or even absent spatial priming effect under those circumstances, since superficial tasks do not engage in situated simulations and hence do not give access to spatial information.

Recent neuroimaging studies have provided further evidence supporting embodied cognition theories. These studies have shown that activation in the motor cortex occurs when processing action related words, e.g. reading the word *kick* activated the leg-area in the motor cortex (for a review see Willems & Hagoort, 2007). A point of criticism raised by Willems and Hagoort is that the reported motor cortex activity might also result from deliberate motor imagery. It seems unlikely that the present spatial priming results are based on motor imagery, since there is no reference to action nor does it impose active spatial body scanning. As such, it might be relevant to study the neural activation of processing of body parts. According to embodied cognition one might expect body parts to generate simulations of these body parts which might result in additional motor resonance while not explicitly involved in action. In addition two areas have been identified that selectively respond to images of human bodies and body parts: the fusiform body area (FBA) in the ventral surface and the extrastriate body area (EBA) on the lateral surface (Downing et al., 2001; Schwarzlose et al., 2005; Schwarzlose et al., 2008). While both areas responded to images of human bodies, the EBA also coded the location of the stimulus. Although this location information referred to the position of the stimulus, rather than the position of the body parts within the body, it might be a possible candidate for coding position information of body parts and might be involved in our set of tasks.

In conclusion, the body image incorporates veridical metric spatial information about distances between body parts. When asked to judge distances on the body, participants are able to use this veridical spatial information. In addition, this spatial information is also accessed in a more implicit, automatic manner during a property verification task, in which participants do not consciously attend to the spatial layout of their body. Coupled to this automatic activation we also found that an influence of spatial body categories emerged. In line with recent theories on embodied cognition (Barsalou, 1999, 2008; Barsalou et al., 2008; Zwaan, 1999; Zwaan & Radvansky, 1998) these results show that hearing a verbal referent to a body part automatically evokes a mental simulation of the body containing categorical and more metric spatial information similar to how our physical body is composed.

Chapter 9

Spatial iconicity in blind and sighted individuals

M. E. Struiksma, M. L. Noordzij, R. A. Zwaan & A. Postma

(in preparation)

Abstract

The analogue property between perceptual simulations and interactions with the environment predicts that words referring to specific spatial locations, e.g. *attic-basement*, activate these locations and affect semantic relatedness judgments: an effect known as spatial iconicity. Vision is the prime modality for spatial information processing. Hence, blind individuals might differ with regard to the spatial organization of their semantic knowledge. The present study examined this issue. Early blind and blindfolded sighted participants performed an auditory semantic judgment task. The presence of a spatial iconicity effect for both blind and sighted participants suggests that information from visual and nonvisual modalities can provide adequate spatial information sources. Spatial information is extracted from different input modalities resulting in a supramodal representation. In the current task this supramodal spatial information affects semantic relatedness decisions about words referring to specific spatial locations. These results provide further support for the idea that cognition is grounded.

Introduction

People's knowledge about the world is represented by the conceptual system. Classic theories of cognition have argued that knowledge is represented in an amodal semantic or propositional way (Fodor, 1975; Pylyshyn, 1981). However, a more recent alternative view claims that conceptual representations are grounded in a modality-specific perceptual symbols system (Barsalou, 1999). In the amodal system the external referent is represented by an abstract symbol, whereas in the perceptual symbols system the external referent is represented by perceptual simulations, which are analogue to their referent. In the past few years there has been a growing body of evidence supporting grounded cognition, especially in understanding language (e.g. Barsalou et al., 2003; Pecher et al., 2009; e.g. Pecher et al., 2003; Stanfield & Zwaan, 2001; Wu & Barsalou, 2009; Zwaan & Madden, 2005; Zwaan et al., 2002).

Perceptual simulations can be seen as experiential traces of perception and interactions with the environment which share similar characteristics with its referents (Zwaan & Madden, 2005). This analogue property predicts that changes in the referent will be reflected in the perceptual symbol (Barsalou, 1999). The work from Zwaan and colleagues has shown with multiple paradigms that this prediction holds for the analogue relationships between orientation and shape and the perceptual simulations (Stanfield & Zwaan, 2001; Wassenburg & Zwaan, 2010; Zwaan et al., 2002). These findings provide evidence supporting the idea that cognition is grounded and that participants generated visual simulations in order to understand the sentence.

Not only object orientation and shape information are part of perceptual simulations, but spatial information can be activated as well. The spatial registration hypothesis put forward by Coslett (1999) assumes that all cognitive processes are registered with respect to the spatial map. Spatial information about stimuli is automatically coded even when not directly relevant to the task at hand. Moreover, the activation of spatial information is not limited to processes with an inherent spatial nature, e.g. sensory and motor functions, but extends to all processing modules including linguistic processing. The spatial registration hypothesis predicts that also words referring to specific spatial locations, for example *attic-basement*, activate these locations and affect language function. This effect is also known as spatial iconicity (Zwaan & Yaxley, 2003), where an analogue relation exists between the linguistic sign and its spatial referent. Zwaan and Yaxley (2003) showed that semantic relatedness decisions were faster when the two words were placed in an iconic spatial arrangement consistent with their referents. For example, they were faster when *attic* was placed above *basement* in an iconic orientation than when *basement* was placed above *attic* in a reverse-iconic orientation. Thus the spatial arrangement of the semantically related words affected the interpretation, which provides evidence that participants perceptually simulate the location information implied by the presented words. Converging evidence was obtained in studies that presented single words at the top or bottom of the screen (Pecher et al., 2010; Šetić & Domijan, 2007).

If conceptual knowledge indeed is grounded in the perceptual modality by which it has been obtained, one would expect that the quality of the perceptual system has a direct effect on an

individual's knowledge representations. For example, the sensory/perceptual loss that elderly persons suffer could lead to subtle changes in their semantic system. Even stronger differences would be expected for people blind from birth. For example, their concept of *tomato* would be dominated by how a tomato feels or tastes rather than by the hue of its color. Because perceptual symbols are associated with experiential traces one could argue that direct experience modulates the strength of simulations. Familiar concepts will have access to richer simulations compared to unfamiliar or novel concepts (Barsalou, 2003). Two studies recently have shown that the lack of experience with visual properties indeed affects the conceptual organization in early blind participants (Connolly et al., 2007; Marques, 2010). Connolly et al. (2007) reported that sighted but not early blind participants used color knowledge for implicit similarity judgments of fruits and vegetables. In addition, Marques (2010) found that in a recall task the blind participants were not hindered by visual categorization. The salience of visual cues was diminished in blind participants. The salience of nonvisual cues was not affected and could be used by the blind in a similar way as by the sighted participants. Both studies provide strong evidence in favor of grounded cognition. Notably, access to visual information not only facilitates, but may also hinder cognition, as was demonstrated by Knauff and May (2006). In a deductive reasoning task sighted participants were impeded by visual relations irrelevant to the task. In contrast, performance of the blind participants was not affected by these visual relations. These three examples demonstrate that the salience or accessibility of visual information depends on the role that this modality plays in the acquisition of knowledge (Marques, 2010).

An important question that thus arises is to what extent the accessibility of visual information affects the degree to which semantic knowledge is spatially organized. The evidence discussed above was obtained with visual inputs and manipulations of spatial position in the visual domain. Vision is the prime modality for spatial information processing. Hence, blind individuals might have weaker spatial organization of their semantic knowledge. Alternatively, no clear differences might exist as other sensory inputs provide ample spatial information. The fact that blind people are able to successfully perform a wide variety of spatial tasks suggests that other modalities can provide adequate spatial information sources, eventually leading to a supramodal representation (Struiksma et al., 2009).

The research by Connolly et al. (2007) and Marques (2010) has shown that blind participants do have knowledge about visual characteristics of concepts, however, they do not use that knowledge in a way that sighted participants do. The lack of visual experience modulates the accessibility of that knowledge. How much spatial information is available to blind participants is also subject to experience. Space can be divided into near space and far space. Near space involves the personal space occupied by the body, but also reaching distance, while far space occupies extrapersonal space at walking distance (Halligan & Marshall, 1991). One can imagine that experience with far space depends much more on visual input than near space.

The strong visual hypothesis predicts that visual input is essential for building up a spatial organization which would result in a spatial iconicity effect for sighted, but not for blind participants.

On the other hand, spatial representations are central to our interaction with the environment (Coslett, 1999). Spatial information could be automatically linked to our simulations when understanding language. This would predict no differences between near and far space in blind participants and provide further support in favor of the supramodal representation of space. A third possibility combines the supramodal representation of spatial information with the fact that experience modulates grounded cognition. Therefore, based on a strong embodied supramodal representation we might predict that blind participants will have less experience with far space resulting in a reduced spatial iconicity effect. In order to test these predictions we included both near space trials and far space trials in the spatial iconicity task.

The aim of the current study was to examine whether spatial information in conceptual knowledge differs between blind and sighted individuals. We adapted the Zwaan and Yaxley (2003) paradigm to an auditory version with two vertically aligned loudspeakers. Participants had to judge the semantic relatedness of two words: the first word was presented from the top loudspeaker and the second word from the bottom loudspeaker. The strength of the spatial iconicity effect was assessed in a group early blind participants and a group matched, blindfolded sighted participants.

Methods

Participants

Sixteen early blind, with no memory of vision, and sixteen blindfolded sighted control participants participated in this experiment. One blind and one sighted participant had trouble understanding and following the instructions and were excluded from the analyses resulting in a group of fifteen early blind and fifteen sighted participants. The blind participants and healthy controls were matched for sex, education and age (blind: $M = 44$, $SD = 13$; sighted: $M = 44$, $SD = 14$; for details and etiology of the blind participants see Table 1). Verbal IQ was measured with the Information and Similarities subtests of the Dutch WAIS-III and did not differ between the two groups ($F(1,28) = .13$, $p = .724$). Prior to the experiment participants signed an informed consent, which was approved by the Medical Ethical Board (METC-protocolnumber 05/186-E).

Stimulus material

The task consisted of 56 word pairs (see Appendix): 28 word pairs were semantically related (“yes” response: right-arrow; experimental trials), the other 28 word pairs were semantically unrelated (“no” response: left-arrow). The experimental trials consisted of names of common objects or parts of objects that are in a canonical vertical relation to each other. For example, in a canonical situation of a house the attic is above the basement. The experimental pairs were equally divided into near space (14 pairs) and far space (14 pairs). For each type of space half of the pairs were iconic (7 pairs: *attic-basement*) and the other half were reverse-iconic (7 pairs: *basement-attic*). Two lists of word pairs were created, which were counterbalanced across subjects. All words were recorded

using Audacity (audacity.sourceforge.net) and had a duration of approximately 1000 ms.

A trial started by a beep (100 ms) to warn the subject, followed by an interval of 500 ms. The first word (± 1000 ms) was presented from the top speaker followed by a silent interval of 150 ms and subsequently the second word (± 1000 ms) was presented from the bottom speaker. After the second word there was a 4000 ms response interval which ended as soon as the participant responded. No feedback was given. The next trial started 500 ms later.

Table 1 Details and etiology of the blind participants.

Participant	Education	Sex	Handedness	Age	Etiology*	Age of onset
1	Higher Education	M	L	52	B	4
2	Higher Education	F	L	30	CG	0
3	Higher Education	F	L	49	LCA	0
4	Higher Education	M	A	44	ROP	0
5	University	F	R	42	ROP	0
6	Vocational Education	F	L	38	CG	3
7	Vocational Education	M	L	71	ONC	3
8	Higher Education	M	R	53	LCA	0
9	University	F	R	39	Rb	1
10	University	F	L	34	LCA	0
11	Higher Education	M	R	22	ND	0
12	Higher Education	M	R	53	ROP	0
13	Higher Education	F	R	49	ROP	0
14	Vocational Education	F	R	53	ROP	0
15	Higher Education	M	L	22	LCA	0

*definitions of etiology: A = Anophthalmia, B= Buphtalmus, CG = Congenital Glaucoma, CRS = Congenital Rubella Syndrome, LCA = Leber's Congenital Amaurosis, ND=Norrie Disease, ONC = Optic Nerve Calcification, Rb = Retinoblastoma, ROP = Retinopathy of Prematurity.

Design and procedure

Sighted participants were blindfolded outside the experimental room. All participants were guided to the experimental set-up. The experiment was part of a larger task-battery. The spatial iconicity task was the first task of the battery. The participant was seated 95 cm in front of two loudspeakers which were placed in a vertical position. The top speaker was positioned at an elevation of 160cm and the bottom speaker at 35cm. The participants received an auditory explanation of their task, which was to judge whether or not the two words were semantically related. Prior to the experiment participants received six practice trials with feedback. After finishing the task-battery participants were tested on the two subtests of the WAIS-III.

Data analysis

The responses and reaction times with respect to the onset of the second word were recorded. Response times that exceeded 3000ms were excluded. Subsequently, for the iconic and the reverse-iconic conditions response times that deviated more than two standard deviations from each participant's condition mean were excluded from the analysis. This affected less than 5% of the total number of observations.

The mean response times of correct experimental trials were analyzed using SPSS (SPSS for Windows, Rel. 16.0.2 2008. Chicaco: SPSS Inc.) with a $2 \times 2 \times 2$ mixed Analysis of Variance (ANOVA). The within-subject factors were Space (iconic or reverse-iconic) and Type (near-space or far-space). The between-subject factor was Group (blind or sighted).

Results

Accuracy in both groups on the experimental trials (semantically related) in the spatial iconicity task was very high for the blind ($M = .96$, $SE = .009$) and sighted ($M = .97$, $SE = .009$) and did not differ between the groups ($t(28) = -.19$, $p = .848$). Accuracy reached a ceiling effect in half of the conditions. Therefore, no further analyses were performed.

We found a significant spatial iconicity effect in the response times (Space: $F(1,28) = 23.62$, $p < .001$, see Figure 1). Responses to iconic items were faster ($M = 1279$, $SE = 32$) compared to reverse-iconic items ($M = 1377$ ms, $SE = 34$). There was also a significant main effect of Type ($F(1,28) = 5.44$, $p = .027$); responses to near-space items were faster ($M = 1293$, $SE = 32$) than responses to far-space items ($M = 1338$, $SE = 36$). The interaction effects were not significant ($F_s < 2.57$) neither was the Group difference ($F(1,28) = .05$, $p = .825$). Accordingly, blind and sighted participants showed a similar spatial iconicity effect.

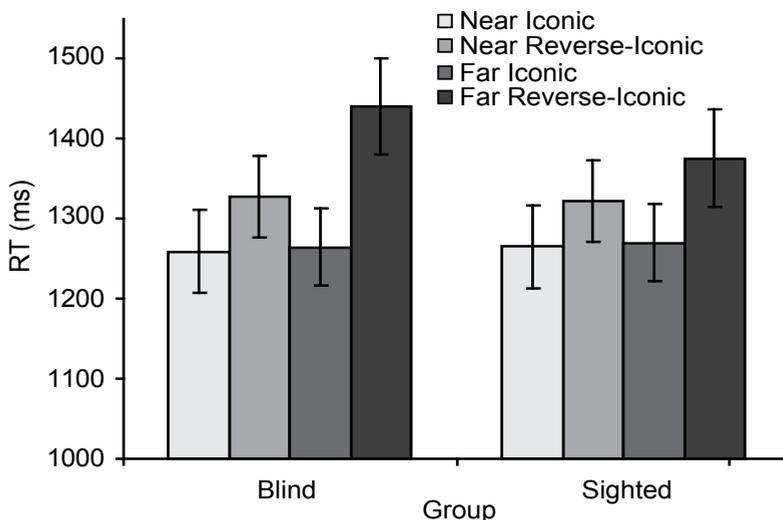


Figure 1 Mean response times as a function of Space, Type and Group (\pm SEM)

Discussion

The current study aimed to examine whether spatial information in conceptual knowledge, resulting in a spatial iconicity effect, differs between blind and sighted individuals. Our response time results showed that there exists an analogue relation between words and their spatial referent in both blind and sighted individuals. The spatial arrangement of the semantically related words affected the interpretation of these words resulting in a spatial iconicity effect. Participants were faster when judging semantic relatedness in iconic pairs; where the first word presented from the top speaker also referred to a word occupying the top position in the pair, such as *attic-basement*.

The presence of the spatial iconicity effect in the auditory setting provides supporting evidence that participants perceptually simulate location information associated with the words. The fact that early blind subjects also demonstrated a spatial iconicity effect refutes the strong visual hypothesis. This spatial information cannot be visual, given that all blind participants had no visual experience. As such it is more in line with the notion of a supramodal representation of spatial information (Struiksma et al., 2009). In the absence of vision blind people obtain spatial information about the world through non-visual modalities. Accordingly they also link spatial information about default locations to lexical items. Importantly this applies to both near and far space items. Although spatial information from the near space can very well be obtained by touch, it is less clear how this could have happened for the far space domain. We may speculate that blind people obtain knowledge about vertical alignment from haptic exploration and extrapolate this knowledge to far space. Another source could be auditory and locomotion information. Despite the result that response times to the near items were faster than to the far items we have to be careful in interpreting these results since the two conditions consisted of different words.

Louwerse (2008) provided a complementary explanation for the effect reported by Zwaan and Yaxley (2003). He argued that iconic word pairs occur more frequently in natural language than reverse-iconic word pairs. Interactions with the real world are reflected in linguistic conceptualizations, for example by means of word order frequency. Louwerse found a spatial iconicity effect, but higher word order frequencies for iconic word pairs predicted this effect better than the perceptual simulation account. This pattern was replicated by Tse et al. (2010) in a surprise recognition task. These findings suggest that both perceptual and linguistic conceptualizations are available during the task at hand. Suppose the spatial iconicity effect is merely linguistic, then we would not expect to find any differences in performance between blind and sighted individuals according to the word order argument. There are no a priori reasons to assume that blind participants would have a different experience of word order frequencies. Both blind and sighted showed a spatial iconicity effect, which can also be explained by the linguistic representation argument. Nevertheless, Šetić and Domijan (2007) and Borghi et al. (2004) have shown that word order cannot explain all effects related to spatial iconicity. Šetić and Domijan (2007) have demonstrated an interaction between single words and their implicit spatial position. They found a response time advantage for words associated with an upper position which were presented in a top position compared to when presented

in the bottom position. The word order argument cannot account for this finding. The words used were very common and probably will not differ very much in frequency of occurrence. Moreover, the response time advantage was specific to a certain spatial position. If word frequency could have caused the effect, then there would have been an overall response time advantage regardless of spatial position. Furthermore, Borghi et al. (2004) have demonstrated that the verification times of object parts depended on the perspective used to describe the object (e.g. “You are driving a car” versus “You are fueling a car”). In particular, they found that for a given perspective responses were faster to near parts than to far parts. On the contrary, distance had no effect in the other perspective. The authors also used the same procedure to obtain word association frequencies as Louwerse and tested the possibility that co-occurrence in related contexts explained the results. For example *fueling* and *tire* might be higher associated in car fueling situations than *fueling* and *antenna*. Interestingly, there were no significant differences between near and far parts. Therefore, the authors concluded that it was unlikely that differences in frequency of occurrence in similar contexts yielded the distance effects. Rather their findings support a link between conceptual knowledge and fine-grained spatial information.

One final aspect that needs to be considered follows from the auditory presentation of the stimuli. During the visual paradigm by Zwaan and Yaxley (2003) the two stimuli were presented simultaneously on the screen. In an auditory setting presenting the stimuli simultaneously would have yielded a situation where none of the words would be audible. By presenting the words sequentially the stimuli remained audible. The first word was presented from the top speaker and the second word from the bottom speaker in alignment with the scanning direction (Zwaan & Yaxley, 2003). A possible confound could be the temporal order of the stimuli. This confound was also present in the original visual paradigm; it was likely that participants read the words from top to bottom and as a consequence the reverse-iconic items would violate the default reading. Zwaan and Yaxley refuted this confound by including a horizontal version of the task. Importantly, the spatial iconicity effect only emerged in the vertical version indicating that temporal order did not cause the spatial iconicity effect. The results from Šetić and Domijan (2007) further support this interpretation. They found a consistent interaction between spatial position and word type (upper or lower spatial location words) for single words. Additionally, the serial presentation in the auditory task does not immediately give away the word pair. The words are presented in isolation which does not automatically determine whether they are “top” or “bottom” words. However, the information about the location of the first word does indicate “top”. When the location subsequently matches the iconic configuration of the second word a spatial iconicity effect is observed.

In conclusion, the present study contributes to the body of evidence supporting the idea that cognition is grounded. Our results extend this idea to blind participants. The fact that blind participants also demonstrated a spatial iconicity effect supports the supramodal representation of spatial information. Notably, spatial location is not only triggered by a visual configuration, but also by an auditory configuration, strengthening the idea that spatial information is available in multiple modalities and can be accessed through a supramodal representation.

Appendix

Stimulus list translated from Dutch.

Near-space iconic words:

head-foot, sweater-trousers, chest-belly, tap-sink, handrail- stairs, hood-stove, curtain rail-baseboard.

Near-space reverse-iconic words:

ankle-knee, chin-forehead, scarf-bonnet, vase-flower, mattress-duvet, pedal-steering-wheel, floor-ceiling.

Far-space iconic words:

smoke-chimney, heron-ditch, rider-horse, skier-snow, cow-meadow, viaduct-road, floater-fish hook.

Far-space reverse-iconic words:

river-bridge, canal-dike, nest-bird, pond-duck, rudder-sail, conservatory-dormer, safety-net-trapeze.

Unrelated words:

apartment-nail, cap-flame, hat-canyon, airplane-saddle, ground-monkey, rain-finger, bike-towel, sea-sock, skate-letter, castle-plant, coffee-pig, paper-hair, bread-boulder, cup-eyebrow, lamp-dog, iceberg-spider, toe-oven, glasses-tea, hand-cloud, plug-mill, coat-pillar, shark-bottle, train-cork, fountain-mosquito, city-mug, frog-violin, rock-pie, pier-bath.

Chapter 10

Summary and conclusions

Summary

Molyneux's question asked whether a blind man who knew the difference between a sphere and a cube from touch, with his sight restored would be able to distinguish between them based on vision. Molyneux's question referred to whether processing spatial information is innate, or based on experience. Another interpretation of Molyneux's question is whether spatial representations are modality-specific, or rather converted into an encompassing supramodal representation that communicates with all our input senses (vision, touch, language etc.). In a series of experiments we have studied the nature of spatial mental representations and the link with spatial language and imagery. In Chapter 2 we have reviewed the current literature on this topic. Spatial information can be obtained through modality-specific input (such as vision, touch and audition, etc.), but also through amodal verbal input. We describe a model in which these inputs feed into a supramodal representation of spatial information. A supramodal representation contains the purely spatial information, but maintains links with input modalities. This model explains how different sources of input can yield behaviorally similar results. In line with this, results from early blind people, without any visual memories, indicate that they perform quite well on various spatial and navigation tasks, suggesting that vision is not the only modality that can convey spatial information. Moreover, only subtle behavioral differences between blind and sighted participants exist, implying that different sources of input may result in functionally equivalent supramodal representations.

Following the model put forward in chapter two a series of empirical studies were conducted. These studies were divided into three parts. First, we investigated how simple spatial sentences, such as *"the ball is above the shoe"*, are represented using, verbal, visual and haptic comparisons (Chapters 3, 4 and 5). The second part focused on complex spatial descriptions and their neural representations (Chapters 6 and 7). The final part studied spatial representations in relation to grounded cognition (Chapters 8 and 9). Three chapters (3, 5 and 9) also included early blind participants. This special participant group enabled us to determine the role of visual experience in representing spatial information.

Simple spatial sentences

The first part of this thesis investigated processing of simple spatial sentences. In three studies we compared simple spatial sentences to another sentence, a picture and a haptic scene. The rationale was based upon the classic sentence-picture verification paradigm (Clark & Chase, 1972). In Chapter 3 we used a sentence-sentence verification task to examine the neural correlates of processing simple spatial sentences. Participants had to determine whether two spoken sentences described the same situation (e.g. *"Wies left of Max"* and *"Max right of Wies"*). Neuropsychological and imaging studies have shown that the left supramarginal gyrus (SMG) is specifically involved in processing spatial terms (e.g. *above, left of*). The current fMRI study focused on the nature and specificity of representing spatial language in the left SMG by combining behavioral and neuronal activation

data in blind and sighted individuals. We found that the left SMG was activated during processing of spatial (*left/right, taller/smaller*) and other relational dimensions (*older/younger*) in both blind and sighted individuals. This study with auditory stimuli replicated and extended the findings from Noordzij et al. (2008) who used a visual paradigm. Thus, the left SMG is activated while processing relations that can be represented on an imaginary axis and regardless of the input modality (visual or auditory). The activation of the left SMG in the blind participants implies that this activation cannot be related to visual processes in this group, supporting a supramodal representation of spatial information.

An important aspect of successful communication is to ensure that the point of view of the interlocutor and the listener align. In other words, the reference frames of both communicators must be in agreement. The default reference frame in Dutch and English is the relative reference frame (Levinson, 1996). This frame employs the communicator as a point of reference and uses the body-axis as a coordinate system. Alignment of reference frames is important when the objects of communication are directly perceivable, but also when the scene is not directly perceivable. In a potentially ambiguous situation, such as communicating directions, it is important to realize which reference frame is being used to communicate the message in order to successfully interpret this message. Such an ambiguous situation stresses the existence of other reference frames.

In Chapter 4 we focused on processing different reference frames of simple spatial sentences and examined what happened when visual and linguistic information was asynchronous. In three sentence-picture and picture-sentence verification experiments we investigated to what extent participants could prepare mapping reference frame alignment and whether it was subject to control. The first experiment showed that after having seen a picture activation of multiple reference frames occurred. In contrast this did not occur after having read a sentence, which suggests that only an ambiguous visual stimulus triggers the need to activate multiple reference frames. On the other hand, reading a sentence did not immediately assume an ambiguous situation. However, when the ambiguous situation was stressed by cueing participants to use a specific reference frame they were able to activate these frames already after having read a sentence. In the third experiment the expectancy of a specific cue was manipulated. This tested whether the non-cued reference frame was still available, or whether participants would adjust to a more efficient strategy, in which only a single reference frame was activated. The results suggested that after reading a sentence the absolute/relative reference frames were default, while the availability of the intrinsic reference frame was strategic. On the other hand, after having seen a picture there was clear evidence of a strategy focusing on a specific reference frame as indicated by significantly longer response times to the unexpected reference frame. The set of experiments described in Chapter 4 provide evidence that participants could activate reference frames from both verbal and visual input, which presumably allows communication between these modalities.

Chapter 5 extended the foregoing to reference frame use in the haptic modality. Given a supramodal representation it might be expected that reference frames play an important role in multiple modalities, including the haptic modality. In this chapter participants were tested on a haptic

sentence-verification task. The task was to judge whether a ball was “above” or “in front of” a shoe following haptic exploration. Results from the spatial relation “above”, in the vertical plane, showed that multiple reference frames were available also in the haptic domain. There were three groups of participants: sighted, low-vision and blind. The pattern of results was similar for all three groups and resembled patterns found for the sighted on visual sentence-verification tasks, which is another confirmation of the idea that reference frames are available in different modalities. In contrast, when judging the spatial relation “in front”, in the horizontal plane, the blind showed a markedly different response pattern. The sighted and visually impaired did not show a clear preference for either the absolute/relative or the intrinsic reference frame when these frames were dissociated. Yet, the blind showed a clear preference for the intrinsic reference frame. In the absence of a dominant cue, such as gravity in the vertical plane, the blind might have emphasized the functional relationship between the objects due to enhanced experience with haptic exploration of objects.

In this section on processing simple spatial sentences we have shown that it proceeds in a very similar manner in blind and sighted participants, both on a behavioral and on a neural level. We have also shown that an important aspect of spatial language processing, i.e. the use of reference frames, is available in multiple modalities. Moreover, evidence on the availability of multiple reference frames was also found in blind participants. Together these results provide converging evidence that spatial information is represented supramodally and can be accessed through spatial language in order to verify information from various modalities.

Interestingly, in Chapter 4 the patterns of results for sentence-picture and picture-sentence verification were different. This stresses the different roles that the context played and shows the flexibility that allowed participants to focus on different reference frames depending on the situation. This flexibility is also present when dealing with more complex spatial descriptions. In a communicative situation where, for example, you are giving a route description it seems more natural to use the absolute/relative reference frame. By using this perspective the receiver can later link this information to his or her own position and match the instruction with the environment. This idea is supported by research on different perspectives in giving route descriptions on the basis of a visual map. Hund et al. (2008) showed that participants demonstrated flexibility and adapted their directions to the needs of their listeners. Participants used a route perspective, with left and right turns, when the receiver was thought to drive through the city: On the other hand, when the receiver was also in front of a map, participants were more inclined to use a survey perspective, with cardinal directions. The second part of this thesis investigated the neural correlates of these different perspectives.

Complex spatial descriptions

Previous research on processing route and survey descriptions has revealed mixed results. There is both evidence in favor of a perspective dependent and a perspective independent representation. However, these conclusions have been made in the context of experiments with greatly varying task aspects, such as number of exposures and type of tasks. Building up a spatial mental representation from a complex description seems to be a progressive process. At first, perspective dependent information is available. However, after over-learning a perspective independent supramodal spatial mental representation can be built up. Chapter 6 reported an fMRI study that investigated this progressive process.

Two groups of participants learned the configuration of a zoo from either a route or a survey description. The route description took the listener on a mental tour through the zoo and explained the locations of the animals with respect to the listener, which stressed the egocentric relations between the listener and the locations. The survey description provided a hierarchical overview of the zoo and explained the locations of the animals with respect to each other, which stressed the allocentric relations between them. A general overlapping network of fronto-parieto-temporal activation was found for the route and the survey group associated with the processing of spatial descriptions. A parametric modulation was used to determine which areas showed accumulating activation after increasing exposures to the description. This analysis revealed an increasing BOLD response in the caudate nucleus due to the growing familiarity with the description in both groups. An important structure involved in processing allocentric spatial information is the hippocampus. The parametric modulation revealed hippocampal activation for the survey and route group which suggested that both groups were able to process allocentric components of the environment. Importantly, all subjects could draw the correct configuration of the zoo after the experiment had finished suggesting that the route and the survey group were able to learn the configuration. Based on these results we concluded that route and survey descriptions are encoded in a hierarchical fashion where allocentric components can also be inferred from egocentric components resulting in a functional equivalent mental representation for both encoding perspectives. Importantly, the current neurocognitive results from complex verbal descriptions overlap to a great extent with previous results from visual paradigms. These findings again support the idea that spatial information is represented in a supramodal manner where spatial information is processed regardless of the input modality.

Chapter 6 showed considerable overlap during *learning* from different perspectives. Yet another possibility is that perspective related differences only arise during retrieval. In chapter 7, we therefore contrasted *retrieval* and usage of route distance information against that of survey distance information. After having learned the configuration of the zoo from either a route or a survey perspective, participants were tested on a route and survey distance comparison task. These tasks provided insight in the mental representation that had been built up. They assessed whether the representation contained spatial metric detail and was functionally equivalent for the route and

survey group. In the route task participants were asked to compare route distance differences and in the survey task they were asked to compare bird-flight distance differences. Potentially, the type of question that was asked determined which type of knowledge was accessed.

The behavioral results showed high similarity between the route and the survey group and a similar global pattern for the route and survey task. A distance comparison effect was observed where smaller distance differences were more difficult to compare which resulted in longer response times and more errors. This effect indicates that participants used metric information to solve the task and supports the idea that over-learning yields a supramodal representation which is perspective independent. On a neural level we observed activation of bilateral precuneus, linked to processing spatial information, and bilateral middle occipital gyrus and left parahippocampal gyrus, involved in processing landmarks for both groups on both tasks. No differences between the groups or the type of task that they performed were found, nor was there evidence of distinct egocentric or allocentric knowledge. The current neural network was involved in both spatial tasks for the route and survey group. We may question whether this activation could be linked to a supramodal representation, or rather represents general spatial task related activity such as imagery processes. Clearly the relevance of landmarks was evident in all conditions, which implies that the use of landmarks is not limited to route descriptions. Further research is needed to corroborate a supramodal representation based on complex verbal descriptions. A potentially interesting approach could be to focus on a parametric design, as was employed in Chapter 6. The distance comparison effect found in the behavioral data provides a handle to examine the neural correlates modulated by these distance differences. A preliminary conclusion would be that after over-learning of a complex spatial description participants have built up a perspective independent spatial mental representation and show comparable neural activation predominantly related to processing landmarks.

Together, these two neuroimaging chapters demonstrate that processing complex spatial descriptions from different perspectives is linked to a large neural network involved in processing spatial language. Regardless of the learning perspective a large fronto-parieto-temporal network of activation was found. Building up a spatial mental representation from a complex verbal description is an intricate process during which multiple brain areas are involved. In addition to overlapping neural activation the parametric modulation revealed distinct neural correlates. The left posterior hippocampus was activated in the survey group by processing allocentric information. Interestingly, the route group activated bilateral anterior hippocampus, which suggests that the allocentric relations between the animal locations could be inferred from the route description. The fact that the route group could also correctly draw the configuration of the zoo and performed well on the survey task further corroborates this statement. By the time the configuration had been learned participants were able to flexibly access this information. Both groups performed equally, on a behavioral and neural level, on the survey task, but also on the route task. Once the initial learning differences had been dealt with the resultant spatial representation was equivalent and allowed for flexible use. The conclusions from Chapter 6 and 7 further indicate that spatial information is represented supramodally, although additional research is needed.

Grounded cognition

The results from the last two studies described above add to the body of literature demonstrating that metric information can be conveyed through verbal descriptions. Even though the complex descriptions about the zoo did not contain explicit metrical information the observed distance comparison effects indicate that participants had built up a spatial mental representation and had access to metric information about the locations of the animals. A question that follows from these observations is how compelling this spatial information is. This question brings us to the final section of this thesis. The grounded cognition account implies that language is understood by linking words to perceptual simulations. One dimension of these perceptual simulations is the spatial dimension. If cognition really is 'grounded', then processing of verbal items would activate that spatial dimension even when no explicit spatial instructions are given. This section describes two studies that tested whether processing words could have access to spatial information.

The body plays a prominent role in the perception of space. Our actions are to a large extent determined by the abilities of our body and consequently, cognition can be shaped by these abilities. Chapter 8 described two experiments that examined the extent to which mental body representations contain spatial information. Participants were asked to compare distances between various body parts. Similar to what happens when people compare distances on a real visual stimulus, they were faster as the distance differences between body parts became larger (Experiment 1), and this effect could not (only) be explained by the crossing of major bodily categories (*umbilicus* to *knee* vs. *knee* to *ankle*, Experiment 2). It indicates that our mental representation of our body preserves the proportions of our body. In addition, participants also performed simple animate/in-animate verification on a set of nouns. The nouns describing animate items were names of body parts. A spatial priming effect was found: Verification was faster for body part items preceded by body parts in close spatial proximity. This suggests automatic activation of spatial body information and supports the grounded cognition view. In order to understand the meaning of body part names the mental body representation was simulated including the locations of these parts within the body. As a result a spatial priming effect was observed, based on the spatial locations of the body parts.

Although we instructed participants to imagine their own body in the distance comparison task we cannot control for the fact that they might have conjectured a mental image of a standard body. However, the simple verification task, which was conducted prior to the distance comparison task, did not explicitly require activation of the mental body representation. Rather, this task showed that by merely understanding the meaning of the presented words participants were compelled to simulate a mental body representation. Most likely, they simulated their own body, since the own bodily experience is probably responsible for the mental body representation. In light of the foregoing it might be interesting to test this set of tasks in people with unusual body proportions, such as people with abnormal growth.

Given that our mental body representation is based on our own bodily experience one might wonder whether a change in perceptual experience has an effect on this representation. This leads

us back to our interpretation of Molyneux's question: Is haptic and auditory information sufficient to convey spatial information resulting in a supramodal representation similar to visual information? In order to answer this question further research could focus on blind individuals.

The final chapter of this part made a start at investigating spatial information processing in relation to grounded cognition in blind individuals. This chapter described a task where semantic relatedness judgments had to be made about word pairs. The perceptual simulations that are thought to be activated when understanding language can be seen as experiential traces of perception and interactions with the environment which share similar characteristics with its referents, such as orientation and shape. This analogue property also predicts that words referring to specific spatial locations, e.g. *attic-basement*, activate these locations and hence could affect semantic relatedness judgments: an effect known as spatial iconicity. In Chapter 9 blind and blindfolded sighted participants were compared in order to determine the influence of visual experience on the simulation of spatial locations. Both groups performed an auditory semantic relatedness judgment task. In this task participants heard two words, the first word was presented from the top loudspeaker and the second word from the bottom loudspeaker. In line with the grounded cognition account the location of presentation of the word was implicitly linked to the meaning of the word. As a result responses were facilitated for word pairs that were presented in their iconic spatial orientation, i.e. *attic* above *basement*, and hindered in their reverse-iconic orientation. The presence of this spatial iconicity effect for both blind and sighted participants suggests that information from visual and nonvisual modalities can provide adequate spatial information sources. Apparently, the haptic experience that blind people have with “above” and “below” can be used to extrapolate this ordering to far space that is beyond reach. The resulting link with spatial information does not directly depend on their lack of experience with these far space items. Again we found evidence supporting the idea that spatial information is extracted from different input modalities resulting in a supramodal representation.

Spatial iconicity or enhancement of relatedness judgments when the words were arranged canonically has been shown for concrete concepts, but also for abstract concepts such as valence and power (Meier & Robinson, 2004; Schubert, 2005). Valence and power are associated with verticality. It is appealing to link this also to the results of Chapter 3. One of the aims of Chapter 3 was to establish the specificity of the left SMG activation. In order to determine whether the left SMG was activated only by spatial prepositions, or rather more general spatial relations three conditions were added: *together with* (nonspatial preposition), *taller/smaller* (spatial adverb), and *older/younger* (nonspatial adverb). Although *older/younger* was labeled a nonspatial adverb the Metaphoric Mapping Theory (Boroditsky, 2000) states that time is imagined horizontally, therefore, this condition could also be considered spatial. Indeed, the behavioral and neural results suggested that *left/right*, *taller/smaller*, and *older/younger* all activated the left SMG. We may speculate that horizontal spatial iconicity might occur for *left/right* and *older/younger*, and vertical spatial iconicity for *taller/smaller*. Given the spatial analogy of abstract concepts such as valence and power we might expect left SMG activation for these words as well.

Conclusions

This thesis examined whether spatial representations are modality-specific, or rather convert into an encompassing supramodal representation that communicates with all our input senses. In this thesis we have shown that different output modalities could show distinct experience modulations. The differences between blind and sighted in the haptic domain can be explained by the enhanced experience of the blind. The results from the verbal studies, on the other hand, show that blind and sighted perform very similar. Moreover, perspective differences in complex descriptions yield functionally equivalent mental representations after over-learning. In short, spatial information can be obtained from multiple different input modalities and is represented supramodally. This supramodal representation can be addressed using language and spatial language in particular.

These results bear on Molyneux's question, as described at the beginning of this thesis. We have shown that spatial representations are modality-general and supramodal, which allow communication with all different input senses. The original interpretation of Molyneux's question referred to the innateness of spatial information. We have argued that spatial information is represented supramodally, which assumes that links with input modalities are maintained. We have shown that vision is not a prerequisite to build up these supramodal representations, but the question remains whether automatic transfer to the visual modality is possible. Evidence from case-studies on patients born blind, who had their vision restored, reported profound difficulties with visual perception (Pascual-Leone & Hamilton, 2001; Von Senden, 1932). They were able to distinguish between objects, but could not identify them. In order to 'tell' what the object was, the new visual input had to be matched to the established supramodal representation. Apparently, the visual input needed to be calibrated, possibly by touch. In conclusion, spatial information from different input modalities is converted into a supramodal representation. Importantly, links with the input modalities are maintained and are modulated by experience. Although the brain unifies space, its supramodal mechanisms are not fully innate.

Nederlandse samenvatting

Summary in Dutch

De vraag van Molyneux

Op 7 juli 1688 schreef William Molyneux (1656-1698) een brief aan John Locke (1632-1704) waarin hij hem een vraag stelde over een vroeg-blinde blinde man. Molyneux vroeg zich af of een blinde man die op de tast het verschil had geleerd tussen een bol en een kubus, dit onderscheid ook zou kunnen maken op basis van zijn gezichtsvermogen indien dat door een wonder hersteld zou zijn. Deze vraag is vandaag de dag nog steeds relevant; hij gaat over het verwerken van ruimtelijke informatie. Is dit verwerkingsproces aangeboren, of gebaseerd op ervaring, en dan in het bijzonder op visuele ervaring? In een bredere zin luidt de vraag of ruimtelijke waarneming en begrip dus modaliteitspecifiek zijn of dat er altijd een vertaalslag plaatsvindt naar een representatie die communiceert met de verschillende modaliteiten (zicht, tast, taal)?

De vraag van Molyneux vormt de basis van dit proefschrift, dat een theoretisch kader aanreikt en waarin nieuw experimenteel werk wordt beschreven. Dit experimenteel werk onderzoekt de aard van mentale representaties die worden opgebouwd aan de hand van simpele of meer complexe beschrijvingen. Zijn deze representaties propositioneel, dat wil zeggen vormen ze een set van abstracte verwerkingsinstructies of hebben ze analoge, direct ruimtelijke kwaliteiten? In een serie van experimenten hebben we onderzocht wat de aard is van ruimtelijke mentale representaties en hoe deze gekoppeld zijn aan ruimtelijke taal en voorstellingsvermogen. In hoofdstuk 2 wordt de bestaande literatuur over dit onderwerp besproken. Ruimtelijke informatie kan worden verkregen door modaliteitspecifieke informatie (zoals zicht, tast, gehoor, etc.), maar ook door abstracte verbale informatie. Hoofdstuk 2 beschrijft een model waarin deze verschillende soorten zintuiglijke informatie de basis vormt voor een supramodale representatie van ruimtelijke informatie. Een supramodale representatie bestaat uit pure ruimtelijke informatie en behoudt een koppeling met de verschillende zintuiglijke informatiebronnen. Dit model beschrijft hoe verschillende zintuiglijke bronnen toch een vergelijkbaar gedrag als resultaat kunnen leveren, ook wel functioneel gelijkwaardig genoemd. Zo laten resultaten van vroeg-blinde personen, zonder visuele herinneringen, zien dat zij behoorlijk goed presteren op verschillende soorten ruimtelijke taken. Dit suggereert dat zicht niet de enige informatiebron is die ruimtelijke informatie over kan brengen. Daarnaast bestaan er slechts subtiele verschillen tussen blinde en ziende proefpersonen, wat erop wijst dat verschillende informatiebronnen resulteren in een supramodale representatie die functioneel gelijkwaardig is.

Volgend op het model dat in hoofdstuk 2 is beschreven, is er een serie onderzoeken uitgevoerd. Deze studies kunnen in drie delen worden onderscheiden. Het eerste deel van dit proefschrift gaat over simpele ruimtelijke zinnen, zoals “*de bal is boven de schoen*”. Daarnaast gaat het over referentiekaders, die bepalend zijn voor het gezichtspunt in een conversatie en zeer belangrijk voor succesvolle communicatie. Het zou zo kunnen zijn dat de referentiekaders communicatie tussen informatiebronnen mogelijk maken, zoals Molyneux stelde in zijn beroemde vraag. Dit hebben we onderzocht met verbale, visuele en tactiele vergelijkingen (hoofdstuk 3, 4, en 5).

Ruimtelijk taal komt haast nooit in isolatie voor, maar moet in een bredere context bekeken worden, zoals het praten over een ruimte. Een ander voorbeeld is om van de ene plaats naar de

andere plaats te komen. Het tweede deel van dit proefschrift (hoofdstuk 6 en 7) is gericht op complexe ruimtelijke beschrijvingen. Referentiekaders maken duidelijk waar voorwerpen zijn geplaatst en ondersteunen verschillende perspectieven in complexe beschrijvingen. Mogelijk leidt dit tot verschillende perspectiefafhankelijke representaties, zowel op gedragsniveau als in het brein. In dit deel beschrijven we twee studies die deze representaties onderzoeken in termen van gedragsverschillen en verschillen in neurale activiteit.

Het laatste deel van dit proefschrift (hoofdstuk 8 en 9) behandelt de vraag hoe dwingend ruimtelijke informatie is wanneer we taal begrijpen. Hebben we ook toegang tot ruimtelijke informatie aan de hand van taal die niet per definitie ruimtelijk is? Deze vraag koppelt ruimtelijke representaties aan de invloedrijke theorie over “grounded cognition” waarbij het idee is dat taal begrepen wordt doordat we die koppelen aan een perceptuele ervaring. In drie hoofdstukken (3, 5 en 9) beschrijven we onderzoek met vroeg-blinde proefpersonen. Deze speciale groep proefpersonen vormt een belangrijke schakel in het bepalen wat de rol van visuele ervaring is in het vormen van een ruimtelijke representatie.

Simpele ruimtelijke zinnen

In essentie wordt ruimtelijke taal gebruikt om de relatie tussen twee voorwerpen te beschrijven. De meest voor de hand liggende manier is het beschrijven van de relatie tussen twee voorwerpen in een scène. Deze scène kan direct waarneembaar zijn, maar kan ook een beschrijving zijn van een toekomstige situatie, zoals bijvoorbeeld bij het geven van een route beschrijving. De relatie tussen de voorwerpen wordt beschreven met een ruimtelijk voorzetsel zoals *boven*, *voor* en *links van*. Ruimtelijke voorzetsels zijn onderdeel van een kleine klasse van, in de meeste talen, ± 80 woorden (Landau & Jackendoff, 1993).

In drie studies hebben we simpele ruimtelijke zinnen vergeleken met een andere zin, een plaatje of een tactiele scène. De opzet van deze studies is gebaseerd op het klassieke zin-plaatje vergelijkings-paradigma (Clark & Chase, 1972). In hoofdstuk 3 is een zin-zin vergelijkingstaak gebruikt om te onderzoeken wat de neurale basis is van het verwerken van simpele ruimtelijke zinnen. De proefpersonen moesten bepalen of twee gesproken zinnen dezelfde situatie beschreven (e.g. “*Wies links van Max?*” en “*Max rechts van Wies?*”). Eerder neuropsychologisch onderzoek met patiënten en onderzoek naar hersenactiviteit heeft aangetoond dat de linker supramarginale gyrus (SMG) specifiek betrokken is bij het verwerken van ruimtelijke voorzetsels als *boven* en *links van*. Het huidige functionele MRI onderzoek richt zich op de aard en gevoeligheid voor het representeren van ruimtelijke taal in de linker SMG. Dit werd gedaan door gedragsdata te combineren met hersenactiviteit van blinde en ziende personen. We hebben gevonden dat de linker SMG geactiveerd werd tijdens het verwerken van ruimtelijke (*links/rechts*, *groter/kleiner*) en andere relaties (*ouder/jonger*) bij zowel blinden als zienden. Deze studie gebruikte alleen auditieve stimuli en sluit aan bij de resultaten van Noordzij et al. (2008) met een visuele taak. We concluderen dat de linker SMG wordt geactiveerd tijdens het verwerken van relaties die op een denkbeeldige as gezet kunnen worden en

dat dit niet afhankelijk is van de informatiebron (visueel of auditief). Het feit dat we ook linker SMG activiteit vonden in de blinde proefpersonen suggereert dat deze activiteit niet veroorzaakt kan worden door visuele processen en ondersteunt het idee van een supramodale representatie van ruimtelijke informatie.

Een belangrijk aspect van succesvolle communicatie is het afstemmen van het referentiekader tussen spreker en luisteraar. Een referentiekader maakt gebruik van een gezichtspunt om de ruimte in verschillende locaties op te delen. Er zijn drie verschillende soorten gezichtspunten die een verschillend referentiekader beschrijven. Het absolute referentiekader (zie figuur 1, hoofdstuk 1) gebruikt eigenschappen van de omgeving, zoals zwaartekracht en windrichtingen, om te bepalen wat bijvoorbeeld links en rechts is. Het relatieve referentiekader (zie figuur 1, hoofdstuk 1) is gebaseerd op de spreker en gebruikt de lichaamsas om locaties te bepalen. Bij het intrinsieke referentiekader (zie figuur 1, hoofdstuk 1) worden links en rechts bepaald vanuit het referentieobject gezien. In het Nederlands en Engels is het relatieve referentiekader standaard (Levinson, 1996). Niet alleen wanneer objecten waarneembaar zijn is het belangrijk dat de referentiekaders van gesprekspartners overeenkomen, maar ook wanneer deze niet direct waarneembaar zijn. Bijvoorbeeld tijdens het geven van route beschrijvingen is het van essentieel belang dat beide gesprekspartners weten welke kant links is en welke rechts om verdwalen te voorkomen.

In hoofdstuk 4 hebben we ons gericht op het verwerken van verschillende referentiekaders van simpele ruimtelijke zinnen. We hebben onderzocht wat er gebeurt wanneer visuele en verbale informatie niet tegelijkertijd wordt aangeboden. In drie experimenten is bekeken of proefpersonen zich konden voorbereiden op het toepassen van een specifiek referentiekader en of ze daar controle over hadden. Uit het eerste experiment bleek dat mensen na het zien van een plaatje wel, maar na het lezen van een zin niet meerdere referentiekaders activeren. Aan de andere kant, de zin suggereerde niet direct een ambigue situatie waardoor men dus ook niet direct de behoefte had om zich voor te bereiden. In experiment twee werd er de nadruk gelegd op de ambigue situatie door middel van een hint. Hieruit bleek dat wanneer mensen een ambigue situatie verwachten, ze ook na het lezen van een zin wel degelijk meerdere referentiekaders activeren. In het derde experiment werd de verwachting van een specifieke hint gemanipuleerd. Hiermee konden we onderzoeken of het referentiekader dat niet onder de aandacht was gebracht, nog steeds werd geactiveerd of dat het uit strategisch oogpunt werd genegeerd. De resultaten lieten zien dat na het lezen van een zin het relatieve referentiekader standaard geactiveerd werd, terwijl het intrinsieke referentiekader strategisch geactiveerd werd. Aan de andere kant, na het zien van een plaatje gebruikten proefpersonen altijd een strategische activatie. Reactietijden voor het onverwachte referentiekader waren significant langer dan voor het verwachte referentiekader. Deze drie experimenten laten zien dat het mogelijk is om referentiekaders te activeren aan de hand van verbale en visuele informatie. Dit maakt het mogelijk dat er communicatie tussen de informatiebronnen is.

In hoofdstuk 5 hebben we gekeken naar het gebruik van referentiekaders in het tactiele domein. Uitgaand van een supramodale representatie kunnen we verwachten dat referentiekaders ook in de tactiele modaliteit een rol spelen. Dit hoofdstuk beschrijft een experiment waarbij proefpersonen

werden getest op een tactiele beoordeling van een ruimtelijke zin zoals “*de bal is boven de schoen*”. De resultaten voor de ruimtelijke relatie boven, in het verticale vlak, laten zien dat er meerdere referentiekaders beschikbaar zijn, ook in het tactiele domein. Er waren drie groepen proefpersonen: zienden, slechtzienenden en blinden. De resultaten zijn vergelijkbaar voor alle drie de groepen en komen overeen met die van zienden binnen het visuele zin-plaatje vergelijkings-paradigma. Deze bevinding sterkt het idee dat referentiekaders beschikbaar zijn in verschillende modaliteiten. Echter, de resultaten van de blinden zijn wezenlijk anders dan die van de zienden en slechtzienenden in het horizontale vlak, voor de ruimtelijke relatie voor. Wanneer het mogelijk was verschillende antwoorden te geven volgens het relatieve en het intrinsieke referentiekader, lieten de zienden geen duidelijke voorkeur blijken. De blinden daarentegen hadden een sterke voorkeur voor het intrinsieke referentiekader. Wanneer een duidelijk referentie ontbreekt, zoals zwaartekracht in het verticale vlak, hechten de blinden waarschijnlijk sterker aan de gebruiksfunctie van het voorwerp. Het is niet onwaarschijnlijk dat dit komt door een grotere bekendheid met het betasten van voorwerpen.

In dit deel over het verwerken van simpele ruimtelijke zinnen menen we aangetoond te hebben, dat het gedrag en de hersenactiviteit in blinden en zienden vergelijkbaar is, dat een belangrijk aspect van ruimtelijk taalverwerking, namelijk het gebruik van referentiekaders, beschikbaar is in verschillende modaliteiten en dat meerdere referentiekaders ook beschikbaar zijn voor blinde proefpersonen. Dit bij elkaar genomen levert steun voor het idee dat ruimtelijke informatie supramodaal gerepresenteerd wordt. Door middel van ruimtelijke taal hebben we toegang tot die representatie en kunnen we deze vergelijken met informatie van andere modaliteiten.

Een interessante bevinding in hoofdstuk 4 is dat het resultatenpatroon voor zin-plaatje en plaatje-zin vergelijkingen anders is. Dit verschil benadrukt de invloed van de situatie en laat zien dat proefpersonen flexibel met verschillende situaties om kunnen gaan. Deze flexibiliteit vinden we ook terug wanneer proefpersonen meer complexe ruimtelijke beschrijvingen verwerken, bijvoorbeeld in een situatie waarin iemand een route beschrijving geeft, ligt het meer voor de hand om het relatieve referentiekader te gebruiken. Dit referentiekader stelt de luisteraar in een later stadium in staat om de instructie te koppelen aan zijn of haar eigen positie. Dit idee wordt ondersteund door onderzoek naar het gebruik van verschillende perspectieven tijdens het geven van route beschrijvingen op basis van een kaart. Hund et al. (2008) lieten zien dat proefpersonen flexibel waren en hun beschrijvingen aanpasten aan de behoefte van de luisteraar. Als men dacht dat de luisteraar door een stad reed, gebruikte men een routeperspectief, met *linksaf* en *rechtsaf*. Aan de andere kant, als men dacht dat de luisteraar een kaart voor zich had, dan gebruikte men een vogelvluchtperspectief, met windrichtingen. Het tweede deel van dit proefschrift onderzocht de hersenactiviteit behorende bij deze verschillende perspectieven.

Perspectief in complexe ruimtelijke beschrijvingen

Dagelijkse communicatie beperkt zich niet alleen tot simpele zinnen. Ruimtelijke taal speelt een belangrijke rol in het vinden van de weg, zoals hierboven beschreven. De beschrijving vanuit het route perspectief vormt een mentale wandeling, waarbij de verschillende plaatsen en gebouwen beschreven worden ten opzichte van de luisteraar. Kenmerkend voor een egocentrisch referentiekader is dat locaties serieel gekoppeld worden aan de luisteraar, vergelijkbaar met een relatief referentiekader zoals eerder beschreven. Het vogelvluchtperspectief wordt gebruikt wanneer de locaties hiërarchisch beschreven worden ten opzichte van elkaar. Dit wordt ook wel een allocentrisch referentiekader genoemd en vertoont parallellen met het absolute of intrinsieke referentiekader zoals gebruikt bij simpele zinnen.

Voor succesvolle navigatie is het waarschijnlijk dat een luisteraar een geïntegreerde ruimtelijk mentale representatie van de omgeving opbouwt (Taylor & Tversky, 1992). De verschillende perspectieven hierboven beschreven suggereren, dat de bijbehorende ruimtelijk mentale representaties ook verschillend kunnen zijn. Aan de andere kant is eerder beargumenteerd, dat ruimtelijke informatie vanuit verschillende modaliteiten verkregen kan worden en wordt opgeslagen in een enkele ruimtelijke representatie. Voorgaand onderzoek naar het verwerken van route- en vogelvluchtbeschrijvingen laat verschillende resultaten zien. Er is bewijs voor zowel een perspectiefafhankelijke als een -onafhankelijke representatie. Een kanttekening daarbij is, dat deze conclusies zijn gebaseerd op onderzoeken met zeer verschillende taakaspecten, waaronder het aantal herhalingen en het type taak. Het opbouwen van een ruimtelijk mentale representatie van een complexe beschrijving is een geleidelijk proces. Allereerst is er perspectiefafhankelijke informatie beschikbaar. Echter, na een uitvoerig leerproces kan een perspectiefonafhankelijke supramodale ruimtelijk mentale representatie worden opgebouwd. In hoofdstuk 6 rapporteren we een fMRI studie die dat geleidelijke proces onderzoekt.

Twee groepen proefpersonen leerden een plattegrond van een dierentuin door middel van een route- of een vogelvluchtbeschrijving. De routebeschrijving nam de luisteraar mee op een mentale wandeling door de dierentuin. De locaties van de verschillende dieren werden uitgelegd ten opzichte van de luisteraar. Dit benadrukte de egocentrische relaties tussen de luisteraar en de locaties. De vogelvluchtbeschrijving gaf een hiërarchisch overzicht en beschreef de locaties van de dieren ten opzichte van elkaar. Dit benadrukte de allocentrische relaties. Voor zowel de route- als de vogelvluchtgroep werd er een overlappend netwerk gevonden van activiteit in de frontale, pariëtale en temporele cortex. Deze activiteit werd gevonden tijdens het verwerken van ruimtelijk beschrijvingen in het algemeen. Een parametrische modulatie werd gebruikt om te achterhalen in welke gebieden de activiteit toenam met het aantal herhalingen van de beschrijving. Deze para-metrische modulatie liet zien dat de caudate nucleus steeds sterker geactiveerd werd naarmate de proefpersonen meer bekend werden met de beschrijving. Een belangrijke structuur die betrokken is bij het verwerken van allocentrische ruimtelijk informatie is de hippocampus. De parametrische modulatie liet voor beide groepen activiteit in de hippocampus zien. Dit suggereert

dat de proefpersonen uit de vogelvlucht-, maar ook die uit de routegroep in staat waren om allocentrische elementen van de omgeving te verwerken. Belangrijk was ook dat alle proefpersonen in staat waren om aan het einde van het experiment een correcte plattegrond van de dierentuin te tekenen. Dus beide groepen waren in staat om de plattegrond van de dierentuin te leren aan de hand van een ruimtelijke beschrijving. Op basis van deze bevindingen kunnen we concluderen dat route- en vogelvluchtbeschrijvingen worden gecodeerd op een hiërarchische manier. Op basis van egocentrische informatie kunnen allocentrische relaties worden bepaald, met als eindresultaat een mentale representatie die functioneel hetzelfde is voor beide groepen. Daarnaast overlappen de huidige neurocognitieve resultaten op basis van complexe verbale beschrijvingen met de eerdere resultaten op basis van visuele taken. Deze bevindingen ondersteunen wederom het idee dat ruimtelijke informatie wordt opgeslagen in een supramodale mentale representatie.

In hoofdstuk 6 vonden we een aanzienlijke overlap tussen het *leren* vanuit verschillende perspectieven. Een andere mogelijkheid is dat de verschillen voornamelijk ontstaan tijdens het ophalen van informatie. In hoofdstuk 7 hebben we daarom het *ophalen* van route- en vogelvluchtinformatie met elkaar vergeleken. Nadat de proefpersonen de omgeving hadden geleerd vanuit een route- of een vogelvluchtperspectief, werden ze getest op een route- en een vogelvlucht-afstandsvergelijkingstaak. Deze taken geven inzicht in de mentale representatie die is opgebouwd. Het doel was om uit te zoeken of de mentale representatie ruimtelijke details bevatte en of deze functioneel hetzelfde was voor beide groepen, dat wil zeggen of beide groepen hetzelfde presteerden op een taak. In de routetaak moesten proefpersonen routeafstanden vergelijken en in de vogelvluchttask vogelvluchtafstanden. Mogelijk bepaalt het type vraag welk type kennis wordt gebruikt.

De gedragsresultaten lieten een grote gelijkenis zien voor de route- en de vogelvluchtgroep en een vergelijkbaar patroon voor beide taken. Er was sprake van een afstands-vergelijkingseffect waarbij het moeilijker is om een klein afstandsverschil te beoordelen dan een groot afstandsverschil. Dit effect wijst erop dat proefpersonen metrische informatie gebruikten om de taak op te lossen en steunt het idee dat na overmatig leren een supramodale representatie wordt opgebouwd die perspectiefonafhankelijk is. Op het neurale niveau vonden we activiteit in de linker en rechter precuneus, gekoppeld aan het verwerken van ruimtelijke informatie. Daarnaast vonden we activiteit in de linker en rechter middelste occipitale gyrus en linker parahippocampus, gekoppeld aan het verwerken van herkenningpunten in beide groepen voor beide taken. Er waren geen verschillen tussen de groepen of tussen de taken en ook geen aanwijzingen voor specifiek egocentrische of allocentrische kennis. Het neurale netwerk werd zowel voor de route- als voor de vogelvluchtgroep gevonden. We kunnen ons afvragen of deze activiteit geassocieerd kan worden met een supramodale representatie, of dat deze voortkomt uit meer algemene taak gerelateerde processen, zoals het maken van een mentale voorstelling. Het gebruik van herkenningpunten speelde een belangrijke rol in alle condities, wat erop wijst dat herkenningpunten niet alleen in route beschrijvingen van belang zijn. Meer onderzoek is nodig om te kunnen concluderen dat een supramodale representatie ook bestaat voor complexe beschrijvingen. Een interessante benadering is de parametrische modulatie zoals beschreven in hoofdstuk 6. Het afstands-vergelijkingseffect dat werd gevonden in

de gedragsdata kan een handvat zijn om dit ook op neuraal niveau te onderzoeken. Een voorlopige conclusie is, dat na het overmatig leren van een complexe ruimtelijke beschrijving proefpersonen een perspectiefonafhankelijke ruimtelijke mentale representatie opbouwen wat gepaard gaat met vergelijkbare neurale activiteit gerelateerd aan het verwerken van herkenningpunten.

Samengenomen laten deze twee hoofdstukken zien dat het verwerken van complexe ruimtelijke beschrijvingen vanuit verschillende perspectieven gekoppeld is aan een uitgebreid neuraal netwerk dat betrokken is bij het verwerken van ruimtelijke taal. Onafhankelijk van het geleerde perspectief werd een uitgebreid fronto-pariëto-temporeel netwerk actief. Het opbouwen van een ruimtelijke mentale representatie aan de hand van een complexe verbale beschrijving is een ingewikkeld proces waarbij meerdere hersengebieden betrokken zijn. Naast overlap liet de parametrische modulatie specifieke neurale gebieden zien. In de vogelvluchtgroep vonden we activiteit in het achterste deel van de linker hippocampus tijdens het verwerken van allocentrische informatie. Interessanter is het feit dat in de routegroep het voorste deel van de linker en rechter hippocampus geactiveerd werd. Dit suggereert dat de allocentrische relaties tussen de locaties van de dieren kunnen worden afgeleid uit de route beschrijving. Dit wordt ondersteund door het feit dat de routegroep ook in staat was om de plattegrond te tekenen en goed presteerde op de vogelvluchttaak. Nadat de proefpersonen ruim de tijd hadden gekregen om de omgeving van de dierentuin te leren, waren ze in staat om deze informatie flexibel te gebruiken. Beide groepen presteerden vergelijkbaar op gedragsniveau en op neuraal niveau, niet alleen in de vogelvluchttaak, maar ook in de routetaak. In eerste instantie (hoofdstuk 6) bestonden er verschillen tussen het leren vanuit een verschillend perspectief, maar deze verschillen verdwenen naarmate de proefpersonen de beschrijvingen vaker hoorden. Het resultaat was een vergelijkbare ruimtelijke representatie die flexibel gebruikt kon worden (hoofdstuk 7). De resultaten van hoofdstuk 6 en 7 laten zien, dat ook complexe ruimtelijke beschrijvingen supramodaal gerepresenteerd kunnen worden, maar meer onderzoek is nodig om dit verder uit te zoeken.

Grounded cognition

De resultaten van de laatste twee hoofdstukken dragen bij aan de bestaande literatuur waaruit blijkt dat metrische informatie overgebracht kan worden door verbale beschrijvingen. Hoewel de complexe beschrijvingen van de dierentuin geen expliciet metrische informatie bevatten, werd er toch een afstands-vergelijkingseffect gevonden. Dit wijst erop dat de proefpersonen een ruimtelijke mentale representatie hadden opgebouwd met afstanden tussen de locaties van de dieren. Een vraag die voortvloeit uit deze bevindingen is hoe dwingend het ontstaan van deze ruimtelijke informatie is. Het laatste deel van dit proefschrift gaat over die vraag. De invloedrijke theorie over “grounded cognition” stelt dat taal begrepen wordt door middel van een perceptuele simulatie (Barsalou et al., 2003; Stanfield & Zwaan, 2001). Dit houdt in dat situaties die in een zin beschreven worden, bijvoorbeeld “*de hond is nat*” gekoppeld worden aan een herbeleving van de situaties in de zin zoals “*de hond*” (dus hoe deze er uit ziet, voelt en ruikt) en “*nat zijn*” (wat nat zijn doet met bijvoorbeeld

de vacht en de geur). Volgens deze definitie is “grounded cognition” modaal, omdat de herbeleving plaatsvindt in een specifieke modaliteit, zoals zicht, geur etc. Ruimtelijke en temporele informatie aan de andere kant vormen hier een uitzondering op. Het is duidelijk dat deze informatie niet beperkt is tot een enkele modaliteit. Het zou kunnen zijn dat ruimtelijke informatie een aparte dimensie vormt. Op basis van “grounded cognition” zou verwacht worden dat deze dimensie ook gesimuleerd wordt bij het begrijpen van taal, zelfs wanneer deze taal niet expliciet ruimtelijk is. In dit deel van het proefschrift beschrijven we twee studies die onderzochten of proefpersonen bij het verwerken van woorden toegang hebben tot ruimtelijke informatie.

Een belangrijke rol in de waarneming van ruimte is weggelegd voor het lichaam. De acties die wij uitvoeren, worden voor een groot deel bepaald door de mogelijkheden van ons lichaam en logischerwijze wordt onze cognitie beïnvloed door deze mogelijkheden. Hoofdstuk 8 beschrijft twee experimenten waarbij werd onderzocht in hoeverre de mentale lichaamsrepresentatie ruimtelijke informatie bevat. Proefpersonen werd gevraagd om twee afstanden van lichaamsdelen met elkaar te vergelijken, bijvoorbeeld welke afstand is groter: *knie-knie* of *heup-heup*? Wederom werd een afstands-vergelijkingseffect gevonden. Proefpersonen presteerden beter wanneer de afstandsverschillen groter werden (Experiment 1) en dit effect kon niet (alleen) verklaard worden door het passeren van een belangrijke lichaamscategorie (*navel-knie* vs. *knie-enkel*, experiment 2). Dit wijst erop dat onze mentale lichaamsrepresentatie de verhoudingen tussen lichaamsdelen correct weergeeft. Proefpersonen deden ook een taak waarbij een simpele beslissing over items genomen diende te worden (is dit levend of niet?). De levende items verwezen steeds naar lichaamsdelen. De resultaten lieten een ruimtelijk “priming” effect zien: beslissingen waren sneller wanneer een lichaamsdeel werd voorafgegaan door een lichaamsdeel dichtbij dan ver af. Dit wijst op het automatisch activeren van ruimtelijke lichaamsinformatie en sluit aan bij de “grounded cognition” theorie. Om de betekenis van lichaamsdelen te kunnen begrijpen wordt een mentale lichaamsrepresentatie geactiveerd die informatie bevat over de locaties van deze lichaamsdelen op het lichaam.

Hoewel de proefpersonen in de afstands-vergelijkingstaak werden geïnstrueerd om zich hun eigen lichaam voor te stellen, was niet te controleren of ze zich wellicht toch een voorstelling van een standaard lichaam hebben gemaakt. Aan de andere kant, de simpele beslissingstaak werd eerst afgenomen en vroeg niet expliciet naar de mentale lichaamsrepresentatie. Sterker nog, deze taak laat zien dat simpelweg door het begrijpen van woorden de proefpersonen automatisch toegang hadden tot een mentale lichaamsrepresentatie. Het meest waarschijnlijk is, dat ze wel hun eigen lichaamsrepresentatie gebruikten, omdat de eigen lichaamservaring gebruikt wordt om die representatie op te bouwen. Dit vermoeden leidt tot een interessante suggestie om deze taken te onderzoeken bij personen met ongebruikelijke lichaamsproporties, zoals personen met een groeistoornis.

Gegeven het feit dat onze mentale lichaamsrepresentatie gebaseerd wordt op onze eigen ervaringen met ons lichaam, kunnen we ons afvragen of een verandering in perceptuele ervaring een effect heeft op deze representatie. Hiermee komen we weer terug bij de interpretatie van de vraag van Molyneux: is tactiele en auditieve informatie voldoende om ruimtelijke informatie over te brengen naar een supramodale representatie, vergelijkbaar met visuele informatie? Om dit te

kunnen beantwoorden is het interessant om onderzoek te doen bij blinde personen.

In het laatste hoofdstuk van dit deel hebben we een begin gemaakt met het onderzoeken van ruimtelijke informatieverwerking in relatie tot “grounded cognition” bij blinde personen. Dit hoofdstuk beschrijft een taak waarbij een beslissing genomen moest worden over een woordpaar: bijvoorbeeld horen *kelder* en *zolder* bij elkaar? Zoals hierboven genoemd zouden we ruimtelijke informatie kunnen beschouwen als een aparte modaliteit die ook gesimuleerd kan worden tijdens het begrijpen van taal. Dit zou betekenen dat woorden zoals *kelder* en *zolder* ook informatie over een gebruikelijke locatie zouden activeren. In hoofdstuk 9 vergeleken we blinde en gebliindoekte ziende proefpersonen om te onderzoeken wat de invloed is van visuele ervaring op het simuleren van ruimtelijke locaties. Beide groepen deden de auditieve beslissingstaak. Proefpersonen hoorden elke keer twee woorden: het eerste woord kwam uit een hoog geplaatste luispreker en het tweede woord uit een laag geplaatste. De locatie van het aangeboden woord werd impliciet gekoppeld aan de betekenis van het woord. Hierdoor werd de prestatie van proefpersonen positief beïnvloed: ze reageerden sneller wanneer de woorden in een standaard situatie werden aangeboden (*zolder* hoog, *kelder* laag). Wanneer de woorden in een niet standaard situatie werden aangeboden (*kelder* hoog, *zolder* laag) reageerden proefpersonen langzamer. Dit effect werd gevonden bij de ziende en de blinde proefpersonen wat erop wijst dat informatie van visuele en niet-visuele modaliteiten voldoende ruimtelijke informatie kan overbrengen. Blijkbaar is de tactiele ervaring die blinde personen hebben met “hoog” en “laag” of “boven” en “onder” beschikbaar om te gebruiken voor een ruimte die buiten armbereik is. De koppeling met ruimtelijke informatie hangt niet direct af van een gebrek aan ervaring met deze items buiten armbereik. Wederom hebben we ondersteunend bewijs gevonden voor het feit dat ruimtelijke informatie van verschillende modaliteiten kan worden gebruikt om een supramodale mentale representatie op te bouwen.

Het zojuist beschreven effect waarbij de positie van de woorden invloed heeft op de beslissing, ook wel “spatial iconicity” genoemd, is niet alleen voor voorwerpen aangetoond, maar ook voor abstracte concepten als macht welke op een verticale as wordt uitgedrukt (Meier & Robinson, 2004; Schubert, 2005). Een van de doelen van hoofdstuk 3 was de gevoeligheid van de linker SMG te bepalen. Hiervoor hadden we verschillende relaties: *links/rechts* (ruimtelijk voorzetsel), samen met (niet ruimtelijk voorzetsel), *groter/kleiner* (ruimtelijk bijwoord) en *ouder/jonger* (niet ruimtelijk bijwoord). Hoewel *ouder/jonger* een niet ruimtelijk bijwoord is, stelt de Metaforische Mapping Theorie (Boroditsky, 2000) dat tijd op een horizontale as wordt uitgedrukt. Volgens die redenering zou *ouder/jonger* ook als ruimtelijk kunnen worden beschouwd. Dit lijkt een aannemelijke interpretatie die ondersteund wordt door linker SMG activiteit voor alle drie de onderzochte relaties. We kunnen speculeren dat horizontale “spatial iconicity” plaatsvindt voor *links/rechts* en *ouder/jonger* en verticale “spatial iconicity” voor *groter/kleiner*. Daarnaast zouden we kunnen verwachten dat de linker SMG ook actief wordt tijdens het verwerken van abstracte concepten zoals macht.

Conclusies

In dit proefschrift is onderzocht of ruimtelijke representaties modaliteits specifiek zijn of omgezet worden in een supramodale representatie die communiceert met alle verschillende modaliteiten. We hebben aangetoond dat afhankelijk van de modaliteit waarin gedrag wordt onderzocht er subtiele verschillen bestaan tussen blinde en ziende proefpersonen. De verschillen tussen blinde en ziende proefpersonen in het haptische domein kunnen worden verklaard op basis van de ruime ervaring die blinden hebben met tactiele waarneming. Aan de andere kant, de resultaten van de verbale studies laten zien dat blinde en ziende proefpersonen juist zeer vergelijkbaar presteren. Daarbij verdwijnen perspectief verschillen tussen complexe beschrijvingen na overmatig leren met als gevolg een functioneel gelijke mentale representatie. In het kort, ruimtelijke informatie kan vanuit verschillende modaliteiten verkregen worden en wordt supramodaal gerepresenteerd. Deze supramodale representatie kan aangesproken worden door middel van taal en in het bijzonder ruimtelijke taal.

Deze resultaten verwijzen terug naar de vraag van Molyneux. We hebben aangetoond dat ruimtelijk representaties supramodaal zijn, wat een wisselwerking met verschillende modaliteiten mogelijk maakt. De originele lezing van de vraag van Molyneux ging over de vraag in hoeverre ruimtelijke informatieverwerking aangeboren is. We hebben beargumenteerd dat ruimtelijke informatie supramodaal gerepresenteerd wordt, wat inhoudt dat er een koppeling met de oorspronkelijke informatiemodaliteit blijft bestaan. We hebben aangetoond dat zicht niet noodzakelijk is om zo'n supramodale representatie op te bouwen, maar de vraag blijft of het mogelijk is om deze informatie automatisch naar het visuele domein om te zetten. Er is een aantal patiënten dat blind geboren is, bij wie het gezichtsvermogen is hersteld. Juist deze patiënten ervoeren aanzienlijke problemen met visuele waarneming (Pascual-Leone & Hamilton, 2001; Von Senden, 1932). Ze konden onderscheid maken tussen verschillende voorwerpen, maar konden ze niet identificeren. Om een voorwerp te kunnen identificeren moest de nieuwe visuele informatie gekoppeld worden aan de bestaande supramodale representatie. Kennelijk moest de visuele informatie gekalibreerd worden, bijvoorbeeld door middel van tast.

Ruimtelijke informatie vanuit verschillende modaliteiten wordt omgezet in een supramodale mentale representatie. Daarbij blijft de koppeling bestaan met de informatiebron. Deze koppeling blijkt gevoelig voor ervaring. Het brein combineert ruimtelijke informatie tot een geheel, maar deze supramodale processen zijn niet volledig aangeboren en afhankelijk van ervaring.

References

- Ackermann, H., Wildgruber, D., Daum, I., & Grodd, W. (1998). Does the cerebellum contribute to cognitive aspects of speech production? A functional magnetic resonance imaging (fMRI) study in humans. *Neuroscience Letters*, *247*(2-3), 187-190.
- Aguirre, G. K., & D'Esposito, M. (1997). Environmental Knowledge Is Subservied by Separable Dorsal/Ventral Neural Areas. *Journal of Neuroscience*, *17*(7), 2512-2518.
- Aguirre, G. K., Detre, J. A., Alsup, D. C., & D'Esposito, M. (1996). The parahippocampus subserves topographical learning in man. *Cerebral Cortex*, *6*(6), 823-829.
- Aleman, A., Van Lee, L., Mantione, M. H., Verkoijen, I. G., & De Haan, E. H. F. (2001). Visual imagery without visual experience: evidence from congenitally totally blind people. *NeuroReport*, *12*(11), 2601-2604.
- Alloway, T. P., Corley, M., & Ramsar, M. (2006). Seeing ahead: Experience and language in spatial perspective. *Memory & Cognition*, *34*(2), 380-386.
- Amedi, A., Floel, A., Knecht, S., Zohary, E., & Cohen, L. G. (2004). Transcranial magnetic stimulation of the occipital pole interferes with verbal processing in blind subjects. *Nature Neuroscience*, *7*(11), 1266-1270.
- Amedi, A., Raz, N., Pianka, P., Malach, R., & Zohary, E. (2003). Early 'visual' cortex activation correlates with superior verbal memory performance in the blind. *Nature Neuroscience*, *6*(7), 758-766.
- Amorapanth, P. X., Widick, P., & Chatterjee, A. (2010). The neural basis for spatial relations. *Journal of Cognitive Neuroscience*, *22*(8), 1739-1753.
- Andersen, E. (1978). Lexical universals of body-part terminology. In J. H. Greenberg (Ed.), *Universals of human language* (Vol. 3 Word structure, pp. 335-368). Stanford: Stanford University Press.
- Ashburner, J., & Friston, K. J. (2005). Unified segmentation. *NeuroImage*, *26*(3), 839-851.
- Avraamides, M. N., Loomis, J. M., Klatzky, R. L., & Golledge, R. G. (2004). Functional equivalence of spatial representations derived from vision and language: evidence from allocentric judgments. *Journal of Experimental Psychology: Learning, Memory and Cognition*, *30*(4), 801-814.
- Baciu, M., Koenig, O., Vernier, M. P., Bedoin, N., Rubin, C., & Segebarth, C. (1999). Categorical and coordinate spatial relations: fMRI evidence for hemispheric specialization. *NeuroReport*, *10*(6), 1373-1378.
- Barsalou, L. W. (1999). Perceptual symbol systems. *Behavioral and Brain Sciences*, *22*(04), 577-660.
- Barsalou, L. W. (2003). Situated simulation in the human conceptual system. *Language and Cognitive Processes*, *18*(5), 513 - 562.
- Barsalou, L. W. (2008). Grounded cognition. *Annual Review of Psychology*, *59*(1), 617-645.
- Barsalou, L. W., Santos, A., Simmons, W. K., & Wilson, C. D. (2008). Language and simulation in conceptual processing. In M. De Vega, A. M. Glenberg & A. C. Graesser (Eds.), *Symbols and embodiment: debates on meaning and cognition* (pp. 245-283). Oxford, UK: Oxford University Press.
- Barsalou, L. W., Simmons, W. K., Barbey, A. K., & Wilson, C. D. (2003). Grounding conceptual knowledge in modality-specific systems. *Trends in Cognitive Sciences*, *7*(2), 84-91.
- Baumgart, F., Kaulisch, T., Tempelmann, C., Gaschler-Markefski, B., Tegeler, C., Schindler, F., et al. (1998). Electrodynamical headphones and woofers for application in magnetic resonance imaging scanners. *Medical Physics*, *25*(10), 2068-2070.
- Binder, J. R., Frost, J. A., Hammeke, T. A., Bellgowan, P. S., Springer, J. A., Kaufman, J. N., et

- al. (2000). Human temporal lobe activation by speech and nonspeech sounds. *Cerebral Cortex*, *10*(5), 512-528.
- Bonino, D., Ricciardi, E., Sani, L., Gentili, C., Vanello, N., Guazzelli, M., et al. (2008). Tactile spatial working memory activates the dorsal extrastriate cortical pathway in congenitally blind individuals. *Archives Italiennes de Biologie*, *146*, 133-146.
- Bookheimer, S. Y. (2002). Functional MRI of language: new approaches to understanding the cortical organization of semantic processing. *Annual Review of Neuroscience*, *25*(1), 151-188.
- Borghgi, A. M., Glenberg, A. M., & Kaschak, M. P. (2004). Putting words in perspective. *Memory & Cognition*, *32*(6), 863-873.
- Boroditsky, L. (2000). Metaphoric structuring: understanding time through spatial metaphors. *Cognition*, *75*(1), 1-28.
- Brunyé, T. T., Rapp, D. N., & Taylor, H. A. (2008). Representational flexibility and specificity following spatial descriptions of real-world environments. *Cognition*, *108*(2), 418-443.
- Brunyé, T. T., & Taylor, H. A. (2008a). Extended experience benefits spatial mental model development with route but not survey descriptions. *Acta Psychologica*, *127*(2), 340-354.
- Brunyé, T. T., & Taylor, H. A. (2008b). Working memory in developing and applying mental models from spatial descriptions. *Journal of Memory and Language*, *58*(3), 701-729.
- Bryant, D. J. (1997). Representing space in language and perception. *Mind & Language*, *12*(3-4), 239-264.
- Burgess, N. (2006). Spatial memory: how egocentric and allocentric combine. *Trends In Cognitive Sciences*, *10*(12), 551-557.
- Burgess, N. (2008). Spatial cognition and the brain. *Annals of the New York Academy of Sciences*, *1124*(1), 77-97.
- Burgess, N., Maguire, E. A., & O'Keefe, J. (2002). The human hippocampus and spatial and episodic memory. *Neuron*, *35*(4), 625-641.
- Burton, H., & McLaren, D. G. (2006). Visual cortex activation in late-onset, Braille naive blind individuals: an fMRI study during semantic and phonological tasks with heard words. *Neuroscience Letters*, *392*(1-2), 38-42.
- Burton, H., Snyder, A. Z., Conturo, T. E., Akbudak, E., Ollinger, J. M., & Raichle, M. E. (2002a). Adaptive changes in early and late blind: a fMRI study of Braille reading. *Journal of Neurophysiology*, *87*(6), 589-611.
- Burton, H., Snyder, A. Z., Diamond, J. B., & Raichle, M. E. (2002b). Adaptive changes in early and late blind: a fMRI study of verb generation to heard nouns. *Journal of Neurophysiology*, *88*(6), 3359-3371.
- Byrne, R. W., & Salter, E. (1983). Distances and directions in the cognitive maps of the blind. *Canadian Journal of Psychology*, *37*, 293-299.
- Carlson-Radvansky, L. A., Covey, E. S., & Lattanzi, K. M. (1999). "what" effects on "where": functional influences on spatial relations. *Psychological Science*, *10*(6), 516-521.
- Carlson-Radvansky, L. A., & Irwin, D. E. (1993). Frames of reference in vision and language: Where is above? *Cognition*, *46*, 223-244.
- Carlson-Radvansky, L. A., & Irwin, D. E. (1994). Reference frame activation during spatial term assignment. *Journal of Memory and Language*, *33*(5), 646-671.
- Carlson-Radvansky, L. A., & Jiang, Y. (1998). Inhibition accompanies reference-frame selection. *Psychological Science*, *9*(5), 386-391.
- Carlson-Radvansky, L. A., & Logan, G. D. (1997). The influence of reference frame selection on spatial template construction. *Journal of Memory and Language*, *37*(3), 411-437.
- Carlson-Radvansky, L. A., & Radvansky, G. A. (1996). The influence of functional relations on

- spatial term selection. *Psychological Science*, 7(1), 56-60.
- Carlson, L. A. (1999). Selecting a reference frame. *Spatial Cognition and Computation*, 1(4), 365-379.
- Carlson, L. A., & Kenny, R. (2006). Interpreting spatial terms involves simulating interactions. *Psychonomic Bulletin & Review*, 13(4), 682-688.
- Carlson, L. A., & Logan, G. D. (2001). Using spatial terms to select an object. *Memory & Cognition*, 29(6), 883-892.
- Carlson, L. A., & Van Deman, S. R. (2004). The space in spatial language. *Journal of Memory and Language*, 51(3), 418-436.
- Carlson, L. A., & Van Deman, S. R. (2008). Inhibition within a reference frame during the interpretation of spatial language. *Cognition*, 106(1), 384-407.
- Carlson, L. A., West, R., Taylor, H. A., & Herndon, R. W. (2002). Neural correlates of spatial term use. *Journal of Experimental Psychology: Human Perception and Performance*, 28(6), 1391-1408.
- Carpenter, P. A., Just, M. A., Keller, T. A., Eddy, W. F., & Thulborn, K. R. (1999). Time course of fMRI-activation in language and spatial networks during sentence comprehension. *NeuroImage*, 10(2), 216-224.
- Cattaneo, Z., & Vecchi, T. (2008). Supramodality effects in visual and haptic spatial processes. *Journal of Experimental Psychol: Learning, Memory and Cognition*, 34(3), 631-642.
- Cattaneo, Z., Vecchi, T., Cornoldi, C., Mammarella, I., Bonino, D., Ricciardi, E., et al. (2008). Imagery and spatial processes in blindness and visual impairment. *Neuroscience & Biobehavioral Reviews*, 32(8), 1346-1360.
- Chatterjee, A. (2001). Language and space: some interactions. *Trends in Cognitive Sciences*, 5(2), 55-61.
- Chatterjee, A. (2008). The neural organization of spatial thought and language. *Seminars in Speech and Language*, 29(03), 226-238.
- Chomsky, N. (2000). *New horizons in the study of language and mind*. Cambridge: Cambridge University Press.
- Clark, H. H., & Chase, W. G. (1972). On the process of comparing sentences against pictures. *Cognitive Psychology*, 3(3), 472-517.
- Cohen, L. G., Celnik, P., Pascual-Leone, A., Corwell, B., Faiz, L., Dambrosia, J., et al. (1997). Functional relevance of cross-modal plasticity in blind humans. *Nature*, 389(6647), 180-184.
- Colby, C. L., & Goldberg, M. E. (1999). Space and attention in parietal cortex. *Annual Review of Neuroscience*, 22(1), 319-349.
- Committeri, G., Galati, G., Paradis, A. L., Pizzamiglio, L., Berthoz, A., & LeBihan, D. (2004). Reference frames for spatial cognition: different brain areas are involved in viewer-, object-, and landmark-centered judgements about object location. *Journal of Cognitive Neuroscience*, 16(9), 1517-1535.
- Connolly, A. C., Gleitman, L. R., & Thompson-Schill, A. L. (2007). Effect of congenital blindness on the semantic representation of some everyday concepts. *Proceedings of the National Academy of Sciences*, 104(20), 8241-8246.
- Cornoldi, C., & De Beni, R. (2007). Imagery and blindness. In S. Della Sala (Ed.), *Tall tales about the mind & brain: Separating fact from fiction* (pp. 369-379). New York, NY, USA: Oxford University Press.
- Cornoldi, C., De Beni, R., Giusberti, F., & Massironi, M. (1998). Memory and imagery: a visual trace is not a mental image. In M. Conway, S. Gathercole & C. Cornoldi (Eds.), *Theories of memory* (pp. 87-110). Hove, UK: Psychology Press.
- Cornoldi, C., Tinti, C., Mammarella, I. C., Re, A. M., & Varotto, D. (2009). Memory for an imagined pathway and strategy effects in sighted and in totally congenitally blind individuals.

- Acta Psychologica*, 130(1), 11-16.
- Coslett, H. B. (1999). Spatial influences on motor and language function. *Neuropsychologia*, 37(6), 695-706.
- Coventry, K. R., & Garrod, S. C. (2004). *Saying, seeing and acting. The psychological semantics of spatial prepositions*. Hove and New York: Psychology Press Taylor & Francis Group.
- Creem, S. H., & Proffitt, D. R. (2001). Defining the cortical visual systems: “what”, “where”, and “how”. *Acta Psychologica*, 107(1-3), 43-68.
- Crinion, J., Ashburner, J., Leff, A., Brett, M., Price, C. J., & Friston, K. J. (2007). Spatial normalization of lesioned brains: Performance evaluation and impact on fMRI analyses. *NeuroImage*, 37(3), 866-875.
- Cuenod, C. A., Bookheimer, S. Y., Hertz-Pannier, L., Zeffiro, T. A., Theodore, W. H., & Le Bihan, D. M. D. (1995). Functional MRI during word generation, using conventional equipment: A potential tool for language localization in the clinical environment. *Neurology*, 45(10), 1821-1827.
- Damasio, H., Grabowski, T. J., Tranel, D., Ponto, L. L., Hichwa, R. D., & Damasio, A. R. (2001). Neural correlates of naming actions and of naming spatial relations. *NeuroImage*, 13(6), 1053-1064.
- De Beni, R., & Cornoldi, C. (1988). Imagery limitations in totally congenitally blind subjects. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 14(4), 650-655.
- De Lillo, C. (2004). Imposing structure on a Corsi-type task: Evidence for hierarchical organisation based on spatial proximity in serial-spatial memory. *Brain and Cognition*, 55(3), 415-426.
- De Vega, M., Cocude, M., Denis, M., Rodrigo, M. J., & Zimmer, H. D. (2001). The interface between language and visuo-spatial representations. In M. Denis, R. H. Logie, C. Cornoldi, M. De Vega & J. Engelkamp (Eds.), *Imagery, language, and visuo-spatial thinking* (pp. 109-136). Hove, England: Psychology Press.
- De Vignemont, F. (2009). Body schema and body image - Pros and cons. *Neuropsychologia*, 48(3), 669-680.
- Degnaar, M. J. L. (1992). *Molyneux's problem: Three centuries of discussion on the perception of forms*. Erasmus University Rotterdam, Rotterdam.
- Denis, M. (2008). Assessing the symbolic distance effect in mental images constructed from verbal descriptions: A study of individual differences in the mental comparison of distances. *Acta Psychologica*, 127(1), 197-210.
- Denis, M., & Cocude, M. (1992). Structural properties of visual images constructed from poorly or well-structured verbal descriptions. *Memory & Cognition*, 20(5), 497-506.
- Denis, M., Pazzaglia, F., Cornoldi, C., & Bertolo, L. (1999). Spatial discourse and navigation: an analysis of route directions in the city of Venice. *Applied Cognitive Psychology*, 13(2), 145-174.
- Denis, M., & Zimmer, H. D. (1992). Analog properties of cognitive maps constructed from verbal descriptions. *Psychological Research*, 54(4), 286-298.
- Deyzac, E., Logie, R. H., & Denis, M. (2006). Visuospatial working memory and the processing of spatial descriptions. *British Journal of Psychology*, 97(2), 217-243.
- Dijkerman, H. C., & De Haan, E. H. F. (2007). Somatosensory processing subserving perception and action: Dissociations, interactions, and integration. *Behavioral and Brain Sciences*, 30(02), 224-230.
- Dodds, A. G., Howarth, C. L., & Carter, D. C. (1982). The mental maps of the blind: The role of previous visual experience. *Journal of Visual Impairment & Blindness*, 76, 5-12.
- Doeller, C. F., King, J. A., & Burgess, N. (2008). Parallel striatal and hippocampal systems for

- landmarks and boundaries in spatial memory. *Proceedings of the National Academy of Sciences*, 105(15), 5915-5920.
- Downing, P. E., Jiang, Y., Shuman, M., & Kanwisher, N. (2001). A cortical area selective for visual processing of the human body. *Science*, 293(5539), 2470-2473.
- Emmorey, K., Damasio, H., McCullough, S., Grabowski, T. J., Ponto, L. L. B., Hichwa, R. D., et al. (2002). Neural systems underlying spatial language in american sign language. *NeuroImage*, 17(2), 812-824.
- Epstein, R. A., Parker, W. E., & Feiler, A. M. (2007). Where am I now? Distinct roles for parahippocampal and retrosplenial cortices in place recognition. *J. Neurosci.*, 27(23), 6141-6149.
- Farah, M. J. (1989). The neural basis of mental imagery. *Trends in Neuroscience*, 12(10), 395-399.
- Farah, M. J., Hammond, K. M., Levine, D. N., & Calvanio, R. (1988a). Visual and spatial mental imagery: dissociable systems of representation. *Cognitive Psychology*, 20(4), 439-462.
- Farah, M. J., Péronnet, F., Gouon, M. A., & Giard, M. H. (1988b). Electrophysiological evidence for a shared representational medium for visual images and visual percepts. *Journal of Experimental Psychology: General*, 117(3), 248-257.
- Fauconnier, G. (1999). Creativity, simulation, and conceptualization. *Behavioral and Brain Sciences*, 22, 615-615.
- Fodor, J. A. (1975). *The language of thought*. Cambridge, MA: Harvard University Press.
- Franklin, N., & Tversky, B. (1990). Searching imagined environments. *Journal of Experimental Psychology: General*, 119(1), 63-76.
- Freksa, C., Barkowsky, T., & Klippel, A. (1999). Spatial symbol systems and spatial cognition: A computer science perspective on perception-based symbol processing. *Behavioral and Brain Sciences*, 22, 616-617.
- Friedman, A. (2009). The role of categories and spatial cuing in global-scale location estimates. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 35(1), 94-112.
- Friston, K. J. (1997). Testing for anatomically specified regional effects. *Human Brain Mapping*, 5(2), 133-136.
- Friston, K. J., Holmes, A., Poline, J. B., Price, C. J., & Frith, C. D. (1996). Detecting activations in PET and fMRI: levels of inference and power. *NeuroImage*, 4(3), 223-235.
- Friston, K. J., Penny, W. D., & Glaser, D. E. (2005). Conjunction revisited. *NeuroImage*, 25(3), 661-667.
- Galati, G., Lobel, E., Vallar, G., Berthoz, A., Pizzamiglio, L., & Le Bihan, D. (2000). The neural basis of egocentric and allocentric coding of space in humans: a functional magnetic resonance study. *Experimental Brain Research*, 133(2), 156-164.
- Galati, G., Pelle, G., Berthoz, A., & Committeri, G. (2010). Multiple reference frames used by the human brain for spatial perception and memory. *Experimental Brain Research*, 206(2), 109-120.
- Ganis, G., Thompson, W. L., & Kosslyn, S. M. (2004). Brain areas underlying visual mental imagery and visual perception: an fMRI study. *Cognitive Brain Research*, 20(2), 226-241.
- Gernsbacher, M. A., & Kaschak, M. P. (2003). Neuroimaging studies of language production and comprehension. *Annual Review of Psychology*, 54(1), 91-114.
- Ghaem, O., Mellet, E., Crivello, F., Tzourio, N., Mazoyer, B., Berthoz, A., et al. (1997). Mental navigation along memorized routes activates the hippocampus, precuneus, and insula. *Neuroreport*, 8(3), 739-744.
- Giudice, N. A., Bakdash, J. Z., & Legge, G. E. (2007). Wayfinding with words: spatial learning and navigation using dynamically updated verbal descriptions. *Psychological Research*, 71(3), 347-358.
- Gizewski, E. R., Gasser, T., de Greiff, A., Boehm, A., & Forsting, M. (2003). Cross-modal plasticity for sensory and motor activation patterns in blind subjects. *NeuroImage*, 19(3), 968-

- 975.
- Glaser, W. R. (1992). Picture naming. *Cognition*, *42*(1-3), 61-105.
- Glenberg, A. M. (2010). Embodiment as a unifying perspective for psychology. *Wiley Interdisciplinary Reviews: Cognitive Science*, *1*(4), 586-596.
- Glover, G. H., Li, T.-Q., & Ress, D. (2000). Image-based method for retrospective correction of physiological motion effects in fMRI: RETROICOR. *Magnetic Resonance in Medicine*, *44*(1), 162-167.
- Goyal, M. S., Hansen, P. J., & Blakemore, C. B. (2006). Tactile perception recruits functionally related visual areas in the late-blind. *NeuroReport*, *17*(13), 1381-1384.
- Grefkes, C., Weiss, P. H., Zilles, K., & Fink, G. R. (2002). Crossmodal processing of object features in human anterior intraparietal cortex: an fMRI study implies equivalencies between humans and monkeys. *Neuron*, *35*(1), 173-184.
- Halligan, P. W., & Marshall, J. C. (1991). Left neglect for near but not far space in man. *Nature*, *350*(6318), 498-500.
- Hamilton, R., Keenan, J. P., Catala, M., & Pascual-Leone, A. (2000). Alexia for Braille following bilateral occipital stroke in an early blind woman. *NeuroReport*, *11*(02), 237-240.
- Hartley, T., Maguire, E. A., Spiers, H. J., & Burgess, N. (2003). The well-worn route and the path less traveled: Distinct neural bases of route following and wayfinding in humans. *Neuron*, *37*(5), 877-888.
- Haspelmath, M. (1997). *From space to time: temporal adverbials in the world's languages*. Newcastle, UK: Lincom Europa.
- Head, H., & Holmes, G. (1912). Sensory disturbances from cerebral lesions. *Brain*, *34*, 102-254.
- Heine, B., Claudi, U., & Hünemeyer, F. (1991). *Grammaticalization*. Chicago, IL: University of Chicago Press.
- Henson, R. N. A. (2003). Analysis of fMRI time series. In R. S. J. Frackowiak, K. J. Friston, C. Frith, R. Dolan, C. J. Price, S. Zeki, J. Ashburner & W. D. Penny (Eds.), *Human Brain Function* (2nd ed.). London: Academic Press.
- Herholz, K., Thiel, A., Wienhard, K., Pietrzyk, U., von Stockhausen, H. M., Karbe, H., et al. (1996). Individual functional anatomy of verb generation. *NeuroImage*, *3*(3), 185-194.
- Hommel, B., Gehrke, J., & Knuf, L. (2000). Hierarchical coding in the perception and memory of spatial layouts. *Psychological Research*, *64*(1), 1-10.
- Hubona, G. S., Everett, S., Marsh, E., & Wauchope, K. (1998). Mental representations of spatial language. *International Journal of Human-Computer Studies*, *48*(6), 705-728.
- Hund, A. M., Haney, K. H., & Seanor, B. D. (2008). The role of recipient perspective in giving and following wayfinding directions. *Applied Cognitive Psychology*, *22*(7), 896-916.
- Huttenlocher, J., Hedges, L. V., & Duncan, S. (1991). Categories and particulars: Prototype effects in estimating spatial location. *Psychological Review*, *98*(3), 352-376.
- Iaria, G., Petrides, M., Dagher, A., Pike, B., & Bohbot, V. D. (2003). Cognitive strategies dependent on the hippocampus and caudate nucleus in human navigation: variability and change with practice. *Journal of Neuroscience*, *23*(13), 5945-5952.
- Jackendoff, R. (2002). *Foundations of language: brain, meaning, grammar, evolution*. Oxford University Press.
- Jager, G., & Postma, A. (2003). On the hemispheric specialization for categorical and coordinate spatial relations: a review of the current evidence. *Neuropsychologia*, *41*, 504-515.
- Janzen, G., & Weststeijn, C. G. (2007). Neural representation of object location and route direction: An event-related fMRI study. *Brain Research*, *1165*, 116-125.
- Jeannerod, M. (1995). Mental imagery in the motor context. *Neuropsychologia*, *33*(11), 1419-1432.

- Kaski, D. (2002). Revision: is visual perception a requisite for visual imagery? *Perception*, *31*(6), 717-731.
- Kemmerer, D. (2005). The spatial and temporal meanings of English prepositions can be independently impaired. *Neuropsychologia*, *43*(5), 797-806.
- Kemmerer, D. (2006). The semantics of space: Integrating linguistic typology and cognitive neuroscience. *Neuropsychologia*, *44*(9), 1607-1621.
- Kemmerer, D., & Tranel, D. (2000). A double dissociation between linguistic and perceptual representations of spatial relationships. *Cognitive Neuropsychology*, *17*(5), 393-414.
- Kerr, N. H. (1983). The role of vision in “visual imagery” experiments: evidence from the congenitally blind. *Journal of Experimental Psychology: General*, *112*(2), 265-277.
- Klarhöfer, M., Dilharreguy, B., van Gelderen, P., & Moonen, C. T. (2003). A PRESTO-SENSE sequence with alternating partial-Fourier encoding for rapid susceptibility-weighted 3D MRI time series. *Magnetic Resonance in Medicine*, *50*(4), 830-838.
- Klatzky, R. L., & Golledge, R. G. (1995). Performance of blind and sighted persons on spatial tasks. *Journal of Visual Impairment & Blindness*, *89*(1), 70-82.
- Knauff, M., & May, E. (2006). Mental imagery, reasoning, and blindness. *Quarterly Journal of Experimental Psychology*, *59*(1), 161-177.
- Kobayashi, M., Takeda, M., Hattori, N., Fukunaga, M., Sasabe, T., Inoue, N., et al. (2004). Functional imaging of gustatory perception and imagery: “top-down” processing of gustatory signals. *NeuroImage*, *23*(4), 1271-1282.
- Kosslyn, S. M. (1987). Seeing and imagining in the cerebral hemispheres: a computational approach. *Psychological Review*, *94*(2), 148-175.
- Kosslyn, S. M. (1988). Aspects of a cognitive neuroscience of mental imagery. *Science*, *240*(4859), 1621-1626.
- Kosslyn, S. M., Ball, T. M., & Reiser, B. J. (1978). Visual images preserve metric spatial information: evidence from studies of image scanning. *Journal of Experimental Psychology: Human Perception and Performance*, *4*(1), 47-60.
- Kosslyn, S. M., Ganis, G., & Thompson, W. L. (2001). Neural foundations of imagery. *Nature Reviews Neuroscience*, *2*(9), 635-642.
- Kosslyn, S. M., Ganis, G., & Thompson, W. L. (2003). Mental imagery: against the nihilistic hypothesis. *Trends in Cognitive Sciences*, *7*(3), 109-111.
- Kosslyn, S. M., Koenig, O., Barrett, A. M., Cave, C. B., Tang, J., & Gabrieli, J. D. (1989). Evidence for two types of spatial representations: hemispheric specialization for categorical and coordinate relations. *Journal of Experimental Psychology: Human Perception and Performance*, *15*(4), 723-735.
- Kosslyn, S. M., Margolis, J. A., Barrett, A. M., Goldknopf, E. J., & Daly, P. F. (1990). Age differences in imagery abilities. *Child Development*, *61*(4), 995-1010.
- Kosslyn, S. M., Pascual-Leone, A., Felician, O., Camposano, S., Keenan, J. P., Thompson, W. L., et al. (1999). The role of area 17 in visual imagery: convergent evidence from PET and rTMS. *Science*, *284*(5411), 167-170.
- Kriegeskorte, N., Simmons, W. K., Bellgowan, P. S. F., & Baker, C. I. (2009). Circular analysis in systems neuroscience: the dangers of double dipping. *Nature Neuroscience*, *12*(5), 535-540.
- Kupers, R., Fumal, A., de Noordhout, A. M., Gjedde, A., Schoenen, J., & Ptito, M. (2006). Transcranial magnetic stimulation of the visual cortex induces somatotopically organized qualia in blind subjects. *Proceedings of the National Academy of Sciences*, *103*(35), 13256-13260.
- Lacey, S., Campbell, C., & Sathian, K. (2007). Vision and touch: Multiple or multisensory representations of objects? *Perception*, *36*(10), 1513-1521.

- Laeng, B. (1994). Lateralization of categorical and coordinate spatial functions: A study of unilateral stroke patients. *Journal of Cognitive Neuroscience*, *6*(3), 189-203.
- Laeng, B., Chabris, C. F., & Kosslyn, S. M. (2003). Asymmetries in encoding spatial relations. In K. Hugdahl & R. Davidson (Eds.), *The Asymmetrical Brain* (pp. 303-339). Cambridge, MA: The MIT Press.
- Lambert, S., Sampaio, E., Mauss, Y., & Scheiber, C. (2004). Blindness and brain plasticity: contribution of mental imagery? An fMRI study. *Cognitive Brain Research*, *20*(1), 1-11.
- Landau, B., & Jackendoff, R. (1993). "What" and "where" in spatial language and spatial cognition. *Behavioral and Brain Sciences*, *16*, 217-265.
- Latini-Corazzini, L., Nesa, M. P., Ceccaldi, M., Guedj, E., Thinus-Blanc, C., Cauda, F., et al. (2010). Route and survey processing of topographical memory during navigation. *Psychological Research*, *74*(6), 545-559.
- Levelt, W. J. M. (1984). Some perceptual limitations on talking about space. In A. J. van Doorn, W. A. van der Grind & J. J. Koenderink (Eds.), *Limits in perception* (pp. 323-358). Utrecht: VNU Science Press.
- Levinson, S. (1996). Frames of reference and Molyneux's question: Crosslinguistic evidence. In P. e. a. e. Bloom (Ed.), *Language and space* (pp. 109-169). Cambridge, MA: MIT Press.
- Linde, C., & Labov, W. (1975). Spatial networks as a site for the study of language and thought. *Language*, *51*(4), 924-939.
- Logan, G. D. (1994). Spatial attention and the apprehension of spatial relations. *Journal of Experimental Psychology: Human Perception and Performance*, *20*(5), 1015-1036.
- Logan, G. D. (1995). Linguistic and conceptual control of visual spatial attention. *Cognitive Psychology*, *28*(2), 103-174.
- Logan, G. D. (1999). Selection for cognition: cognitive constraints on visual spatial attention. *Visual Cognition*, *6*(1), 55 - 81.
- Logan, G. D., & Compton, B. J. (1996). Distance and distraction effects in the apprehension of spatial relations. *Journal of Experimental Psychology: Human Perception and Performance*, *22*(1), 159-172.
- Logan, G. D., & Sadler, D. D. (1996). A computational analysis of the apprehension of spatial relations. In P. Bloom, M. A. Peterson, L. Nadel & M. Garrett (Eds.), *Language and space* (pp. 493-529). Cambridge, MA: MIT Press.
- Longo, M. R., Azañón, E., & Haggard, P. (2009). More than skin deep: Body representation beyond primary somatosensory cortex. *Neuropsychologia*, *48*(3), 655-668.
- Loomis, J. M., Klatzky, R. L., Avraamides, M., Lippa, Y., & Golledge, R. G. (2007). Functional equivalence of spatial images produced by perception and spatial language. In F. Mast & J. Jäncke (Eds.), *Spatial processing in navigation, imagery, and perception* (pp. 29-48). New York: Springer.
- Loomis, J. M., Klatzky, R. L., Golledge, R. G., Cicinelli, J. G., Pellegrino, J. W., & Fry, P. A. (1993). Nonvisual navigation by blind and sighted: assessment of path integration ability. *Journal of Experimental Psychology: General*, *122*(1), 73-91.
- Louwerse, M. M. (2008). Embodied relations are encoded in language. *Psychonomic Bulletin & Review*, *15*(4), 838-844.
- MacLeod, C. M., Hunt, E. B., & Mathews, N. N. (1978). Individual differences in the verification of sentence-picture relationships. *Journal of Verbal Learning & Verbal Behavior*, *17*(5), 493-507.
- Maguire, E. A., Burgess, N., Donnett, J. G., Frackowiak, R. S., Frith, C. D., & O'Keefe, J. (1998a). Knowing where and getting there: a human navigation network. *Science*, *280*(5365), 921-924.
- Maguire, E. A., Frackowiak, R. S. J., & Frith, C. D. (1996). Learning to find your way: A role for

- the human hippocampal formation. *Proceedings: Biological Sciences*, 263(1377), 1745-1750.
- Maguire, E. A., Frith, C. D., Burgess, N., Donnett, J. G., & O'Keefe, J. (1998b). Knowing where things are: Parahippocampal involvement in encoding object locations in virtual large-scale space. *Journal of Cognitive Neuroscience*, 10(1), 61-76.
- Mahon, B. Z., Anzellotti, S., Schwarzbach, J., Zampini, M., & Caramazza, A. (2009). Category-specific organization in the human brain does not require visual experience. *Neuron*, 63(3), 397-405.
- Maki, R. H. (1981). Categorization and distance effects with spatial linear orders. *Journal of Experimental Psychology: Human Learning and Memory*, 7(1), 15-32.
- Marques, J. F. (2010). The effect of visual deprivation on the organization of conceptual knowledge: testing the grounded cognition hypothesis. *Experimental Psychology*, 57(2), 83-88.
- Matteau, I., Kupers, R., Ricciardi, E., Pietrini, P., & Ptito, M. (2010). Beyond visual, aural and haptic movement perception: hMT+ is activated by electrotactile motion stimulation of the tongue in sighted and in congenitally blind individuals. *Brain Research Bulletin*, 82(5-6), 264-270.
- Mazard, A., Laou, L., Joliot, M., & Mellet, E. (2005). Neural impact of the semantic content of visual mental images and visual percepts. *Cognitive Brain Research*, 24(3), 423-435.
- McMullen, P. A., & Jolicoeur, P. (1990). The spatial frame of reference in object naming and discrimination of left-right reflections. *Memory & Cognition*, 18(1), 99-115.
- McNamara, T. P. (1986). Mental representations of spatial relations. *Cognitive Psychology*, 18(1), 87-121.
- Meier, B. P., & Robinson, M. D. (2004). Why the sunny side is up. *Psychological Science*, 15(4), 243-247.
- Mellet, E., Bricogne, S., Crivello, F., Mazoyer, B., Denis, M., & Tzourio-Mazoyer, N. (2002). Neural basis of mental scanning of a topographic representation built from a text. *Cerebral Cortex*, 12(12), 1322-1330.
- Mellet, E., Bricogne, S., Tzourio-Mazoyer, N., Ghaëm, O., Petit, L., Zago, L., et al. (2000). Neural correlates of topographic mental exploration: The impact of route versus survey perspective learning. *NeuroImage*, 12(5), 588-600.
- Melzer, P., Morgan, V. L., Pickens, D. R., Price, R. R., Wall, R. S., & Ebner, F. F. (2001). Cortical activation during Braille reading is influenced by early visual experience in subjects with severe visual disability: A correlational fMRI study. *Human Brain Mapping*, 14(3), 186-195.
- Merabet, L. B., Thut, G., Murray, B., Andrews, J., Hsiao, S., & Pascual-Leone, A. (2004). Feeling by sight or seeing by touch? *Neuron*, 42(1), 173-179.
- Millar, S. (1994). *Understanding and representing space. Theory and evidence from studies with blind and sighted children*. Oxford: Clarendon Press.
- Miller, G. A., & Johnson-Laird, P. N. (1976). *Language and perception*. Cambridge: Cambridge University Press.
- Moyer, R. S., & Bayer, R. H. (1976). Mental comparison and the symbolic distance effect. *Cognitive Psychology*, 8, 228-246.
- Neggers, S. F. W., Hermans, E., & Ramsey, N. F. (2008). Enhanced sensitivity with fast three-dimensional blood-oxygen-level-dependent functional MRI: comparison of SENSE-PRESTO and 2D-EPI at 3 T. *NMR in Biomedicine*, 21(7), 663-676.
- Neggers, S. F. W., Van der Lubbe, R. H. J., Ramsey, N. F., & Postma, A. (2006). Interactions between ego- and allocentric neuronal representations of space. *NeuroImage*, 31(1), 320-331.
- Nichols, T., Brett, M., Andersson, J., Wager, T., & Poline, J.-B. (2005). Valid conjunction inference with the minimum statistic. *NeuroImage*, 25(3), 653-660.
- Noordzij, M. L., Neggers, S. F. W., Ramsey, N. F., & Postma, A. (2008). Neural correlates of

- locative prepositions. *Neuropsychologia*, 46(5), 1576-1580.
- Noordzij, M. L., & Postma, A. (2005). Categorical and metric distance information in mental representations derived from route and survey descriptions. *Psychological Research*, 69(3), 221-232.
- Noordzij, M. L., van der Lubbe, R. H. J., Neggers, S. F. W., & Postma, A. (2004). Spatial tapping interferes with the processing of linguistic spatial relations. *Canadian Journal of Experimental Psychology*, 58(4), 259-271.
- Noordzij, M. L., Van der Lubbe, R. H. J., & Postma, A. (2005). Strategic and automatic components in the processing of linguistic spatial relations. *Acta Psychologica*, 119(1), 1-20.
- Noordzij, M. L., Van der Lubbe, R. H. J., & Postma, A. (2006a). Electrophysiological support for strategic processing of spatial sentences. *Psychophysiology*, 43(3), 277-286.
- Noordzij, M. L., Zuidhoek, S., & Postma, A. (2006b). The influence of visual experience on the ability to form spatial mental models based on route and survey descriptions. *Cognition*, 100(2), 321-342.
- Noordzij, M. L., Zuidhoek, S., & Postma, A. (2007). The influence of visual experience on visual and spatial imagery. *Perception*, 36(1), 101-112.
- Noppeney, U. (2007). The effects of visual deprivation on functional and structural organization of the human brain. *Neuroscience and Biobehavioral Reviews*, 31, 1169-1180.
- Noppeney, U., Friston, K. J., Ashburner, J., Frackowiak, R. S. J., & Price, C. J. (2005). Early visual deprivation induces structural plasticity in gray and white matter. *Current Biology*, 15(13), R488-490.
- Noppeney, U., Friston, K. J., & Price, C. J. (2003). Effects of visual deprivation on the organization of the semantic system. *Brain*, 126(7), 1620-1627.
- Ofan, R. H., & Zohary, E. (2007). Visual cortex activation in bilingual blind individuals during use of native and second language. *Cerebral Cortex*, 17(6), 1249-1259.
- Oldfield, R. C. (1971). The assessment and analysis of handedness: The Edinburgh inventory. *Neuropsychologia*, 9(1), 97-113.
- Ota, H., Fujii, T., Suzuki, K., Fukatsu, R., & Yamadori, A. (2001). Dissociation of body-centered and stimulus-centered representations in unilateral neglect. *Neurology*, 57(11), 2064-2069.
- Paivio, A. (1971). *Imagery and verbal processes*. New York: Holt, Rinehart, & Winston.
- Paivio, A. (1978). Comparisons of mental clocks. *Journal of Experimental Psychology: Human Perception and Performance*, 4(1), 61-71.
- Pascual-Leone, A., Amedi, A., Fregni, F., & Merabet, L. B. (2005). The plastic human brain cortex. *Annual Review of Neuroscience*, 28, 377-401.
- Pascual-Leone, A., & Hamilton, R. (2001). The metamodal organization of the brain. *Progress in Brain Research*, 134, 427-445.
- Pecher, D., Van Dantzig, S., Boot, I., Zanolie, K., & Huber, D. E. (2010). Congruency between word position and meaning is caused by task induced spatial attention. *Frontiers in Psychology*, 1(30), doi: 10.3389/fpsyg.2010.00030.
- Pecher, D., van Dantzig, S., Zwaan, R. A., & Zeelenberg, R. (2009). Language comprehenders retain implied shape and orientation of objects. *The Quarterly Journal of Experimental Psychology*, 62(6), 1108-1114.
- Pecher, D., Zeelenberg, R., & Barsalou, L. W. (2003). Verifying different-modality properties for concepts produces switching costs. *Psychological Science*, 14(2), 119-124.
- Pederson, E. (2003). How many reference frames? In C. F. e. al. (Ed.), *Spatial Cognition III: Routes and Navigation, Human Memory and Learning, Spatial Representation and Spatial Learning* (Vol. 2685, pp. 287-304): Springer Verlag.
- Pederson, E., Levinson, S., Danziger, E., Kita, S., Wilkins, D., & Senft, G. (1998). Semantic

- typology and spatial conceptualization. *Language*, 74(3), 557-589.
- Péruch, P., Chabanne, V., Nesa, M.-P., Thinus-Blanc, C., & Denis, M. (2006). Comparing distances in mental images constructed from visual experience or verbal descriptions: The impact of survey versus route perspective. *The Quarterly Journal of Experimental Psychology*, 59(11), 1950-1967.
- Pietrini, P., Furey, M. L., Ricciardi, E., Gobbin, M. I., Wu, W. H. C., Cohen, L. G., et al. (2004). Beyond sensory images: Object-based representation in the human ventral pathway. *Proceedings of the National Academy of Sciences of the United States of America*, 101(15), 5658-5663.
- Poirier, C., Collignon, O., Scheiber, C., Renier, L., Vanlierde, A., Tranduy, D., et al. (2006). Auditory motion perception activates visual motion areas in early blind subjects. *NeuroImage*, 31(1), 279-285.
- Poldrack, R. A., Clark, J., Pare-Blagoev, E. J., Shohamy, D., Creso Moyano, J., Myers, C., et al. (2001). Interactive memory systems in the human brain. *Nature*, 414(6863), 546-550.
- Postma, A., Huntjens, R. J. C., Meuwissen, M., & Laeng, B. (2006). The time course of spatial memory processing in the two hemispheres. *Neuropsychologia*, 44(10), 1914-1918.
- Postma, A., Zuidhoek, S., Noordzij, M. L., & Kappers, A. M. L. (2007). Differences between early-blind, late-blind, and blindfolded-sighted people in haptic spatial-configuration learning and resulting memory traces. *Perception*, 36, 1253-1265.
- Postma, A., Zuidhoek, S., Noordzij, M. L., & Kappers, A. M. L. (2008). Haptic orientation perception benefits from visual experience: Evidence from early-blind, late-blind, and sighted people. *Perception & Psychophysics*, 70(7), 1197-1206.
- Ptito, M., & Kupers, R. (2005). Cross-modal plasticity in early blindness. *Journal of Integrative Neuroscience*, 4(4), 479-488.
- Polyshyn, Z. W. (1981). The imagery debate: Analogue media versus tacit knowledge. *Psychological Review*, 88(1), 16-45.
- Raz, N., Amedi, A., & Zohary, E. (2005). V1 activation in congenitally blind humans is associated with episodic retrieval. *Cerebral Cortex*, 15(9), 1459-1468.
- Regier, T., & Carlson, L. A. (2001). Grounding spatial language in perception: an empirical and computational investigation. *Journal of Experimental Psychology: General*, 130(2), 273-298.
- Reichle, E. D., Carpenter, P. A., & Just, M. A. (2000). The neural bases of strategy and skill in sentence-picture verification. *Cognitive Psychology*, 40(4), 261-295.
- Reisberg, D. (1997). *Cognition: exploring the science of the mind*. New York: W.W. Norton & Company.
- Ricciardi, E., Bonino, D., Gentili, C., Sani, L., Pietrini, P., & Vecchi, T. (2006). Neural correlates of spatial working memory in humans: A functional magnetic resonance imaging study comparing visual and tactile processes. *Neuroscience*, 139(1), 339-349.
- Ricciardi, E., Bonino, D., Sani, L., & Pietrini, P. (2008). Functional exploration studies of supramodal organization in the human extrastriate cortex. In A. Bicchi, M. Buss, M. O. Ernst & A. Peer (Eds.), *The Sense of Touch and its Rendering* (Vol. STAR 45, pp. 7-24). Heidelberg: Springer Berlin.
- Ricciardi, E., Bonino, D., Sani, L., Vecchi, T., Guazzelli, M., Haxby, J. V., et al. (2009). Do we really need vision? How blind people “see” the actions of others. *Journal of Neuroscience*, 29(31), 9719-9724.
- Ricciardi, E., Vanello, N., Sani, L., Gentili, C., Scilingo, E. P., Landini, L., et al. (2007). The effect of visual experience on the development of functional architecture in hMT+. *Cerebral Cortex*, 17(12), 2933-2939.
- Rieser, J. J., Guth, D. A., & Hill, E. W. (1986). Sensitivity to perspective structure while walking without vision. *Perception*, 15, 173-188.

- Röder, B., Kusmirek, A., Spence, C., & Schicke, T. (2007). Developmental vision determines the reference frame for the multisensory control of action. *Proceedings of the National Academy of Sciences*, *104*(11), 4753-4758.
- Röder, B., Stock, O., Bien, S., Neville, H., & Rösler, F. (2002). Speech processing activates visual cortex in congenitally blind humans. *European Journal of Neuroscience*, *16*(5), 930-936.
- Rosch, E., Mervis, C. B., Gray, W. D., Johnson, D. M., & Boyes-Braem, P. (1976). Basic objects in natural categories. *Cognitive Psychology*, *8*(3), 382-439.
- Rosenbaum, R. S., Ziegler, M., Winocur, G., Grady, C. L., & Moscovitch, M. (2004). "I have often walked down this street before": fMRI Studies on the hippocampus and other structures during mental navigation of an old environment. *Hippocampus*, *14*(7), 826-835.
- Sadato, N., Okada, T., Honda, M., & Yonekura, Y. (2002). Critical period for cross-modal plasticity in blind humans: a functional MRI study. *NeuroImage*, *16*(2), 389-400.
- Sadato, N., Pascual-Leone, A., Grafmani, J., Ibanez, V., Deiber, M. P., Dold, G., et al. (1996). Activation of the primary visual cortex by Braille reading in blind subjects. *Nature*, *380*(6574), 526-528.
- Saito, D. N., Okada, T., Morita, Y., Yonekura, Y., & Sadato, N. (2003). Tactile-visual cross-modal shape matching: a functional MRI study. *Cognitive Brain Research*, *17*(1), 14-25.
- Santiago, J., Lupiáñez, J., Pérez, E., & Funes, M. J. (2007). Time (also) flies from left to right. *Psychonomic Bulletin & Review*, *14*(3), 512-516.
- Sathian, K. (2005). Visual cortical activity during tactile perception in the sighted and the visually deprived. *Developmental Psychobiology*, *46*(3), 279-286.
- Schinazi, V. R., & Epstein, R. A. (2010). Neural correlates of real-world route learning. *NeuroImage*, *53*(2), 725-735.
- Schubert, T. W. (2005). Your Highness: Vertical positions as perceptual symbols of power. *Journal of Personality and Social Psychology*, *89*(1), 1-21.
- Schwarzlose, R. F., Baker, C. I., & Kanwisher, N. (2005). Separate face and body selectivity on the fusiform gyrus. *Journal of Neuroscience*, *25*(47), 11055-11059.
- Schwarzlose, R. F., Swisher, J. D., Dang, S., & Kanwisher, N. (2008). The distribution of category and location information across object-selective regions in human visual cortex. *Proceedings of the National Academy of Sciences*, *105*(11), 4447-4452.
- Schwoebel, J., & Coslett, H. B. (2005). Evidence for multiple, distinct representations of the human body. *Journal of Cognitive Neuroscience*, *17*(4), 543-553.
- Semenza, C., & Goodglass, H. (1985). Localization of body parts in brain injured subjects. *Neuropsychologia*, *23*(2), 161-175.
- Šetić, M., & Domijan, D. (2007). The influence of vertical spatial orientation on property verification. *Language and Cognitive Processes*, *22*(2), 297-312.
- Shelton, A. L., & Gabrieli, J. D. E. (2002). Neural correlates of encoding space from route and survey perspectives. *Journal of Neuroscience*, *22*(7), 2711-2717.
- Shelton, A. L., & McNamara, T. P. (2004). Orientation and perspective dependence in route and survey learning. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *30*(1), 158-170.
- Shimony, J. S., Burton, H., Epstein, A. A., McLaren, D. G., Sun, S. W., & Snyder, A. Z. (2006). Diffusion tensor imaging reveals white matter reorganization in early blind humans. *Cerebral Cortex*, *16*(11), 1653-1661.
- Smeets, M. A. M., Klugkist, I. G., Van Rooden, S., Anema, H. A., & Postma, A. (2009). Mental body distance comparison: A tool for assessing clinical disturbances in visual body image. *Acta Psychologica*, *132*(2), 157-165.
- Stanfield, R. A., & Zwaan, R. A. (2001). The effect of implied orientation derived from verbal

- context on picture recognition. *Psychological Science*, 12(2), 153-156.
- Stevens, A., & Coupe, P. (1978). Distortions in judged spatial relations. *Cognitive Psychology*, 10(4), 422-437.
- Stevenson, R. J., & Case, T. I. (2005). Olfactory imagery: A review. *Psychonomic Bulletin & Review*, 12, 244-264.
- Stilla, R., Hanna, R., Hu, X., Mariola, E., Deshpande, G., & Sathian, K. (2008). Neural processing underlying tactile microspatial discrimination in the blind: A functional magnetic resonance imaging study. *Journal of Vision*, 8(10), 1-19.
- Struiksma, M. E., Noordzij, M. L., & Postma, A. (2009). What is the link between language and spatial images? Behavioral and neural findings in the blind and sighted. *Acta Psychologica*, 132(2), 145-156.
- Taylor, H. A., Faust, R. R., Sitnikova, T., Holcomb, P. J., & Naylor, S. J. (2001). Is the donut in front of the car? An electrophysiological study examining spatial reference frame processing. *Canadian Journal of Experimental Psychology*, 55, 177-186.
- Taylor, H. A., Naylor, S. J., Faust, R. R., & Holcomb, P. J. (1999). "Could you hand me those keys on the right?" Disentangling spatial reference frames using different methodologies. *Spatial Cognition and Computation*, 1(4), 381-397.
- Taylor, H. A., & Tversky, B. (1992). Spatial mental models derived from survey and route descriptions. *Journal of Memory and Language*, 31(2), 261-292.
- Thinus-Blanc, C., & Gaunet, F. (1997). Representation of space in blind persons: Vision as a spatial sense? *Psychological Bulletin*, 121(1), 20-42.
- Tranel, D., & Kemmerer, D. (2004). Neuroanatomical correlates of locative prepositions. *Cognitive Neuropsychology*, 21(7), 719-749.
- Trojano, L., Conson, M., Maffei, R., & Grossi, D. (2006). Categorical and coordinate spatial processing in the imagery domain investigated by rTMS. *Neuropsychologia*, 44(9), 1569-1574.
- Trojano, L., Grossi, D., Linden, D. E. J., Formisano, E., Goebel, R., Cirillo, S., et al. (2002). Coordinate and categorical judgements in spatial imagery. An fMRI study. *Neuropsychologia*, 40(10), 1666-1674.
- Tse, C.-S., Kurby, C. A., & Du, F. (2010). Perceptual simulations and linguistic representations have differential effects on speeded relatedness judgements and recognition memory. *The Quarterly Journal of Experimental Psychology*, 63(5), 928-941.
- Tversky, B. (1991). Spatial mental models. In G. H. Bower (Ed.), *The psychology of learning and motivation: Advances in research and theory* (Vol. 27, pp. 109-145). San Diego: Academic Press.
- Uhl, F., Franzen, P., Lindinger, G., Lang, W., & Deecke, L. (1991). On the functionality of the visually deprived occipital cortex in early blind persons. *Neuroscience Letters*, 124(2), 256-259.
- Uhl, F., Kretschmer, T., Lindinger, G., Goldenberg, G., Lang, W., Oder, W., et al. (1994). Tactile mental imagery in sighted persons and in patients suffering from peripheral blindness early in life. *Electroencephalography and clinical Neurophysiology*, 91(4), 249-255.
- Van der Lubbe, R. H. J., Van Mierlo, C. M., & Postma, A. (2010). The involvement of occipital cortex in the early blind in auditory and tactile duration discrimination tasks. *Journal of Cognitive Neuroscience*, 22(7), 1541-1556.
- Vanlierde, A., De Volder, A. G., Wanet-Defalque, M.-C., & Veraart, C. (2003). Occipito-parietal cortex activation during visuo-spatial imagery in early blind humans. *NeuroImage*, 19(3), 698-709.
- Vanlierde, A., & Wanet-Defalque, M.-C. (2004). Abilities and strategies of blind and sighted subjects in visuo-spatial imagery. *Acta Psychologica*, 116, 205-222.
- Vecchi, T. (2001). Visuo-spatial processing in congenitally blind people: is there a gender-related

- preference? *Personality and Individual Differences*, 30, 1361-1370.
- Vecchi, T., Tinti, C., & Cornoldi, C. (2004). Spatial memory and integration processes in congenital blindness. *NeuroReport*, 15(18), 2787-2790.
- Von Senden, M. (1932). *Raum und Gestalt: Auffassung bei operierten Blindgeborenen vor und nach der Operation*. Leipzig: Barth.
- Wallentin, M., Roepstorff, A., & Burgess, N. (2008a). Frontal eye fields involved in shifting frame of reference within working memory for scenes. *Neuropsychologia*, 46(2), 399-408.
- Wallentin, M., Weed, E., Østergaard, L., Mouridsen, K., & Roepstorff, A. (2008b). Accessing the mental space - Spatial working memory processes for language and vision overlap in precuneus. *Human Brain Mapping*, 29(5), 524-532.
- Wang, R. F., & Spelke, E. S. (2002). Human spatial representation: insights from animals. *Trends in Cognitive Sciences*, 6(9), 376-382.
- Wassenburg, S. I., & Zwaan, R. A. (2010). Readers routinely represent implied object rotation: The role of visual experience. *The Quarterly Journal of Experimental Psychology*, 63(9), 1665-1670.
- Willems, R. M., & Hagoort, P. (2007). Neural evidence for the interplay between language, gesture, and action: A review. *Brain and Language*, 101(3), 278-289.
- Wolbers, T., & Buchel, C. (2005). Dissociable retrosplenial and hippocampal contributions to successful formation of survey representations. *Journal of Neuroscience*, 25(13), 3333-3340.
- Worsley, K. J., Marrett, S., Neelin, P., Vandal, A. C., Friston, K. J., & Evans, A. C. (1996). A unified statistical approach for determining significant signals in images of cerebral activation. *Human Brain Mapping*, 4(1), 58-73.
- Wu, L.-L., & Barsalou, L. W. (2009). Perceptual simulation in conceptual combination: Evidence from property generation. *Acta Psychologica*, 132(2), 173-189.
- Yoo, S. S., Freeman, D. K., McCarthy III, J. J., & Jolesz, F. A. (2003). Neural substrates of tactile imagery: a functional MRI study. *NeuroReport*, 14(4), 581-585.
- Yoo, S. S., Lee, C. U., & Choi, B. G. (2001). Human brain mapping of auditory imagery: event-related functional MRI study. *NeuroReport*, 12(14), 3045-3049.
- Zachle, T., Jordan, K., Wüstenberg, T., Baudewig, J., Dechent, P., & Mast, F. W. (2007). The neural basis of the egocentric and allocentric spatial frame of reference. *Brain Research*, 1137, 92-103.
- Zuidhoek, S., Kappers, A. M. L., Van der Lubbe, R. H. J., & Postma, A. (2003). Delay improves performance on a haptic spatial matching task. *Experimental Brain Research*, 149(3), 320-330.
- Zwaan, R. A. (1999). Embodied cognition, perceptual symbols, and situation models. *Discourse Processes*, 28(1), 81 - 88.
- Zwaan, R. A., & Madden, C. J. (2005). Embodied sentence comprehension. In D. Pecher & R. A. Zwaan (Eds.), *Grounding Cognition: The Role of Perception and Action in Memory, Language, and Thought* (pp. 224-245). New York: Cambridge University Press.
- Zwaan, R. A., & Radvansky, G. A. (1998). Situation models in language comprehension and memory. *Psychological Bulletin*, 123(2), 162-185.
- Zwaan, R. A., Stanfield, R. A., & Yaxley, R. H. (2002). Language comprehenders mentally represent the shapes of objects. *Psychological Science*, 13(2), 168-171.
- Zwaan, R. A., & Yaxley, R. H. (2003). Spatial iconicity affects semantic relatedness judgments. *Psychonomic Bulletin & Review*, 10(4), 954-958.

Publications and conference abstracts

Publications

Struiksma, M. E., Noordzij, M. L., & Postma, A. (2009). What is the link between language and spatial images? Behavioral and neural findings in the blind and sighted. *Acta Psychologica*, 132(2), 145-156.

Struiksma, M. E., Noordzij, M. L., & postma, A. (in press). Embodied representation of the body contains veridical distance information. *Quarterly Journal of Experimental Psychology*.

Manuscript under revision

Struiksma, M. E., Noordzij, M. L., & Postma, A. (under revision). Reference frame acceptability in haptics differs for the blind and sighted in the horizontal but not in the vertical plane. *Perception*.

Manuscripts submitted/in preparation

Struiksma, M. E., Noordzij, M. L., Neggers, S. F. W., & Postma, A. (submitted). Spatial language processing in the blind: Evidence for supramodal representations and cortical reorganization.

Struiksma, M. E., Noordzij, M. L., & Postma, A. (in preparation). The level of preparation and control in reference frame processing.

Struiksma, M. E., Noordzij, M. L., Zwaan, R. A., & Postma, A. (in preparation). Spatial iconicity in blind and sighted individuals.

Struiksma, M. E., Talsma, L. J., Noordzij, M. L., Neggers, S. F. W., & Postma, A. (in preparation). Neural correlates of learning route and survey knowledge from complex spatial descriptions.

Struiksma, M. E., Talsma, L. J., Noordzij, M. L., Neggers, S. F. W., & Postma, A. (in preparation). Neural correlates of retrieving route and survey knowledge from a spatial mental representation built up from different perspectives.

Peer reviewed abstracts – oral presentations

Struiksma, M. E., Noordzij, M. L., Neggers, S. F. W., & Postma, A. (2007). Understanding space from verbal descriptions: An fMRI study on spatial sentence comprehension in blind and sighted individuals. *6th Endo-Neuro-Psycho Meeting*, Doorwerth, the Netherlands.

- Struiksma, M. E., Noordzij, M. L., Neggers, S. F. W., & Postma, A. (2007). Processing spatial sentences: an fMRI study with blind and sighted individuals. *11th European Workshop on Imagery and Cognition*, Utrechts, the Netherlands.
- Struiksma, M. E., Noordzij, M. L., Neggers, S. F. W., & Postma, A. (2007). Neural correlates of spatial sentence processing in blind and sighted. *11th NVP Winter Conference*, Egmond aan Zee, the Netherlands.
- Struiksma, M. E., Noordzij, M. L., Neggers, S. F. W., & Postma, A. (2008). How do we process text with spatial information? *Conference on Research and Training in Spatial Intelligence*, Evanston, USA.
- Struiksma, M. E., Noordzij, M. L., & Postma, A. (2008). Timing properties of reference frame activation in spatial language. *Helmholtz PhD-day*, Utrecht, the Netherlands.
- Struiksma, M. E., Noordzij, M. L., & Postma, A. (2009). Reference frame activation in spatial language. *ESF-CNCC Workshop "Motor Representations and Language of Space"*, Lille, France.
- Struiksma, M. E., Noordzij, M. L., & Postma, A. (2009). Temporal dynamics of activating and selecting reference frames in spatial language. *16th ESCoP Conference*, Crakow, Poland.

Peer reviewed abstracts – poster presentations

- Struiksma, M. E., Noordzij, M. L., & Postma, A. (2008). Timing properties of reference frame selection. *3rd International Conference on Spatial Cognition*, Freiburg, Germany.
- Struiksma, M. E., Noordzij, M. L., & Postma, A. (2009). Temporal dynamics of reference frame selection in spatial language. *Workshop on Mental Imagery, Spatial Cognition, and Language*, Padua, Italy.
- Struiksma, M. E., Noordzij, M. L., & Postma, A. (2009). Activation and selection of reference frames in spatial language: temporal dynamics. *4th International Conference on Spatial Cognition*, Rome, Italy.
- Struiksma, M. E., Noordzij, M. L., & Postma, A. (2009). The processing of verbal information on body parts depends on a veridical body image. *12th NVP Winter Conference*, Egmond aan Zee, the Netherlands.

List of co-authors and their affiliations

Wendy M. Bosker, Drs.

Department of Neuropsychology and Psychopharmacology
Maastricht University
Maastricht, The Netherlands

Sebastiaan F. W. Neggers, Dr.

Department of Psychiatry, Division of Brain Research
Utrecht University Medical Center
Utrecht, The Netherlands.

Matthijs L. Noordzij, Dr.

Department of Cognitive Psychology and Ergonomics
University of Twente
Enschede, The Netherlands

Albert Postma, Prof. dr.

Experimental Psychology, Helmholtz Institute
Utrecht University
Department of Neurology
Utrecht University Medical Center
Utrecht, The Netherlands

Lotte J. Talsma, BSc.

Experimental Psychology, Helmholtz Institute
Utrecht University
Utrecht, The Netherlands

Rolf A. Zwaan, Prof. dr.

Department of Psychology
Erasmus University Rotterdam
Rotterdam, The Netherlands

Curriculum Vitae

Marijn Struiksma werd geboren op 23 oktober 1980 te Kingston (Jamaica). In 1999 behaalde ze het tweetalig VWO diploma aan de O.S.G. Wolfert van Borselen in Rotterdam. In datzelfde jaar ging zij naar het University College Utrecht. In 2002 haalde zij haar Science Bachelor met een minor in linguïstiek. Hierna werkte ze een aantal maanden om vervolgens een lange reis in Azië te kunnen maken. In 2004 begon ze met de onderzoeksmaster Neuroscience & Cognition aan de Universiteit Utrecht. Tijdens deze masteropleiding liep ze stage bij de afdeling Fysica van de Mens onder leiding van prof. dr. Raymond van Ee. Voor haar tweede stage deed ze onderzoek bij de afdeling Psychologische Functieleer onder leiding van prof. dr. Albert Postma. Na het afronden van haar tweede stage werkte zij verder aan het project onder leiding van prof. dr. Albert Postma. Het project werd onderdeel van haar promotieonderzoek. Vanaf februari 2011 werkt Marijn als post-doc onderzoeker bij het UiL-OTS in Utrecht binnen het Vici-project “De logica van woorden” onder leiding van dr. Yoad Winter.

Marijn Struiksma was born on 23 October 1980 in Kingston (Jamaica). In 1999 she completed her bilingual education at the O.S.G. Wolfert van Borselen in Rotterdam. That same year she started at Utrecht University College. In 2002 she received her Science Bachelor with a minor in linguistics. Afterwards she worked for several months before she travelled in Asia. In 2004 she started the research master Neuroscience & Cognition at Utrecht University. During her masters she performed an internship at the Physics of Man department supervised by prof. dr. Raymond van Ee. She performed her second internship at the Experimental Psychology department under supervision of prof. dr. Albert Postma. After she finished her second internship she continued to work on the project supervised by prof. dr. Albert Postma. The project became a part of her PhD thesis. From Februari 2011 Marijn works as a post-doc researcher at the UiL-OTS in Utrecht on the Vici-project “The logic of words” supervised by dr. Yoad Winter.

Dankwoord / Acknowledgments

Dit proefschrift is het resultaat van vier jaar werk, maar eigenlijk begon het vijf jaar geleden al met een idee voor een tweede stage tijdens mijn masteropleiding. Op een borrel sprak ik met prof. dr. Albert Postma en vertelde hem dat ik graag met fMRI wilde werken. Ik viel met mijn neus in de boter, want hij was juist samen met Wendy Bosker een fMRI project aan het opzetten. Albert betrok dr. Matthijs Noordzij en dr. Bas Neggers ook bij het 'blindenproject' en zo gingen we van start. Ik had het geluk dat het project zodanig ambitieus was, dat het niet mogelijk bleek om het af te ronden binnen mijn stage. Dit was een geluk, omdat ik daardoor de gelegenheid kreeg aio te worden.

Albert, dankjewel voor alle mogelijkheden die je me hebt gegeven, niet alleen tijdens mijn stage, maar gedurende mijn hele aio-traject. Het was fijn om met je te werken en het was bijzonder om samen de wereld van de blinden te ontdekken. Wie had gedacht dat je op een beurs voor blinden en slechtzienden ook onderzoek kon doen. Vanaf het begin heb ik me onderdeel van het 'space-team' gevoeld. Naast het blindenproject liep er ook nog het Wayfinding project, een groot Europees project dat zich bezighield met onderzoek naar navigatie. Hoewel ik daar officieel niet bij betrokken was, werd ik vanaf het begin van mijn aio-carrière in de groep opgenomen en ging ik mee naar workshops en conferenties. Bedankt voor de steun die je me altijd hebt gegeven en bedankt dat je me hebt laten kennismaken met jouw netwerk. Ik hoop dat onze wegen elkaar nog menigmaal zullen kruisen.

Matthijs, ook jij was vanaf het begin al bij ons project betrokken. Dankjewel voor je helderheid en je adviezen. Het was altijd prettig om bij je langs te gaan voor intensief en telkens zeer nuttig overleg. Zelfs toen je van Nijmegen naar Twente ging, bleef het de treinreis meer dan waard. Samen met Bas hebben jullie me wegwijs gemaakt in de wondere wereld van de fMRI. Bas, dankjewel voor alle specialistische begeleiding. Dankzij jou ben ik nu bevriend met Matlab. De uren dat we op vrijdagavond bij de scanner zaten, waren niet alleen leerzaam, maar ook erg gezellig.

Het scannen was een groot deel van mijn promotieonderzoek. Wendy, dankjewel dat je het eerste project met me wilde delen. We hebben een mooie tijd gehad samen en er een vriendschap aan overgehouden.

Vele vrijdagavonden heb ik in het ziekenhuis bij de scanner doorgebracht. Ik wil alle mensen bedanken die mij daarbij hebben geholpen. In het bijzonder Erika, Peter en Estrella als flow-managers, en Lotte als stagiaire. Lotte, je had een pittig project, maar je deed het super. Ik ben erg blij dat je na je stage nog een paar maanden als onderzoeksassistent hebt willen werken. Dankzij jou heeft dit proefschrift twee hoofdstukken over complexe routebeschrijvingen.

Ook wil ik alle studenten bedanken die bij mij onderzoek hebben gedaan: Angela, Brenda, Caroline, Sterre, Heidi, Janice, Jasper, Lotte, Janne, Charlotte, Miriam, Pieter, Rosalie, Jorrit, Jeanette, Joyce, Sascha en Sanja. Van het begeleiden van jullie bachelor- en masteronderzoeken heb ik veel geleerd.

Dan bedank ik alle (ex-)collega's van de afdeling Psychologische Functieleer voor de fijne tijd die ik in het Van Unnik heb gehad, in het bijzonder Jelmer, mijn kamergenoot. Het was fijn om met

jou een kamer te delen. Je hebt het knap vier jaar lang met mij uit weten te houden. Het was vooral mooi, dat we zo goed ons verhaal aan elkaar kwijt konden. Dankzij jou kan ik nu meepraten over oogbewegingen, je huis verbouwen en allerlei soorten afstandsbedieningen. En niet te vergeten het al eerder genoemde ‘space-team’: Maartje, Jessie, Ineke en Mijke. De afgelopen jaren als ‘space-ladies’ waren super. We hebben veel aan elkaar gehad op het gebied van de wetenschap, maar ook op het gebied van sociale contacten. We hebben veel plezier gehad op congressen over de hele wereld en het weekendje Friesland was een mooie afsluiter. Katinka, dankjewel voor de gezelligheid tijdens de lunch, maar ook voor de samenwerking voor de redactie van de Helmholtz Herald.

De workshop in Lille in 2009 was zeer vruchtbaar. Ik wil prof. dr. Rolf Zwaan bedanken voor de fijne samenwerking. Ik hoop dat ons project nog een vervolg krijgt.

I would also like to thank prof. dr. Marianne Gullberg and prof. dr. Susan Goldin-Meadow for the creative collaboration. It was fun designing the gesture experiment for the blind and I am looking forward to discussing the results.

I am also grateful to prof. dr. Lawrence Barsalou. Thank you for the very nice talks we had when you were in Holland and thank you for your input on the project with the blind.

Zonder alle proefpersonen was dit proefschrift nooit tot stand gekomen. Ik wil iedereen die aan mijn experimenten heeft deelgenomen bedanken voor hun inzet. In het bijzonder alle blinde proefpersonen, die de moeite hebben genomen om naar Utrecht te komen en zelfs om in een scanner te gaan liggen. Het was heel bijzonder om jullie te ontmoeten en ik heb veel van jullie geleerd, niet alleen op wetenschappelijk gebied. Daarbij wil ik ook graag de organisatie van de ZieZo beurs bedanken. Dankzij jullie kon ik vier jaar lang op de beurs staan en heb ik niet alleen nieuwe proefpersonen ontmoet, maar ook nieuwe onderzoeken kunnen doen.

Ten slotte bedank ik mijn familie en vrienden. Het schrijven van een proefschrift doe je uiteindelijk alleen, maar zonder de steun van de mensen om je heen lukt het niet. Dankjewel voor de ontspanning en de gezelligheid. Mijn ouders, Chris en Esther, bedankt dat jullie me altijd hebben gestimuleerd om te leren. Mijn zussen, Lotte en Anne, dankjewel dat jullie 6 april achter me staan, we zijn een mooi team. Astrid, de vrijheid die jij me gaf om met Vasco te rijden, gaf me de gelegenheid om me te ontspannen. Even het hoofd leegmaken, er oplet met het paard en buiten zijn, danjewel daarvoor. En als laatste, Joor, zonder jou had ik het allemaal niet kunnen behappen. Dankjewel voor je goede zorgen, de peptalks en de mooie vriendschap.