The Influence of Language on Spatial Memory and Visual Attention

By

Harmen B. Gudde

A thesis submitted in partial fulfilment of the requirements of the University of East Anglia for the degree of Doctor of Philosophy.

Research undertaken in the School of Psychology, University of East Anglia.

May 2017

The Influence of Language on Spatial Memory and Visual Attention

Abstract

This thesis examines the relationship between language and non-linguistic processes. The experimental work presented, focusses on the influence of language on two non-linguistic processes: spatial memory and visual attention.

In the first series of experiments, the influence of spatial demonstratives (*this/that)* and possessives (*my/your)* on memory for object location was examined in four experiments, using an adapted version of the memory game procedure (Coventry et al., 2008, 2014). The experiments were designed to test between different models regarding how language affects memory: the Expectation model, the Congruence model, and the Attention-allocation model. Over a series of experiments, our data supports the Expectation model, which suggests, consistent with models of predictive coding (cf., Lupyan & Clark, 2015), that memory for object location is a concatenation of the actual location and the expected location. The expectation of a location can be elicited by language use (e.g., demonstrative or possessive pronouns).

The second series of experiments examined demonstratives and memory in English and Japanese. We chose Japanese, because it purportedly employs a three-demonstrative system, compared to a binary system as in English (*this, that)*. Three-way systems can be used to explicitly encode parameters that are not encoded in English, for example the position of a conspecific. In four experiments, we wanted to test whether a system as different as the Japanese demonstrative system is from English, has a similar influence on non-linguistic cognition. To this aim, we had to first experimentally establish which parameters are encoded in the Japanese demonstrative system. Second, we tested how this three-term demonstrative system acted in light of the Expectation model. The idea that Japanese demonstratives encode the position of a conspecific, which we confirmed in this study, poses an interesting problem for the Expectation model. The Expectation model works via the idea of an expected location; but the expected location calculated from a speaker gives a contradicting expectation value to the expected location from a hearer. Our memory data did not completely support any of the current models. However, interestingly, the position effect found in Japanese was also apparent in English. This

might suggest that demonstrative pronoun systems, despite the fact that they seem different, could be based on universal mechanisms. However, the effects we found were stronger in Japanese, suggesting the weight of a parameter (such as position) might be influenced by whether or not a language explicitly codes the parameter.

In the last experiment, we considered the influence of language on visual attention. Specifically, we examined if language expressing different spatial frames of reference affect how people look at visual scenes. The results showed different eye-movement patterns for different frames of reference (i.e., intrinsic vs. relative). These eye-movement *signatures* were consistent with participants' verbal descriptions and persisted throughout the trials. We show for the first time that different reference frames, expressed in language, elicit distinguishable eye-movement patterns.

The work presented in this thesis shows effects of language on memory for object location and visual attention. Effects of language on memory for object location were consistent with models of predictive coding. Furthermore, despite the fact that English and Japanese employ different demonstrative systems, results for both languages were remarkably similar. These results could indicate universal parameters underlying demonstrative systems, but perhaps parameters differentially weighted, as a function of whether or not they are explicitly encoded in a language. Finally, we showed that spatial language (prepositions) guide visual attention. To our knowledge this is the first time frames of reference are associated with identifyable eye-movement patterns. The results are discussed and situated in current literature, with theoretical implications and directions for future research highlighted.

Contents

List of Tables

List of Figures

Acknowledgements

First, I would like to thank my examiners for taking the time out of their schedules to read and assess this body of work.

Second, I would like to thank the following people for making this work possible (although not necessarily in the following order):

My friends at the school of Psychology among staff and PhD students, who provided entertainment and ample opportunities to finish any given week (or day) off with a pint. Next to those close by I would like to thank my friends and colleagues in the LanPercept network. The trips across Europe surely made these three years far more lively and exciting. This project would not have been possible without the generous support of the European Committee, under the flag of the Marie Skłodowska-Curie Actions (no. 316748!).

Mao and Kevin, whose bitchie-events made for a lot of happy memories.

My family, and specifically Jinnie Ooi.

Jackie Orford, and others from the brilliant support staff: Amanda Dorr, Jeff Smith, Scott Steward, Katy Quigley, and Roanne Ephithite, for their help with all sorts of random requests (e.g., arranging tables, curtains, and a washing machine), in the name of science.

Special thanks to Shunei Yasuda, Heloise Ungless, and Jennifer Bowler for their assistance in data collection.

Finally I thank my supervisors, instrumental in the process towards my PhD. Kenny Coventry for his ideas and advice, pushing me further on all aspects of this research, even at times when he had far more serious issues to deal with, and Paul Engelhardt for his continuous support and insights, and the occasional slow-smoked barbeque.

I would like to express my sincere thanks for their help during this project, it made my time and work considerably better.

Author's declaration

I declare that the work contained in this thesis has not been submitted for any other award and that it is all my own work. I also confirm that this work fully acknowledges opinions, ideas and contributions from the work of others.

The research presented in Chapter 2 has been published in *Cognition* and in the Proceedings of the 37th annual conference of the Cognitive Science Society (Gudde, Coventry, & Engelhardt, 2016; Gudde, Coventry, & Engelhardt, 2015). These papers have been added respectively in Appendices A and B. The research presented in Chapter 3 has been accepted for the 39th annual conference of the Cognitive Science Society (Gudde & Coventry, in press). The published abstract has been added in Appendix C.

Parts of this work has been presented at conferences:

Oral presentations

Gudde, H.B., & Coventry, K.T. (September, 2015). *Spatial Demonstratives and Perceptual Space: Describing and remembering object location.* Presented at the International Conference on Spatial Cognition, Sapienza University. Rome, Italy.

Gudde, H.B., Coventry, K.R., & Engelhardt, P.E. (May, 2014). *The Influence of Language on Spatial Memory.* Presented at the Nordforsk network Closing Conference, NTNU. Trondheim, Norway.

Poster Presentations

Gudde, H.B., Coventry, K.R., & Engelhardt, P.E. (July, 2015). *The Influence of Language on Spatial Memory.* Poster presented at the Annual Meeting of the Cognitive Science Society. Pasadena, CA, US.

Gudde, H.B., Coventry, K.R., & Engelhardt, P.E. (February, 2015). *The Influence of Language on Spatial Memory*. Poster presented at the workshop Language and Perception: Theoretical Connections and Foundations. Norwich, UK

Gudde, H.B., Coventry, K.R., & Engelhardt, P.E. (August, 2014). *The Influence of Demonstratives and Possessives on Spatial Memory.* Poster presented at the Embodied and Situated LanguageProcessing Conference (ESLP). Rotterdam, the Netherlands.

Gudde, H.B., Coventry, K.R., Engelhardt, P.E., & Vulchanova, M.D. (November, 2013*). The Role of Language in Visual Prediction and Spatial Memory*. Poster presented at the Cognition andLanguage in Developmental Disorders Workshop. Seville, Spain.

Any ethical clearance for the research presented in this thesis has been approved. Approval has been sought and granted by the School of Psychology Ethics Committee at the University of East Anglia.

Name:

Signature:

Date:

SECTION 1
–
The Influence of Language on Non-Linguistic Cognition

Chapter 1 – On the Relation between Language and Non-Linguistic Cognition

The relationship between language and non-linguistic processes is a fundamental topic in cognitive science. The fact that we spend the vast majority of our waking time using language marks the importance of this potential relationship: if language has an effect on non-linguistic processes, this influence could be very persistent. Hitherto, there has been limited consideration regarding the mechanism via which language can affect cognition and perception. In this thesis, we use the domain of space to explore the relationship between different types of language and cognition – specifically spatial memory and visual attention – and tease apart mechanisms driving potential influences.

Space is fundamental in cognition and it can be used as a structuring tool for domains such as time, sound, number cognition, emotion, and metaphors in language (cf. Casasanto & Boroditsky, 2008; Dehaene, Bossini, & Giraux, 1993; Lakoff & Johnson, 2008). Therefore, space is widely regarded a fundamental building block of language and cognition (Bowerman, 1996). At the same time, there is a radical diversity in how space is encoded in languages around the world (Evans & Levinson, 2009). In testing the relationship between language and cognition, it makes sense to test language that is both high frequency and distinctive between languages. In this thesis, we use two types of languages that fit these criteria: demonstrative pronouns and spatial prepositions, arguably the most important types of spatial language.

Demonstrative pronouns occur in every language, they are suggested to have developed early on in any given language, are among the earliest words infants acquire. Demonstratives are among the most frequently used words within a language and are commonly used to create a joint focus of attention (Burenhult, 2003; Clark, 1978; Clark & Sengul, 1978; Deutscher, 2005; Diessel, 2006; Tomasello, 1999). Like demonstratives, spatial prepositions are among the first words children learn (Bowerman, 1996). Spatial communication requires selecting a spatial term from a range of available options, often with the additional choice of an underlying conceptual reference frame that guides the interpretation of spatial directions.

This leads us to our central question: is there any evidence that language affects (non-linguistic) cognition and perception? In this chapter we first review literature regarding the relationship between language and non-linguistic cognition and perception in the context of different theoretical perspectives on language (learning) which assume to varying degrees that language can affect non-linguistic processes. After this we will cover research in different domains such as memory, number cognition, and colour, both within and between languages. We will then review the types of language that we use in the experimental chapters in this thesis, demonstrative pronouns and spatial prepositions, in more detail. The chapter will conclude with a brief précis of the rest of the thesis.

1.1 Networks for Perception

Perception has often been considered a bottom-up system, in which different areas in the brain are activated by domain specific information (cf. Kveraga, Ghuman, & Bar, 2007). However, for language to be able to affect cognition and perception, there needs to be an integration of information processing, in which top down processes can influence cognition and perception. For example, Fodor (1989) suggested modularity of the brain, in which different modalities (e.g., visual and auditory perception) are innately based in different brain areas, defined by their functional role, and operate independently of each other. These individual brain areas would provide input for central computational processes that combine the modal areas. A cognitive system like this would not allow for top-down influences. However, this modularity-perspective is problematic, since the noise and clutter in natural perception (e.g., vision: differences in lighting, shadows, occlusions, reflections, etc.), for instance, would make it near impossible to identify objects in any but the simplest circumstances (Bullier, 2001a, 2001b; Kveraga et al., 2007). For successful perception (and cognition), humans need to be able to bind knowledge of concepts with perceived information. Indeed, research has suggested that input in the visual system is rapidly distributed to a large number of visual areas and feedback networks in the brain (Lamme & Roelfsema, 2000; Lupyan, Thompson-Schill, & Swingley, 2010). These networks span across the brain and integrate different areas associated with the processing of different modalities (cf., Barsalou, 2008; Meteyard, Cuadrado, Bahrami, & Vigliocco, 2012; Meyer & Damasio, 2009).

It has been proposed that in these types of feedback systems, memory consists of multiple modal ‘data points’ stored in their respective brain areas (e.g., Bird & Burgess, 2008; Byrne, Becker, & Burgess, 2009). Remembering is suggested to happen via the activation of at least one of those data points, after which the brain activates a full representation, via a process of pattern completion in a feedback system. For example, if one brews a cup of coffee, there is a co-occurrence of sensory-information, say smell and taste. Each brain area processes its respective input and sends the information in both forward and backward direction between convergence-divergence zones (CDZs), such that the combination of sensory input constructs a full representation of the experience of coffee, including smell and taste (Meyer & Damasio, 2009). The next time one smells coffee, the olfactory area processes the smell and forwards this information to higher-level processing areas, which in turn send signals to lower-level areas of other modalities, so that the entire representation of coffee is activated, triggered by only the smell of coffee.

These feedback systems have two components specifically interesting for this thesis. First, these networks integrating different types of information are instrumental for the potential of language influencing cognition and perception. Second, networks like this can explain that cognition works via predictive coding; the pattern completion involved requires anticipation of prediction of data points that are not (yet) available (Bar, 2009; Bird & Burgess, 2008; Friston, 2009). We will take a closer look at the idea of predictive coding in discussing the results of our experimental chapters.

In this chapter, we will explore how language might affect cognition – spatial memory and visual attention. We next review literature that tested the relationship between language and cognition via cross-linguistic studies. If language influences cognition, it might be that different languages do so in a different way. After this review, we will look at studies that tested differences within a language.

1.2 Cross-linguistic differences in language and cognition/ perception

There is much debate regarding if and how language affects non-linguistic cognition. One long-standing debate regards the influences of one's native language on cognition and perception. Do speakers of different languages have different (non-linguistic) cognition and perception as a function of the language they speak, as proposed by the *linguistic-relativity hypothesis* (also known as the Sapir-Whorf hypothesis) (Sapir, 1929; Whorf, 1956)? It is important to note that there are many different variants of the linguistic relativity theory, ranging from a weak version (suggesting that linguistic structure influences the way we think about things under specific circumstances when language may be used as a tool in a task) to a strong version (the language one speaks directly influences perception and cognition) (cf., Wolff & Holmes, 2011). The strong version of the hypothesis is controversial since it denies the possibility of a universality in (human) cognition. In contrast to linguistic relativity, the theory of universal language assumes that even though languages are superficially different (e.g., different words and grammar), their underlying structure is universal (e.g., Chomsky & Foucault, 2006). In the universalist account, perceptual information is encoded by the perceptual system and affected by later reconstructions, when higher-order cognitive processes (e.g., contextual information, categories) play a role (Fodor, 1989). Any differences in memory caused by contextual information should occur at retrieval in this view. Linguistic relativity, on the other hand, proposes that language affects perception directly, at the level of encoding.

It was originally proposed that speaking a different language means one experiences (e.g., perceives, analyses, and acts in) the world differently (Whorf, 1956), based on the fact that languages differ strongly in terms of how they encode the world (e.g., space, time, substances, and objects) (cf. Kemmerer & Tranel, 2000; Levinson, 2004). For example, if a language contrasts the colours light blue and dark blue, as is the case in languages such as Russian, Korean, and Greek, linguistic relativity suggests that one perceives blue differently compared to when one's language does not encode this contrast (Gilbert, Regier, Kay, & Ivry, 2006; Winawer et al., 2007). To give another example, data from the domain of sound perception suggested that differences in how tone pitches are verbalised (Dutch: high/low; Farsi: thick/ thin) formed different non-linguistic pitch representations consistent with these linguistic metaphors for pitch (Dolscheid, Shayan, Majid, & Casasanto, 2013). After training in the Farsi thin/thick representation, Dutch speakers performed similar to Farsi speakers. We will discuss linguistic differences across domains in depth in sections 1.3 (between languages) and 1.4 (within languages).

Whereas (strong) linguistic relativity suggests that language affects our perception of the world (cf. Gumperz & Levinson, 1991, 1996; Slobin, 1996; Wolff & Holmes, 2011), there is a mediating account suggesting 'pragmatic inference'occurs through language (Frank & Goodman, 2014). In this account, cognition is not (re)shaped; language merely alters a speaker's interpretation of a task. On the borderline of these two accounts one can find the *thinking for speaking* hypothesis (Slobin, 1996). Slobin suggested that while we are constructing utterances in discourse, thoughts are processed in a linguistic form. However, this linguistic form cannot fit perfectly onto the original representation, since each language has a limited number of words and grammatical forms to characterise a situation model (e.g., the representation of an object or event). Thinking for speaking therefore suggests that while engaging in discourse, one needs to pick the specific characteristics of a situation model that helps the conceptualisation of that model while being available in the language that is used (Slobin, 2003). In other words, when investigating how language shapes our thoughts, we need to evaluate how we think while we use language for communication. For example, if one wants to say that their Uncle John took Aunt Mary for dinner at a nice restaurant on 10th street, then the language of expression affects how the event is communicated. In English, the verb specifies whether it is in the past or the future, but in Mian (a language from Papua New Guinea) the verb specifies whether it happened just now, yesterday, or longer ago than that. In Indonesian the verb would not specify whether it had already happened or has yet to happen and in Turkish the description would specify whether one witnessed the event themselves. In Russian, the verb would encode the speaker's gender, in Mandarin one would specify whether the uncle is paternal or maternal and whether he is the first, second, etc., or youngest brother. In some indiginous languages it would be difficult to specify precisely where it was, as they reportedly lack linguistic methods for expressing exact quantity (e.g., Pirahă employs a 'one-two-many' system, Mundurukú has number words up to 5, whereas Warlpiri and Anindilyakwa completely lack counting words) (example based on Boroditsky, 2011; Butterworth & Reeve, 2008; Frank, Everett, Fedorenko, & Gibson, 2008; Gordon, 2004; Pica, Lemer, Izard, & Dehaene, 2004; Slobin, 1996).

When describing Uncle John's event in different languages, a speaker has to think about differences between languages to be able to encode them appropriately. That does not necessarily mean however, that language actually changes their cognition. Despite the fact that, in this example, English speakers do not specify from which side of the family Uncle John is, they usually have this information – they just do not explicitly encode it in language. However, there is evidence that language does shape spatial cognition. For example, it has been argued that children are influenced by spatial language in the development of their spatial cognition (Choi

& Bowerman, 1991; Choi, McDonough, Bowerman, & Mandler, 1999). Choi and colleagues tested how children in English and Korean talked about motion events. In English, motion and path are usually expressed separately, whereas Korean makes this distinction based on whether the motion was caused or spontaneous. They found that English speaking children, from as early as 17-20 months generalised spatial terms, whereas Korean children kept words for spontaneous and caused motion separate. They argue this challenges the idea that children map spatial words onto non-linguistic spatial concepts, and that they are instead influenced by the organisation of their spatial language as they learn it.

In this thesis, we aim to contribute to this discussion. We use spatial language as a way to explore the relationship between language and non-linguistic cognition, specifically spatial memory and visual attention. In the next section, we will discuss how cross-linguistic differences might influence cognition and perception. This review is not intended to be exhaustive, but rather will illustrate the range of domains in which the influence of language on cognition and perception has been considered, often with decidedly mixed results.

1.3 Cross-linguistic differences and non-linguistic cognition and perception

1.3.1 The influence of motion expression on memory for events

One way to test the influence of language is by exploring differences in non-linguistic cognition comparing speakers of different languages that make different distinctions (Gennari, Sloman, Malt, & Fitch, 2002). Gennari and colleagues presented videos of events (e.g., a man carrying a board into a room), while Spanish and English participants were equally divided over three conditions: an encoding condition, in which they were encouraged to verbally encode the events, a free encoding condition in which participants just watched the videos, and a ‘shadow’ condition, in which participants repeated nonsense syllables as a verbal interference task. Results showed that participants did verbally encode the videos differently: Spanish speakers used the same verb for actions that shared path, and used more path-verbs, while English speakers used the same verb for actions that shared manner and used more manner verbs. There was no evidence that there was a recognition difference between languages. This is important, as this would be predicted by linguistic relativity. Similarity judgments in the linguistic encoding condition did differ between languages, consistent with the pattern of descriptions in each language, but only after linguistic coding.

Another example of differences between languages in describing events is that there are languages in which the verb encodes the *manner* of motion (e.g., English: strolling, running, etc.) and languages where the verb instead encodes the *direction* of motion (e.g., Greek: descend, cross). Papafragou, Massey, and Gleitman (2002) explored the influence of this linguistic difference in two separate tasks, run over two separate days. On the first day, participants engaged in a linguistic task, in which they described events depicted in different pictures (e.g., a boy jumping over a log, See Figure 1.1). On the second day, participants engaged in a non-linguistic task, in which they were presented with pictures that were similar or different from the first task (e.g., a boy jumping over a log or a boy stumbling over a log), and had to judge whether they had seen that picture before. Results showed that linguistic preferences differed significantly between languages; Greek participants verbally coded the direction and English participants the manner of movement. But even though participants showed clear linguistic preferences on the linguistic task, their (non-linguistic) memory performance was identical (Papafragou et al., 2002). Similarly, when Greek and English participants listened to verbal descriptions of events, their eye-movements followed the linguistic pattern of their native language, but in a free-inspection task (not cued by language), attention allocation was similar across languages (Papafragou, Hulbert, & Trueswell, 2008).

Figure 1.1. Example of the manner variation stimuli used by Papafragou et al. (2002): “the boy jumping over a log” vs. “the boy stumbling over a log”.

These findings are consistent with 'thinking for speaking' (Slobin, 1996, 2003) and other accounts suggesting language does influence non-linguistic processing, but only in linguistic tasks (Frank et al., 2008; Gilbert et al., 2006; Hermer-Vazquez, Spelke, & Katsnelson, 1999). However, it has been suggested that language can have an effect despite verbal interference. For example, Athanasopoulos and colleagues (Athanasopoulos et al., 2015), suggested differences between language use in bilinguals, after they found that German-English bilingual participants matched motion-events using constraints of the specific language they were using at that time. Participants speaking in German matched motion-events based on motion completion moreoften than when speaking in English. Moreover, the authors claim to have found a double dissociation: when participants engaged in an English verbal interference task (repeat a string of three two-digit numbers in English), their matching behaviour was congruent with German, but if they did a German verbal interference task (repeat a string of three two-digit numbers in German), their matching behaviour fitted the English language categorization. This would indicate the influence of language on human cognition is strongly context-bound and transient.

1.3.2 Numbers

Not all languages have cardinal number terms. However, in the domain of number processing, evidence shows that a lack of language does not necessarily inhibit number processing in an online task, but does influence memory for numbers. For example, Gordon (2004) conducted a matching task with Pirahă speakers, a language known for its lack of quantity expressions (cf., Everett, 2005). Whereas Gordon (2004) claimed that Pirahă language employed a one-two-many structure, Frank et al. (2005) suggest it lacks any quantity expression, even 'one'. In the matching task by Gordon, participants were asked to match two arrays of objects (ranging from simple linear arrays to more advanced clusters of objects, see Figure 1.2). It was found that the lack of a linguistic counting system limited the ability of Pirahă speakers to establish arrays exceeding two or three items, although they showed evidence of chunking when presented with larger set sizes. Gordon (2004) concluded that humans who are not exposed to a number system through language cannot represent exact quantities for medium-sized arrays.

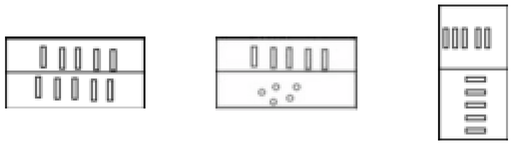


Figure 1.2. Three examples of the matching task by Gordon (2004). He presented participants with an array of batteries (lower half) and participants were asked to match their own array (top half).

In a follow-up study, Frank et al. (2008) retested Pirahă speakers on the matching tasks but found participants were accurate in tasks where participants could see the original array while recreating it – despite not having a word for the concept ‘one’, they had an understanding of it and appeared to understand calculations including the addition or subtraction of the number ‘one’. The authors report that the difference in performance between their study and Gordon (2004) could be due to an improved control of their testing environment. For example, Gordon’s participants were tested with AA batteries on an uneven surface. This could have added distraction and difficulty to the calculation tasks (Frank et al., 2005). Furthermore, Frank et al. tested participants on memory matching tasks, in which the experimenter covered the array after presenting a number of objects in a line and the participant had to reconstruct the array from memory. Results showed decreased performance, suggesting that language might facilitate memory to compare information across time, space, and modality. Whereas visual and auditory short-term memory have a limited capacity, the ability to verbally encode events allows for categorization and therefore increased memory performance.

An alternative explanation, for the differences between the effects of languages in number cognition, consistent with these previous findings, suggests that approximate number representation might be universal, whereas exact number representation (specifically large numbers), might depend on a linguistic counting system (Pica et al., 2004). This might suggest that there might be a two-tier memory system for number processing, one in which language is important and one in which language does not play a part.

Another line of research suggesting a relationship between language and number is the *mental number line* theory (Restle, 1970), consistent with the more elaborate Spatial Numerical Association of Response Codes (SNARC) theory

(Dehaene, 1992; Dehaene et al., 1993). These theories suggest that numbers of lower magnitude are coupled with preferentially leftwards responses, whereas numbers with higher magnitude are related to rightwards responses (see also Casarotti, Michielin, Zorzi, & Umiltà, 2007; Fischer, Castel, Dodd, & Pratt, 2003). This mental number line is associated with linguistic culture, for example writing direction (cf. Göbel, Shaki, & Fischer, 2011). Moreover, the effects diminish or even reverse for participants from writing cultures where writing is not from left to right, like Hebrew and Arabic (Dehaene et al., 1993; Shaki, Fischer, & Petrusic, 2009). The cultural effect has not only been found between participants, but also within participants. Shaki and Fischer (2008) tested bilingual speakers of Cyrillic (left-to-right script) and Hebrew (right-to-left script), and showed that the left-to-right SNARC effect is stronger after bilingual speakers read Cyrillic and is reduced after reading Hebrew.

Neurological studies have further supported the idea of the spatial integration of number processing on a *mental number line*. Studies with left-neglect patients show that participants with neglect systematically misplace the midpoint of a numerical interval (e.g., claiming that 5 is halfway between 2 and 6) (Umiltà, Priftis, & Zorzi, 2009; Zorzi, Priftis, & Umiltà, 2002), whereas healthy participants usually have a slight bias towards the left side of the number line (Longo & Lourenco, 2007).

1.3.3 Colour

Another classic domain in which the discussion on linguistic influences has traditionally been a hot topic is colour. A large and growing body of research is looking at the influence that different languages have on colour perception. Again, in this domain, the discussion is between linguistic relativity (cf. Drivonikou et al., 2007; Gilbert et al., 2006; Roberson, Davies, & Davidoff, 2000) and a universal account of colour perception (cf. Kay & McDaniel, 1978; Kay & Regier, 2003).

For example, there has been a debate on whether the possession of colour terms affects the way colours are categorised. Heider (1972a), tested speakers of Dani, a Papuan language with a two-term colour system based on brightness rather than hue (loosely translated to mean ‘dark’ and ‘light’). In a naming task, in which a coloured chip was presented and participants were asked to name the colour, Heider found that half her participants used words in addition to these two basic terms, the additional words showing consistency with red, yellow, and blue, suggesting that if a language does not allow direct encoding of colour differences, speakers might seek strategies to circumvent these limitations. In another study, Heider (1972b) explored whether there are focal colours, defined as the best example of a certain colour category, which are universally the most codable and easy to remember. Across 4 experiments, speakers of over 25 different languages were tested in a naming and a memory task. In the naming task, participants were shown colours and asked to name them in their own language. Results showed that focal colours were given shorter names and were named quicker than other colours. In the memory task, participants had 5 seconds to memorize a specific colour. After an interval of 30 seconds they were asked to pick the remembered colour out of an array of 160 different colours. Focal colours were remembered more accurately, although, as Heider reports, the memory data might be confounded: in trials in which a participant doesn't remember a specific colour, they are more likely to guess a focal colour. This leads to a memory bias towards focal colours, which could alternatively explain the memory advantage. In a third experiment, Heider taught participants pairs of colours (focal or non-focal) and a separate response word. Participants engaged in multiple training days to learn these pairs. Results showed that the focal-colour pairs were learned faster.

In a replication study, Roberson et al. (2000) tested speakers of Berinmo, a language with five basic colour terms. First they confirmed that the five colour terms were used consistently and with consensus to describe the colour of a variety of objects. Then they tried to replicate the findings by Heider (1972b), but found no support for the superior learning effects of focal coloured pairs. Furthermore, Roberson et al. found that colour terms available in language affect the categorization of colours, suggesting an important role of language instead of categorization based on innate neurophysiology. In experiments in which participants had to judge to which category different colours belonged, clear differences were found between English and Berinmo speakers. When it was tested whether linguistic distinctions help in categorization tasks, English and Berinmo speakers were asked to split two sets of stimuli, consisting of either green and blue (different colour names in English but not in Berinmo) stimuli, or two shades of green (no different colour name in either English or Berinmo). English participants found the green vs. blue division, colours with different colour names, easier than the division of two shades of green. Berinmo speakers performed, in line with the lack of linguistic colour distinction, equally on both tasks (Roberson et al., 2000). This suggested that colour distinctions are easier when languages make a distinction between the two colours.

In a study comparing English speakers with Russian speakers, Winawer et al. (2007) found that Russian speakers had a language advantage in the categorization of tones of blue. In Russian, speakers make an obligatory distinction between light-blue (goluboy) and dark blue (siniy), whereas in English this distinction is not obligatory. First, both English and Russian participants made categorical distinctions between light and dark blue. Results from the two language groups were very similar; they drew nearly the same boundary between the 20 different tones of blue (See Figure 1.3). Then, in an online task, participants saw three squares, one on top, and two below. Participants had to state which of the lower two squares had the same colour as the top square. Results showed that Russian participants had faster response times if they had to discriminate between two colours encoded in different linguistic categories in Russian (light blue or dark blue), an effect absent in English. The Russian advantage disappeared when participants had to perform a language interference task. These effects suggest that language affects basic perceptual colour discrimination tasks, and that this is an online effect, since the effect disappears when participants engage in a verbal interference task. In an EEG study, Thierry and colleagues (Thierry, Athanasopoulos, Wiggett, Dering, & Kuipers, 2009) found that

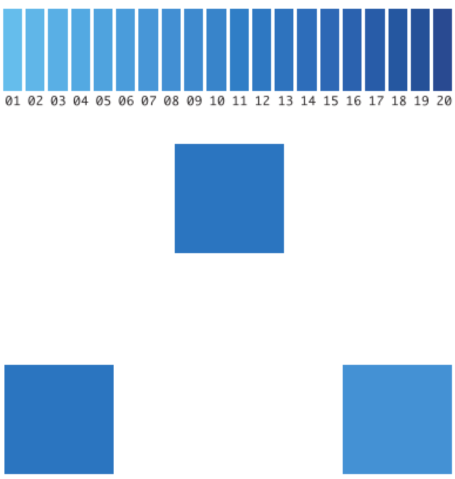


Figure 1.3. The shades of blue (above) used by Winawer et al (2007), and the three squares used in the colour matching task (below)

when Greek speakers (Greek is another language with an obligatory distinction between light and dark blue) were presented with succeeding trials of different shades of blue, there was a larger mismatch negativity if the shade of blue deviated compared to the presentation of different shades of green. English participants did not show a distinction (Thierry et al., 2009). However, this study did not include a verbal interference task, so it cannot be distinguished whether this effect is apparent only while using language, or whether the cognition of speakers of different languages is different, regardless of whether they are using language.

1.3.4 Reference frames as expressed in language

A debate similar to that of colour has been held in the literature on frames of reference. Reference frames are required to assign direction to space when using language. For example, to *the left of the car* could be to the left side from the speaker's perspective (aligned with the left side of the speaker's body), or the left side with reference to the axes of the car. It has been argued that there is a systematic variation between spatial reference frames in language use (Pederson et al., 1998). Pederson et al. ran a cross-cultural and cross-linguistic analysis of the use of spatial reference frames and found, consistent with Levinson (1996), that there are only three distinctive frames of reference, of which they only focussed particularly on the absolute and relative reference frame. In a series of tests, they tested speakers of different languages on their reference frame use, concluding systematic differences between languages and the behavioural responses of their speakers. For example, in the 'animals in a row' task (Brown & Levinson, 1993), a participant has to memorize a sequence of three animals placed on a line, all facing a specific direction (see Figure 1.4). Then the participant turns around 180 degrees before reconstructing the array on a second table. Participants can reconstruct the animal row in two ways: based on an absolute frame of reference (based on north/south) and a relative frame of reference (based on left/right). Pederson et al. found that speakers of Tzeltal, a language that uses the absolute frame to describe object location even in tabletop space, remembered the array in accordance with their language (i.e. reproducing the exact positions of the animals with respect to compass points) while speakers of Dutch and English in contrast reproduced arrays from an egocentric point of view. This specific test of reference frame used has been food for debate over the years, where some have argued that the results show evidence for clear cross-cultural differences in (non-linguistic) reference frame use (Brown & Levinson, 1993; Levinson, Kita, Haun, & Rasch, 2002; Pederson et al., 1998). However, it has been argued that for humans, language may not be the decisive factor in the choice of spatial perspective, but strategy, based on the availability and suitability of landmark cues (Li & Gleitman, 2002). Not every language employs every reference frame, but speakers of languages can nevertheless acquire and use different reference frame systems with ease (Li, Abarbanell, Gleitman, & Papafragou, 2011; Majid, Bowerman, Kita, Haun, & Levinson, 2004). In other words, people use language as a tool to remember locations.

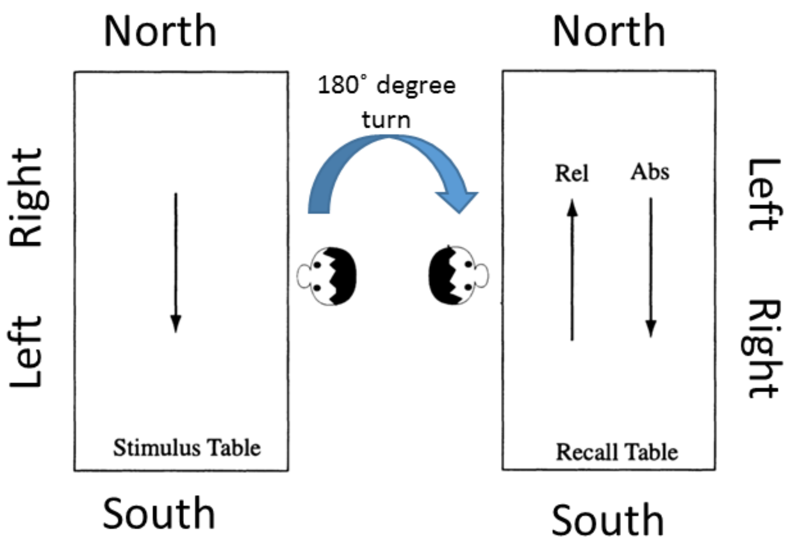


Figure 1.4. The ‘animals in a row’ task. Participants memorize the location of three animals standing in a row, facing a specific direction. After turning 180 degrees participants reconstruct the memorized array. They can do this according to a relative reference frame (based on the animals’ position in respect to their body, or the absolute reference frame (based on the animals’ position to the environment) (Brown & Levinson, 1993).

1.3.5 Summary: cross-linguistic differences

To summarize, despite the large amount of research on linguistic relativity, results provide evidence in favour of and against the general theory. Moreover, the field suffers from replication failures. Recent publications, providing a promising avenue to reconcile some of these differences, argue that *probabilistic inference* could take away the contradiction between linguistic relativity and the universality of cognition (Cibelli, Xu, Austerweil, Griffiths, & Regier, 2016; Regier & Xu, 2017). In their work on colour, Cibelli et al. propose a probabilistic model of color memory, based on two assumptions: there is a universal *colour space*, and languages have specific categories to classify colours in this space. A perceived colour is inferred from the two sources of evidence, as a two-tier system: the actual fine-grained representation of a colour and the language category it is fitted into (See Figure 1.5). Colour memory involves the probabilistic combination of the two sources. The weight of each tier is a function on the amount of uncertainty (for example induced by a delay between presentation and reconstruction in a colour match task) about the actual fine-grained representation. By underscoring the role *uncertainty* plays in cognition, effects showing linguistic relativity are explained without rejecting universals in language or cognition. When there is a high certainty of perception,

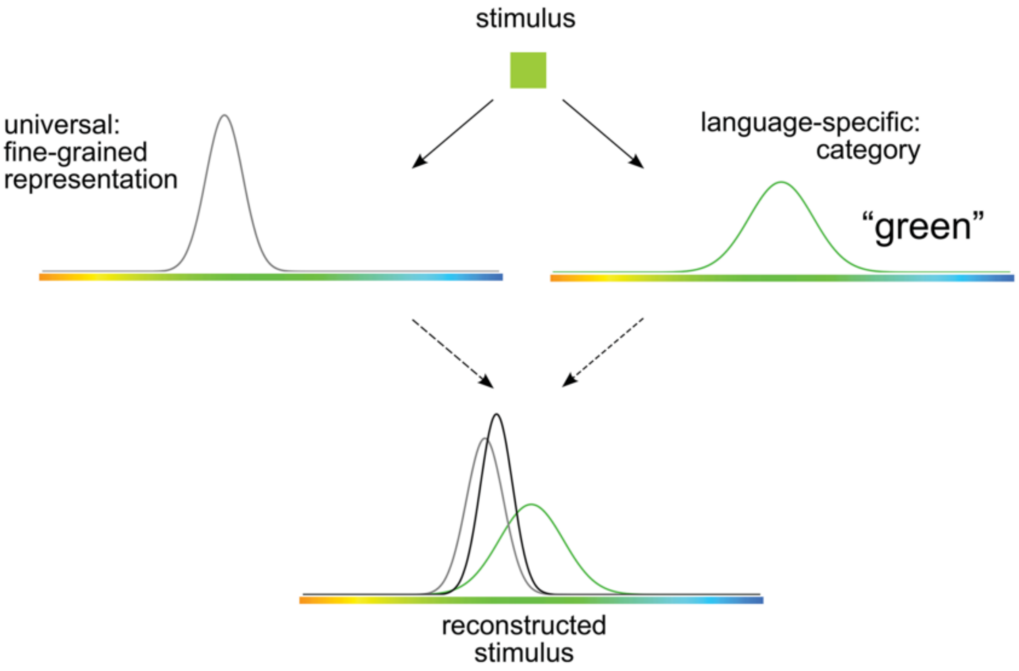


Figure 1.5. Colour memory works via the actual representation in colour space (left) and the language category it is fitted into. Colour memory is the probabilistic combination of these two sources (Cibelli et al., 2016)

there is a low influence of different parameters on the memory for, in Cibelli et al. (2016), colour. However, when the perceptual certainty decreases, the influence of contextual information (i.e., category membership) becomes stronger. The authors suggest that the non-replications in research on the Sapir-Whorf hypothesis could be due to the use of high-certainty stimuli. We will compare our data to this theory in the discussion. Therefore, the fundamental question is: do differences in linguistic encoding change perception or memory for events, and if so, how does this work?

1.4 The influence of language on cognition and perception within a language

In the previous sections, we reviewed literature on cross-linguistic/cultural differences and the influence of different languages on (non-)linguistic tasks. We will now look at how cognition and perception might be influenced within a language. For example, is performance on spatial memory tasks enhanced when spatial language is presented at encoding vs. a no-language condition? We will review literature on attention, memory, verbal overshadowing, and colour perception.

1.4.1 Spatial cognition

There are many studies examining the influence of language on various aspects of spatial cognition. Theoretically, spatial language in communication is thought to help to create a joint focus of attention. For example spatial expressions, such as *these coins* or *the cup is on the table,* serve to direct the attention of a hearer to regions of space (Miller & Johnson-Laird, 1976; Spivey, Tyler, Eberhard, & Tanenhaus, 2001). Studies using eye-trackers showed that, when a verbal auditory description is presented with a visual array, participants tend to attend to the objects presented on the screen following the referring words (e.g., focus on the table when the table is mentioned in the description) (Eberhard, Spivey-Knowlton, Sedivy, & Tanenhaus, 1995). This effect is already found in 18-23-month old children (Choi et al., 1999). Choi et al. found that presenting target words focussed childrens' gaze at different (and language-appropriate) aspects of visual scenes.

Another domain of spatial cognition in which an essential role of spatial language was found is human navigation (Hermer-Vazquez et al., 1999), see Figure 1.6. In their study, an object was hidden in a rectangular room (in corner C) while participants were watching. In one condition, there are no spatial cues (Figure 1.6A, only geometrical information), in the second condition there was one blue wall (Figure 1.6B, geometrical and non-geometrical information). Following the object placement, participants had to reorient themselves in the room and find the object. Participants in the geometrical information condition (1.6A), infants and adults alike, searched the two geometrically potential corners at chance (either corner C or R). In the condition in which both geometrical and non-geometrical information was available (1.6B), adults were able to combine the two types of information to successfully identify corner C, but infants – like other mammals – could not.

In a follow-up study into the development of this skill, children were found to develop the ability to combine the types of information around age 4. Success was related to the production of expressions using *left* and *right*, suggesting that flexible reorientation in the task was linked to emerging spatial language abilities (Hermer-Vazquez, Moffet, & Munkholm, 2001). Moreover, when adult participants engaged in the task while performing a language interference task, their ability to combine geometrical and non-geometrical information declined. This decline was specifically related to linguistic interference tasks; not with other, non-linguistic, working memory tasks (Hermer-Vazquez et al., 1999).

The pairing of language with visual events and images also affects recollection about the spatial world. For example, Loewenstein and Gentner (2005) found that preschool children performed better in a mapping task when spatial relations were combined with spatial language at encoding. In this series of experiments, an object was hidden in a 'hiding box', and children were asked to find a matching item in the corresponding place of a 'finding box' (see Figure 1.7).

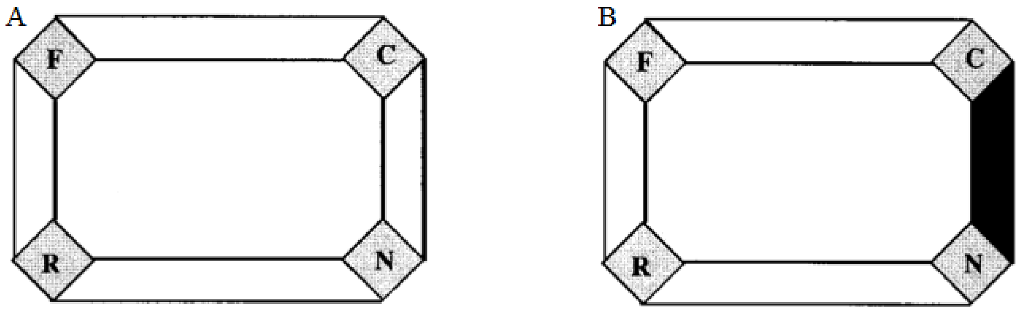


Figure 1.6. Example from Hermer-Vazquez et al. (1999). The room with only geometrical information (1A) and the room with a coloured wall on the right (1B), as a marker to aid participants orientate themselves and find corner C.

Different sets of spatial terms were used, *on, in, under* and *top, middle, bottom*. After the object placement, the children were asked to find the corresponding object in the finding box. Over five experiments, they found that children performed better on spatial relational mapping tasks if they had heard an informative description of the hiding event and this linguistic facilitation was better for the 'easier' linguistic descriptions like *top, middle, bottom*. Loewenstein and Gentner argue that relational language fosters the development of representational structures that facilitate cognitive processing (see also Hermer-Vazquez et al., 1999).

Furthermore, it was shown that children that have no spatial language perform poorly on non-linguistic spatial tasks (Gentner, Özyürek, Gürcanli, & Goldin-Meadow, 2013). In a study with deaf Turkish children without any sign language, deaf children not exposed to a conventional (sign) language rarely produced gestures to indicated spatial relations between two objects. They also had a poor performance on a spatial mapping task compared to cognitively-matched, hearing children (Gentner et al., 2013).

A series of developmental studies further investigated whether language facilitates spatial performance (Dessalegn & Landau, 2008, 2013; Farran & O'Leary, 2015). For example, 4-year olds performed a task in which they had to find a target (e.g., a square split in half by two different colours) in an array. The aim was to see whether language has a role in forming or maintaining a representation of visual features of an object, for example a split square, coloured red on the left and green on the right (Dessalegn & Landau, 2008). During the trials, the children were told to

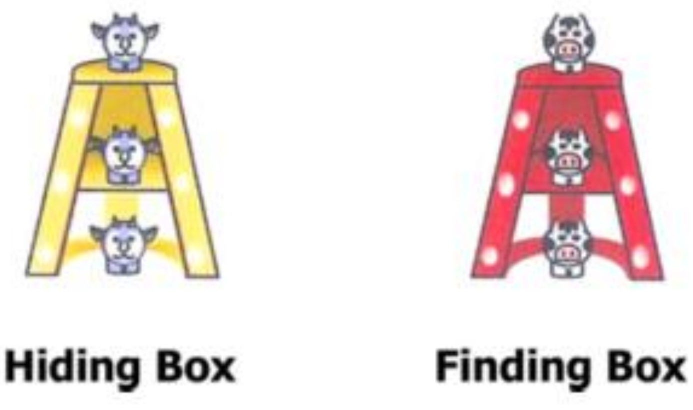


Figure 1.7. Example from Loewenstein and Gentner (2005). Children observed an object being hidden and had to find the corresponding place in a second space.

look at the target object and help the experimenter find the exact same one in an array of three objects presented after a 1 second delay: the target's match, its reflection, and a differently split square. Results showed that the layout of the objects was better retained when directional language was presented (e.g., the red is on the left/right/top/bottom), compared to a number of other conditions (e.g., when no description was given, when the object name was described (after the target was given a name, e.g., *dax, wazzle*), the description was a non-directional term (e.g., "the red is *touching* the green"), or when the child pointed at the object). The fact that none of the other conditions worked as well as directional language suggests that it is not a general effect of attention, but that descriptive spatial language has a particularly powerful facilitation on encoding. Even more surprising is the finding that children did not need to have a fixed understanding of directional language for their task performance to benefit; for example children that did know that *left* and *right* were opposites, but did not know which one was which. Dessalegn and Landau (2008) suggested that the combination of seeing the target and hearing the language (e.g., "the red is on the left") could have led to a temporary understanding of the directional cue (Dessalegn & Landau, 2008, 2013). Follow up studies looking at development replicated the effect for 4 year olds, but found that 6 year olds performed at ceiling for both a 'no label' and 'directional label' condition, although by that age children may be automatically encoding the stimuli using spatial language which counters the different conditions (Dessalegn & Landau, 2013). Performance was enhanced when the target was accompanied by spatial cues (e.g., "yellow is on top"). There was no additional benefit for children verbalizing the linguistic cue themselves over just hearing the cue, as long as they had an understanding of the spatial terms (Farran & O'Leary, 2015).

1.4.2 Colour perception

In addition to the influence that object knowledge has on memory, object knowledge can influence colour perception. Participants had to match colours between a stimulus patch and a variable colour mixer (Bruner, Postman, & Rodrigues, 1951). Participants were presented with object images in *gray*-scale, objects that usually have a stereotypical colour (e.g., banana, tomato, orange). They then moved a colour-wheel to match the represented object's colour, and judged the object-image to be stronger coloured (e.g., more yellow) than it actually was, in both a colour matching and a memory matching task. This effect could be a function of colour diagnosticity. If an object has a high colour diagnosticity (e.g., yellow for a banana) then the object knowledge effects on colour memory and colour perception were stronger than when the object has a low colour diagnosticity (e.g., yellow for a lamp). In a conceptual replication, Delk and Fillenbaum (1965) showed that this overestimation towards the typical colour of an object is not just apparent in colour memory, but also in online tasks. Colour retrieval is suggested to be a blend between actual colour and typical colour (Belli, 1988). It was shown, and replicated, that people overestimate the similarity of colours within a category compared to the colours of objects between different categories even if all stimuli are presented while participants make their judgements (Goldstone, 1995). For example,in an array of letters and numbers, participants judged symbols within a respective category to be more similar in colour than between categories, even if the two, between categories, target symbols were identically coloured (See Figure 1.8).

Bruner et al. (1951) suggested the effect of expectation was the result of a ‘three-step’ mechanism of perception. They hypothesised that the first step of perception is to make a hypothesis (an expectation) about a perception, then to acquire information before one can confirm or disconfirm the initial hypothesis. The effect of object knowledge would be elicited by this hypothesis. The misperception of colour as a function of object knowledge supports the perspective that colour perception is not just a result of incoming sensory data, but is significantly modulated by high-level visual memory (Hansen, Olkkonen, Walter, & Gegenfurtner, 2006). In their study, participants adjusted the colour of pictures of

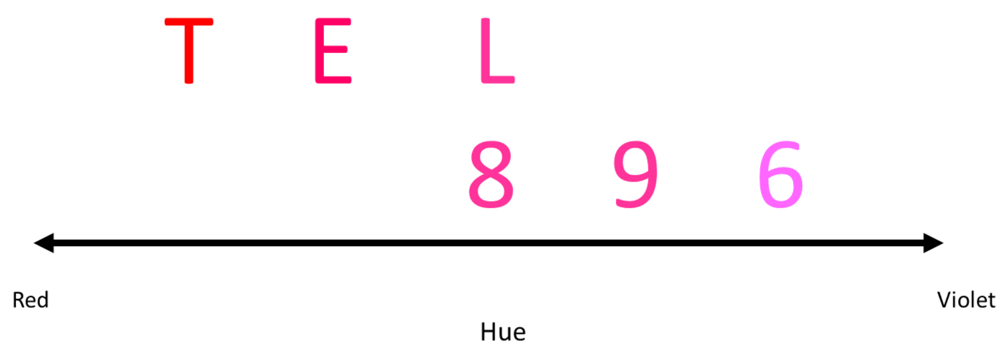


Figure 1.8. Participants judged symbols within a category (T, E, L vs. 8, 9, 6) as being more similar in colour, even though ‘L’ and ‘8’ were identical in colour (Goldstone, 1995).

fruit (e.g., banana) until they appeared to be gray (see Figure 1.9). However, results showed that participants went too far with this adjustment and moved the colour over the actual gray point in the direction opposite to the typical colour of the fruit. The authors suggest that these results show that colour perception is not determined purely by incoming sensory information.

An effect of language on colour perception has also been found, both in cases where language is explicit during a task, and in cases where language is not explicit, but may affect non-linguistic performance. There is a positive correlation between linguistic categorization of colour (the possibility to code a specific colour in a specific linguistic category) and the recognition of a given colour, suggesting that the possession of colour words could affect colour categorization (Brown & Lenneberg, 1954; Roberson et al., 2000). Furthermore, when participants engage in a verbal interference task, the linguistic facilitation disappears (Roberson & Davidoff, 2000), although in a replication study it was argued that whereas language facilitates performance, verbal interference does not necessarily prevent colour categorical perception (Pilling, Wiggett, Ozgen, & Davies, 2003). However, it is disputed whether tasks on colour-judgements can inform the debate on colour perception (cf., Firestone & Scholl, 2015). Firestone and Scholl argue that in interpreting these results, authors often confuse perception with judgment. Although it is not always clear whether top down cognition influences what we see or instead how we infer or judge based on what we see. The fact that participants judge an object to be more colourful than it actually is, whether in an online or a memory task, does not mean

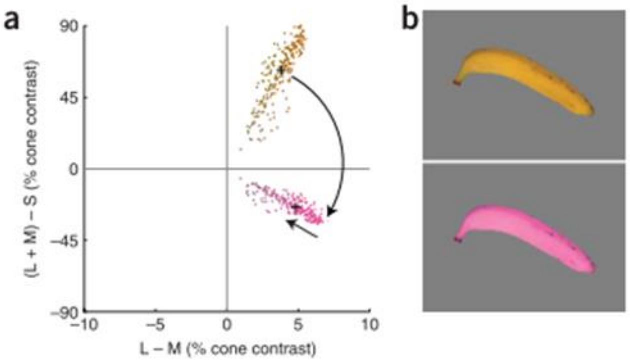


Figure 1.9. The 0-point on the Y-axis is the actual grey point. When participants adjusted the colour of the banana so that it appeared achromatic to them, they generally went over the actual gray-point, suggesting that their perception was influenced by the expected colour of the fruit (Hansen et al., 2006).

that they actually perceive colour differently – one can merely state that they judged colour differently.

Another issue identified by Firestone and Scholl (2015) is that overly confirmatory research designs can identify effects that are not necessarily caused by the experimental manipulation. To control for this, a design must include both conditions in which an effect should occur, but also in which the effect should not occur - apart from finding an effect when one would expect an effect, Firestone and Scholl argued, one must not find the effect when it would not be expected.

In language research, an example of this dissociation is in the effect of language on the different brain hemispheres. In most typically developing individuals, language is more strongly represented in the left hemisphere and linguistic tasks are processed preferentially by the left hemisphere (Carey & Johnstone, 2014; Mazoyer et al., 2014; Pulvermüller, 2012; Rasmussen & Milner, 1977). Therefore, it would be expected that the influence language has on colour perception would be different for the left and the right visual field, as they are processed by respectively the right and the left side of the brain (Brown et al., 2011; Drivonikou et al., 2007; Gilbert et al., 2006). In colour discrimination tasks, language could disproportionately influence colour discrimination in the right visual field (RVF, processed by the typically more linguistic left side of the brain) compared to the left visual field (LVF). Each trial showed two graded tones on a spectrum from green to blue in a circle of squares. In the circle, all squares had the same colour except for the target (See Figure 1.10). Participants then indicated

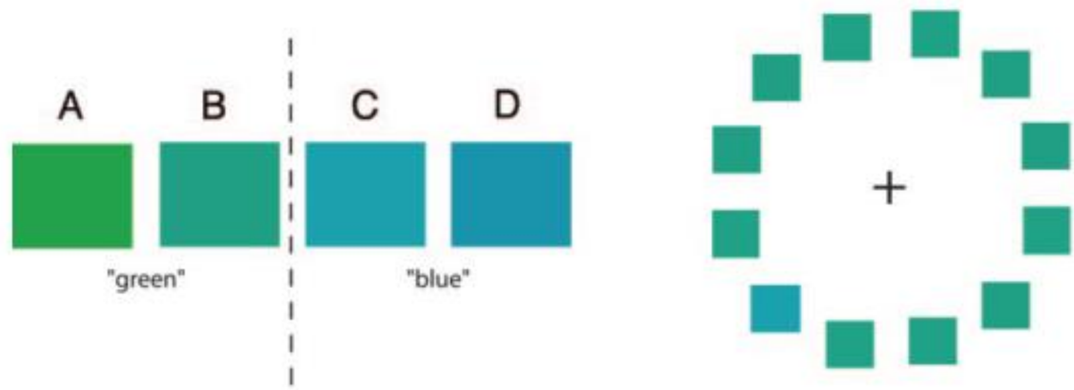


Figure 1.10. Examples of four colours (distinction between light/dark green and blue) on the left. On the right, an overview of the visual search task. Participants had to respond with a left or right button press, to indicate at which side the target was (Gilbert et al., 2007).

whether the target was at the left or the right side of the circle. Results showed a strong, exclusive influence of language in the RVF; colour discrimination of colours with different names is quicker in the RVF than the LVF. In a replication of Gilbert et al. (2006), Drivonikou et al. (2007) found similar results, with the difference that whereas Gilbert et al., found significant category effects in the RVF but not the LVF, Drivonikou et al. (2007) found significant effects in both sides of the brain, although effects were stronger in the RVF. Drivonikou et al. speculate the effects in the LVF effect could be a result of information travelling across the corpus callosum to the right hemisphere or that the LVF category effect reflects a universal categorical distinction. However, multiple studies have focussed on the idea that language affects colour perception differently in the two brain hemispheres since, finding mixed results (cf., Brown et al., 2011). And studies with pre-linguistic infants show categorical effects as well (Ozturk, Shayan, Liszkowski, & Majid, 2013). In their study, 4-6 month old infants were faster and more accurate at fixating on different coloured targets compared to targets from the same colour category, even when they controlled for chromatic separation, suggesting the existence of pre-linguistic colour categories. These results might suggest that categories are not dependent on language.

As mentioned above, linguistic relativity and universals in language are not necessarily mutually exclusive. Cibelli et al. (2016) suggested that there are universals in cognition, and uncertainty can increase the influence of contextual information. In other words, linguistic influences could come into play when perceptual information allows for uncertainty.

1.4.3 The influence of language on motion perception

A key feature in cognition is that people are able to predict events based on information they perceive, for example by engaging in ‘mental animation’ in which they anticipate future states of the world. For example, when in traffic, one continuously predicts what other road users will do, and whether one has time to cross the road or needs to wait until a car has passed. This implies that there is a mapping between perception at any given time, and knowledge stored in memory (Bar, 2009).

In a study presenting static scenes to participants, it was shown how language implying motion can affect memory for object location (Vinson, Engelen, Zwaan, Matlock, & Dale, 2017). Participants saw a picture of a car either facing uphill or downhill on a slope, visually implying that the car moves forward or backward downhill. The picture presentation was accompanied by one of three conditions: language congruent or incongruent with the visual motion implication (congruent: the car moves further down; incongruent: the car moves further up) and a no language condition. In the no-language condition, participants misremembered the cars' location to be further down on the slope. When language implied a downward motion, this effect was enhanced, with participants misremembering the car to be significantly further down. But if language implied an upward motion, there was no significant memory difference from the actual (originally presented) location of the car. These results suggest that motion language can provide a cue to anticipate on motion.

Meteyard, Bahrami, and Vigliocco (2007), found that language can also inhibit the ability to see motion. In their study, participants engaged in a motion detection task, while passively listening to two types of verbs, either implying motion or not. Results showed that when the direction of a motion verb was incongruent with the presented motion, participants performance on the motion detection task deteriorated. Although there was no difference in the detection task when words were congruent or non-motion-related, these results suggest that linguistic information directly influences perception.

Coventry and colleagues (Coventry et al., 2013) tested the influence of object knowledge, situational knowledge, and language on the extent to which people mentally animate static images by measuring middle temporal/middle superior temporal motion processing activation using fMRI (see Figure 1.11). Results showed that mental animation is not driven by individual objects (e.g., a cereal box and a bowl on their own do not elicit mental animation), but suggested that mental animation can be driven by an expectation elicited regarding how objects usually interact (e.g., 'a cereal box above a bowl' does elicit mental animation where 'a bowl above a cereal box' does not). Language presented before images also affected motion processing. When the language verbalised the spatial relation between two objects (e.g., "the glass over the bottle"), more mental animation activation was found than in a comparative-adjective condition (e.g., "the packet is larger than the pan"). This shows that there is an influence of language that highlights a functional

Language conditions: Spatial preposition: “the packet is over the pan”
Proximity preposition: “the packet is near the pan”
Comparative adjective: “the packet is bigger than the pan”

Figure 1.11. Examples of the spatial lay-out used in different trials. Left, functionally congruent items (packet over a pan), in the middle functional incongruent layout (the pan over the packet), and on the right objects that are not functionally related. Participants had to judge whether the verbal description was true or false. (Coventry et al., 2013).

relation, which is absent when the functional relation is not elicited by language. Furthermore, consistent with Gilbert et al. (2006), it was found that the language activation only worked in the left hemisphere, the hemisphere responsible for language processing in most typically developed individuals (Carey & Johnstone, 2014).

Revisiting the criticism of Firestone and Scholl (2015), rightfully stating that a research design should include conditions in which the effect should occur and conditions in which the effect should not occur, we see that Coventry et al. (2013) do this twice. First, they report that there is an effect in the left hemisphere – which predominantly processes language –, but not in the right, which supports the perspective that specifically language can induce this mental animation. If the effect is caused by language, the effect is expected to be weaker (if at all) in the right hemisphere. Second, the fact that language that highlights a functional relation does induce mental animation, but language that does not highlight this functional relation does not, shows that there is not just a language effect, but that it is an effect that occurs only in specific types of language.

1.4.4 The influence of higher order cognition on perception: A cautionary note

Despite the large body of literature on cognitive penetration of perception (i.e., Goldstone, 1995; Hansen et al., 2006; Lupyan et al., 2010; Meteyard et al., 2007; Pinto, Gaal, Lange, Lamme, & Seth, 2015; Thierry et al., 2009), it has been suggested that none of the many findings are evidence of top-down effects of cognition on perception (e.g., Firestone & Scholl, 2015; Pylyshyn, 1999). Firestone and Scholl identified a number of pitfalls to which they claim many – if not all – of the studies fall prey to. Previously, we mentioned potential confounds related to specific studies. Here we take a closer look at potential pitfalls. Among these pitfalls are mistakes by the experimenter, such as creating a design lacking a condition in which the effect they hypothesise should not occur (as we saw being addressed by for example Coventry et al. (2013) and Gilbert et al. (2006)). Without this condition, research designs are overly confirmatory - any time an effect occurs one can conclude that the condition caused the effect. Another example, specifically salient in the domain of colour, is confusing perception and judgment (cf. Goldstone, 1995; Hansen et al., 2006; Roberson et al., 2000). When participants have to judge whether two colours match, this is not a direct measure of their perception, but merely their judgement. Related to this, is failing to identify the factor memory plays in recognition tasks. When participants are quicker to recognize a stimulus in a specific situation, this could be due to a top-down facilitation of visual processing, but could be driven by memory (or a lack thereof) (cf., Lupyan et al, 2010). Other pitfalls regard the participants, for example if participants fall subject to demand or response biases, in which features of the experiments lead subjects to adjust their responds to suit the experimenters' expected aim. Firestone and Scholl suggested that researchers use a checklist based on these pitfalls, to consider whether the effects do in fact suggest top-down effects of cognition on perception, or whether they can be attributed to any of the identified pitfalls (see Table 1.1). In the studies presented in

Table 1.1. *The checklist as proposed by Firestone and Scholl (2015).*

1. Uniquely disconfirmatory predictions: allow conditions both in which the effect appears and disappears.
2. Perception versus judgment: disentangle postperceptual judgment from online processing.
3. Demand and response bias: mask the purpose of manipulations and test whether participants discovered the aim upon debrief.
4. Low-level differences (and amazing demonstrations): rule out explanations involving lower-level visual features by matching stimuli carefully.
5. Peripheral attentional effects: measure patterns of attention directly.
6. Memory and recognition: distinguish online perception from off-line memory

this thesis, we aim to avoid these pitfalls. In Chapter 5, section 5.3.1-5.3.3 we discuss this checklist with respect to the experimental chapters.

1.4.5 Spatial memory

Before we look at the influence of language on spatial memory, we review spatial memory itself. Our ability to memorize information is fundamental for cognition, all behaviour and every act is a product of the memory of a previous experience with a given situation (Thelen, 2005). Without spatial memory we would continuously be engaged in searching for our keys, phone, or glasses. Memory for object location can be decomposed in at least three different mechanisms: object processing, spatial-location processing, and the binding between object and location (Postma & Haan, 1996; Postma, Kessels, & van Asselen, 2008). For object identity processing, the object needs to be compared to representations of known objects and subsequent semantic properties. Then spatial location processing is needed to gain exact knowledge of object position, both via allocentric and egocentric information (King, Trinkler, Hartley, Vargha-Khadem, & Burgess, 2004). Space needs to be organised hierarchically so that object locations can be estimated (Huttenlocher, Hedges, & Duncan, 1991; Huttenlocher, Hedges, Corrigan, & Crawford, 2004). This creates a more general idea of locations (e.g., the cup is on the right side of the shelf).

The binding of objects and locations has long been debated. For example, whether memory for object location is merely a feature remembered in object knowledge, or whether object location is a memory trace, separate from object knowledge (cf., Treisman & Gelade, 1980). Findings that locations and features can be stored interdependently suggest that they are separate systems (Treisman & Zhang, 2006). Participants depend more on object location when object bindings are tested, compared to when object features were tested (Hollingworth & Rasmussen, 2010; Treisman & Zhang, 2006). Object locations are not remembered as absolute coordinates, but objects are remembered to be in the array-relative locations, relative to other objects or anchors in the array (Hollingworth & Rasmussen, 2010). The relational binding leads to misremembering the location of multiple objects as closer to each other (Jiang, Olson, & Chun, 2000; Pertzov, Dong, Peich, & Husain, 2012), or closer to a salient cue like a landmark or the border of a stimulus array (Nelson & Chaiklin, 1980).

Object location might be important in maintaining a visual memory. Pertzov and Husain (2014) found that location was memorized whether it was relevant for a specific task or not. In their study, participants were asked to memorise object features. Participants exhibited more feature identification errors when multiple objects were presented at the same location, suggesting that the binding of non-spatial object features might be mediated through an object's location. Furthermore, it has been argued that spatial memory is conceptualized as a motor plan, so that memory for object location is coded as a motor plan from a specific body position to the target object (cf. Smith & Samuelson, 2010).

Memory for object location might be represented at two levels of detail, the absolute, perceived location of an object or event and the category it belongs to (Huttenlocher et al., 1991), a suggestion that relates to the representation of absolute coordinates and categorical knowledge in spatial cognition (Postma et al., 2008). The fine-grained stimulus values (absolute location) is inexact, leaving room for contextual information like category membership to add to the memory. Furthermore, memory seems to be biased to the centre of a category it belongs to under the influence of this contextual information (Crawford, Huttenlocher, & Engebretson, 2000; Hollingworth, 1910). This extends the probabilistic model we discussed in the colour domain to spatial memory (Cibelli et al., 2016; Regier & Xu, 2017). In this probabilistic model, spatial memory consists of the actual perception of an objects' location and the category the object belongs to (Huttenlocher et al., 1991; Huttenlocher, Hedges, & Vevea, 2000). For example, memorizing the location of a dot in a circle, participants misplaced the dot towards a central location within the respective quadrant when the circle was divided into quadrants (Huttenlocher et al., 1991). The details of fine-grain values of memory can wear off, for example by memory loss over time or by interference from contextual information, for example, a longer delay before recall leads to an increased estimation error (Pertzov et al., 2012). A conceptual replication of these studies supported the influence of category on spatial memory (Holden, Newcombe, & Shipley, 2013). A location memory task using a 3D environment showed memory to be biased towards the centre of a category, and that this bias increased with longer retention intervals. Different sources can affect memory by averaging integrated information from different sources, for example: haptic information (Ernst & Banks, 2002), categorical knowledge (cf. Coventry, Griffiths, & Hamilton, 2014; Hemmer & Persaud, 2014) or linguistic information (see Chapter 2). The use of contextual knowledge could be to help maintain accurate memories (Huttenlocher et al., 2000, 2004).

1.4.6 Language and Memory

There are many examples of the effects of language on memory. It has been found that language can elicit memory effects consistent with the 'self-reference effect'. This self-reference effect entails that memory is enhanced when one feels s/he owns an object (Cunningham, Turk, Macdonald, & Macrae, 2008). Cunningham and colleages had participants engage in a picture sorting task. The participants worked with a confederate in a task sorting cards, which represented shopping items, into their own or the confederate's basket. At the end of the sorting task, they got a surprise memory task at which point their memory performance was better for items sorted in their basket. Object ownership hereby showed to elicit the self-memory advantage (cf., Cunningham, Vergunst, Macrae, & Turk, 2013; Turk, van Bussel, Waiter, & Macrae, 2011).

This effect has been extended by merely presenting possessive pronouns in combination with objects (Shi, Zhou, Han, & Liu, 2011). Shi et al. manipulated whether Chinese nouns describing items were followed by a first-person pronoun (*my)* or third-person (*his*) pronoun. Then, using key responses, half the participants had to judge whether they liked the item or not (contextual encoding) and the other half had to judge whether the nouns were presented in green or blue (perceptual encoding). This distinction was added to emphasize ownership in the contextual encoding condition compared to the perceptual encoding condition. Directly following the trials, participants completed a surprise memory test in which they were asked to recall the nouns. It was found that participants responded quicker to, and memory performance was better for, nouns presented in the *my* condition than in the *his* condition. The effects were stronger in the contextual encoding condition compared to the perceptual encoding condition, explained by the fact that ownership was more salient in the contextual encoding condition. Shi et al. (2011) concluded that the use of language – the possessive pronoun – effectively developed a sense of ownership, creating a 'self-related' or 'other-related' content to the noun referred to by the pronoun.

Language can also lead to memory errors (Feist & Gentner, 2007; Gentner & Loftus, 1979). For instance, Feist and Gentner (2007) showed that recognition memory for spatial scenes was shifted in the direction of spatial relational language (spatial prepositions) presented with scenes at encoding (see Figure 1.12). In their study, they presented participants with ambiguous pictures depicting spatial relations accompanied with or without spatially related sentences. For example, a picture of a block was presented ambiguously located on top of a building, as in Figure 1.12. In a spatial language condition, participants heard the sentence "*the block is on the building".* When participants responded in a later yes-no recognition task, spatial language at encoding was associated with more false positives (in cases where the spatial language at encoding was associated with a more prototypical version of the spatial relation than the relation in the picture). In the example, the picture on the left would have more false positives. Feist and Gentner (2007) argued this is a result of an interactive encoding of language and visual memory, in which language influences the way people encode visual scenes. More broadly, language can be used as a task-related tool to aid memory and/or processing of spatial information (see for example Frank, Everett, Fedorenko, & Gibson, 2008; Li, Abarbanell, Gleitman, & Papafragou, 2011) consonant with weaker versions of linguistic relativity (for review, see Wolff & Holmes, 2011).

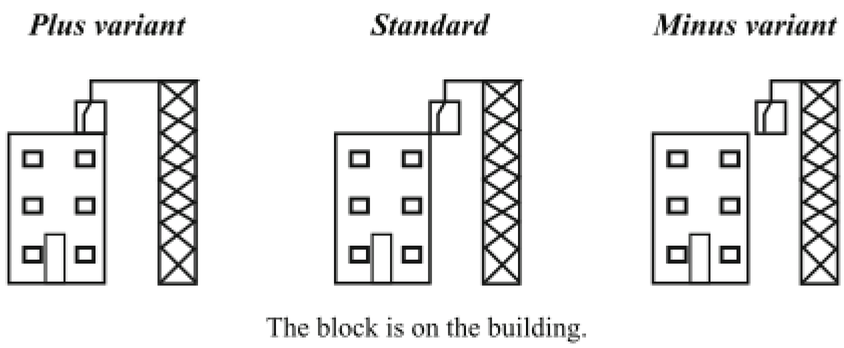


Figure 1.12. An example of a spatial scene combined with a verbal description from Feist and Gentner (2007). Memory data showed that participants reported more false positives in the direction of the verbal description (e.g., the block is on the building when the standard image does in fact not show this).

1.4.7 Verbal overshadowing

Other effects of language on memory are found in *verbal overshadowing* (Schooler & Engstler-Schooler, 1990). A body of literature on verbal overshadowing – recently replicated in a many-labs replication (Alogna et al., 2014) – showed that whereas repetition typically improves memory, it is found that verbally describing a situation decreases memory accuracy. In the original study, participants saw a 30s video of a bank robbery, followed by a 20-min filler task. After the filler task, participants were randomly assigned to a face verbalisation or a control group. The face verbalisation group were given 5 minutes to write detailed descriptions of the bank robber's face, and were encouraged to describe each facial feature as detailed as they could. The control group were given 5 minutes for an unrelated task. Afterwards, both groups were presented with a slide containing eight similar faces, plus an option to say the robber was 'not present' on the slide. The group that had to describe the individual performed worse on a recognition task than a control group that did not engage in memory verbalization. Follow up studies tested whether the effect was caused by language, or whether face visualization has a similar influence. Visualization did not impair face recognition (Schooler & Engstler-Schooler, 1990).

Verbal overshadowing might interfere in memory recall, but also changes memory. In a series of studies, the concept of verbal overshadowing was extended (e.g., Gentner & Loftus, 1979; Loftus, 1975; Loftus, Miller, & Burns, 1978; Loftus & Palmer, 1974). For example, participants saw videos of car accidents, after which they had to answer questions. The manipulation in these questions focussed on which words described the impact of the two cars, for example: contacted, bumped, or smashed. If the question had been "*about how fast were the cars going when they smashed into each other?",* participants estimated higher speeds and were more likely to report seeing broken glass in a follow-up question (Loftus & Palmer, 1974). Other participants saw a video of a car driving around a neighbourhood, in which at some point the car drives by a yield sign. After the visual information was presented, participants were asked different questions, for example whether another car passed by when the car stopped at the stop/ yield sign. After the questions, participants saw slides pairs, for example depicting the car next to a stop or a yield sign, and participants had to pick the slide they had seen before. The accurate selection of either slide decreased from 75% to 41% based on the consistent vs. misleading information, suggesting a semantic integration of the verbal information into the visual memory. This is consistent with the perspective we discussed above, in which information from different modalities is processed in a feedback system, in which pattern completion generates a full representation of an event.

1.4.8 Summary of the Influence of Language on Cognition and Perception

In this section we reviewed the influence that language can have on cognition and perception. We saw that language can do so indirectly, by focussing attention or assigning an object (e.g., colour) to a specific category. Upon recollection of a memory, general objects knowledge, like category membership, can merge with the actual perception memory and thereby influence memory for colour (Cibelli et al., 2016) and object location (Coventry et al., 2014; Huttenlocher, 2004).

In this thesis, we will explore the influence of pronouns (demonstratives and possessives) on spatial memory, and test whether the mere use of language to assign an object to a category can influence memory similar to the actual category membership (e.g., ownership) (see Chapter 2 and 3). Furthermore, we tease apart different models predicting effects of language on memory for object location. In Chapter 4, we test whether reference frames as expressed in language (spatial prepositions) affect eye-movements in a structural way. Before we do so, we will discuss the language we focus upon, pronouns and spatial prepositions, in some more detail.

1.5 Demonstrative pronouns and spatial prepositions

So far, we have reviewed research on the relationship between language and non-linguistic cognition. In the experimental chapters of this thesis, we will explore the influence that language has on non-linguistic cognition. We use demonstrative pronouns and spatial prepositions as a vehicle, because these types of language are both very frequent, clearly map onto a non-linguistic domain, space, and vary considerably across languages. In Chapter 2, we will test how demonstrative pronouns affect memory for object location. In Chapter 3, we compare how demonstratives are used in English and Japanese, and how this relates to their influence on memory for object location. In Chapter 4, we test the effect of reference frames as expressed in spatial prepositions on visual attention. Before moving to the experimental chapters, we will introduce demonstratives and spatial prepositions.

Spatial demonstratives (e.g., this, that, here, there, etc.) are a small class of referential expressions, but a growing body of research shows the important role they play in language use. Their importance can be inferred from findings that every language has demonstratives, that there is no evidence these were derived from content words (such as verbs or nouns) – suggesting they are an early, individually developed linguistic class – and the fact that they are among the most frequent words in language use (Diessel, 1999, 2006, 2014; Heine & Kutteva, 2002).

Demonstratives did not only emerge early in language evolution; they are also acquired very early on in language learning. Deictic words are among the first words children acquire (in the form of sounds) - although the proximal/non-proximal contrasts (e.g., here, there, this, and that) take longer to master and are usually present by age two-and-a-half years of age (Clark & Sengul, 1978). Demonstratives are often accompanied by a specific deictic pointing gesture, which is usually acquired when children are 1 year old (Capirci, Iverson, Pizzuto, & Volterra, 1996; Cooperrider, 2016). Deictic terms, such as *that* are used to pick out many different kinds of objects (such as *general purpose* terms like *do* are used to indicate many different actions) (Clark & Sengul, 1978). With this early use of demonstratives, and the accompanying deictic pointing behaviour, a joint focus of attention can be established (Tomasello, Carpenter, Call, Behne, & Moll, 2005). Joint attention is one of the most basic aspects of social cognition, in which conspecifics overlap focus of attention to achieve a common understanding of a referent, but it is also argued that this joint attention is necessary for the acquisition of language (Bruner, 1983). It is of note, however, that there is little empirical research that has experimentally tested the function of spatial demonstratives.

There are different theoretical views on of the function of demonstratives. Spatial demonstratives might contrast discrete zones of peri-personal (near) and extra-personal (far) perceptual space (Bowden, 2014; Clark & Sengul, 1978; Coventry, Valdés, Castillo, & Guijarro-Fuentes, 2008; Diessel, 2006; Enfield, 2003; Peeters, Hagoort, & Ozyürek, 2014; Stevens & Zhang, 2013; Talmy, 1975), although there is no consensus yet on this deictic contrast (Kemmerer, 1999; Peeters et al., 2014). The proposed contrast between peri-personal and extra-personal space is flexible and graded. Near space can be contracted by weight use (Longo & Lourenco, 2006), and the use of *this* is similarly extended when participants use a stick to point to objects (Coventry et al., 2008). However, the choice of demonstrative can be based on multiple dimensions (e.g., spatial, temporal, and task performance) (Byron & Stoia, 2005). Alternatively, it has been argued that demonstrative use cannot be explained by a mere *egocentric account* (Jarbou, 2010; Peeters et al., 2014; Peeters & Özyürek, 2016), and it has been claimed that there are more parameters that are important in demonstrative use (Burenhult, 2003; Coventry et al., 2014; Jarbou, 2010; Özyürek, 1998). For example, Özyürek (1998) reported that in Turkish, a language with a three-way demonstrative system, the ‘medial’ demonstrative is used to encode an object in a *hearer’s* territory (in this thesis we will use the word ‘territory’ to describe an interlocutors’ personal space).

Another view of demonstratives is that of the *schema of focus for demonstrative reference* (Oh, 2001; Strauss, 2002) which suggests that demonstratives are used to indicate different levels of importance, see Figure 1.13. In the schema, *this* indicates the highest importance to a referent, thereby signalling a hearer should pay more attention compared to *that* (medium focus of attention) and *it* (low focus of attention).

Sampling over 200 languages, Diessel (2005) found that, while most languages have a demonstrative system expressing a two-way contrast (54.4%) (e.g., English: *this/ that*), a large group of languages employ a three-way system (37.4%). An example three-way system is Spanish, which is often assumed to have a distance based demonstrative system (‘*este/ese/aquel’).* The different demonstratives are used to indicate referents at different distances from the speaker: *este* is used to refer to objects close by, *ese* for objects at medium distance (a place between close by and far away), and *aquel* for objects far away (empirically tested in Coventry et al., 2008). Coventry et al. found that both English and Spanish languages showed a relation between distance and demonstrative use, in which distance of an object from the speaker is verbally encoded. However, while Spanish may be distance based like English, it is important to test whether languages maybe fundamentally different in

	Form	**MEANING SIGNAL**	**Hearer**	**Referent**
Degree of the attention to pay to the referent	This	HIGH FOCUS	New information (not shared)	Important (to speaker)
	That	MEDIUM FOCUS	↓	↓
	It	LOW FOCUS	shared information	Unimportant

Figure 1.13. The ‘Schema of Focus for Demonstrative Reference’ (Strauss, 2002), suggesting that demonstratives can be used to ascribe different levels of importance to a referent.

how they use demonstratives and/or structure space, as it has been argued that the principles underlying demonstratives are similar despite superficial differences (Luz & van der Sluis, 2008). In Chapter 3, we compare English with Japanese. It has been proposed that Japanese is among a group of languages that are ‘person-centred’ (Diessel, 1999), although this is still subject of debate.

The other type of language we will focus on are spatial adpositions (e.g., in front of, behind, left, right). Just like demonstrative pronouns, this type of language is acquired early on in language and is high in frequency. Furthermore, across languages, there are only few prepositions that can be used to encode a wide variety of spatial relations. Next to that, languages vary massively in how they encode spatial relations using prepositions (Bowerman, 1996; Landau & Jackendoff, 1993). For example, in English the word ‘on’ would be used to encode many relations (e.g., the plate *on* the table, the painting *on* the wall, the passenger *on* the bus, the fly *on* the window, the napkinring *on* the napkin). In German and Dutch, these relations would be marked with different prepositions (e.g., the plate *auf/op* the table, the painting *an/aan* the wall, the passenger *auf/in* de bus, the fly *an/op* the window, the napkinring *om/um* the napkin) (example based on: Bowerman, 1996; Deutscher, 2005). Another variation in preposition use is how relations are categorized. For example, English distinguishes whether an object is placed in containment or on a surface, whereas Korean makes a distinction on whether the relation is a tight-fit or loose-fit regardless of containment or support (Choi et al., 1999).

In Chapter 4, we focus on how spatial prepositions influence visual attention by expressing a frame of reference (cf., Levinson, 1996), as discussed in Section 1.3.4. Spatial prepositions can encode space based on different reference frames: intrinsic, relative, absolute (see Figure 1.14). For example, the marble can be located based on the intrinsic axis of the ladybird (“in front of”), the marble can be located relative to the ladybird (“to the left”), or by using an absolute reference frame (“the marble is south of the ladybird”). In Section 1.4.1, we reviewed how language can focus attention. In Chapter 4, we will present a study exploring whether there are differences between how spatial prepositions focus attention on a visual array.

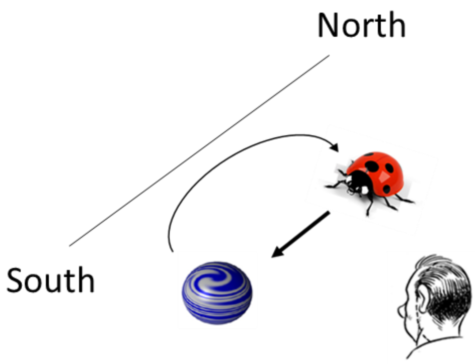


Figure 1.14. Example of different reference frames. The marble is 'in front of' the ladybird, from an intrinsic perspective of the ladybird. However, the marble is to the left relative to the ladybird based on the perspective of the man. Using an absolute reference frame, based on cardinal directions, the marble is south of the ladybird.

1.6 Précis of thesis

So far we have seen that more research on the influence of language on non-linguistic cognition and perception is needed. Not only are results mixed, there are also many pitfalls in the relevant studies (cf., Firestone & Scholl, 2015, see Table 1.1). Even though it seems language can influence non-linguistic cognition, hitherto there has been limited consideration regarding *how* language affects cognition and perception. In this thesis, we explored the relationship via spatial cognition. In Chapter 2, we identified three models offering a mechanism as to how language might influence memory for object location. Over four experiments, we tested predictions of models suggesting the mechanism is driven by the expectation that language elicits (the Expectation model); that the influence is a function of congruence, in which performance is enhanced when language and perception are congruent and diminished when there is incongruence (the Congruence model); and that a difference is driven by the difference in the focus of attention elicited by language (the Attention Allocation model). To preview our results, across four experiments we find evidence supporting the Expectation model, which suggests that language affects spatial memory by the *expectation* of object location it elicits. However, as different languages carve up the spatial world in different ways, it is important to see how these supposed expectation values affect a language in which the use of demonstratives is not as clear; for example languages employing a three-way demonstrative system.

In Chapter 3, we compare English and Japanese demonstratives, as Japanese is one of the languages with a three-way demonstrative system. In four experiments, we tested how Japanese demonstratives are used and what the influence of demonstratives on spatial memory is. Testing between four models explaining the use of Japanese demonstratives, we show that Japanese demonstratives encode both *distance* and *territory.* This ambiguity in language use manifests in the memory data, as the data does not reveal a clear pattern as we found with English in Chapter 2. However, interestingly, we did find an effect of the position of a conspecific in English, even though this parameter is not explicit in English. This could suggest that there are non-linguistic universal parameters underlying demonstrative use.

In Chapter 4, we test the influence of language (specifically reference frames) on visual attention. We tested whether different reference frames (intrinsic: the marble is *in front of/ behind*; relative: the marble is to the *left/rig*ht; and included a neutral condition: the marble is *blahblah*) expressed in English, influence the way we scan the world. We find distinctive eye-movement patterns of fixations between the two objects in a visual array, elicited by the different reference frames. These different patterns are consistent across prime and succeeding probe trials, suggesting a close relation between attention and verbal descriptions of visual scenes. Chapter 5 is the concluding chapter in which we situate the findings in existing literature, discuss limitations of the presented work and suggest future directions.

SECTION 2

-

Experimental Chapters

Chapter 2 – The Influence of Demonstratives and Possessives on Spatial Memory; Four Experiments

2.1 Introduction

The relationship between language and non-linguistic representations is a fundamental topic in the cognitive sciences. Often this relationship is approached from the standpoint of the extent to which non-linguistic representations are necessary for language comprehension (e.g., within the framework of 'embodied' cognition; cf. Barsalou, 1999). However, as reviewed in Chapter 1, the extent to which language can influence non-linguistic processes is equally important (Coventry, Christophel, Fehr, Valdés-Conroy, & Herrmann, 2013). Language can direct the attention of a conspecific to the spatial world; spatial expressions, such as *these coins* or *the cup is on the table* serve to direct the attention of a hearer to regions of space (Miller & Johnson-Laird, 1976). The pairing of language with visual events and images also affects memory for the spatial world. For example, Loewenstein and Gentner (2005) found that children performed better in a mapping task when spatial relations were paired with spatial language at encoding (e.g., "I'm putting the book on the shelf"). They argue that relational language fosters the development of representational structures that facilitate cognitive processing (cf., Dessalegn & Landau, 2008, 2013; Farran & O'Leary, 2015; Hermer-Vazquez et al., 1999).

As discussed in Chapter 1, language presented with a spatial scene can lead to memory errors (Feist & Gentner, 2007; Gentner & Loftus, 1979). Feist and Gentner (2007) presented participants with a spatial scene, combined with spatial language and showed that recognition memory for spatial scenes was shifted in the direction of the spatial relational language (spatial prepositions) presented with scenes at encoding. Language can be used as a tool to aid memory and/or process spatial information (see for example Feist & Gentner, 2007; Frank et al., 2008; Li et al., 2011).

The effects of language on memory are not limited to spatial cognition. It has also been found that presenting possessive pronouns in combination with a memory task enhances response times and memory for objects (Shi et al., 2011). Shi et al. presented Chinese nouns preceded by a pronoun (*my/ his*). Participants had to scale the presented nouns for likeability before completing a surprise memory test. In the *my* condition, participants responded faster and showed a better memory performance for the nouns than in the *his* condition.

Although language can influence memory, it has yet to be demonstrated how it does so. In this chapter, our focus is on the (possible) influence of spatial demonstratives and possessives on memory for object location. The continuous nature of spatial memory errors affords testing directly between different potential mechanisms regarding *how* language affects memory for object location.

As discussed in Chapter 1, spatial demonstratives (specifically *that)* are among the earliest words children learn (Diessel, 2006) and have been shown to be associated with discrete zones of peri-personal (near) and extra-personal (far) perceptual space (Clark & Sengul, 1978; Coventry, Valdés, Castillo, & Guijarro-Fuentes, 2008; Diessel, 2006; Lynott & Coventry, 2014; Peeters, Hagoort, & Ozyürek, 2014; Stevens & Zhang, 2013). However, this distinction is flexible and graded. Near space can be extended or contracted by tool or weight use (Longo & Lourenco, 2006), and the use of *this* is similarly extended when participants use a stick to point at objects (Coventry et al., 2008). In Coventry et al. (2008), the memory game was a cover to test how people used demonstratives to refer to objects at different distances. In this memory game, participants sit at a long table without any spatial cues (see Figure 2.15). Objects were placed at the different locations, and participants were told their memory for these object locations was tested. As they were in a *language* condition, they had to 'verbally encode' (with other people – not actually tested – in the no-language condition) the location of an object at different distances, using three words: a demonstrative (*this/that*), the colour, and name of the object (e.g., 'this red circle). The 'memory game' cover meant that participants' demonstrative use could be tested, while participants thought that the study was about spatial memory and were unaware of the significance of their demonstrative choice.

In addition to distance, other variables affect demonstrative choice. Employing variations on this memory game procedure, Coventry et al. (2014) explored the relationship between object knowledge and distance on both demonstrative choice in English and memory for object location. Different parameters of objects knowledge were manipulated (e.g., whether the participant owned the object, was familiar with the object, and could see the object during 'memory encoding'. Furthermore, Coventry et al. (2014) ran an actual memory study. In these experiments, participants memorized the respective location of different objects. On every trial, participants would get time to encode an objects' location, and, after object removal, were asked to match the position of an indication stick to the location they remembered the object had been placed. The influence of object knowledge on memory for object location was tested by comparing the remembered location to the actual object location. Across seven experiments Coventry et al. (2014) found that object familiarity (i.e., familiar versus unfamiliar coloured shapes), object ownership (whether the participant owned the object or not) and object visibility (whether the object was covered with an opaque cover or not) all affected demonstrative choice to describe object location and had a similar effect on (non-linguistic) memory for object location. For example, unfamiliar objects (low frequency colour-shape combinations, such as a viridian nonagon) were preferentially referred to with *that*, and were misremembered as being further away than they actually were relative to familiar objects (e.g. red square). The other parameters provided similar results: when the participant owned or saw (during encoding) the used object they were more likely to say *this* and thought it was closer by compared to when they did not own or see (during encoding) the object. In other

Figure 2.15. Schematic representation of the experimental setup in the memory game. Objects are placed on the coloured location marks on the table, each at 25cm further from the participant (Coventry et al., 2008).

words, results showed that referents that were preferentially referred to with *this* were remembered to be closer to the participant, relative to *that*. Based on these results, Coventry et al. (2014) proposed the expectation model.

In the current studies, we aim to tease apart predictions about how language influences special memory contrasting three models: the Expectation model, the Congruence model, and the Attention allocation model. It is important to note that the Expectation model is presented previously by Coventry et al. (2014), whereas the Congruence model and the Attention allocation model are not in previous literature. We derived these models from previous findings and formulated them to enable comparisons between the respective lines of research.

2.1.1 Expectation model

In order to account for both the demonstrative choice data and the memory data, Coventry et al. (2014) proposed a model of the influence of object knowledge on both measures. In their *Expectation Model,* memory for object location is a concatenation of where an object is located and where an object is expected to be located (see Figure 2.16). The expectation of the objects' location concatenates with the actual object location (with an associated estimation error) in memory, as follows:

$$M_D = f(D_a, D_{exp}, D_{err})$$

where M = signed memory error, D = distance, $_a$ = actual, $_{exp}$ = expected and $_{err}$ = estimation error (Coventry et al., 2014).

The expectation of object location can be elicited by object knowledge, for example from the parameters tested in Coventry et al. (2014). Participants preferentially referred to objects they owned, saw, or were familiar with using *this*, indicating a closeness to the objects, whereas if they did not own, saw, or were familiar with the objects, participants preferentially used *that*, marking a certain distance from the object. These results showed a parallel with the results in memory trials. Objects placed in conditions in which participants prefer to use *this* were remembered to be closer by compared to objects participants did not own, knew, or saw. The expectation model assumes that the information on which demonstrative contrasts are based, also influences the encoding or reconstruction of memory. Rather than memorizing the location of an object based on absolute coordinates derived from visual perception, it seems that different sources of information about an objects' location are combined to store and/or reconstruct this location. One of these sources is the actual perceived distance of an object, another source might be the knowledge one has about the object and where it is expected to be placed from long-term memory. A concatenation of these two sources, combined with an estimation error, is the memorized location of an object. The Expectation model does not specify whether this influence of different variables works during the original perception (and therefore influences memory at the encoding stage) or whether the influence works at the stage of retrieval. However, since the effects of the memory tasks are parallel to the effects of the language production task (Coventry et al., 2014), one could speculate that a top down effect is already present during encoding. We will discuss this in Chapter 5.

It might be noted that an effect of the distance between an observer and the object of which the location is remembered, is not explicitly encoded in the Expectation model (Coventry et al., 2014). This does however not mean that the Expectation model does not allow for an effect of this distance. For example, we could speculate that the distance of an object affects the precision in perception at the level of the actual distance. When objects are in peripersonal space, the need to make accurate moevements to manipulate these objects in reachable space may mean that object locations within peripersonal space should be remembered quite accurately. When objects are in extrapersonal space, the actual encoding of location is more noisy and the exact location of that object will be perceived less accurately (i.e., when one has to remember the location of an object placed at 50cm distance, distance estimation can be precise to the milimeter; when one has to remember the location of an object placed on the other side of a lake, distance estimation will be precise to the meter rather than milimeter). In other words, this scale difference would influence the precision of object location memory at the level of the actual perceived location of an object. Alternatively, the effect of distance might influence at the level of the estimation error: the further away an object is located, the larger the estimation error will be. As there is a general tendency to overestimate object location, distances become more marked as the object is positioned further away.

Relevant to the current studies, is that Coventry et al. (2014) did not examine the influence of language on memory for object location. However, by extension, the expectation model makes predictions regarding how language might affect memory for location. As *this* is associated with near space and *that* with far space, one can assume that the expected distance value associated with *that* would be greater than the expected value distance associated with *this*. Combined with the actual distance, the expectation model therefore predicts a main effect of language on memory for object location, with *that* associated with (mis)memory for objects further away than they actually were compared to *this* (Figure 2.16a).

2.1.2 Congruence model

In contrast to the expectation model, there is a considerable body of work within an 'embodied cognition' framework providing evidence for the importance of (in)congruence effects between language and space, that makes different predictions from the expectation model. For example, it has been shown that participants respond more quickly to positively valenced stimuli in a congruent high location than an incongruent low location, and vice versa for negative stimuli (e.g. Barsalou, 2008; Meier & Robinson, 2004; cf., Lynott & Coventry, 2014). What one might term a 'congruence account' has been extended to movement planning, whereby movements are prepared based on given language (Bonfiglioli et al., 2009; see also Stevens & Zhang, 2013). For example, Bonfiglioli et al. (2009) required participants to grip an object after listening to an instruction that indicated whether the object was near or far. They found an interaction in which reaction times were significantly longer when language was incongruent with space compared to when language and space were congruent. In extending this congruence account to memory for object location, one would predict a similar interaction. Congruence in language and space enhances the accuracy of memory for location, with greater errors (without specification of direction) when there is a mismatch between the demonstrative and location, as follows:

$$M_D = f(D_a, C, D_{err})$$

where M = signed memory error, D = distance, $_a$ = actual, C = congruence of language with location and $_{err}$ = estimation error (Figure 2.16b).

In other words, the Congruence model predicts that language influences memory performance via the congruence between a linguistic (spatial) description and a spatial situation. Using the contrastive nature of demonstratives, this indicates an object is closer by compared to that. When the demonstrative use is appropriate for a spatial situation (e.g., this for objects close by, that for objects further away), the linguistic cue enhances the memory compared to a situation where the used demonstrative is inappropriate for the spatial situation (e.g., that for objects close by, this for objects further away) - the incongruence between demonstrative and the situation distorts the memory and results in a larger memory performance error. The main difference between the Congruence model and the Expectation model is that the Congruence models predicts a specific interaction between language and location. Since this is the appropriate demonstrative for objects close by, but not for objects further away (whereas that is appropriate for objects further away, but not for objects close by), the effect of this will be different at different locations (mirrored by the effect of that). In contrast, the Expectation model predicts that the effects of this and that are similar across locations. Our results can therefore tease apart these two models. If the data show a cross-over interaction (as visualised in Figure 2.16b), this would be inconsistent with the Expectation model, but support the Congruence model. If the data does not show this type of interaction, this would falsify the Congruence model.

2.1.3 Attention allocation model

Distinct from both the Expectation and Congruence models, the effect of language on memory can work via the allocation of attention. A large literature shows that language affects where one looks in a visual scene, for example in terms of fixating particular objects when they are mentioned (e.g., Allopenna, Magnuson, & Tanenhaus, 1998; Tanenhaus, Spivey-Knowlton, Eberhard, & Sedivy, 1995). Building on earlier work, it is possible that language also affects the amount of time one spends looking at an object. *This* has been suggested to indicate high focus of attention, compared to *that,* which would indicate medium attention (Oh, 2001; Strauss, 2002). This importance could be driven by the fact that *this* is associated with proximity to a speaker. One might speculate that participants might look longer at an object at a location when preceded with *this* compared with *that*, as visual attention is allocated preferentially to near objects compared to objects further away (Garrido-Vásquez & Schubö, 2014). Following evidence that longer looking times are associated with better memory performance (e.g., Huebner & Gegenfurtner, 2010), one might then predict better accuracy of recall for trials preceded with *this* compared with *that.*

Alternatively, we can look at the difference between demonstratives and a neutral determiner. Since it is argued that demonstratives encode more information than merely a proximal/distal distinction (Jarbou, 2010; Özyürek, 1998; Peeters & Özyürek, 2016), it would be expected that demonstratives elicit more attention than a determiner would – since the determiner does not convey any information about the referent. If this would be the case, it might be expected that fixation times are longer in demonstrative conditions (*this/ that)* compared to the determiner condition (*the*). In Experiment 4 we used eye tracking during the encoding phase, to investigate whether differences are driven by attention allocation (Figure 2.16c).

Research is often subject to the trade-off between ecological validity and experimental control. Past methods used to determine the function, use, and understanding of demonstratives range from studies with high ecological validity, but low experimental control, to studies with high experimental control but low ecological validity (Cooperrider, Slotta, & Núñez, 2016; Enfield, 2003; Pederson & Wilkins, 1996; Peeters, Hagoort, & Özyürek, 2015; Stevens & Zhang, 2013, 2014; Wilkins, 1999). Observational studies like Cooperrider et al. (2016) and Enfield (2003) are high on ecological validity, but lower on experimental control. A bit further on the scale of experimental control are Wilkins (1999) and Pederson and Wilkins (1996). In these studies a structured questionnaire was used to ask

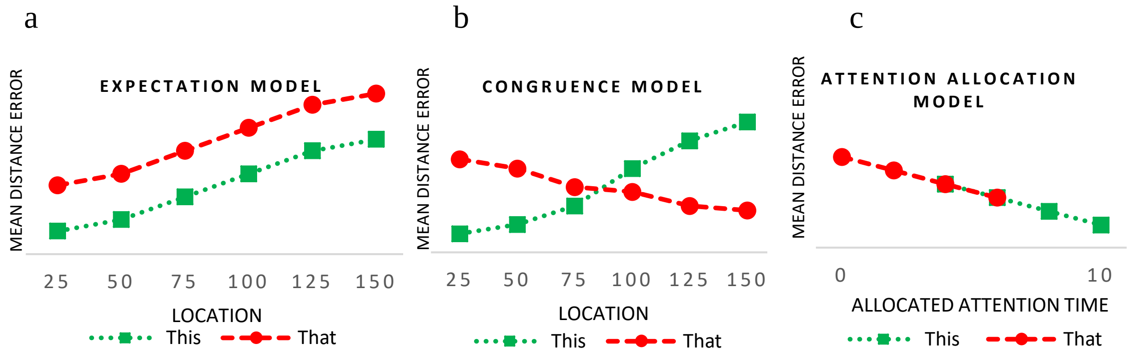


Figure 2.16. Predictions by the different models, from left to right: a. Expectation Model, b. Congruence Model, c. Attention Allocation model. On the y-axis the signed memory error (the difference between the actual location and the remembered location) is presented, a higher value on the y-axis means an object is remembered as being further away than it actually was. The lines represent the performance with different demonstratives (this/that); in Experiment 3 *this* and *that* are replaced respectively by *my* and *your*. In 2.16a and 2.16b, the six distances from the participant (in cm) used in Experiment 1, 2, and 3 are plotted. In 2.16c, the x-axis represents the total possible fixation time (10 seconds) for participants in Experiment 4. In Figure 2.16c more attention leads to a smaller memory error, and *this* is predicted to elicit more attention than *that*.

participants which demonstrative they would use in a number of hypothetical situations. This method increases experimental control, in the situations that are suggested, but are lower in ecological validity since participants have to imagine the hypothetical situations in which they would use demonstratives. Furthermore, asking participants about their intuitions about demonstrative use opens up the possibility of bias, in which participants could generate their own theory about demonstratives or what an experimenter might be looking for, thus not producing demonstratives naturally (cf., Firestone & Scholl, 2015). Studies with a high experimental control are lab studies, for example testing demonstratives while scanning brain activity using EEG (Peeters et al., 2015; Stevens & Zhang, 2013, 2014). These studies score high on experimental control – all images are exactly the same for participants and responses are time-locked per presented situation, while brain activity is measured, but the ecological validity is lower. For example, the use of photographs on which distance is manipulated is arguably not a fair test of the influence of distance on demonstrative use when the whole image on the screen is in peri-personal space. Methods used by the papers mentioned above each have their perks and the memory game we used is not meant to be a substitute any of these methods, but rather complementary, retaining the strengths of various approaches within one single paradigm. Key is that the method retains high experimental control while ensuring that participants use language naturalistically, in a real life, three-dimensional space, without being aware that researchers were exploring language. Our aim was to test whether language affects memory for object location, and to elucidate the mechanism involved. Specifically, we aimed to tease apart the three accounts by examining the effects of demonstrative and possessives on memory for object location. Experiment 1 and 2 tested whether spatial demonstratives affected memory for object location with contrasting predictions from two possible models of how language affects memory: congruence vs. expectation. Experiment 3 tests whether the effects found for demonstratives also occur for possessives (*my/your*) – terms that have also been associated with the peri-personal/ extra-personal space distinction. Experiment 4 tests predictions from the attention allocation model using eye tracking.

2.2 Experiment 1. The influence of demonstratives on spatial memory

2.2.1 Method.

2.2.1.1 Participants.

Eighteen native English-speaking students took part, receiving either course credit or payment for their participation. Stereoacuity was measured using the Randot Stereotest (Stereo Optical Inc. Chicago, USA). Two participants were eliminated from analysis, as they did not have a threshold of 40” (arcseconds). This left 16 participants for the analysis, 8 males and 8 females, with an age range of 19-30 years old (M = 20.5, SD = 2.8).

2.2.1.2 Materials

Six distinguishable, different coloured shapes on plastic discs (e.g. yellow triangle/blue heart), 6.5 cm in diameter, were placed on six different locations. The locations were spaced equidistantly along a midline from the participants’ edge of a large conference table (L = 320, W = 90cm), starting at 25cm from the participant up to 150cm. The three dots that were closest to the participants were located within peri-personal space, while the remaining three dots were within extra-personal space (confirmed for each participant). The table was covered with a black cloth so that no spatial cues were present.

2.2.1.3 Procedure and design.

Participants sat as close to the table as was comfortable, to ensure that all participants were approximately the same distance from the objects. Then, the memory version of the ‘memory game’ (described in section 2.1), pioneered by Coventry et al. (2014) was explained. As in previous variants of the memory game, participants were told the experiment was testing memory for object location. Each trial, the participant was asked to read out an instruction card indicating at which location to place a specific object. In this experiment, we used 6 different objects and combined those with the 6 locations closest to the participant, each 25cm further, starting at 25cm up to 150cm from the participant. In the current adaptation, we manipulated the language – demonstrative pronoun: this/that/the - used to instruct object placement. The instructions all had the form: “Place [demonstrative (3)] [object (6)] on the [colour (6)] dot” (e.g., “Place *this/that/the* red triangle on the blue dot”). Following the instruction, participants closed their eyes while the experimenter placed the object as instructed. The participant had 10 seconds to view the object and to memorize the object location before the object and the dots were removed and the experimenter went behind a curtain to present an indication stick. Next, the participant verbally instructed the experimenter to match the near edge of the indication stick to the remembered near edge of the object location. Participants were then required to verbally indicate the demonstrative used on the instruction card to ensure they had attended to the instructions (see Figure 2.17).

There were two demonstratives (*this/ that)* and a neutral determiner (*the*), six locations and six objects. Participants were presented with six practice trials, after which 54 experimental trials were conducted (consisting of three trials of every term on every location: 3×3×6). The indication stick was presented at a distance of 10cm (counterbalanced to be further or nearer) from the actual location. Within the first 10 trials, there were three filler-trials in which the indication stick was presented at a distance of 20cm from the object location, to prevent the initial placement of the stick becoming a cue for the object location. Trials in which a participants' estimate of the object location was >10cm from the original location were repeated. At the end of the experiment, reaching distance was measured for each participant. Every participant could reach only the first three dots. The 'memory game' cover meant that participants were not aware that we were interested in the differences between demonstratives (confirmed during debrief).

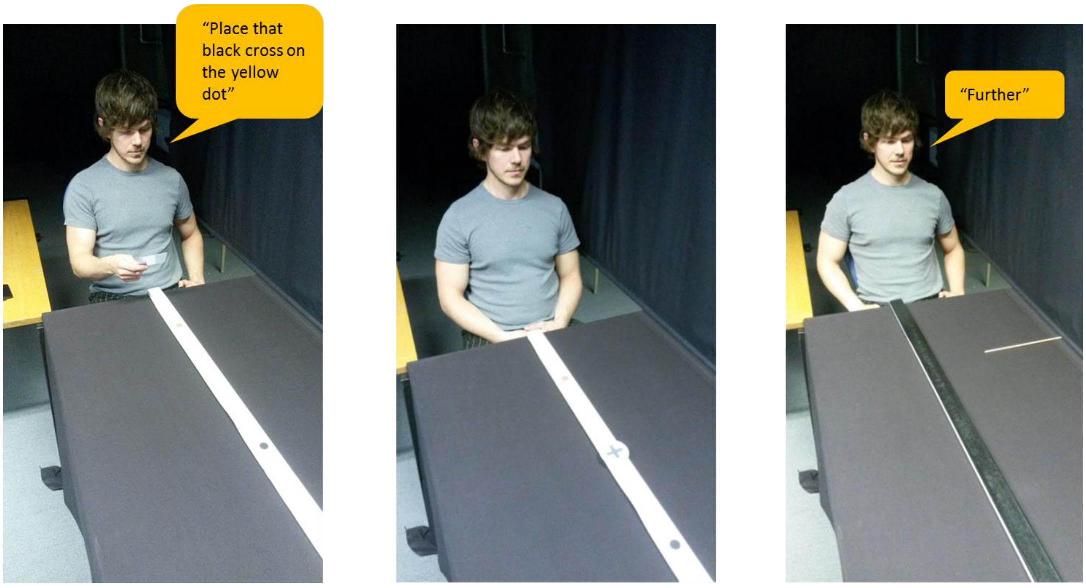


Figure 2.17. The participant reads out the instruction card, memorizes the object location, and subsequently instructs the experimenter to move the indication stick so it is aligned with where the edge of the object was. Finally, the participant recalls the demonstrative used on the instruction card they read out at the start of the trial.

2.2.2 Results

The memory data is presented in Table 2.2. A 3 × 6 (demonstrative × location) ANOVA was performed on the memory data (see Table 2.*3*). The memory data is expressed as the difference between the actual location and the memorized location, or the signed memory error.

Table 2.2. *Mean distance error memory errors (cm) in each condition by distance in Experiment 1.*

Experiment 1	condition	25 cm (SEM)	50cm (SEM)	75cm (SEM)	100cm (SEM)	125cm (SEM)	150cm (SEM)
Demonstratives *	this	0.48	1	2.06	0.74	3.6	2.7
		(0.55)	(0.59)	(0.73)	(1.06)	(0.63)	(1.68)
	that	0.42	1.21	0.99	1.94	1.7	3.93
		(0.55)	(0.68)	(0.96)	(0.92)	(1.06)	(1.02)
	the	0.32	0.63	0.95	2.64	2.17	3.09
		(0.4)	(0.5)	(0.67)	(0.74)	(0.87)	(1.12)

* mean distance error in cm, the 0 value represents the actual object location.

There was a strong main effect of location, $F(5,75) = 4.32$, $p < .01$, $\eta p^2 = .22$, but no effect of demonstrative ($p = .96$) nor an interaction ($p = .44$), indicating that there is no effect of the use of demonstrative pronoun on memory for object location. Alternatively, it is possible that participants did not process the demonstrative on the instruction card. In the current procedure participants read out the instruction card with the demonstrative, but we cannot be sure participants actually attended to the demonstratives as they were reading them on the instruction card. To be sure participants attended the differences in language we ran the study again with a minor alteration. In Experiment 2, participants engaged in exactly the same memory game paradigm, but an extra task is introduced. At the end of each trial, participants recalled the demonstrative as presented on the instruction card. As a cover story, we told participants the additional task increased cognitive load and make the memory game more challenging.

Table 2.3. *Results of the ANOVA on Memory for Object Location in Experiment 1.*

Source	**df and F value**	**MS**	**Significance**	**ηp²**
Demonstrative (D)	F(2,30) = .042	0.407	0.958	0.003
Location (L)	F(5,75) = 4.324	51.463	0.002	0.22
D x L	F(10,150) = .970	8.86	0.472	0.061

2.3 Experiment 2: The influence of demonstratives on spatial memory (with demonstrative recall task)

Experiment 2 is similar to Experiment 1, with the exception for an added demonstrative recall task. This allows us to test whether participants actually remember the different demonstratives used in the instructions.

2.3.1 Method.

2.3.1.1 Participants.

Thirty-six native English speaking students[1] were tested, receiving either course credit or payment for their participation. Stereoacuity was measured using the Randot Stereotest (Stereo Optical Inc. Chicago, USA). Two participants did not have a threshold of at least 40" and were excluded. Two additional participants were excluded because they had more than 10% incorrect answers in the demonstrative recall task. This left 32 participants, 9 males and 23 females, with an age range of 18 – 31 years old (*M* = 20.78, *SD* = 3.14).

2.3.1.2 Procedure and design

The procedure was similar to Experiment 1, but the demonstrative recall task was added. At the end of each individual trial participants recalled the demonstrative indicated on the instruction card. Failed recall of the demonstrative meant the trial was repeated at the end of the experiment[2]. However, if a participant could not recall the demonstrative at >10% of the trials, they were to be excluded from the analysis. Even though the demonstrative recall task emphasized the differences in language on

Table 2.4. *Mean distance errors (cm) in each condition by distance in Experiment 2.*

Experiment 2	Condition	25cm (SEM)	50cm (SEM)	75cm (SEM)	100cm (SEM)	125cm (SEM)	150cm (SEM)
Demonstratives*	This	0.64 (.38)	1.64 (.5)	2.53 (.63)	2.21 (.78)	2.64 (.77)	2.42 (.71)
	That	2.04 (.36)	2.18 (.45)	2.36 (.62)	3.3 (.76)	3.42 (.74)	4.33 (.82)
	The	1.7 (.28)	1.35 (.55)	1.1 (.63)	2.51 (.74)	2.28 (.75)	2.1 (.95)

* Mean distance error in cm, the 0 value represents the actual object location.

[1] Sample size is based on Coventry et al., 2014

[2] Trials were repeated if the recalled location was >10cm from the original location as well.

the instruction cards, participants were not aware of the aim of the study – this was confirmed during debrief.

2.3.2 Results and discussion

The memory displacement data – that is, the difference between the recalled distance and the actual distance between the recalled distance and the actual distance measured in centimetres – are displayed in Table 2.4. Note that the error is signed, a positive value indicates that an object was (mis)remembered as further away than it actually was (and a negative value was closer by). Generally speaking participants overestimated the location of an object, a phenomenon that has been found very early on in spatial cognition (Hollingworth, 1910).

A 3 × 6 (demonstrative × location) ANOVA was performed on the memory displacements (see Figure 2.18). The assumption of sphericity was violated in both the location and the demonstrative × location analysis. We therefore used the Greenhouse-Geisser correction for these analyses. There was a main effect of demonstrative, $F(2,62) = 6.68$, $p < .01$, $\eta p^2 = .18$ (see Table 2.5), showing an effect of language on memory for object location: follow up (LSD) tests showed significant differences between locations accompanied by the *that* ($M = 2.94$, $SE = .42$) compared to both the *this* ($M = 2.01$, $SE = .41$) and the *the* ($M = 1.84$, $SE = .47$) conditions (both *p's* < .01; see Figure 2.18). There was a marginal effect of location, $F(5,155) = 2.33$, $p = .08$, $\eta p^2 = .07$, suggesting that memory for object location deteriorated with distance, consistent with the previous studies. Importantly, there was no interaction between demonstrative and location, $F(10,310) = 1.4$, $p = .21$, $\eta p^2 = .04$[3]. The results therefore support the expectation model rather than the congruence model; *this* leads to more accurate object location memory than *that*, irrespective of the congruence between the specific demonstrative and location. We

Table 2.5. *Results of the ANOVA on Memory for Object Location in Experiment 2.*

Source	***df* and *F* value**	**MS**	**Significance**	**ηp²**
Demonstrative (D)	F(2,62) = 6.682	67.097	0.002	0.177
Location* (L)	F(5,155) = 2.325	59.266	0.078	0.07
D × L*	F(10,310) = 1.399	12.767	0.21	0.043

*Greenhouse-Geisser correction

[3] An exploratory analysis showed no interactions between demonstrative and gender, $F(2,60) = 1.456$, $p = .241$, $\eta p^2 = .046$ nor demonstrative, location, and gender, $F(10,300) = 1.398$, $p = .212$, $\eta p^2 = .045$.

next considered whether the same pattern of results might emerge with a different language manipulation involving possessives.

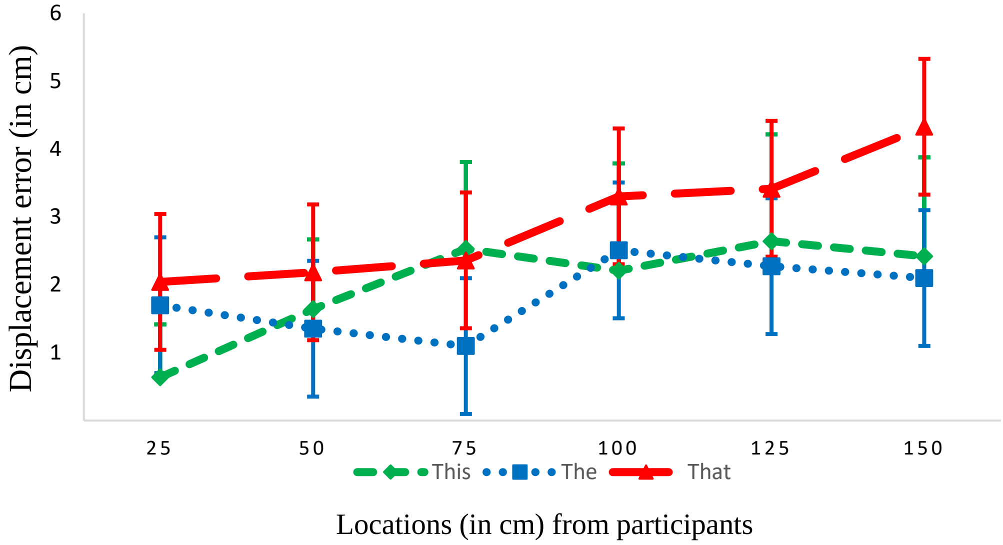


Figure 2.18. Results of Experiment 2, error bars are 95% confidence intervals.

2.4 Experiment 3. The influence of possessives on spatial memory

Some studies have shown that ownership improves memory for objects (S. Cunningham et al., 2008; Shi et al., 2011; Turk et al., 2015) and influences how people physically interact with objects (Constable, Kritikos, & Bayliss, 2011). For example, Cunningham et al. (2008) had a participant and a confederate sort cards with pictures of shopping items into their own basket or the other person’s basket. At the end of the trials participants completed a surprise memory test for the objects depicted on the cards. Participants had more accurate memories for self-owned objects than objects owned by a conspecific.

In another study, specifically targeting memory for object location, Coventry et al. (2014), found that object ownership affected memory for object location (and demonstrative choice). Using the memory game paradigm participants were given a set of coins in payment at the start of the experiment, and the coins placed at different to-be-remembered locations were either those coins or coins owned by the experimenter of the same denominations. Participants misremembered the conspecific’s coins as being further away than their own coins.

One of the problems with the ownership studies described above is that they cannot easily distinguish between an effect of the abstract concept of ownership and an effect of the possessives (*my/your*) used to indicate ownership during task instruction. For example in Coventry et al. (2014), coins were given to participants as participant payment at the start of the task to confer ownership, but language during the task itself involved by necessity the use of possessives (e.g. "Place your coin on the red dot") in order to disambiguate which coin was to be placed during the task. It is therefore unclear whether language indicating ownership (possessives), or the conceptual representation of ownership itself, or a combination of the two drive the effect of ownership. Here we investigated whether possessives have the same influence on memory for object location as did the demonstratives in Experiment 2, whether personal possessives *alone* are able to drive memory effects, and again, whether the expectation vs. congruence models offer a better account as to how possessives affect memory for object location.

2.4.1 Method.

2.4.1.1 Participants.

Thirty-nine native English speaking participants were tested, as in Experiment 2. Five participants were excluded as they did not score above the threshold of 40" (N=2), had more than 10% mistakes in the memory task (N=2) or could not reach the 50cm point (N=1). This left 34 participants; 14 male and 20 female, with an age range of 18 – 44 years old (M = 23.76, SD = 4.87).

2.4.1.2 Procedure and design

The procedure was similar to Experiment 2, with the exception that the demonstratives were replaced with possessives (*my*, *your;* the *the* condition was retained). In order to able to distinguish between an actual ownership effect and a language effect of possessives, participants did not own any of the objects, and all objects were used in all language conditions.

2.4.2 Results and discussion

The memory displacement data are displayed in Table 2.6. A 3 × 6 (possessive × location) ANOVA was performed on the difference (in centimetres) between the actual position of an object and the memorized position (See Table 2.7). There was a main effect of possessive, $F(2,66) = 8.25$, $p = .001$, $\eta p^2 = .2$, showing that objects in the *your* condition (M = 1.89, SE = .43) were remembered as being significantly further away than objects in both the *my* condition (M = .81, SE = .34)

Table 2.6. *Mean distance errors (cm) in each condition by distance in Experiment 3*

Experiment 3	Condition	25cm (SEM)	50cm (SEM)	75cm (SEM)	100cm (SEM)	125cm (SEM)	150cm (SEM)
Possessives*	My	0.2	1.05	0.67	1.08	0.63	1.25
		(.3)	(.55)	(.49)	(.39)	(.61)	(.72)
	Your	0.68	1.14	1.06	3.32	2.35	2.77
		(.27)	(.62)	(.7)	(.72)	(.77)	(.68)
	The	0.4	0.21	1.14	2.27	2.27	0.33
		(.3)	(.53)	(.47)	(.56)	(.72)	(.8)

* mean distance error in cm, a value of 0 represents the actual object location

and the *the* condition (M = 1.11, SE = .34), both p's < .01; see Figure 2.189). A significant effect of location was also found, $F(5,165) = 3.47$, $p = .01$, $\eta p^2 = .1$, showing that accuracy deteriorated as the objects were placed further away. These results are compatible with earlier studies on ownership. However, as all objects were used in all language conditions, there was no actual sense of ownership over any of the objects; the ownership was only marked by the use of possessives. This shows that possessives on their own affect memory for object location.

Additionally there was an interaction between possessive and location, $F(10,330) = 2.25$, $p = .03$, $\eta p^2 = .06$. However, as can be seen in Figure 2.19, the interaction pattern is consistent with the Expectation model and is not consistent with the Congruence account: there is no crossover between peri-personal and extra-personal space as would be expected in the congruence account. However, the effect of distance does seem to vary as a function of language. To further unpack this, we ran three one-way ANOVAs to test location effects by term, revealing that there was only a reliable peri-personal/extra-personal effect in the *your* and *the* conditions ($p < .05$). This effect was absent in the *my* condition ($p > .05$; see Figure 2.19). This

Table 2.7. *Results of the ANOVA on Memory for Object Location in Experiment 3.*

Source	***df* and *F* value**	**MS**	**Significance**	**ηp²**
Possessive (P)	F(2,66) = 8.246	62.857	0.001	0.2
Location* (L)	F(5,165) = 3.471	62.71	0.013	0.095
P × L*	F(10,330) = 2.254	23.383	0.034	0.064

*Greenhouse-Geisser

suggests that memory for owned objects maybe particularly enhanced, overriding any effect of peri-personal versus extra-personal space[4].

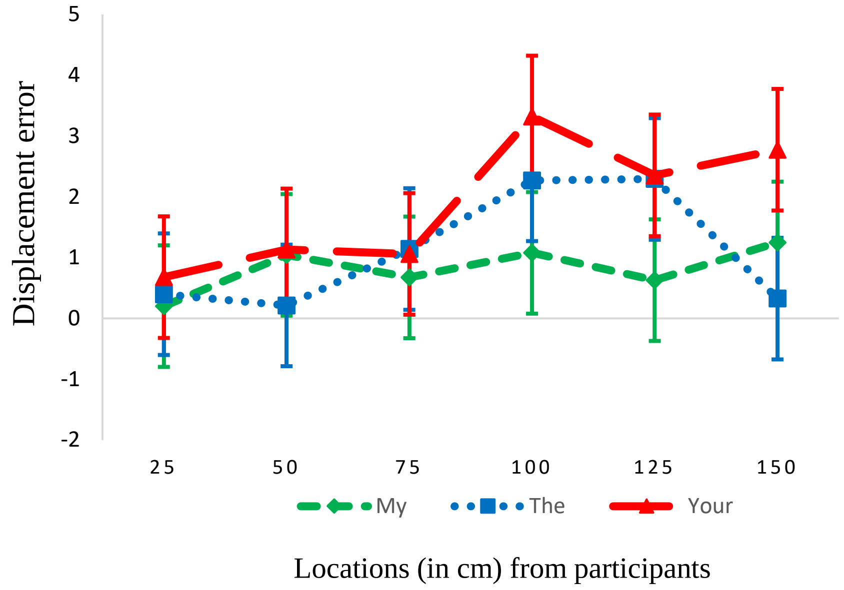


Figure 2.19. Results of Experiment 3, error-bars are 95% confidence intervals.

2.5 Experiment 4. The influence of attention on spatial memory (Eye-tracking method)

So far the results are consistent the Expectation model. However, it is important to also consider the possibility that the results might be caused the influence that language has on attention allocation (Allopenna et al., 1998; Tanenhaus et al., 1995). Visual attention is allocated preferentially to objects nearby, compared to objects further away (Garrido-Vásquez & Schubö, 2014) and longer fixation times lead to better memory performance (Huebner & Gegenfurtner, 2010; Naveh-Benjamin, 1987, 1988). Therefore, the predictions of memory error in the Expectation Model and what we have coined an “Attention Allocation Model” are similar, but differ in underlying mechanism. The Expectation Model predicts that memory for object location is a function of the language used to refer to the object (and the expectation of location associated with that language) combined with actual object location. The Attention Allocation model suggests memory for object location is a function of the fixation time and the object location. An alternative explanation for the results of Experiment 2 could therefore be the Attention Allocation Model, as

[4] An exploratory analysis showed no interactions between possessive and gender, $F(2,64) = 1.241$, $p = .296$, $\eta p^2 = .037$; location and gender, $F(5,160) = .928$, $p = .928$, $\eta p^2 = .006$, nor possessive, location, and gender, $F(10,320) = 1.406$, $p = .208$, $\eta p^2 = .042$.

driven by longer fixation times to objects paired with *this* versus *that* rather than differences in expectation values. In this experiment, we used eye tracking to measure participants' looking time during encoding. That allowed us to measure the time a participant fixated on objects in each language condition to see whether attention might account for the main effect of language reported above.

A second aim of Experiment 4 is to explore the connection between demonstratives and reference frames. As peri-personal space is the area within our grasp, this can be seen as an 'action space' in which objects are mapped onto an egocentric reference frame, compared to extra-personal space which may be mapped onto an allocentric reference frame (ter Horst, van Lier, & Steenbergen, 2011). If the language effect in the first two experiments is driven by the expectation raised by the specific use of language, then this expectation may result in different use of reference frames. We explored whether encoding object location onto an egocentric reference frame resulted in more searching behaviour along the sagittal line, to encode distance from the participant, compared to encoding onto an allocentric reference frame, which could result in more searching behaviour along the coronal line (see Figure 2.20). Results could help distinguish between models that predict solely an influence of egocentric representations on spatial memory versus 'two-system' models that predict a parallel egocentric and allocentric representations in object location memory (see Burgess, 2006).

2.5.1 Method.

2.5.1.1 Participants.

Nineteen participants were tested as in Experiment 2. Three participants were excluded from the analysis as the eye-tracker could not be calibrated. All participants showed a score of at least 40" in the depth perception task. This left 16 suitable participants for the analyses, 5 male and 11 female, with an age range of 18 – 22 (M = 19.19, SD = 1.17).

2.5.1.2 Procedure and design

The procedure was based on Experiment 2, but in this experiment, participants wore SMI eye-tracker glasses (30Hz binocular eye tracking glasses). For this reason, only 4 positions were used – two locations in peri-personal space and two in extra-personal space. (The location at 25cm was too close for the eye-tracker and the location 150cm too distorted). Before the experiment started, the glasses were calibrated using marks on the wall. After that, we validated the calibration four times throughout the experiment by having participants look at the four different locations on the table. The eye-tracking data were coded using semantic gaze mapping[5]. As the angle from the participant to the object was different for every location, the standard error in calibration of the eye-tracker image was slightly different per location. These distortions had to be accounted for in the semantic gaze mapping. Therefore, the coding was slightly less stringent for further locations compared to closer locations. For the furthest location, any fixation within an area of 6.5 cm (equivalent to the diameter of the object discs) around the object was marked as a fixation on the object. In the nearest location any fixation within an area of 3.25 cm (half an objects' diameter) was marked as a fixation on the object (see Figure 2.21b). Although the coding was adjusted for the different distances, this does not detract from the results as the adjustments were conducted across the different language conditions. The gaze mapping data were used in a 3 × 4 (demonstrative × location) design, investigating the differences in total fixation time (ms) on the object.

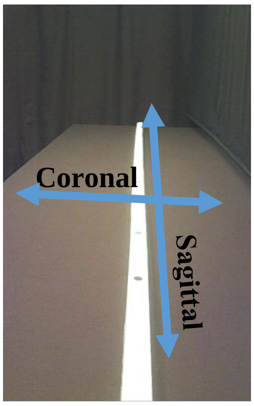


Figure 2.21a. Coronal and Sagittal plane of searching behavior

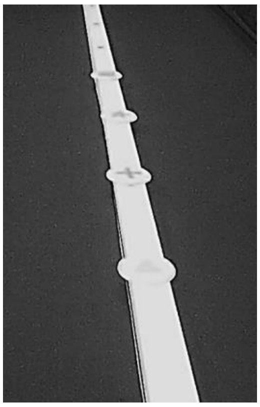


Figure 2.211b. Object area in semantic gaze mapping

[5] This involves the manual coding of video-based eye-tracking data, by which fixations are coded on a gaze map.

2.5.2 Results and discussion

The memory displacement data and the fixation times are displayed in Table 2.8. The memory data were analysed in a 3 × 4 (demonstrative × location) ANOVA.

A main effect of demonstrative was found, $F(2,30) = 5.77$, $p < .01$, $\eta p^2 = .28$, in which recalled distances for object location in the *that* condition ($M = 1.77$, $SE = .68$) were significantly further away than those in the *this* condition ($M = -.07$, $SE = .79$), $p < .05$. The *this* condition distances were also significantly closer than in the *the* condition ($M = 1.2$, $SE = .59$), $p < .05$ (see Figure 2.23). This replicates the result of Experiment 2. There was also a main effect of location, $F(3,45) = 9.69$, $p = .001$, $\eta p^2 = .39$, in which participants' accuracy deteriorated as locations were further away. There was no interaction effect between demonstrative and location, $F(6,90) = 1.61$, $p = .15$, $\eta p^2 = .1$, which means that the effect of language was the same across locations[6].

Table 2.8. *Mean distance errors (cm) in Memory data and fixation time (ms) for Gaze data, by distance*

Experiment 3	Condition	50cm (SEM)	75cm (SEM)	100cm (SEM)	125cm (SEM)
Memory data*	This	-1.54 (.71)	-0.28 (.78)	0.55 (1.59)	1 (.92)
	That	-0.32 (.79)	0.61 (.81)	2.73 (1.19)	4.06 (1.09)
	The	-0.88 (.49)	-1.06 (1)	2.29 (1.41)	4.77 (1.06)
Gaze data**	This	4179.2 (389.08)	5242.07 (406.67)	5525.92 (435.47)	5755.62 (436.49)
	That	4805.17 (348.06)	5216.23 (491.92)	5498.52 (445.73)	5403.98 (471.89)
	The	4826.967 (540.49)	5275.9 (468.50)	5621.11 (487.86)	5416.56 (606.53)

* mean distance error (cm)

** fixation time (ms)

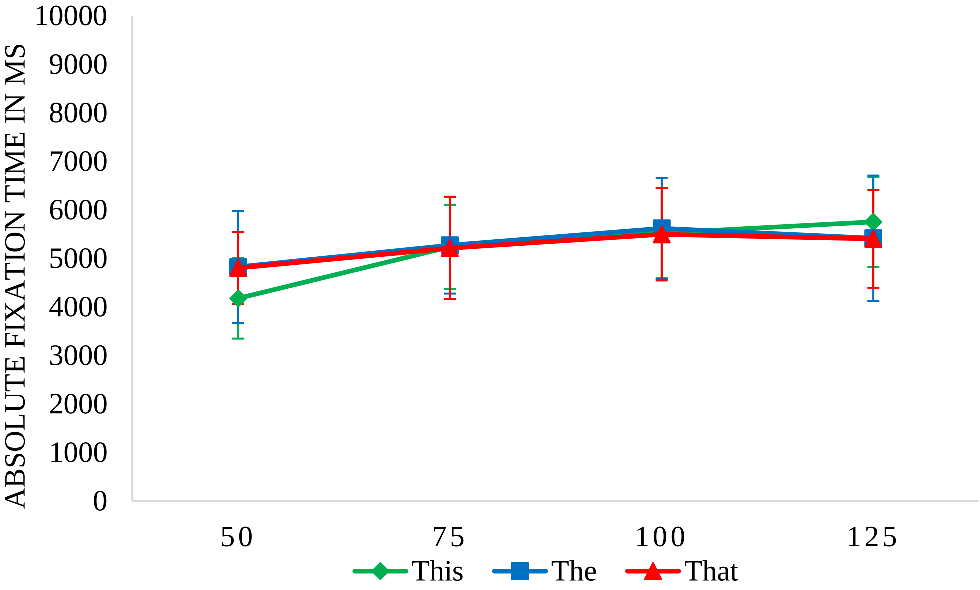
[6] An exploratory analysis showed no interactions between demonstrative and gender, $F(2, 28) = .735$, $p = .488$, $\eta p^2 = .05$, nor demonstrative, location, and gender, $F(6, 84) = 2.313$, $p = .077$, $\eta p^2 = .142$.

To see whether the language effects found were driven by a mechanism as hypothesized by the Expectation Model or the Attention Allocation Model, we next examined the gaze data collected during encoding. A 3 × 4 (demonstrative × location) ANOVA was used to analyse object fixation time. We found no effect of language, $F(2,30) = .13$, $p = .81$, $\eta p^2 = .009$ (*this*, *M* = 5175.70, *SE* = 345.44; *that, M* = 5230.97, *SE* = 257.65; *the*, *M* = 5285.14, *SE* = 416.76) suggesting that the language effect is not driven by differences in attention (See Figure 2.22). There was a location effect, $F(3,45) = 4.66$, $p < .01$, $\eta p^2 = .24$, showing that participants fixated longer on locations further away. However, this location effect could be due to the differences in coding caused by distance, as explained in Figure 2.211b. There was no interaction effect between demonstrative × location, $F(6,90) = .62$, $p = .71$, $\eta p^2 = .04$.

Table 2.9. *Results of the ANOVA on Memory and Gaze data in Experiment 4.*

	Source	***df* and *F* value**	**MS**	**Significance**	**ηp^2**
Memory data	Demonstrative (D)	F(2,30) = 5.771	57.81	0.008	0.278
	Location* (L)	F(3,45) = 9.69	288.487	0.001	0.392
	D × L	F(6,90) = 1.611	14.916	0.153	0.097
Gaze data	Demonstrative* (D)	F(2,30) = .133	261845.06	0.812	0.009
	Location (L)	F(3,45) = 4.655	9297167.3	0.006	0.237
	D × L	F(6,90) = .622	896923.92	0.712	0.04

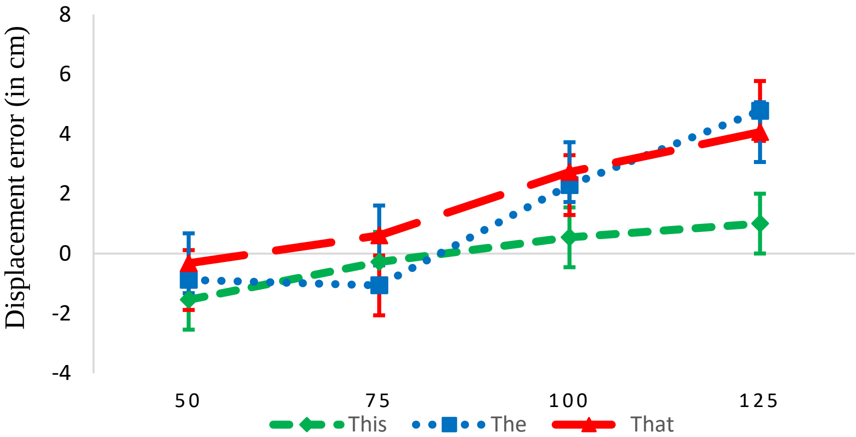
*Greenhouse-Geisser



Location (in cm) from participants

Figure 2.22. Gaze data from Experiment 3, based on summed fixation time on the objects in the respective language condition at the respective locations averaged per trial (in ms). Total looking time is 10000ms. Error-bars are 95% confidence intervals.

In a second analysis, we explored the connection between demonstratives and reference frames, and specifically to test whether people use different coordinate systems to remember object locations, based on spatial language. Fixations were coded as sagittal searching behaviour, if a sequence of two or more fixations fell within a range on either side (left/ right) of the white location stick, the range being



Locations (in cm) from participants

Figure 2.23. Behavioural data of Experiment 3, error-bars are 95% confidence intervals.

3.25cm from the sides for the closest location and 6.5cm from the furthest location. These distances were based on the size of an object on the respective location as represented on the screen (the actual objects had a diameter of 6.5cm). Fixations were coded as searching behaviour along the coronal line if a sequence of two or more fixations fell within a range above or below the object location. The range was half an objects' size for the closest location and one objects' size for the furthest location along the coronal plane. Fixations coded as fixations on the actual object were excluded from this analysis, so no fixation was used twice. After this coding, a ratio of fixations was calculated (coronal / (sagittal + coronal) (see Figure 2.21a.).

A 3 × 4 (demonstrative × location) ANOVA was performed. There was no main effect for demonstrative, F(2,30) = .15, p = .86, ηp^2 = .01; *this M* = .42, *SE* = .07; *that M* = .4, *SE* = .09; *the M* = .42, *SE* = .08), nor location, F(3,45) = .2.25, p = .13, ηp^2 = .13, nor an interaction with distance, $F(6,90)$ = .78, p = .59, ηp^2 = .05, suggesting that the language effect was not caused by differences in search-behaviour[7]. Based on these data, we cannot distinguish between different models for the use of reference frames in memory for object location.

2.6 General discussion

Over four experiments, we tested the influence that language has on memory for object location. The results of Experiments 2, 3, and 4 (where the language manipulation was explicit) showed that language affects memory for object location, with main effects of language in all three studies. In all experiments, the language manipulation was the same for participants. However, in Experiment 1, we did not test whether participants attended sufficiently to the language manipulations as they were reading them out, and effects of language on memory were absent in this experiment. By adding the demonstrative recall task at the end of every trial, we controlled whether participants remembered the language manipulation throughout the trials. Nevertheless, the fact that there is a difference in results between Experiment 1 and the other experiments merits some further discussion. The difference between Experiment 2, 3, and 4, in which we found an effect of language, and Experiment 1, in which we did not, is that we did not include a demonstrative

[7] In an exploratory analysis we did not find any interactions of gender with demonstrative, $F(2,28)$ = .1.035, p = .368, ηp^2 = .069, location, $F(3,42)$ = ..455, p = .715, ηp^2 = .031, demonstrative and location, $F(6,84)$ = .717, p = .637, ηp^2 = .049.

recall task in the latter. The demonstrative recall task adds the possibility to control whether participants attended to the demonstrative on the instruction card. However, this extra control could have led to demand characteristics, in which participants guessed the aim of the study and behaved accordingly. To test for this, we tested upon debrief whether participants knew what the study was about. None of the participants named the possibility that we tested for the influence of demonstrative on memory for object location, indeed most participants seemed to buy into our cover story that the demonstrative recall added cognitive load to the memory task.

The use of both demonstratives (Experiment 2 and 4) and possessives (Experiment 3) affected memory for object location. These results are consistent with previous studies showing an influence of language on memory for spatial relations (Feist & Gentner, 2007; Loewenstein & Gentner, 2005), and also with the effects of object knowledge on object location memory reported by Coventry et al. (2014). We have found robust effects of language on object location memory, together with an effect of distance on memory for object location. We first consider explanations for these results prior to implications for theories of language and memory more generally.

Three possible accounts of the influence of language on object location memory were set out prior to designing the present series of experiments: the congruence account, the expectation account, and an attention allocation model. The difference between the expectation and congruence models is the prediction of an interaction in the latter, and a main effect of language without an interaction in the former. The expectation model, proposed by Coventry et al. (2014), to explain object knowledge effects on memory, maintains that language elicits an expectation about an objects' location, which is concatenated with actual object location, leading to the prediction that the language effect should be the same for objects in near space and far space. In contrast, the congruence account predicts that memory should be more accurate for trials in which language is congruent with the object location, predicting an interaction between language and location; congruent trials (where *this/ that* are respectively combined with *near/ far* space) should be remembered more accurate than incongruent trials (in which *this/ that* are respectively combined with *far/near* space). In Experiments 2 and 4 there was no interaction, supporting the expectation account. In Experiment 3 (possessives) there was an interaction, but this effect was driven by the absence of a location effect for the *my* condition and not by congruence/incongruence contrasts. Thus, as a whole, results of the current experiments all support the expectation model.

Experiment 4 tested the third possibility that different types of language might result in different amounts of attention paid to objects/locations, with associated differences in memory performance. Put simply, the longer one spends looking at an object, the better one's memory for object location. The eye tracking data from this experiment revealed no differences in viewing time as a function of demonstrative. Participants did not present different searching behaviour based on different demonstratives, allowing us to rule out the attention allocation model.

Given that the results support the expectation model, there are three key issues that merit discussion. First, we can consider the relationship between the expectation model and memory models more broadly. Memory for object location is often taken to involve memory for the location in which an object is positioned, memory for the object itself, and a binding between object location and object (see for example Postma & Haan, 1996). Previously, Coventry et al. (2014), finding effects of object knowledge on memory for object location, argued against memory models that prioritize object location over object knowledge (e.g. the model of Jiang et al., 2000, who argued that location may act as an anchor to which object properties are attached). However, the effects of language on memory for object location and the previous effects of object knowledge reported are consistent with variants of *object file* theory (Kahneman & Treisman, 1984; Treisman & Zhang, 2006), in which object location may or may not be one of the features integrated in the file. Location features appear not to be bound to an absolute location but are defined relative to an abstract representation, which leads to memory errors (see Hollingworth & Rasmussen, 2010; Pertzov, Dong, Peich, & Husain, 2012). This focus on relative location can explain how spatial language can cue memory for object location, via the expectation of the objects' location relative to the speaker or another object. Wang and Spelke (2002) suggested that the human representational system depends in some way on language, by which humans can go beyond the limits of orientation systems as are found in animals. This influence of language skills may facilitate more flexible problem solving (Hermer-Vazquez et al., 2001). One can speculate that the advantage is that such a relative, dynamic system enables us to mentally process arrays in different contexts (e.g. a desk or the universe), using the same language and concepts.

Second, one needs to unpack in more detail how the expectation model works, and in particular, how the expectation values form and how they combine with the actual distance information available. Coventry et al. (2014) do not offer detail regarding this, but they assume that the expectation model works via the prediction of object location as a product of the history and context of past bindings between language, objects and location. For example, objects owned by people are more likely to be near people than equivalent objects owned by someone else. This likelihood is then used to predict future encounters with objects: if an object is owned, one would anticipate the object is nearer than if one does not own the object. This anticipation works similarly for the visibility and familiarity parameters Coventry et al. (2014) identified. Respectively, visible objects are usually closer than objects one cannot see, and familiar objects are more likely to be near us than unfamiliar objects. This anticipation-mechanism could be accommodated by correlational learning (see Pulvermüller, 2012) - the process in which neurons that fire together strengthen their connections and become more tightly associated (also known as potentiation). Such a mechanism has been implicated not only in mapping language to perception (Coventry et al., 2013), but also how one learns how words co-occur to form meaningful language structures during language learning (Saffran, Aslin, & Newport, 1996; Saffran, 2002). In the current experiments, language elicited the expectation. When participants memorised object location, they do not remember absolute coordinates, but they remember a construct of what they know about the object location, a construct build of the visual information and object knowledge they have about the object location. If a participant owned the object of which s/he was required to remember the location, the remembered object location is likely to be closer. The expectation raised by the object knowledge drives the difference in memory for object location. The predictive values would therefore be consistent with the concept of probabilistic learning, similar to how people learn language, as discussed at the start of this chapter (Andrews, Vigliocco, & Vinson, 2009; Clark & Sengul, 1978; Li et al., 2011; Saffran et al., 1996).

The Expectation Model can also be extended outside spatial language, both in cases where language is explicit during a task (as in our studies), but also in cases where language may not be explicit, but may nevertheless affect non-linguistic performance. For example evidence from colour perception (Jerome S. Bruner et al., 1951; Delk & Fillenbaum, 1965) has shown similar effects of the influence of object knowledge on memory. Object knowledge influences categorization of objects, so that participants judge objects within a category to have a more similar hue than objects between categories. For example, in an array of letters and numbers, participants judged symbols within a respective category to be more similar in colour than between categories, even if the two, between categories, target symbols were identically coloured (Goldstone, 1995). In another series of studies, it was shown that colour memory and colour perception judgements are influenced by the characteristic colour of an object. These object knowledge effects were stronger in objects with a high colour diagnosticity (e.g. yellow for a banana) than in objects with low colour diagnosticity (yellow for a lamp) (Belli, 1988; Jerome S. Bruner et al., 1951; Hansen et al., 2006; Tanaka & Presnell, 1999). These ‘top down’ effects on colour perception are consistent with the idea that knowledge of expected hue combines with actual hue information leading to categorization errors. Such an account merits further testing in this domain.

More broadly the Expectation Model is consistent with models of predictive coding (Clark, 2013; Friston, 2005; Lupyan & Clark, 2015). Clark (2013) describes a unifying model of perception and action, suggesting that the human brain continuously predicts a future state of the world. The brain prepares a response based on this prediction and only needs to process the *error signal*, the difference between the prediction and the updated visual input, once this new state emerges. This means that instead of processing or ‘creating’ a full response, the brain only needs to adjust the predicted response to be appropriate to the actual input. Friston described this as the *free energy principle*, suggesting that adjustments of the prediction error take less energy than surprise would (Friston, 2009). In the Expectation Model, the prediction is based upon learned associations between language, objects and locations (for example via statistical/ correlational learning), and it is these associations that can then reduce the work needed to process continually changing object location bindings on a moment to moment basis. In terms of spatial memory, the predictive coding account would suggest that by regenerating a spatial position of an object based on parts of information (visual input, verbal descriptions), is a more efficient use of resources than memorizing the absolute coordinates of an object.

A third issue that needs discussion is whether the effects of language operate at the level of encoding or retrieval. One possibility is that *this*, for example, actually activates peri-personal space more when looking at an object than *that*, and therefore that the memory differences are a direct result of differences in peri-personal space activation during encoding. Such a view is consistent with recent models of perception (e.g. Bar, 2009) that incorporate top-down predictions from memory as a mechanism during the act of perceiving. Alternatively, it is possible that the influence of language only occurs at retrieval, with remembered distances migrating in the direction of the remembered demonstrative/possessive. This would be consistent with effects found in verbal overshadowing (Alogna et al., 2014; Schooler & Engstler-Schooler, 1990) and eye-witness testimony (E. Loftus et al., 1978; E. Loftus & Palmer, 1974; McCloskey & Zaragoza, 1985) literature.

Furthermore, to test between alternatives suggesting the effect takes place at encoding vs. retrieval, it is possible to run neuroimaging studies to measure the degree of peri-personal space activation while viewing objects under different object knowledge and/or language conditions (see Coventry et al., 2014 for discussion). A second way one can get at this issue relates to memory decay: if the influence of language operates at retrieval, then the longer the time interval, the greater the effects of language there should be. We are currently exploring these possibilities.

In summary, we found a main effect of language (demonstratives and possessives) on memory for object location across experiments. We teased apart the predictions of three different models explaining this mechanism: the Expectation model, the Congruence model, and the Attention allocation model. Overall, results favoured the Expectation model, suggesting that the expected location of an object – cued by language (e.g., *this* for referents close by; *that* for referents further away) – and the actual location concatenate, leading to (mis)memory for object location.

However, languages vary widely in how they encode space (Munnich, Landau, & Dosher, 2001; E Pederson et al., 1998). Therefore, cross-linguistic research needs to explore whether languages other than English encode via the same Expectation Model, for example, when those other languages use a different demonstrative system. Since 37.4% of the world's languages employ a three-way demonstrative system (Diessel, 2006), it is important to evaluate which both how these contrasts in different demonstratives systems work, and how they they might influence memory. In Chapter 3, we will run a cross-linguistic study in which we compare English with Japanese, a language in which the three-term demonstrative system is thought to encode either egocentric distance (from the speaker), or person centeredness. In two experiments, we will test how demonstratives are used in Japanese and how demonstratives affect memory for object location. We also compare the Japanese data with data from two new English experiments, to be able to compare Japanese and English data using a similar, slightly altered procedure.

Chapter 3 – The Use and Influence of Japanese Spatial Language on Spatial Memory

3.1 Introduction

In Chapter 2, we saw effects of spatial demonstratives on memory for object location in speakers of English. Using the 'memory game' method devised by Coventry and colleagues (Coventry et al., 2014, 2008), participants were presented with objects positioned at various distances in front of them on a table and had to direct a stick to where the object had been positioned following its removal. Instructions indicating which object to place and where to place it - e.g. *place this/that/the red triangle on green dot* – manipulated the demonstrative presented just prior to object placement. The results across experiments indicated strong effects of language on memory for object location. *That* is associated with (mis)remembering objects further away than they actually were (relative to *this*). Critically, however, the data also allowed teasing apart between different possible accounts as to how language affects memory for object location.

Coventry et al. (2014) originally proposed the 'expectation model' in order to account for the effect of object knowledge on memory for object location, and object knowledge effects on choice of language to describe objects placed at different distances. In the expectation model, memory for object location is a concatenation of where the object is and where it is expected to be. For example, when presenting the word *that*, it is assumed in the model that *that* is associated with a larger expected distance value compared to *this*. When combined with actual distance, this model predicts a main effect of demonstrative on memory for object location. In contrast, Gudde et al. (2016) considered two alternative models as to how language may impact upon memory for object location. The 'congruence model' assumes that the mapping between the demonstrative and distance determines memory errors. When there is a mismatch between demonstrative and location (i.e., *this* when an object is out of reach; *that* when an object is within reach), one might expect that there is a greater error associated with memory than when the demonstrative and location match. More specifically, one would expect an effect of demonstrative when incongruence occurs, but not when the term and location are congruent (see Chapter 2, Gudde et al, 2016, for discussion). Finally, an alternative view of demonstratives is that the function of demonstratives is to direct the focus of attention of an interlocutor to a location. This effect could be between contrasting demonstratives, for example, the ‘schema of focus for demonstrative reference’ suggests that *this* elicits the highest focus of attention, compared to *that* (medium focus) and *it* (low focus) (Oh, 2001, based on Strauss, 1993; Strauss, 2002). An alternative, (e.g., Peeters, Hagoort, & Ozyürek, 2014), is that demonstratives encode relative location. In this case, one would expect that the implicit information conveyed by demonstratives would elicit longer looking times compared to a neutral determiner. As memory is more accurate the longer one attends to an object (Huebner & Gegenfurtner, 2010), these hypotheses would predict behavioural results similar to the expectation model, but also predict different looking times either between demonstratives, or between demonstratives and a neutral determiner.

The data in Chapter 2 (Gudde et al., 2016) favour the expectation model. In all experiments, there was a main effect of language on memory for object location, objects paired with *that* were misremembered as being further away than they actually were relative to objects paired with *this* or *the*. Crucially, demonstrative at encoding did not interact with distance or any other variables (as predicted by the congruence model), and demonstratives were not associated with paying differential attention to the placed object (as measured using eye tracking) as one might expect if demonstratives serve to draw attention to object locations or different territories (cf. Peeters et al., 2014).

Coventry et al. (2014) suggested that demonstrative systems across languages might reflect a common set of constraints, underlying non-linguistic perception and memory, which is explicit in the demonstrative systems of some languages but affects demonstrative usage in languages that do not employ these explicit distinctions. While the expectation model fits the data across ten experiments to date (seven in Coventry et al., 2014; three in Gudde et al., 2016 (Chapter 2)), the data thus far are for only one language, English. Given that there are 6,000-8,000 languages used around the world, it would be both ethnocentric and premature (Evans & Levinson, 2009) to assume that language affects memory in the same way across languages. While demonstrative systems may be as good a candidate as any for semantic universal status (Deutscher, 2005; Diessel, 1999, 2006), it is the case that demonstrative systems vary considerably cross-linguistically. In the most comprehensive typological analysis to date, Diessel (2005) sampled over 230 languages, noting that approximately 54% show a binary demonstrative contrast (English among them), and approximately 37% exhibit a three-way demonstrative system, including Spanish (*este, ese, aquel*) (Coventry et al., 2008) and Japanese (*kono, sono, ano*) (Diessel, 2014). The remaining languages possess three or more demonstratives, affected by a range of parameters, including ownership, visibility and slope (uphill/downhill). The focus of this chapter is to explore how language influences memory for spatial location across two languages – Japanese and English.

Talmy (1983) observed two different explanations of how spatial language can code different conceptual domains, dependent on the mutual exclusivity of spatial prepositions. If spatial prepositions define clear boundaries, spatial prepositions could refer to a classification of the relationship between objects, using necessary and sufficient characteristics. This would mean that a spatial preposition is mutually exclusive of other spatial prepositions. If however spatial categories do not clearly define boundaries, Talmy suggested spatial prepositions could be representative of relationships. This would mean that multiple spatial prepositions could be used for the same spatial relation (see also Hayward & Tarr, 1995). This is directly applicable in exploring the difference between English and Japanese – if Japanese demonstratives can encode multiple parameters; Japanese offers a higher degree of uncertainty in demonstrative interpretation.

The variation in demonstrative systems across languages opens up the possibility that different demonstrative systems pose different constraints on memory when demonstratives are presented with objects at encoding. In particular, when a language has three or more lexical items in their demonstrative systems, involving dimensions that are not mutually exclusive, it is quite possible that the expectation model does not apply. For example, Japanese has a demonstrative system that might be either distance based and/or person-centred. The latter construct assumes that demonstratives are affected by the position of a conspecific, in which there would be a distinct speaker space, hearer space and space far from both speaker and hearer (e.g., Diessel, 2014). If this is the case (and we test this empirically in Experiment 5), then this presents an interesting problem for speakers when one considers the expectation model. As egocentric distance and distance computed from the hearer (if the hearer is opposite the speaker, for example) give two different (conflicting) expectation values, speakers may have to choose between egocentric versus othercentric perspectives in order to retrieve a single expectation value for the model to be operable. Alternatively, it is possible that speakers with more complex demonstrative systems do not behave according to the expectation model, with language either not affecting memory for object location at all, or alternatively affecting memory in accordance with one of the other models we discounted for English (see Chapter 2, Gudde et al., 2016).

In summary, our main goals are two-fold. First, we aim to understand how a three-term demonstrative system works using the memory game procedure. Japanese is the language of choice as it might be distance centred, person centred, or a combination of the two. Second, we test the influence of Japanese demonstratives on memory for object location in order to test whether the expectation model fits the pattern of data for a language with a more complex demonstrative system. To preview the results, we show that Japanese is both distance and person centred. Interestingly, the position of a conspecific also influenced English participants. Even though English does not explicitly encode the position of an interlocutor, English and Japanese data were consistent. Although we appreciate more cross-linguistic research is necessary before actual claims can be made, this tendency could suggest that there are universals underlying demonstrative use.

3.2 Experiment 5. Spatial demonstrative use in Japanese

Japanese uses three demonstratives: *kono*, *sono*, *ano* (built from the prefixes ko-, so-, a-), but it is unknown how these demonstratives map onto the world. They could encode the *distance* from a speaker, or the *territory* in which an object is located (near the speaker, near the hearer, or far from both). However, they might encode both distance and territory, as a function of the spatial configuration of interlocutors, or even contrast non-deictic parameters. We first discuss different possible models regarding the function of Japanese demonstratives before testing between them.

Model 1: The *distance model.* This suggests that different demonstratives, as in English and Spanish (see Coventry et al., 2008), are used based on the referents' (egocentric) distance from the speaker. *Kono* refers to an entity close to the speaker, *sono* is used for an entity at a greater distance, and *ano* is used to refer to an entity at a greater distance still (see Figure 3.24) (e.g., Ootsuki (1889), as read in Nakamura, 2012).

Model 2: the *territory model*. This suggests that the Japanese demonstrative system encodes a referents' distance based on interlocutors' territories: encoding the speakers' (*kono)* or hearers' (*sono)* territory, or that the object out of both interlocutors' territories (*ano*), (Sakuma, 1936; Sakata, 1971, as cited by Hasegawa, 2012; Hattori, 1992; Niimura & Hayashi, 1994). The Japanese demonstrative system is often referred to as 'person-centred' assuming this underlying model (i.e., Diessel, 1999).

Model 3: *Dual system (distance plus territory).* In this model, speakers encode both distance and territory using demonstratives, choosing distance or territory as a function of the spatial configuration of interlocutors. When interlocutors are located opposite each other, demonstratives encode territory. But when they share a territory, demonstratives encode egocentric distance of a referent from their shared territory, as in the "theory of territory of information" (Kamio, 1994). This theory argues that psychological proximity between speaker/hearer and the referent determines the choice of one model over the other (cf. Hoji, Kinsui, Takubo, & Ueyama, 2003). Furthermore, an adapted version of this model (Yoshimoto, 1992, 1986) suggests a contrast between personal space (directly surrounding the speaker) and interactional space (the space around interlocutors) space (cf. Wilkins, 1999). In this model, *sono* is used to refer to reference in interactional space (irrespective of distance from speaker or hearer), instead of hearer's territory.

Model 4: *Double binary system.* Mikami (1970, 1992, as cited by Hasegawa, 2012) proposed that the three demonstratives contrast different parameters. In this model, space is carved up based on territory, speaker's (*kono*) and hearer's (*sono*), but *ano* cannot be used in this situation. When the speaker and hearer face the same direction, they perceive themselves together in opposition to others. The joined territory of speaker and hearer is expressed by *kono*, that of others by *ano*. In this model, there are two different contrasts in which *kono* and *sono* are opposites, and *kono* and *ano*, but in which *sono* and *ano* are never contrasted (Hasegawa, 2012).

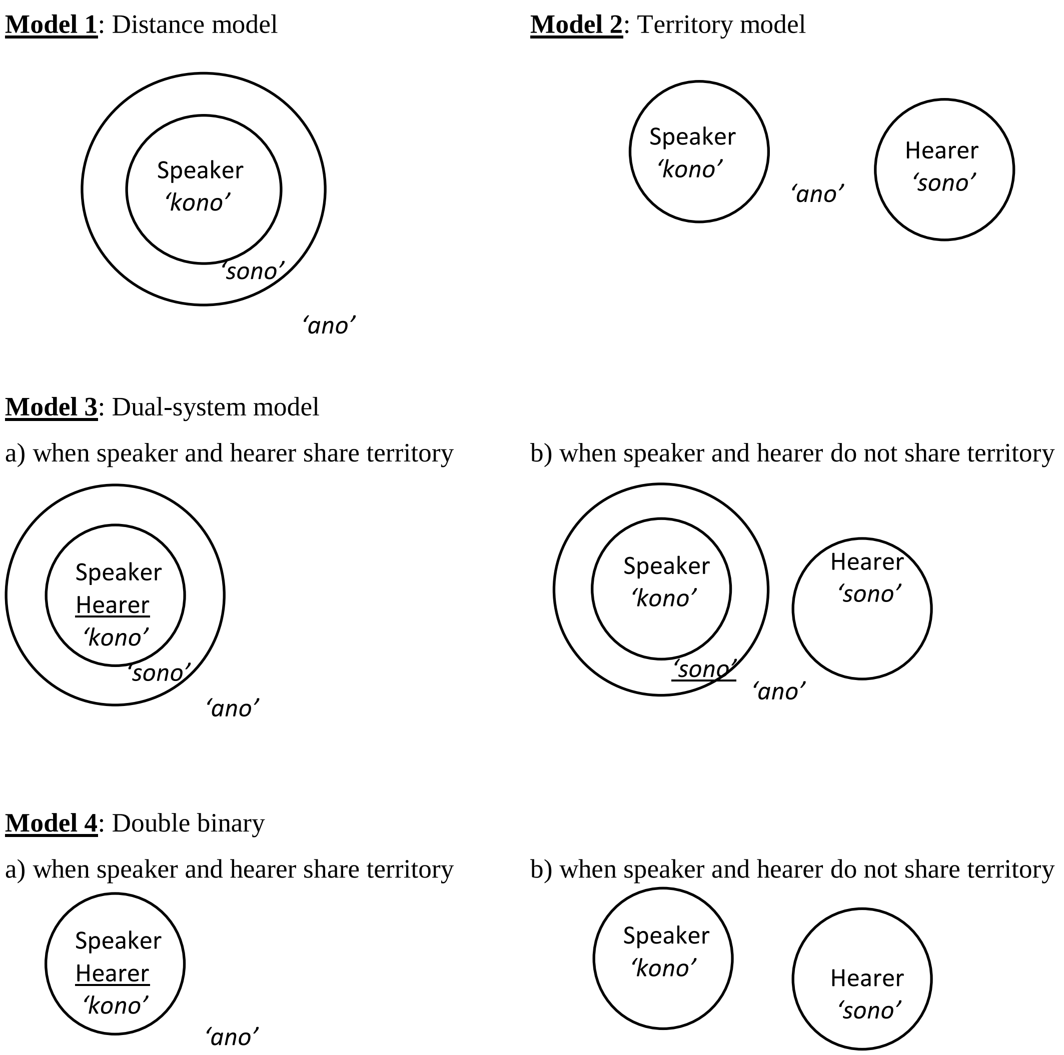


Figure 3.24. The four different models of Japanese demonstratives. In Model 1, demonstratives encode distance from speaker (irrespective of the position of the conspecific), whereas in Model 2 demonstrative is territory-based, encoding the location of an object in relation to hearers' or speakers' territory. Model 3 and 4 both include two possible spatial configurations: when speaker and hearer share territory and when they do not. Model 3 is the dual-system model, in which all three demonstratives can encode both distance and territory, as a function of the spatial configuration of interlocutors – when they share territory (a), demonstratives encode egocentric distance; when interlocutors are facing each other (b), demonstratives encode territory. Model 4 is the double binary model, in which distance (a) and territory (b) can both be contrasted using demonstratives, but in which only 2 demonstratives can be used in each contrast.

Experiment 5 set out to test between these different models of Japanese demonstrative use. We adapted a version of the memory game paradigm to manipulate the position of a conspecific, with locations either within the participant's or the conspecific's territory (and therefore out of conspecific's or participant's peri-personal space respectively), or far from both participant and conspecific. Note that this experimental setup can discriminate between the four different models. Since

Model 1 and 2 are mutually exclusive, the encoding of either distance (Model 1) or territory (Model 2) falsifies the other. If both distance and territory are encoded, we can tease apart Models 3 and 4 by examining demonstratives use in specific trials. If only two demonstratives are used at each of the conspecifics' positions (*kono* and *sono* when territory is shared; *kono* and *ano* when territory is not shared), this is inconsistent with Model 3. If participants us all three demonstratives at both the conspecifics' positions, this falsifies Model 4.

In Experiment 6, we tested a group of English speakers (on their use of English demonstratives), allowing us to compare Japanese and English speakers. Previously, Coventry et al. (2008) found that the use of demonstratives in both English and Spanish is affected by distance from a speaker (peri-personal versus extra-personal space), but not the position of a conspecific. However, Coventry et al. manipulated the position of the conspecific between participants, with relatively small sample sizes, which may have reduced the chances of finding an effect of this variable. In Experiment 6, we manipulate the position of the conspecific within participants for English as well as Japanese.

3.2.1 Method.

3.2.1.1 Participants. Thirty-three native Japanese speakers were tested at the University of East Anglia, receiving payment for their participation. One participant did not engage in the experiment[8] and performed poorly on the memory probe trials, and was therefore excluded from analysis. This left 32 participants (25 female) with an age range of 19-49 years old (M =26.5, SD = 8.25).

3.2.1.2 Materials. Six distinguishable, different coloured shapes on plastic discs (e.g., yellow triangle/blue heart), 6.5 cm in diameter, were placed on 6 different locations. The locations were spaced equidistantly along a midline from the participants' edge of a large conference table (L = 325, W = 122cm), starting at 25cm from the participant up to 300cm (Coventry et al., 2008). We used six locations (see **Error! Reference source not found.**), divided over three spaces, b ased on the territory and distance models describing Japanese demonstrative systems: peripersonal locations at 25cm and 50cm from the participant, in the participants' territory and closest to the participant, extra-personal for both

[8] Upon debrief, one participant stated to have produced demonstrative expressions randomly.

participant and conspecific at 150cm and 175cm, medium far in distance coding and out of both speakers' and conspecifics' territory in any condition, and furthest from participant at 275cm and 300cm, but peripersonal to the conspecific when the conspecific took position opposite the participant. For each participant it was confirmed at the end of the experiment that they could reach (only) the closest two dots. The table was covered with a black cloth to eliminate any spatial cues other than the location marks.

3.2.1.3 Procedure and design. From the moment the Japanese participants entered the lab up until debrief, all communication was in Japanese. Participants were seated at the table and told by the experimenter (who was the conspecific in the experiment) that they would play a 'memory game' and they were in the 'language condition'. This meant that, following object placement, participants were asked to use both body language and verbal language to point (but not touch) and name each object, using a combination of three words: "[demonstrative (Japanese [*kono/sono/ano*]/ English [*this/that*]) [colour] [shape]", for example ["この赤い丸" or "This red circle"]. Participants could only use this three-word structure, so that every participant in the 'language condition' experienced the same amount of verbal coding. They were encouraged to use all demonstratives across the trials. To maintain the 'memory experiment' cover, participants were asked to recall the most

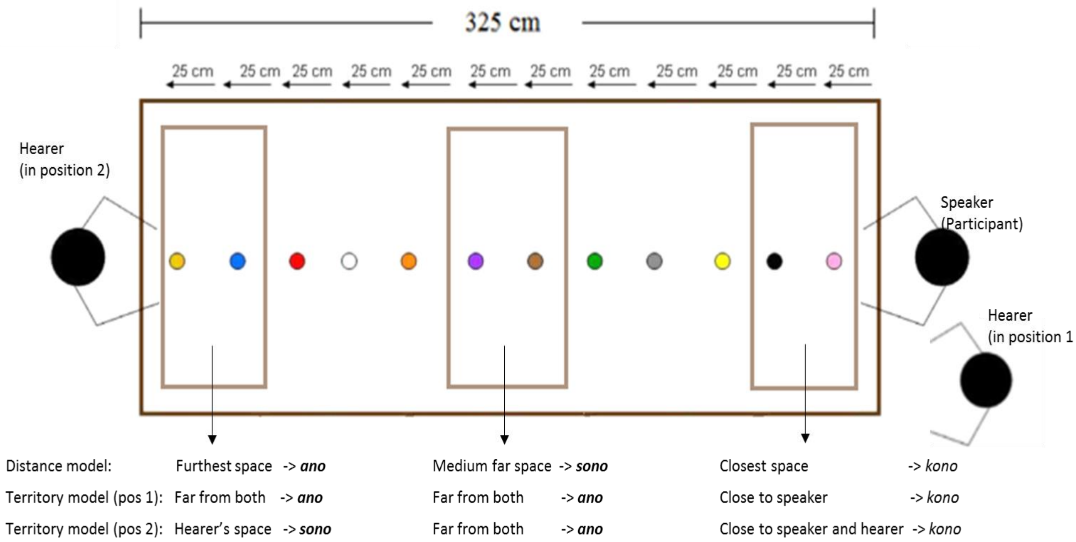


Figure 3.25. Schematic representation of the experimental setup. The participant (speaker) is on the right of the table. The conspecific (hearer) is positioned next to (position 1) or opposite (position 2) the participant. The three spaces are marked with a brown square, accompanied by the definition and predicted demonstrative according to the distance and territory (both in position 1 and position 2) model.

recent location of 4 of the objects, on six different occasions throughout the experiment (a total of 24 memory trials).

On each trial, the conspecific instructed placement of one of the objects on one of the locations. The instructions contained the following information: “[Person (I [conspecific]/ You [participant])] place the [object] on the [location] dot” (e.g., “あなたは緑の点に赤い丸を配置します” or “You place the red circle on the green dot”).

The experiment employed a 2 (agent: participant, conspecific) x 2 (position of conspecific: beside participants, opposite participant) x 6 (location) repeated measures design (with 3 trials per cell of the design, totaling 72 trials).

3.2.2 Results and discussion.

The percentages of demonstrative use in Japanese are shown in Table 3.10. *Kono* was almost exclusively used when the object was placed at the nearest two locations, irrespective of the position of the conspecific. At 25cm and 50cm (within participants’ peri-personal space), *kono* was used in 96% of the trials. The use of *kono* in all other locations was extremely low (1.37% of the trials). Therefore we eliminated *kono* and the first two locations from the analysis, which allowed us to calculate percentages of *ano* use (*ano*/(*sono*+*ano*)* 100). A 2 (agent: conspecific/participant]) × 2 (position of conspecific: next or opposite participant]) × 4 (locations) x 2 (gender) ANOVA was performed on these percentages. We included gender as a variable in this and the later experiment for two reasons. First, there is some evidence of gender differences in memory for objects and object location (see Voyer, Postma, Brake, & Imperato-McGinley, 2007, for a review), with some evidence also that women have a tendency to use *this* in English more than men (Coventry et al., 2014; Expt. 7). Second, as Experiment 1 established that

Table 3.10. *Percentages of Japanese demonstrative use at the different distances by conspecific position.*

	Position	25cm	50cm	150cm	175cm	275cm	300cm
Kono	Next to	98.96%	90.63%	2.08%	0.52%	0.52%	0.52%
	Opposite	98.96%	95.31%	4.69%	1.56%	0.52%	0.52%
Sono	Next to	1.04%	9.38%	38.54%	30.21%	8.33%	8.85%
	Opposite	1.04%	4.69%	58.85%	53.65%	44.27%	43.75%
Ano	Next to	0.00%	0.00%	59.38%	69.27%	91.15%	90.63%
	Opposite	0.00%	0.00%	36.46%	44.79%	55.21%	55.73%

Japanese demonstratives are affected by territory (my space, your space, outside both spaces), we thought that gender might potentially be of relevance given literature suggesting that men have more of a sense of separateness from others, which is suggested to be a quest for independence or part of a competition for status, compared to women (cf. Baumeister & Sommer, 1997; Cross & Madson, 1997). Moreover, Hofstede (1980, 1998) has shown that Japan as a society scores highest on masculinity (out of 50 countries tested), and therefore one might expect more marked gender differences in Japanese than in English.

In the analysis of location, and the interactions of agent × location, position × location, agent × position of conspecific × location the assumption of sphericity was violated, so we used the Greenhouse-Geisser correction. There was a main effect of position of the conspecific, $F(1,30) = 11.737$, $p = .002$, $\eta p^2 = .281$. Overall, *ano* was used more than *sono* (*ano* use $M = 74.4\%$, $SE = 5.6\%$, 95% CI [63%, 85.8%]) when participant and conspecific were side-by-side, with approximately equal use of *ano* and *sono* when the conspecific was at the opposite end of the table ($M = 47.9\%$, $SE = 8.8\%$, 95% CI [30%, 65.8%]) (see also contrasts in Figure 3.26). There was also an effect of location, $F(3,90) = 8.413$, $p < .001$, $\eta p^2 = .219$, showing that the use of *ano* increases the further away from the participant compared to the use of *sono* (see Table 3.10). There was no effect of agent, $p = .16$, $\eta p^2 = .064$. Nor were the

Table 3.11. *Results of the ANOVA on the ratio-use of 'ano' in Experiment 5.*

Source	***df* and *F* value**	**MS**	**Significance**	**ηp²**
Agent (A)	F (1,30) = 2.040	880.734	.164	.064
A × gender	F (1,30) = .512	221.01	.48	.017
Position (P)	F (1,30) = 11.737	61471.466	.002	.281
P × gender	F (1,30) = .431	2255.736	.517	.014
Location* (L)	F (3,90) = 8.413	37380.856	.004	.219
L × gender	F (3,90) = .295	539.925	.829	.01
A × P	F (1,30) = 2.666	894.001	.113	.082
A × P × gender	F (1,30) = 4.064	1362.751	.053	.119
A × L*	F (3,90) = .102	26.212	.922	.003
A × L × gender	F (3,90) = .883	170.2	.453	.029
P × L*	F (3,90) = 1.859	1207.114	.172	.058
P × L × gender	F (3,90) = .698	253.996	.556	.023
A × P × L*	F (3,90) = .885	201.74	.437	.029
A × P × L × gender	F (3,90) = .367	69.679	.777	.012

*Greenhouse-Geisser

interactions between, agent × position, p = .113, ηp^2 = .081; position × location, p = .170, ηp^2 = .059, agent × position × location, p = .437, ηp^2 = .029.

There was one marginally significant interaction involving gender – an agent × position of conspecific × gender interaction, $F(1,30)$ = 4.064, p = .053, ηp^2 = .119 (See Figure 3.26). Follow-up LSD tests showed a reliable effect of agent for male participants, when participant and conspecific were seated side-by-side (Experimenter places, M = 61.9%, SE = 10.5%, *95% CI* [40.5%, 83.3%], Participant places, M = 73.8%, SE = 10%, *95% CI* [53.5%, 94.1%]), but not when the conspecific was at the opposite side of the table (Experimenter places, M = 47.6%, SE = 15.6%, *95% CI* [15.8%, 79.4%], Participant places, M = 45.2%%, SE = 15.9%, *95% CI* [12.8%, 77.6%], p = .02). In contrast, there was no effect of agent for female participants in either position of the conspecific.

Returning to the four models we summarised earlier, the data help us tease apart the validity of each of them. When the participant and conspecific are sitting side by side, *ano* is preferred to *sono* when the object is out of reach, with *ano* increasingly preferred the further away the object is from the participant. On its own, one might argue that Japanese is distance based, with *kono* preferred for peri-

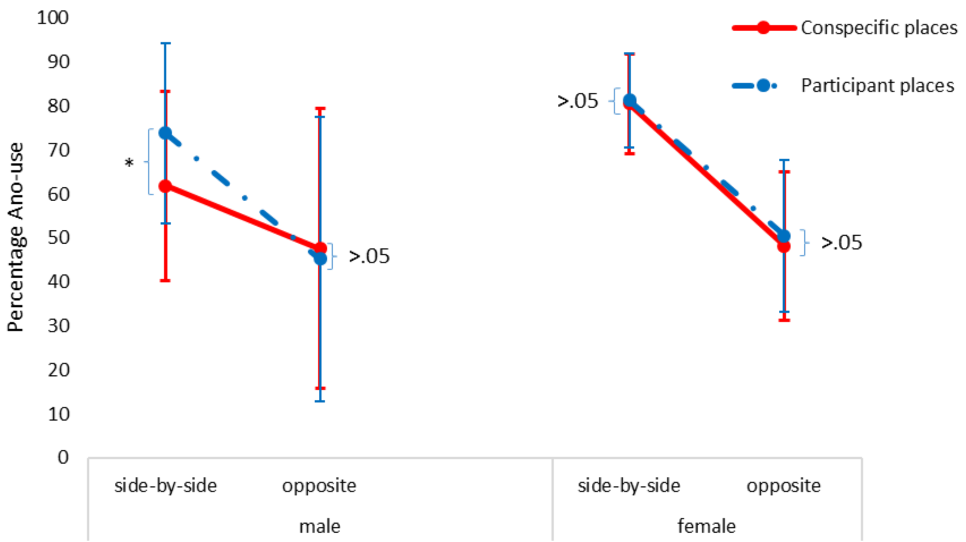


Figure 3.26. The marginal interaction between agent × position of conspecific × gender. Female participants show the same percentage of *ano* use when the experimenter places the object or when they place it themselves, whereas men use *sono* more when the experimenter places. Error bars are 95% confidence intervals, contrasts are LSD.

personal space, and *sono* and *ano* for extra-personal space, with an increased preference for *ano* over *sono* as distance increases (distance model). In that regard, Japanese appears similar to Spanish when the participants are sitting side-by-side. However, there was a main effect of position of conspecific, and an interaction between agent, position of conspecific, and gender, showing that the Japanese demonstrative system is not just distal (distance model), but is also affected by territory (territory model). When the participant and conspecific are opposite each other, the effect of distance from the participant is smaller compared to when they are sitting side by side. The use of *sono* is larger when the two interlocutors are opposite each other, with approximately equal use of *sono* and *ano* when the objects are placed within the conspecifics' peri-personal space and in the participants' extra-personal space. The results falsify Model 1, Model 2, and Model 4.

The double binary distinction, Model 4, in which Japanese demonstratives function via two contrasts, *kono* can contrast territory with *sono* and distance with *ano*¸but *sono* and *ano* do not form a contrast. Most participants used all three demonstratives throughout the experiment, falsifying this double binary model. Our data are consistent with the third model, the dual-system model (Niimura & Hayashi, 1994; Okazaki, 2011; Takahashi, 1992), in which Japanese demonstratives can encode both distance and territory, as a function of the spatial configuration of interlocutors. However, the use of *sono* was higher when the interlocutor sat opposite the participant, regardless of whether the object was placed within the interlocutor's reach (note that all position differences in Figure 3.26 are highly significant). This suggests that our data fits the dual-system model with the adaptation of a *interactional* space, in which the space between interlocutors is coded with *sono*, best (Yoshimoto, 1992, 1986).

3.3 Experiment 6. Spatial demonstratives in English

In Experiment 5, we found main effects of location and position. The effect of location replicates earlier studies (Coventry et al., 2008; 2014), but the effect of position of a conspecific had not been tested before. In this chapter, we aim to explore potential cross-linguistic differences. Therefore, Experiment 6 employs the same experimental procedure as we used in Experiment 5, but this time with English participants.

3.3.1 Method.

3.3.1.1 Participants. Thirty-five native English-speaking students were tested, receiving either course credit or payment for their participation. Two participants were excluded as their produced expressions were not in line with task instructions[9], one participant chose to withdraw. This left 32 participants for the analysis (22 female), with an age range of 18-67 years old (M = 24.83, SD = 11.32).

3.3.1.2 Materials. All materials were similar to Experiment 5.

3.3.1.3 Procedure and design. The procedure was similar to the procedure in Experiment 5, the only difference being that this experiment was run in English.

3.3.2 Results and discussion.

We analysed the English data using a 2 (agent [conspecific/ participant]) × 2 (position of conspecific [next or opposite participant]) × 6 (locations) x 2 (gender) ANOVA (see also Table 3.12). As participants could only use *this* or *that*, the use of *this* is the complement of *that.* As in Japanese, we therefore calculated the percentage use of *that*. The assumption of sphericity was violated for location and the interaction between position × location. For these analyses we used the Greenhouse-Geisser correction. There was a significant effect of agent , F(1,30) =

Table 3.12. *Ratio of 'that' use for each space, per gender, actor, and position of the conspecific.*

			25cm (SEM)	50cm (SEM)	150cm (SEM)	175cm (SEM)	275cm (SEM)	300cm (SEM)
Male	Conspecific places	next to	20.00% (10.18%)	23.33% (11.17%)	73.33% (8.31%)	70.00% (15.28%)	93.33% (4.44%)	96.67% (3.33%)
		opposite	16.67% (11.39%)	36.67% (13.56%)	70.00% (10.48%)	86.67% (7.37%)	90.00% (5.09%)	90.00% (7.11%)
	Participant places	next to	10.00% (7.11%)	10.00% (5.09%)	60.00% (12.96%)	63.33% (13.56%)	83.33% (8.96%)	80.00% (10.18%)
		opposite	10.00% (7.11%)	10.00% (7.11%)	53.33% (13.33%)	73.33% (11.97%)	80.00% (10.18%)	73.33% (11.97%)
Female	Conspecific places	next to	6.06% (2.80%)	9.10% (3.90%)	74.20% (7.30%)	87.90% (4.70%)	98.50% (1.50%)	93.90% (3.60%)
		opposite	10.60% (4.00%)	16.70% (6.10%)	78.80% (6.80%)	95.50% (2.50%)	89.40% (6.40%)	92.40% (4.90%)
	Participant places	next to	9.10% (4.50%)	9.10% (3.20%)	72.70% (7.20%)	89.40% (4.60%)	93.90% (2.80%)	98.50% (1.50%)
		opposite	6.10% (2.80%)	9.10% (5.50%)	77.30% (6.70%)	95.50% (3.30%)	93.90% (4.70%)	92.40% (4.90%)

[9] Upon debrief, one participant stated to have used demonstratives counter-intuitively on purpose, another participant said to have strictly alternated demonstratives between trials instead of using them normally.

6.489, p = .016, ηp^2 = .178. Participants used that more often when the experimenter placed (M = 63.3%, SE = 2%, 95% CI [59.1%, 67.5%]) compared to when the participant placed (M = 56.4%, SE = 2.4%, 95% CI [51.5%, 61.3%]). There was also a significant main effect of location, F(5,150) = 99.567, p < .001, ηp^2 = .768; people mostly used this in the closest two, reachable, locations. Between the 50cm and 150cm point, the border between peri-personal and extra-personal space, the use of that quadruples from 15.5% at 50cm to 70.0% at 150cm, see also Figure 3.27.Figure 3.

The results indicate that there is no effect of the position of the conspecific (*p* = .53, ηp^2 = .013), nor gender (*p* = .143, ηp^2 = .070) on the use of English demonstratives in the memory game task. There was an interaction between agent and gender, *F*(1,30) = 5.576, *p* = .025, ηp^2 = .157, in which male participants used *that* less frequently when they placed the object themselves (*M* = 50.6%, *SE* = 4%, 95% CI [42.5%, 58.6%]) compared to when the experimenter placed (*M* = 63.9%, *SE* = 3.4%, 95% CI [57%, 70.8%]); there was no difference for women. There was also an interaction between location and position, *F*(3.382, 101.471) = 2.649, *p* = .046, ηp^2 = .081, in which participants used *that* less frequently for furthest locations when the objects were placed in near space of the conspecific. There was no interaction between agent and position, *p* = .318, ηp^2 = .033, agent and location, *p* = .326, ηp^2 =

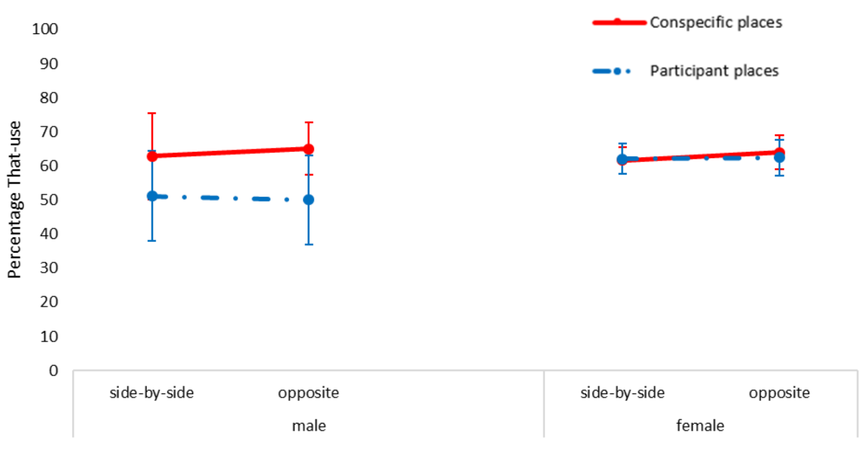


Figure 3.27. The interaction between agent × gender. Female participants show the same percentage of 'that' use when the experimenter places the object or when they place it themselves, whereas men use 'that' more when the experimenter places. Error bars are 95% confidence intervals, contrasts are LSD.

.038, nor a three-way interaction between agent × position × location, $p = .709$, $\eta p^2 = .019$.

Table 3.13. *Results of the ANOVA on use of 'that' in Experiment 6.*

Source	***df* and *F* value**	**MS**	**Significance**	**ηp²**
Agent (A)	F (1,30) = 6.489	.79	.016	.178
A × gender	F (1,30) = 5.576	.679	.025	.157
Position (P)	F (1,30) = .403	.014	.53	.013
P × gender	F (1,30) = .061	.002	.807	.02
Location* (L)	F (2.560,76.813) = 99.567	29.161	<.001	.768
L × gender	F (3,90) = 2.116	.317	.067	.066
A × P	F (1,30) = 1.033	.03	.318	.033
A × P × gender	F (1,30) = .062	.002	.805	.002
A × L	F (3,90) = 1.171	.023	.326	.039
A × L × gender	F (3,90) = .633	.012	.675	.021
P × L*	F (3.382,101.471) = 2.649	.136	.046	.081
P × L × gender	F (5,150) = .613	.021	.69	.02
A × P × L	F (5,150) = .589	.016	.709	.019
A × P × L × gender	F (5,150) = .358	.01	.876	.012

*Greenhouse-Geisser

These results replicate the results of Coventry et al. (2008, 2014). There were effects of distance and agent on demonstrative use in which the use of *this* diminishes when the object is placed further from the participant, and *this* is preferred when the participants places. However, we also found interactions between location and position and between agent and gender, showing that participants in English consider position of a conspecific in demonstrative use, and evidence that the agent effect occurs for male participants and not females.

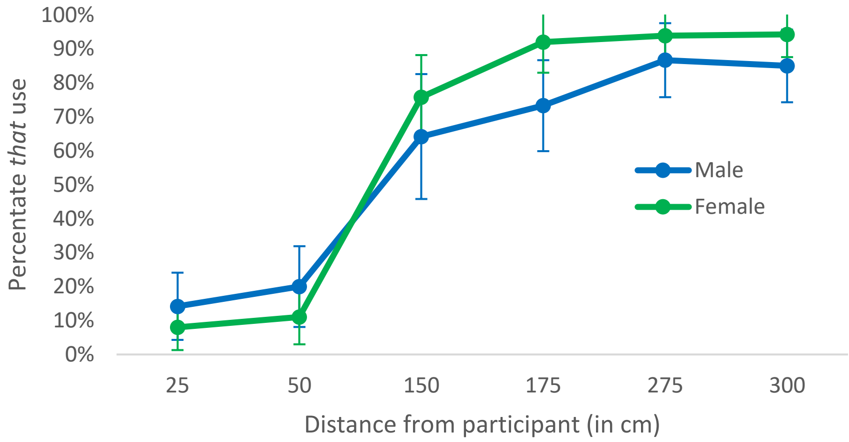


Figure 3.28. The marginal significant interaction between location and gender, indicating a tendency in which the location effect is stronger in female participants. Error bars are 95% confidence intervals. (Error bars for females scores are cut off at the 100% maximum value).

The Japanese data (Experiment 5), reveal that Japanese demonstrate use is affected by both distance and the position of the conspecific. These results confirm that Japanese has both distance and person-centred components. The results of Experiments 6, replicate both earlier studies (Coventry et al., 2008, 2014) providing strong evidence for a main effect of distance and agent, consistent with the Expectation model (which Coventry et al., 2014 argue underlies memory for object

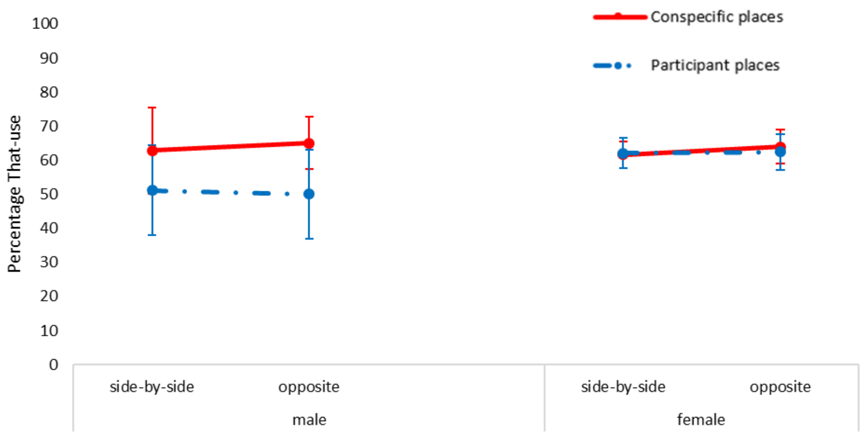


Figure 3.27. The interaction between agent × gender. Female participants show the same percentage of ‘that’ use when the experimenter places the object or when they place it themselves, whereas men use ‘that’ more when the experimenter places. Error bars are 95% confidence intervals, contrasts are LSD.

location and demonstrative choice in English). Interestingly, the results show a conceptual replication of the results in Experiment 5. English speakers seem to be sensitive to the position of a conspecific, even though the English demonstrative system does not make explicit contrasts as a function of hearers' position. This finding is consistent with (Coventry et al., 2014), suggesting the existence of universal parameters affecting demonstratives, even when a language does not explicitly mark a contrast. An important question we cannot answer yet is whether the strength of a parameter is a function of the explicit marking of that parameter in a specific language.

Both language groups show gender differences with respect to the agent manipulation, with men showing an effect of agent, consistent with those reported for the influence of object knowledge on demonstrative choice in Coventry et al. (2014). The effect for men alone for both languages may be a result of the enhanced sensitivity of men to status/competition (consistent with Baumeister & Sommer, 1997; Cross & Madson, 1997).

Next, we turn to the influence of language on memory for object location. Recall that previous studies on English support the expectation model as a model regarding how language affects memory for object location, as well as the effects of object knowledge on both language production and memory (Coventry et al., 2014; Gudde et al., 2016, Chapter 2). However, since Japanese demonstratives encode two parameters at the same time – distance and person-centeredness – it is not clear what expectation values would be associated with Japanese demonstratives. In particular a person-centred encoding means there are two sources from which distance can be coded (hearer and speaker), so person-centred and distance approaches might produce expectation values that conflict. It is therefore important to test the predictions of the Expectation model when considering the influence of Japanese demonstratives on memory for object location.

3.4 Experiment 7. The influence of spatial demonstratives on memory for object location in Japanese

Experiment 5 has shown that Japanese demonstratives map onto space governed by two parameters: person-centeredness and distance. In this study we explore how memory for object location is affected by the Japanese language – a language with a more complex (three-term) demonstrative system than English – and specifically whether the expectation model may potentially account for the influence of language on memory for object location cross-linguistically.

In order to compare Japanese to English speakers, we tested group of English participants in this specific setup (see Experiment 8). This also allowed us to test the possible influence of the position of a conspecific on memory for object location in English speakers for the first time.

3.4.1 Method

3.4.1.1 Participants. Thirty-five Japanese participants were tested. Participants were native Japanese speakers who received payment for their participation. Since the measure in this experiment was spatial memory accuracy, we tested participants for depth perception using the Randot Stereotest (Stereo Optical Inc. Chicago, USA). Participants who did not have a threshold of at least 40" (arcseconds) were excluded from the analysis (N = 1). Two other participants had more than 10% incorrect answers in the memory task and were therefore excluded. This left 32 participants for the analysis, 11 male and 21 female, with an age range of 20-65 years old (M = 28.37, SD = 11.37).

3.4.1.2 Materials. Similar to Experiment 5, six objects were placed at six locations on the table. The six locations were combined in pairs so there were three 'spaces' in the analysis (near space participant, near space conspecific [only when conspecific sat on the other side of the table], space far from both); see Figure 3.25 for an overview of the experimental procedure.

3.4.1.3 Procedure and Design. In this experiment again a 'memory game' was played, but this time the influence of language on memory for object location was actually tested, following the method used in Chapter 2 (Gudde et al., 2016; adapted from Coventry et al., 2014). Participants read out instruction cards directing the conspecific to place one of six objects on one of six locations. Instruction cards contained the following information: "Place [demonstrative (3: Japanese [kono, sono, ano], English [this, that, the])] [object (6)] on the [location (6)]"[10]. The two active manipulations were the demonstrative presented on the instruction card and the position of the conspecific (next to or opposite the participant). Following the placement of the object – during which participants had their eyes closed -

[10] For example, "この赤い丸を緑の点の上に置いてください", or "Place the red circle on the green dot".

participants had 10 seconds to memorize the object location before the participant closed his/her eyes again and the object and dots on which objects were placed were removed. The conspecific, as in Experiment 5, was the experimenter who was instructed by participants (see Figure 3.29). In addition to the experimenter running the trials, there was an assistant behind a screen responsible for producing and moving an indication stick directed by the participant to where s/he thought the edge of the object was. The indication stick started at a distance of 10cm (closer or further) from the actual object location, with the exception of 3 filler trials presented within the first 10 trials, in which the indication stick was presented at a 20cm distance (done to prevent the indication stick being a cue to the location). The assistant was standing permanently behind a curtain to prevent any (unintentional) cues to participants. At the end of every trial, participants recalled the demonstrative on the instruction card to check he/she was paying attention to the instructions. When the participant could not remember the demonstrative, the trial was repeated.

The design was a 3 (demonstrative: *kono, sono, ano*) x 3 (space: near participant, middle, far from participant) x 2 (position of conspecific: beside or opposite participant) within participants design, with 3 trials per condition (totaling 57 trials, including the 3 filler trials).

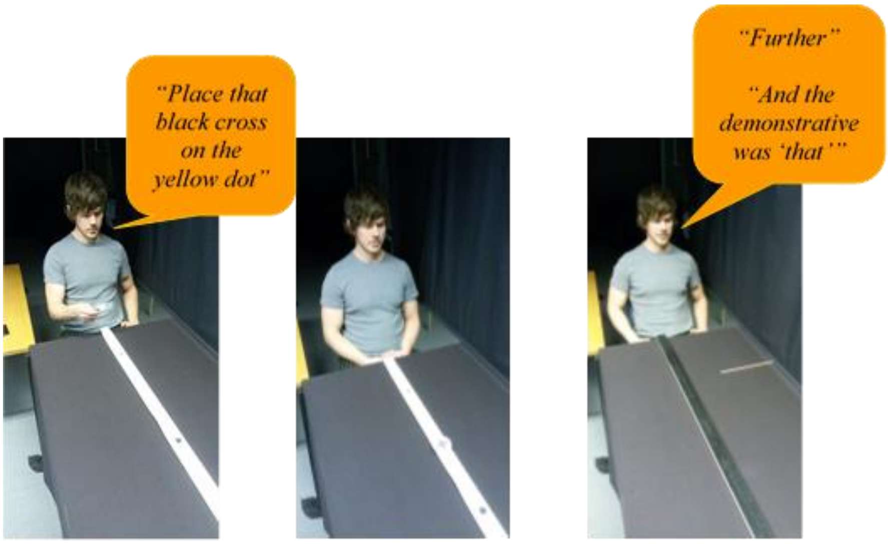


Figure 3.29. Procedure for Experiment 8. The participant reads out an instruction card. While the participant has his/her eyes closed, an object is placed at the instructed location and participants get 10 seconds to memorize the object location. After the 10 seconds, the participants closes his/her eyes again, the object and the locations are taken away and the participant is asked to verbally align the location of an indication stick with the memorized location of the object.

3.4.2 Results and Discussion

A 3 (demonstrative [*kono/sono/ano*]) × 3 (space [25 and 50; 150 and 175; 275 and 300 cm]) × 2 (position of conspecific [next to or opposite the participant]) × 2 (gender) ANOVA was performed on the differences of the memorized location and the actual object location (in cm).

The assumption of sphericity was violated in the analysis of space, the demonstrative × position of conspecific interaction, the space × position of conspecific interaction, and the three-way interaction between demonstrative × space × position of conspecific. In those analyses Greenhouse-Geisser correction was applied.

There were no main effects (demonstrative, p = .315, ηp^2 = .038; space, p = .229, ηp^2 = .048; position, p = .134, ηp^2 = .073). There was however, a significant demonstrative × space interaction, $F(4,120)$ = 4.906 p = .001, ηp^2 = .141 (see Figure 3.30). As can be seen in the figure, in the space nearest the participant, there is no effect of demonstratives (using LSD multiple comparisons all p > .401). For space 2, the pattern of results is consistent with the expectation model, with greater errors for *ano* compared to *kono* (LDS follow-up, p = .012). In space 3, there is no significant difference (all LSD comparisons p > .197). However, there seems to be a tendency in which *kono* is misremembered to be further away than *ano* and *sono*, which would be consistent with the congruence model.

The only other significant interaction was a position of conspecific × demonstrative × space × gender interaction, $F(4,120)$ = 3.232, p = .015, ηp^2= .097. We followed this interaction up with separate analyses for male and female participants.

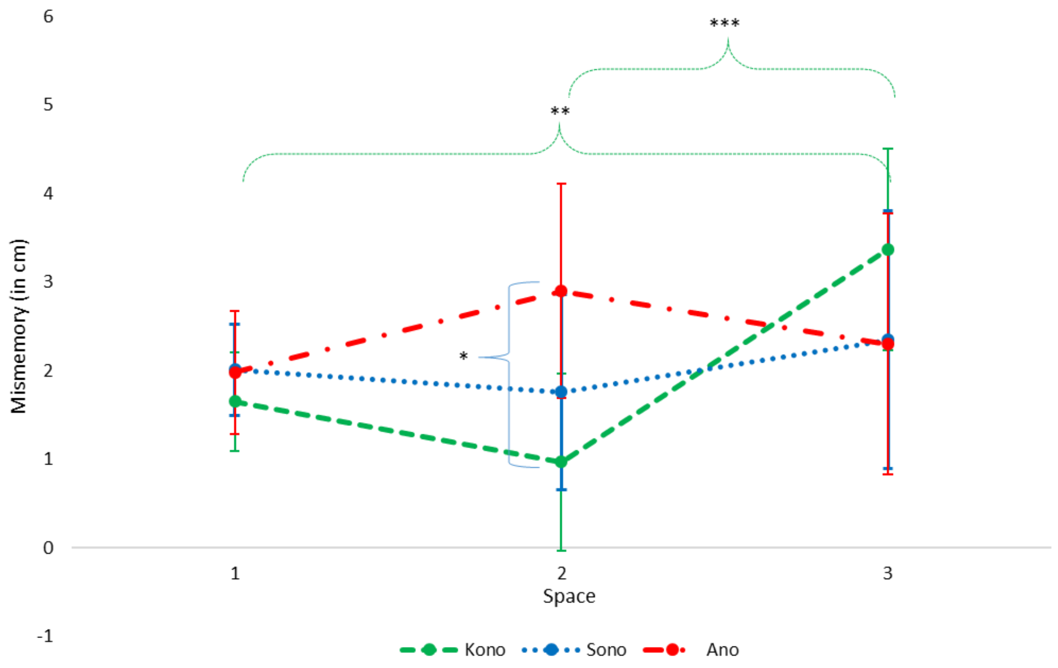


Figure 3.30. Mismemory scores (in cm) of Japanese participants, per demonstrative and space (error bars are 95% confidence intervals). LSD follow up significance values are represented: * $p < .05$, ** $p < .01$, *** $p < .001$.

In the analysis of male participants, the Mauchly's test of sphericity indicated that the assumption of sphericity was violated for the 'space' variable, so we used the Greenhouse-Geisser correction to control for this. There was a strong main effect of position, $F(1,10) = 6.642$, $p = .028$, $\eta p^2 = .399$, showing that objects were misremembered further away from the participant if the conspecific was sitting next to the participant ($M = 2.481$, $SE = .52$, *95% CI* [1.330, 3.633]) compared to opposite the participant ($M = 1.725$, $SE = .67$, *95% CI* [.224, 3.226]). There was also a significant demonstrative × space × position of conspecific interaction, $F(4,40) = 4.208$, $p = .006$, $\eta p^2 = .296$, see Table 3.14.

We followed up this interaction running two separate 3 (demonstrative) × 3 (space) ANOVAs, one for each position. When participant and conspecific sit side-by-side, there were no main effects (space, $p = .579$, $\eta p^2 = .053$; demonstrative, $p = .266$, $\eta p^2 = .124$), but there was a significant space × demonstrative interaction, $F(4,40) = 2.703$, $p = .044$, $\eta p^2 = .213$. This difference was driven by space 2, in which *ano* (M = 3.909, SE = 4.162, 95% CI [1.320, 6.498]) differed from *kono* (M = .648, SE = .912, 95% CI [-1.383, 2.680]), $p = .002$ and *sono* (M = 1.521, SE = .828, 95% CI [-.323, 3.365]), $p = .023$. This effect at space 2 is consistent with the Expectation

Table 3.14. *The mean distance error of male participants, by demonstrative, space, and position*

		space 1		Space 2		Space 3	
	position	Mean (SE)	95% CI	Mean (SE)	95% CI	Mean (SE)	95% CI
Kono	next to	1.79 (.48)	[0.72, 2.87]	.65 (.912)	[-1.38, 2.68]	3.79 (1.57)	[0.29, 7.3]
	opposite	1.18 (.7)	[-0.38, 2.73]	1.59 (1.28)	[-1.25, 4.43]	1.43 (1.51)	[-1.94, 4.8]
Sono	next to	2.8 (.5)	[1.73, 3.96]	1.52 (.83)	[-0.32, 3.37]	2.47 (1.5)	[-.87, 5.82]
	opposite	1.52 (.78)	[-0.22, 3.26]	.53 (.77)	[-1.18, 2.25]	2.79 (1.33)	[-0.16, 5.75]
Ano	next to	1.97 (.6)	[0.64, 3.3]	3.91 (1.16)	[1.32, 6.5]	3.39 (1.29)	[0.52, 6.25]
	opposite	2.75 (.97)	[0.59, 4.91]	1.45 (1.13)	[-1.08, 3.97]	2.29 (1.54)	[-1.15, 5.72]

model. When participant and conspecific sit opposite each other were no significant effects: space, $p = .730$, $\eta p^2 = .031$; demonstrative, $p = .494$, $\eta p^2 = .068$, space × demonstrative, $p = .400$, $\eta p^2 = .094$. Regarding the use and function of demonstrative in Japanese, these results seem consistent with the dual-system model, as the effects show a distance-based result when participant and conspecific are seated side-by-side, but this distance effect disappears when they sat opposite.

In the analysis of data from female participants, the assumption of sphericity was violated in the interactions between position of conspecific and space, position and demonstrative, and the three-way interaction between demonstrative, space, and position of conspecific. These analyses were corrected using Greenhouse-Geisser. There were no main effects of demonstrative ($p = .862$, $\eta p^2 = .007$), space ($p = .338$, $\eta p^2 = .053$), nor position of conspecific ($p = .972$, $\eta p^2 < .001$). There was an interaction between demonstrative and space, $F(4,80) = 8.596$, $p < .001$, $\eta p^2 = .301$. A follow up analysis showed that the pattern of data mirrors the pattern found for this interaction in the main analysis. There is no effect of demonstrative for space 1 (all LSD pairwise comparisons are $p > .715$), in space 2 there is a difference between both *kono* and *sono* (LSD comparison, $p = .016$) and *kono* and *ano* (LSD, $p < .001$). In space 3 there are also significant differences between *kono* and *sono* ($p < .001$) and *kono* and *ano* ($p < .001$), but in contrast to the results in space 2, where *kono* was remembered to be closest by, *kono* is remembered to be furthest away in space 3. There were no other interactions (demonstrative × position, $p = .59$; space × position, $p = .94$; demonstrative × space × position, $p = .32$). These results suggest that, in contrast to male participants, female participants were not influenced by the position of a conspecific. Similar to the main analysis, the female pattern of results is only partly consistent with both the Expectation and Congruence model. In space 2, the pattern is consistent with the Expectation model: objects placed with *kono* were remembered to be closer by than objects placed with *ano* or *sono*. However, in space three the direction has flipped and the results seem to resemble the congruence model. In space 3, objects placed with *kono* are misremembered to be further away. These are also the trials where the incongruence is largest. However, the data seems not overall consistent with Expectation nor Congruence model. It is possible that the ambiguity of Japanese demonstratives diminishes the incongruence effect. As the influence of language works (partly) via the information conveyed in the language, the ambiguity in Japanese demonstratives could mean that demonstratives convey less spatial information. If this is the case, the resulting expectation based prediction (Expectation model) incongruence from that information (Congruence model) could be weaker too.

Table 3.15. *Results of the ANOVA on Memory for Object Location in Experiment 7.*

	Source	***df* and *F* value**	**MS**	**Significance**	**ηp²**
Male	Position (P)	F (1,10) = 6.642	22.977	.028	.399
	Space* (S)	F (1.146,11.465) = .610	38.24	.472	.057
	Demonstrative (D)	F (2,20) = 1.529	13.942	.241	.133
	P × S	F (2,20) = .133	1.786	.876	.013
	P × D	F (2,20) = .012	.067	.988	.001
	S × D	F (4,40) = 0.682	4.961	.609	.064
	P × S × D	F (4,40) = 4.208	15.568	.006	.296
Female	Position (P)	F (1,20) = .059	.598	.811	.003
	Space (S)	F (2,40) = 1.167	34.523	.322	.055
	Demonstrative (D)	F (2,40) = 0.136	1.236	.873	.007
	P × S*	F (1.544,30.882) = 0.174	3.645	.785	.009
	P × D*	F (1.522,30.432) = 0.456	4.079	.585	.022
	S × D	F (4,80) = 8.596	53.729	<.001	.301
	P × S × D	F (4,80) = 1.758	12.886	.145	.081

*Greenhouse-Geisser

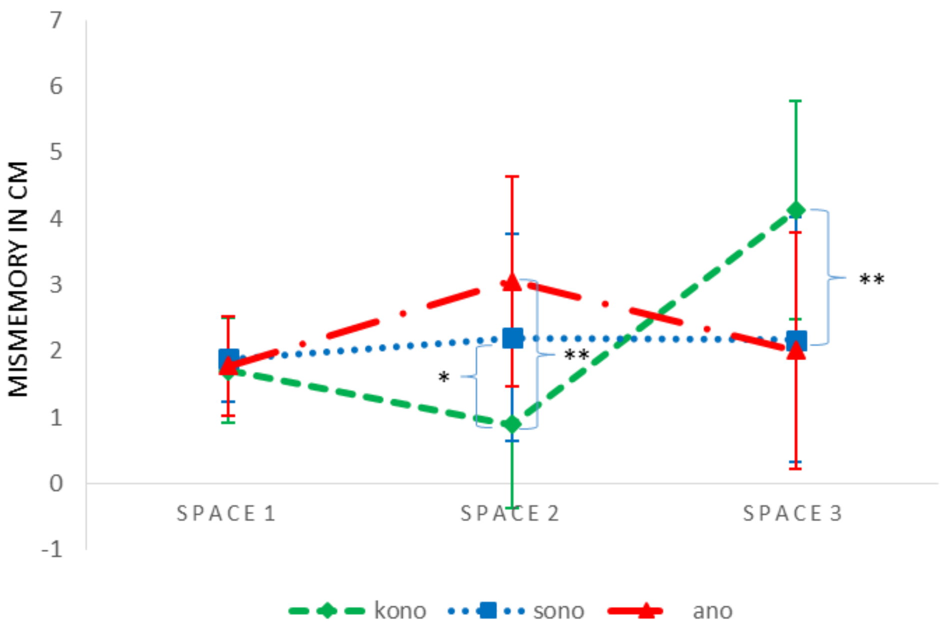


Figure 3.31. The mismemory scores (in cm) of female, Japanese participants, per demonstrative and space (error bars are 95% confidence intervals). Significance values are respresented: * $p < .05$; ** p < .01; *** p < .001. The significant interaction between demonstrative and space shows significant differences between *kono* compared to *sono* and *ano*. In space 2, *kono* is misremembered to closer by, but in space 3 it is misremembered to be further away.

3.5 Experiment 8. The influence of spatial demonstratives on memory for object location in English.

Experiment 7 showed that both the demonstrative and the space (distance) condition interacted with position of a conspecific for Japanese participants. This is consistent with findings in Experiment 5, in which we found that Japanese demonstrative use is influenced by the position of a conspecific. Since we found that the position manipulation also influenced English speakers in Experiment 6, we will test the influence of position on English speakers' memory in Experiment 8. This experiment will be a replication of Experiment 7, but then with English speakers.

3.5.1 Method

3.5.1.1 Participants. The English group comprised 41 native English speakers, studying at the University of East Anglia, receiving payment or student credit for their participation. Four participants did not have a depth perception of 40", three participants had more than 10% incorrect answers in the memory task, and one participant did not engage with the instructions. This left 33 participants for the analysis, 8 male and 25 female, with an age range of 18-55 years old (M = 23.12, SD = 8.26).

3.5.1.2 Materials, Procedure, and design. This English experiment was similar to Experiment 7, with the exception that the entire study was in English.

3.5.2 Results and Discussion.

The English memory displacement data are displayed in **Error! Reference source not found.**. A 3 (demonstrative [*this/that/the*]) × 3 (space [25 and 50 cm; 150 and 175 cm; 275 and 300 cm]) × 2 (position of conspecific [next or opposite participant]) × 2 (gender) ANOVA was performed on the differences of the memorized location and the actual object location (in cm). The assumption of sphericity was violated in the analysis of space, and the demonstrative × space interaction. In those analyses we used the Greenhouse-Geisser correction. There was a main effect of demonstrative, $F(2,62) = 5.759$, $p = .005$, $\eta p^2 = .157$, showing that objects placed in the *this* condition (M = .108, *SE* = .38, 95% CI [-0.668, 0.883]) were remembered to be closer by, than objects in the *that* condition (M = .93, *SE* = .4, 95% CI [0.12, 1.73], *p* = .01) and the *the* condition (M = .77, *SE* = .4, 95% CI [-0.02, -1.57], *p* = .006). There was also a main effect of space, $F(2,62) = 4.252$, $p = .034$, $\eta p^2 = .121$, showing that object were misremembered to be further away in space 3 (M = 1.78, *SE* = .77, 95% CI [0.205, 3.36]), compared to space 2, (M = -.145, *SE* = .343, 95% CI [-0.844, 0.554], (LSD comparison) $p < .001$), and space 1, (M = .168, *SE* = .417, 95% CI [-0.684, 1.019], $p = .003$), but there was no difference between space 1 and 2 ($p > .05$). There was no main effect of position ($p = .196$) nor a significant space × demonstrative interaction (p = .14). There was no effect of gender ($p = .281$), nor

Table 3.16. *Mismemory data representing the main effects of demonstrative and space.*

Experiment 8	Space	1 (SEM)	2 (SEM)	3 (SEM)
Demonstratives	This	.24* (0.41)	-.88 (0.35)	.96 (0.8)
	That	.22 (0.44)	.4 (0.46)	2.16 (0.77)
	The	.04 (0.49)	.05 (0.51)	2.24 (0.93)

* values are represented as difference in cm from the actual location (positive value means the object was misremembered to be further away than it actually was.

were there interactions with gender (gender × demonstrative, p = .514, gender × space, p = .637, gender × position, p = .297).

These results replicated the results from Chapter 2 (Gudde et al., 2016). The main effect of demonstrative and no interaction between demonstrative and location are consistent with the predictions of the Expectation Model (Coventry et al., 2014).

The results from Experiment 5 and 7 show that Japanese demonstrative use is a function of the spatial position of an interlocutor, as well as distance from a speaker. Furthermore, in male participants we found an effect of this position on memory for object location. Even though English demonstratives do not encode an interlocutors' position, we found that in English participants, position affects demonstrative use and memory for object location consistent with the Japanese data (Experiment 6 and 8). This suggests that even though language, or specifically spatial demonstratives, are used to contrast between different parameters, their usage might be based on a universal set of parameters. However, the effect was stronger in Japanese than in English. This could suggest that if demonstratives are influenced by universal parameters, the influence is stronger if a parameter is explicitly coded, compared to when it is implicit in a language. To test the nature of these differences, a future study needs to do a direct comparison between Japanese and English speakers. To test whether the effects are based on language, verbal interference can be used (modelled on Trueswell & Papafragou, 2010). If language is used as a tool to

Table 3.17. *Results of the ANOVA on Memory for Object Location in Experiment 8.*

Source	***df* and *F* value**	**MS**	**Significance**	**ηp²**
Position (P)	F (1,31) = 1.748	15.191	.196	.053
P × gender	F (1,31) = 1.126	9.787	.297	.035
Space* (S)	F (2,62) = 4.252	230.007	.034	.121
S × gender*	F (2,62) = .454	24.58	.562	.014
Demonstrative (D)	F (2,62) = 5.759	27.464	.005	.157
D × gender	F (2,62) = .672	3.205	.514	.021
P × S	F (2,62) = .742	7.647	.48	.023
P × S × gender	F (2,62) = .197	2.028	.822	.006
P × D	F (2,62) = 2.290	10.049	.11	.069
P × D × gender	F (2,62) = .480	2.106	.621	.015
S × D*	F (4,124) = 1.859	13.089	.144	.057
S × D × gender*	F (4,124) = .925	6.513	.43	.029
P × S × D	F (4,124) = 1.296	6.473	.275	.04
P × S × D × gender	F (4,124) = .838	4.182	.504	.026

*Greenhouse-Geisser

remember object location, then the verbal interference should eliminate these effects. However, if the effect of an interlocutors' position is still apparent, one can assume that spatial language and spatial memory are dependent on the same underlying processes.

3.6 General Discussion

In four experiments we examined how distance from a speaker and the position of a conspecific affect demonstrative choice in two differences languages – Japanese and English – and in relation to the influence of that language on memory for object location. Here we first consider the demonstrative systems of these languages, and then models of how language affects memory for object location.

The results for English demonstrative choice, mirrors the results found previously (Coventry et al., 2014, 2008; Gudde et al., 2016). English demonstratives are affected by distance and by agent. Moreover, distance did not interact with any other variables, again consistent with previous results. However, we also uncovered two new effects. First, we found a significant interaction between agent and gender. Given that previous studies did not consider gender differences in language production (save for Experiment 7 in Coventry et al., 2014), the finding that men's use of demonstratives is affected by agent and women's demonstrative use is new. Second, we found an interaction between position and location, showing a larger effect of location when participant and experimenter are side-by-side rather than when they are sitting opposite one another. This pattern is consistent with the view expressed by Coventry et al. (2014) that demonstrative use across languages may be affected by a common set of parameters, with some explicit in a language and some not (a point we develop below).

In relation to Japanese, there have been several proposals regarding how its demonstrative system operates, from claims that Japanese is entirely distance-based to proposals that Japanese is person-centred. The data from Experiment 5 show, under experimentally controlled conditions, that Japanese employs a dual-system model (Takahashi & Suzuki (1982), as read in Niimura & Hayashi, 1994; Okazaki, 2011; Takahashi, 1992), in which demonstratives can encode both distance and person-centeredness, and the choice of demonstrative model is based on the configuration of interlocutors. When speaker and hearer were sitting side-by-side, participants preferred *ano* to *sono* when the object was out of reach, a preference that increased as the object was further away. However, the significant main effect of position of the conspecific shows that distance is not the only parameter affecting Japanese demonstratives; territory affected use as well. In this respect, territory could be *interactional* space, in which the entire space between interlocutors is encoded as a territory (using *sono*), rather than a near 'close to the hearer' territory (Yoshimoto, 1992, 1986). Finally, mirroring the result found for English, the presence of an effect of agent for men but not for women illustrates an increased sensitivity in men regarding who contacts an object. It is possible that this effect is related to territory, and effects of position of conspecific found across both languages (although for both languages we did not find significant gender differences for this variable).

The commonalities across Japanese and English are consistent with the view expressed in Coventry et al. (2014), that demonstrative systems across languages may reflect a common set of constraints on languages, underlying non-linguistic perception and memory. While some of these constraints are explicit in the demonstrative systems of some languages, Coventry et al. (2014) have shown that they nevertheless affect demonstrative usage in languages that do not employ these explicit distinctions. For example, Coventry et al. (2014) found that visibility and ownership affect the choice of *this* and *that*, even though the English demonstrative system does not explicitly encode these parameters. The presence of an effect of position of conspecific in Japanese and English, and the same pattern of gender differences with respect to who places an object support a universal set of underlying non-linguistic parameters, irrespective of whether a parameter is linguistically encoded in both languages. However, the present data also reveal stronger effects of the position of conspecific in Japanese than in English, pointing in the direction of universal constraints, but with differential weightings of these constraints as a function of language. If a language makes an explicit contrast, then the data are consistent with the hypothesis that that language should then exhibit more influence of those parameters than a language that does not make an explicit contrast.

Turning to the memory data, we set out to test whether the influence of Japanese language on memory for object location was similar to the influence of English language, teasing apart two memory models: the Expectation model and the Congruence model. The Expectation model predicts that language affects memory for object location via the expectation that language elicits – *this* refers to objects closer by, object location is remembered to be closer by if the placement instruction used *that.* The Congruence model predicts that the influence works via a congruence/ incongruence of language and space: memory performance is more accurate when the two are congruent – if *this* refers to objects closer by, but is used to instruct the placement of an object further away, there is an incongruence between the instruction and the spatial situation. This incongruence should lead to larger memory errors. It is important to appreciate that the memory game setup can tease the models apart: the Expectation model predicts a main effect of language, whereas the Congruence model predicts an interaction between language and space. The results of English replicate previous studies showing a main effect of demonstrative on memory for object location in English (Coventry et al., 2014; Gudde et al., 2016, Chapter 2). Consistent with the Expectation model, we found a main effect of demonstrative, in which objects placed with *that* were misremembered to be further away. There was no interaction between location and demonstrative, which would have indicated a congruence/incongruence effect (as predicted by the congruence model).

In contrast to the English results in Experiment 8, the influence of Japanese demonstratives on memory for object location (Experiment 7) show a pattern of results that does not sit so well with the Expectation model. In the Japanese analysis, we found intriguing gender differences. For Japanese female participants there was no main effect of demonstrative (as occurs in English), but rather an interaction between demonstrative and space. This interaction is not consistent with the expectation model, but only partly supports the congruence model. When there is a mismatch between the demonstrative and the space, participants misremember objects as being further away than they actually were, although this effect of incongruence was not apparent across locations. It is possible that the ambiguity of Japanese language diminishes the incongruence effect (as they would diminish any effect) - because of the uncertainty that is inherent to Japanese demonstratives, the predictive value of demonstratives and the possible incongruence they would elicit are lower.

For Japanese male participants, on the other hand, the position of the conspecific mediated the influence of language on memory. When the participant and conspecific were seated side-by-side, there was an interaction between space and demonstrative, in which *ano* was misremembered further away from *kono* in space 2, an effect that didn't occur at any other combination of space and demonstratives. However, there was no interaction when the conspecific sat opposite the participant. This finding would suggest that the influence of language disappeared when territory was a factor. This is, although not necessarily indicative, consistent with a dual-system model, in which distance is a less important parameter when demonstratives can code territory. Furthermore, male participants showed larger memory error values when participant and experimenter sat side-by-side, misremembering the location of the object to be further away compared to when the conspecific sat opposite the participants. The dual-system model would predict that in these latter cases, demonstratives would encode territory and not egocentric distance. If male participants encoded the object location partly relative to the conspecific, it might explain the fact that memory error was smaller: distance was relative to both participants' and conspecifics' territory, which could have led to a more balanced spatial memory. An explanation as to why Japanese males are more sensitive to a hearers' territory might be found in social psychological research, showing that males are more focussed to separate themselves from others and therefore more likely to acknowledge interlocutors' territory (Cross & Madson, 1997), a trait that could be stronger in a masculine society such as the Japanese (Hofstede, 1998). However, more research needs to be conducted to confirm any claims on this. This is the first time that we found gender differences in the memory game procedure. Future research could include questionnaires to find underlying causes for these gender differences, for example masculinity, personality, or social intelligence questionnaires.

The results over the four experiments showed interesting different patterns of the use of language and the influence of language use on memory for object location between Japanese and English speakers. Experiment 1 showed that Japanese spatial demonstratives are influenced both by the distance an object has from the speaker, as by the personal territory in which an object is located (speakers' vs. hearers' territory). There was an interaction between position and distance, showing that the distance effect was driven by trials in which participant and conspecific shared their territory. These results confirm the theoretical accounts of a dual-system model, in which Japanese demonstratives can encode both distance and person-centeredness Takahashi and Suzuki (1982, as read in Niimura & Hayashi, 1994; and Okazaki, 2011; Takahashi, 1992). When speaker and hearer have their own territory, language is predominantly based on a person-centred account; when speaker and hearer share a territory, demonstratives are based on distance from the speaker. In these situations, language is not ambiguous for the participant, as the participant is the speaker in these trials, and will therefore know on which parameters s/he based the information conveyed. However, the hearer will experience this ambiguity, since the hearer has to interpret how the speaker encoded demonstratives.

The results of Experiment 7, in which we manipulated linguistic instruction, show this ambiguity. As the distance based and person-centred models are incompatible at any one given time, one has to choose the language parameter to interpret when encoding object position. There is a gender difference in the prioritization of parameters in a memory context. While female participants effectively ignored the person-centred component in their demonstrative system, male participants were influenced by the position of the conspecific when territory was not shared, and showed an interaction between language and space when territory was shared. The female participants however only showed an interaction between space and demonstrative. The reason of this gender effect is unclear; speculative this could be a result of a territorial attitude of men compared to women.

The lack of main effects of Japanese language on spatial memory could be a result of the fact that demonstratives can encode multiple parameters. In English, demonstratives have a clear predictive value: objects referred to with *this* are closer by than objects referred to with *that*. However, since in Japanese this contrast is not as clearly defined using the three different demonstratives, the predictive value of demonstratives – if any – is lower. If one cannot know what a specific demonstrative encodes, there is no information to base an expectation on.

To summarise, we replicated earlier results showing that memory for object location in English is influenced by language via an Expectation Model. Different parameters can affect Japanese demonstrative use (distance and person-centeredness). The memory data of Japanese participants was too unclear to reach any strong conclusions, but a novel finding in this study is that the position of a conspecific affected English male participants similar to Japanese male participants. This suggests that demonstratives might be based on a universal set of parameters, although the strength of the influence could vary based on whether the parameter was encoded explicitly or not.

Future studies need to look further into the difference between Japanese and English. If this difference is facilitated by the differences in language, the behavioural outcome should be more similar when language is not a factor. To test this, the study needs to be replicated, including a verbal interference task. For example, by having participants repeat a sequence of meaningless sounds (Bo, Ba, Bi, etc.). This task was previously used in spatial tasks (Coventry et al., 2014; Garden, Cornoldi, & Logie, 2002). By preventing participants from using language, we can see whether the difference between Japanese and English participants is caused by the differences in their language.

Another line of follow-up studies includes a broad range of different languages. In the current study we found, consistent with Coventry et al. (2014), that English is influenced by the manipulation of parameters that are not encoded in English. This might suggest universal parameters underlying demonstratives across different languages. To test this further, a greater number of languages need to be experimentally tested using the memory game paradigm. Parameters are yet to be determined, but it would be interesting to manipulate at least distance from the participant and the position of a conspecific. If there are universals in language, it would be expected that all languages exhibit similar effects, taking into account that effects might vary in strength as a function of whether or not a specific parameter is encoded in the language that is tested. Currently we are planning to run both these future study proposals.

The past two chapters we have been looking into the relationship between language and spatial memory. However, another type of spatial language studied extensively because of its relation with spatial cognition are spatial prepositions expressing frames of reference. In Chapter 4, we will explore the effect that language (frames of reference) has on visual attention, using eye-tracking methodology.

Chapter 4 – The influence of spatial reference frames on visual perception

In the previous two chapters, we explored the influence of language on spatial memory. Following previous research showing that choice of words – e.g., demonstratives - to describe a situation is affected by seemingly irrelevant cues (e.g., ownership, familiarity, and visibility) (Coventry et al., 2014), we saw that spatial language in turn influences spatial memory. In the final experiment of Chapter 2, we explored whether spatial demonstratives influence how participants focus their attention to an object, but found no influence of language on how participants scanned the environment to encode object location.

In this chapter, we look at how language influences visual attention. Can a speaker's choice of language (e.g., spatial prepositions associated with reference frame) influence how we look at the described visual world? And does a description from a previous speaker impact the language the hearer produces? In this experiment (N = 84), we presented a verbal description of a spatial scene and recorded participants' eye-movements, revealing two findings: eye movement 'signatures' associated with distinct reference frames as expressed in language, and transfer of these eye-movement patterns just prior to spatial description for different (later) picture descriptions. The experiment adds to the current literature on reference frames (cf., Levinson, 1996), examining the relation between reference frames and eye-movement patterns, but will also be presented in a broader 'interactional' perspective.

4.1 Introduction

4.1.1 The influence of language on visual attention

We will explore the relation between perception, memory, and language, via reference frames as expressed in language. Eye-tracking studies have shown that language automatically drives attention to parts of the spatial world. The role overt attention has in the comprehension of spatial language can be tested in visual world tasks. In these tasks, participants are presented with a visual display while hearing utterances describing the array. Objects in the experimental display are usually mentioned in the description. Studies employing this visual world task showed that participants fixate on objects that are mentioned, and make fewer fixations on objects that are not named (e.g., Allopenna et al., 1998; Altmann & Kamide, 2007; Tanenhaus et al., 1995).

In particular, as language unfolds in real time, it directs overt eye-movements towards objects (Burigo & Knoeferle, 2015). As listeners look at a visual array while presented with a verbal, auditory description of that array, their eye-movements are closely time locked to the specific referring words (Eberhard et al., 1995), suggesting a close relationship between language and visual attention. However, this automatic effect does not only occur online, but also in a predictive fashion. For example, there is evidence that people make anticipatory eye-movements to objects they expect to be named in a story (Knoeferle & Crocker, 2006). Consistent with theories of predictive coding, people seem to use language and world knowledge to anticipate future information and fixate on a predicted reference object. For example, a sentence like ‘the man has drunk’ vs. ‘the man will drink’ focusses attention on respectively an empty or a full glass (Altmann & Kamide, 2007).

In the domain of spatial language, people use their world-knowledge to select a reference object in a sentence. For example, when participants were presented with the sentence “*put the cube inside the can”*, they used the information from the preposition (*inside)*, to choose which object would be the object of reference fitting the relevant spatial properties (i.e., container-like objects with specific dimensions to be able to contain the cube - the can) before they heard the actual referent (*can*) (See Figure 4.32) (Chambers, Tanenhaus, Eberhard, Filip, & Carlson, 2002). The idea of an automatic predictive mechanism is supported by research showing participants do not only follow the onset of the actual words, they also ‘mis-predict’ if a presented distractor is (partially) rhyming or a homophone with the reference object (e.g., candy/ candle, beaker/ beetle) (Allopenna et al., 1998; Eberhard et al., 1995; Tanenhaus et al., 1995).

However, attention is not only driven by language. There is a similar effect the other way round, the way attention is directed to a visual scene. When attention

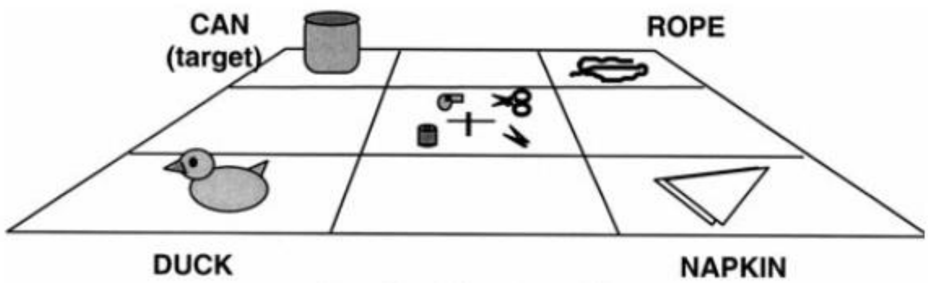


Figure 4.32. Example of the table-top array used by Chambers et al. (2002).

is manipulated towards a specific character (via an attention-capture manipulation), this character is more likely to be the subject of the sentence. For example, when participants describe an array that shows two boxers (see Figure 4.33), attention can be focussed on either boxer by presenting a black square on the location of that boxer prior to showing the boxers. In Figure 4.33, the attention-capture manipulation is focussed on the losing boxer. This increases the likelihood participants use the losing boxer as object in the sentence, and not the winner (e.g., “the man lost the boxing match” vs. “the man won the boxing match”) (Gleitman, January, Nappa, & Trueswell, 2007).

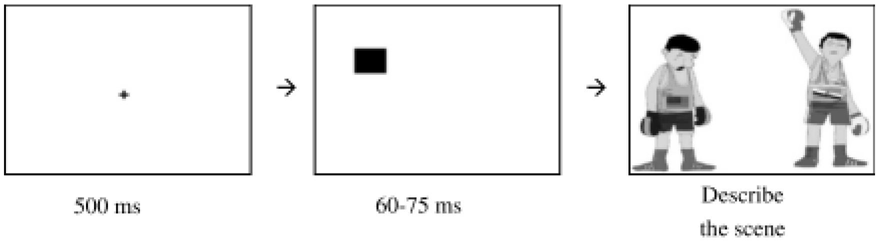


Figure 4.33. Example of the attention-capture manipulation in the timeline in a trial (Gleitman et al., 2007). A fixation cross would be shown, then attention would be focussed by a black square on the position of one of the characters in the scene participants need to describe.

4.1.2 Spatial relations

While looking at an array of two or more objects, one needs to make eye-movements to be able to verify spatial descriptions. For example, if an array is presented consisting of two objects, a marble and a ladybird, a spatial description could be ‘the marble is in front of the ladybird’. In order to verify whether this is the case, one needs to change the focus of attention from the marble to the ladybird. The direction of this attentional shift is predicted in the Attentional Vector Sum (AVS) (Regier & Carlson, 2001). The AVS consists of the summed strength of all attention vectors. This summation is represented in Figure 4.34. A located object (marked with TR, in our study we use a ‘marble’ as located object in the prime trial) is located with respect to a landmark (LM in the example, the ladybird in our study). To comprehend this description, attention needs to shift from the landmark to the located object. This attentional shift consists of different attentional vectors, directed from different locations on the landmark in the direction of the located object (these vectors are represented by arrows). The strength of these vectors depends on the amount of attention each individual vector receives. For example, attention will be

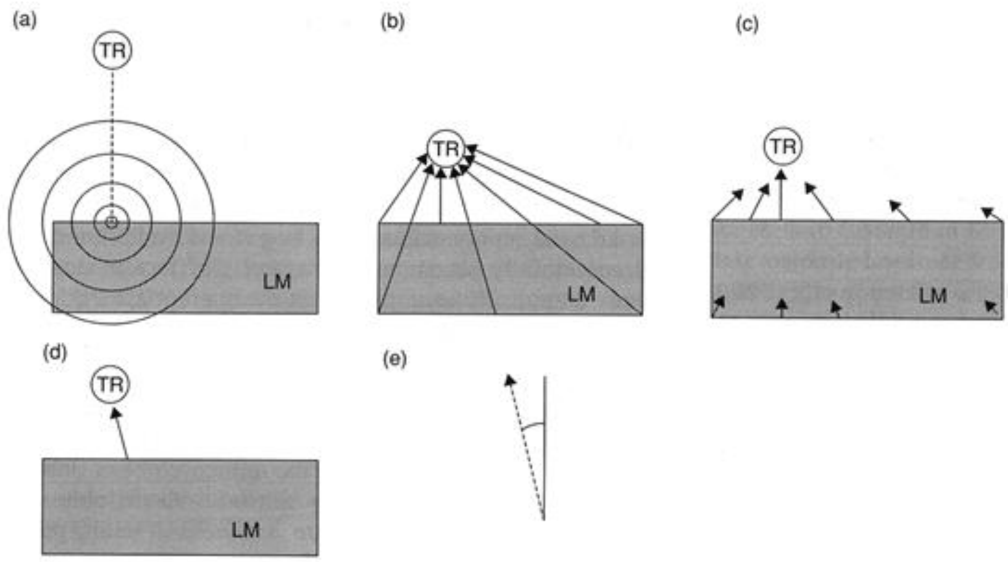


Figure 4.34. Visual representation of the AVS model. The located object (or trajector 'TR') is located with respect to a landmark (LM). Panel b shows the vectors from the different points of the landmark, whereas panel c shows the strength of attention for each vector. The total vector sum is represented in panel d, leading to the orientation of the attentional shift in panel e. (Regier & Carlson, 2001).

mostly focussed from the nearest part of the landmark to the located object, so the vectors closer to the located object are stronger than the vectors on the other side of the landmark. The overall attention bias is the combination of strength and directions of the different vectors.

4.1.3 Frames of Reference

In the use of spatial language, interlocutors need to make sure they are talking about the same object, event, or, in the case of reference frames, direction. Therefore, both speaker and hearer have to choose a frame of reference, which they can express in language (see Figure 4.35). Across different fields of science (e.g., neuroscience, perception, development, linguistics), reference frames have been classified differently: speaker centred (deictic) vs. intrinsic (non-speaker centred) (e.g., Miller & Johnson-Laird, 1976), body-centred (egocentric) versus environment-centred (allocentric) coding (Burgess, 2008; O'Keefe & Nadel, 1978), or viewer- centred versus object-centred (Marr, 1982). Based on these categorizations across disciplines, Levinson (1996) identified three types of reference frames: intrinsic (object-centred: "the marble is *in front* of the ladybird"), relative (viewer-centred: "the marble is *to the left* of the ladybird"), and absolute (environment-centred: "the marble is *north* of the ladybird") (See Figure 4.34).

a) Intrinsic
The marble is *in front* of the ladybird

b) Relative
The marble is *to the left* of the ladybird

c) Absolute
The marble is *north* of the ladybird

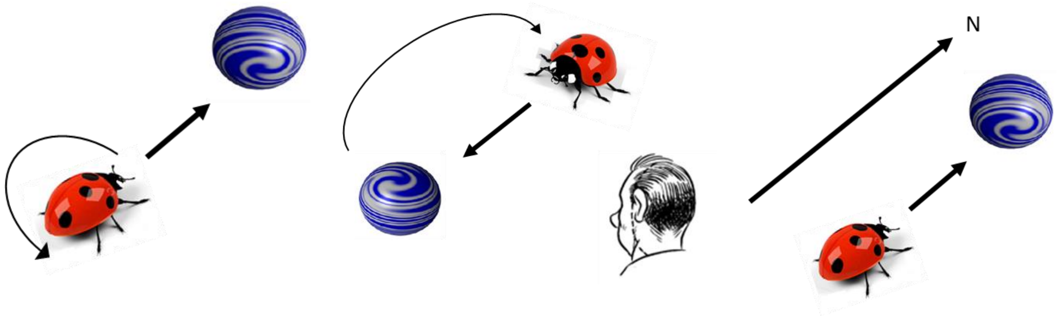


Figure 4.35. In Figure a), *the marble is in front of the ladybird* locates the marble using the intrinsic axes of the ladybird. In Figure b), *the marble is to the left of the ladybird*, the marble is located with respect to the viewer looking at the picture. Finally, the objects can be located based on objective terms, such as cardinal direction *the marble is north of the ladybird* (although this is infrequent in English in small-scale space), see c). This figure is based on (Levinson, 1996).

Carlson-Radvansky and Irwin (1994; 1993) showed that different frames of reference are actively competing during spatial term assignment. In a sentence/picture verification task, participants were slower to respond if reference frames did not align. For example, in 4.35A), the tree in the middle square is presented upright, so the statement ‘the square is *above* the tree’ is appropriate for the same squares in an environment-centered reference frame and an object-centered (intrinsic) reference frame. However, in B) and C), the tree is rotated. The appropriateness of the term *above* now differs between the environment-centered reference frame (see Figure 4.35B) and the object-centered reference frame (see Figure 4.35C).

Results of studies testing the use of reference frames indicated that participants build multiple spatial templates as a result of simultaneous activation (Carlson-Radvansky & Logan, 1997) and only certain components of non-selected reference frames can be inhibited (Carlson & Van Deman, 2008). In both of these studies, participants took part in a sentence verification task, in which they had to verify whether a presented sentence was an acceptable description of a picture (like the example in Figure 4.36). Carlson and Van Deman (2008) found effects of negative priming in which for example the use of front/back was inhibited when above/below/left/right were presented before. Their results suggest that during spatial language comprehension, participants activate multiple spatial representations based on different possible reference frames. The presentation of a specific reference frame expressed in language, inhibited the mental availability of the spatial representations that were not expressed.

In the present chapter, we focus on how the relationship between the use of reference frames in language and attention, whether language affects how participants look at spatial scenes and in turn whether the way participants look at special scenes corresponds with how they describe later scenes. We hypothesized that reference frames, expressed through spoken language, might affect how attention is allocated to a located object versus a reference object in a spatial array. In our study, we used three prime sentences to describe a spatial array with two objects (see Figure 4.35): intrinsic (*the marble is in front of the ladybird)*, relative (*the marble is to the left of the ladybird)*, and neutral (*the marble is blah blah the ladybird)*. Since intrinsic descriptions locate an object in relative to intrinsic axes of a reference object (e.g., ladybird) - whereas relative descriptions provide information about the two objects relative to each other - we predicted that participants in the intrinsic condition would fixate more on the reference object to process the intrinsic axes of the reference object, compared to the relative description condition.

If distinct eye movement patterns are present following the example description, we also wanted to test whether such eye movement patterns are "routinized" – that is if eye movement patterns are "replayed" when looking at a new

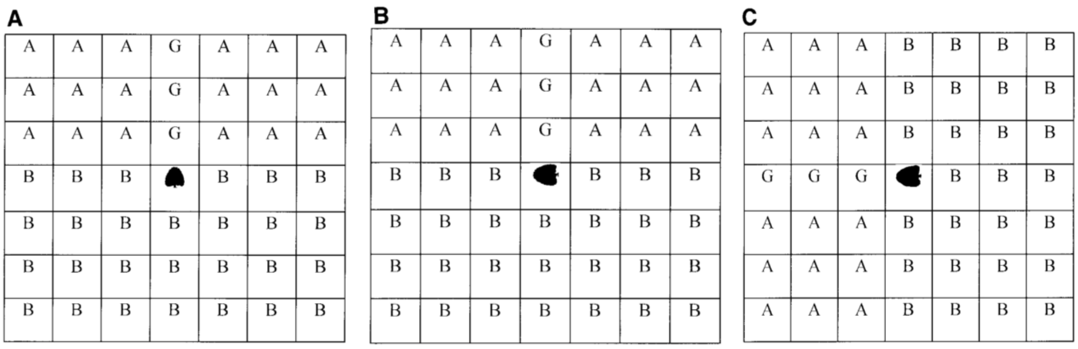


Figure 4.36. Participants judged whether *above the tree* was an appropriate description for the different squares in the grid. In A), the tree is upright, all squares directly above the tree are good (g), all squares higher than the 'tree square' are acceptable (a), and all squares in the grid lower than the tree are bad (b). In B), the tree is rotated. From an environment centered reference frame, the appropriateness of 'above the tree' is the same as in A). However, using a reference frame intrinsic to the tree, the appropriateness of the description changes, see C). Example by Carlson-Radvansky & Logan (1997).

scene prior to producing a new description. This would provide evidence that language drives eye movement patterns (i.e., visual attention), which in turn might drive future spatial description choice.

4.1.4 Dialogue

A second goal of Experiment 9, was to examine the extend to which the use of reference frames in language are aligned to the use of previous reference frame use, and explore one possible mechanism by which it does so: attention. Research has shown that interlocutors tend to align linguistic representations during conversation, for example, lexical and syntactic representations, but it has been much debated how this alignment works.

In conversation, answerers are more likely to use the utterance of the prepositional form of the question. For example, following the question ‘what time do you close’, responses are more likely to be ‘5 o’clock’, wherease a question like ‘at what time do you close’ is answered with ‘at 5 o’clock’ (Levelt & Kelter, 1982). Similarly, participants in a picture description task were more likely to use a similar syntax as a previous descripition uttered by a conspecific. When an interlocutor would describe a situation (e.g., “the cowboy giving the banana to the burglar”), participants would copy this sentence structure, and use a prepositional object (e.g., “the girl gave the book to the boy”) instead of a double object (e.g., “the girl gave the boy the book”) (Branigan, Pickering, & Cleland, 2000). This dynamic is also found when the events participants need to describe are unrelated to prime events (Bock, Dell, Chang, & Onishi, 2007).

It has been suggested that alignment is an automatic process (Garrod & Pickering, 2004; Pickering & Garrod, 2004, 2006). Pickering and Garrod proposed an interactive alignment account, in which production and comprehension are tightly coupled. The advantage of a system of automatic alignment is that it simplifies language processing. If a speaker repeats expressions previously uttered by a previous speaker, there is a very low cost of processing. For example in Garrod and Anderson (1987), participants communicated about mazes, in which boxes were connected by horizontal and vertical pathways (see Figure 4.37). Participants were each presented with their own maze (e.g., A and B in Figure 4.37), and had to communicate to get from their individual starting location (X) towards their goal-box. In this communication, participants quickly agreed on terms to use, without explicitly discussing their use. Garrod and Anderson concluded that this alignment is

a direct result of the need for joint action. For successful communication, interlocutors need to align their situation models (multi-dimensional representations, including dimensions of space, time, causality, intentionality (cf., Zwaan & Radvansky, 1998)).

For example, if one tells a story about two uncles, they need to make sure that the hearer has the right *situation model* for each uncle so that s/he understands which uncle is doing what in the story. Therefore, it is usually not useful to refer to the two uncles as 'uncle', but one could for example specify Uncle John and Uncle Steve. If the story would have been about only about one of the uncles, it would not be problematic to refer to him as uncle. Because interlocutors are aware of the interactive nature of dialogue, they appreciate the practical need for a joint effort to successfully communicate (Garrod & Pickering, 2004). This has to be an implicit process, because it is not feasible to explicitly agree on all concepts and terms ahead of a conversation while engaging in that conversation (Garrod & Anderson, 1987; Pickering & Garrod, 2004).

These effects are consistent with effects of priming or cueing, in which context, or a previous stimulus, activates a specific, related response on a succeeding stimulus. There is evidence that people align the spatial frames of reference they use

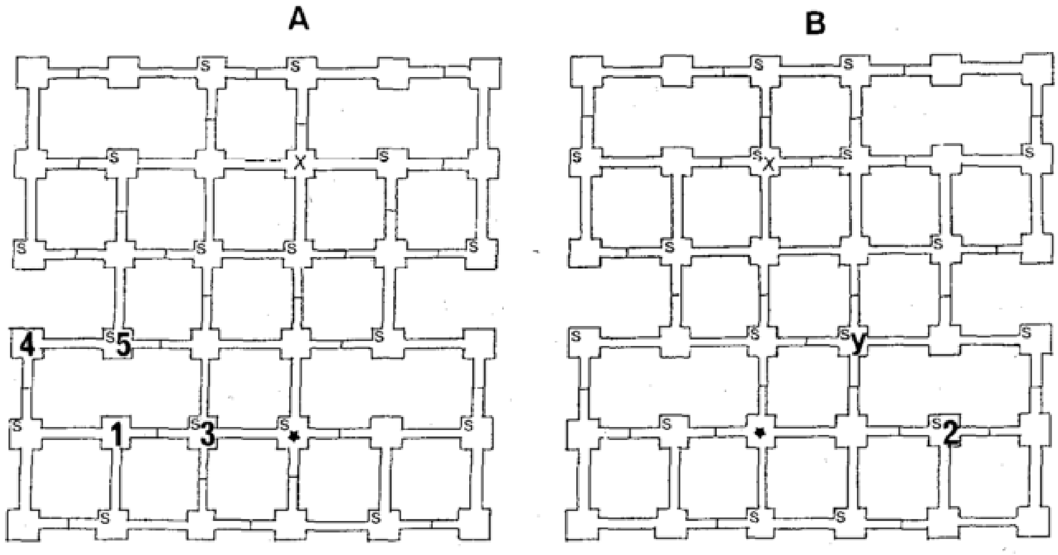


Figure 4.37. Overview of the maze participant A and participant B saw in Garrod & Anderson (1987). Participants started at the X and had to communicate to each reach their respective goal position.

when communicating about object positions (Carlson-Radvansky & Jiang, 1998; Dobnik, Kelleher, & Koniaris, 2014; Johannsen & De Ruiter, 2013; Johannsen & Ruiter, 2013; Schober, 1993; Watson, Pickering, & Branigan, 2004). For example, participants use spatial perspectives differently when they are interacting with people vs. imaginary addressees (Schober, 1993). In a simple task, participants would point out the location of one object relative to another. Interestingly, participants used fewer words and locative expressions, and more egocentric perspectives when talking to an interlocutor, compared to when they were solo speakers. In line with

Garrod and Anderson (1987), participant pairs usually kept using the type of descriptions that were first used. In a different study, response times where lower in a sentence verification task when a frame of reference at any given trial matched the frame of reference in the preceeding trial (Carlson-Radvansky & Jiang, 1998), supporting the account of interactive alignment.

Watson et al. (2004) tested whether interlocutors align reference frames, using a confederate priming paradigm (see Figure 4.38). The confederate manipulates the frame of reference s/he uses in specific trials, to explore the likelihood of participants copying the used frame of reference. It was found that interlocutors align reference frames when describing objects' locations; participants were 10% more likely to use the intrinsic frame if they heard a confederate use that

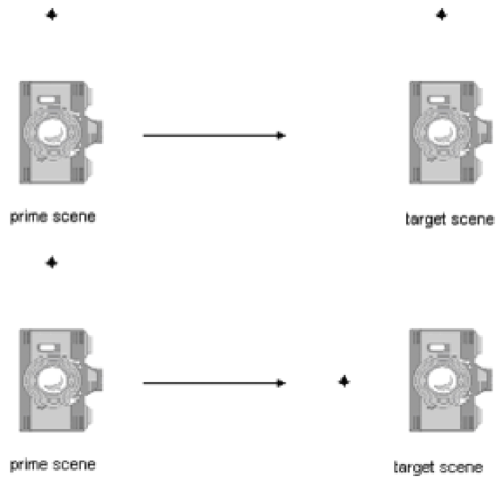


Figure 4.38. A participant and a conspecific were shown an array with a camera and a dot, each on their own screen. The confederate would give a description of the location of the dot in relation to the camera (e.g., "the dot above the camera" or "the dot right of the camera"). Participants then had to choose between two camera – dot pairs which of the two (left or right) fitted the description best. In the next trial, the participant would describe the situation, Watson et al. (2004)

reference frame on the preceeding trial, compared to when the confederate used a relative frame of reference. This effect was robust; the priming was still apparent when a speaker used a different preposition compared to a previous utterance from an interlocutor, showing it was not merely an effect of lexical alignment (Watson et al., 2004). In this type of priming, it was found that successful communication is a function of the extent to which interlocutors aligned their methods to disambiguate frames of reference (Johannsen & De Ruiter, 2013).

There are however theories that alignment effects are explicit and not based on automatic repetition (Brennan & Clark, 1996; Clark, 1996; Healey, Purver, & Howes, 2014). Moreover, speakers sometimes use references that their hearers would not agree to, enabling them to specify and control whether they used the same conceptualization, or situation model (Brennan & Clark, 1996). Only after interlocutors have established that they have a similar situation model do they tend to stick to a specific reference. This *interactive misalignment*, was supported in two large spoken dialogue corpora analyses (Healey et al., 2014). Healey and colleagues found that people tend to repeat syntactic constructions below chance level, and that people instead are more likely to diverge from one-another in syntactic constructions, enabling speakers to clarify misunderstandings (Healey, 2008; Healey et al., 2014). Healey and colleagues suggested that successful communication is dependent on the ability to selectively repeat interlocutors' expressions. In multiple studies, including Garrod and Anderson (1987), results show progression in the use of descriptions, in which initial descriptions are dropped after participants find better ways to describe locations (Mills & Healey, 2008). This is consistent with Brennan and Clark's (1996) suggestion that interlocutors use deviating descriptions to confirm they use the same situation model.

Combining the findings that people might align or mimic during successful conversation, and that language can drive visual attention, we hypothesised that if participants hear a different frame of reference describing a prime trial, they mimic this description and use a similar a reference frame in the succeeding probe trials. Previous studies using a confederate-priming paradigm typically employ a within-participant design (Branigan et al., 2000; Watson et al., 2004). However, following Firestone and Scholl (2015), participants may understand the aims or manipulation of studies. Participants in these studies go through multiple trials, they might be able to identifiy the manipulation that is used in those trials - the manipulation is the only thing that changes between trials. Therefore, it needs to be tested upon debrief whether participants had an understanding of the manipulation and/or aims of the study. If they do, participants might respond based on their beliefs of what the study is about or what the experimenter wants them to do. To prevent this from happening, we choose to use a between-participants design. Since participants only see one prime-trial, they cannot compare between trials, decreasing the likelihood of a participant discovering the manipulation and aim of the study.

Our study aims to text one potential mechanism driving alignment, attention. Eye-movements are a potential realisation of automatic alignment. Our main aim in this study was to examine influence of prior linguistic descriptions on eye-movements in a spatial description task. Participants were instructed to describe the spatial relation of two objects (e.g., the marble is to the left of the ladybird). They were presented with a prime-trial, in which either an intrinsic (in front of/behind), a relative (to the left/right), or no reference frame (blahblah) was expressed in language. It is important to note that our study, like in some of the reviewed papers, has a limitation regarding the free language use of participants. Since the prime-trial was prerecorded, there was no actual interaction and the instructions did prevent participants from using free spoken dialogue.

4.2 Experiment 9: The influence of frames of reference as expressed in spatial language on visual attention

The study we present in this chapter explores the effect of language (expressing frame of reference) on visual attention. We presented simple spatial scenes containing two objects (e.g., a marble and a ladybird), and primed participants with a verbal description, either matching an intrinsic ("the marble is in front of the ladybird") or relative ("the marble is to the left of the ladybird") reference frame, or a neutral description ("the marble is blahblah the ladybird"). We predicted that hearing an intrinsic description would draw attention towards the reference object, while hearing a relative description would lead to attention being divided more equally between located and reference objects.

4.2.1 Method.

4.2.1.1 Participants. Participants were 84 students from the University of East Anglia (56 female, age range 18-50), who took part for course credits or monetary payment.

4.2.1.2 Materials and Setup. Participants sat in front of the eye tracker. On the screen, simple spatial arrays consisting of two objects were shown (for the object-pairs used in the probe trials, see Figure 4.39). Every object had four interest areas (IA) assigned (Figure 4.399). An IA for the whole object, and a further division into three equal parts (left, middle, right). This allowed us to monitor both the objects and the areas within the objects participants were looking at prior to producing their spatial descriptions.

Three different versions were used in an example trial: a relative reference frame (“The marble is to the left/right of the ladybird”), an intrinsic reference frame (“The marble is in front of/behind the ladybird”) or no reference frame (neutral condition: “The marble is ‘blah-blah’ the ladybird”). After the example, prime trial, there were six probe trials with different spatial scenes. The participants verbally described each array while their response was recorded and saved as a .wav file.

Eye movements were recorded with an SR research Eyelink 1000 eye tracker sampling at 1000 Hz (SR Research Ltd., Ontario, Canada). Viewing was binocular, but we only tracked the position of one eye per participant. Stimulus presentation was programmed using SR research Experiment Builder software. The eye tracker and a 19” CRT display monitor (refresh rate of 140 Hz) were interfaced with a 3-

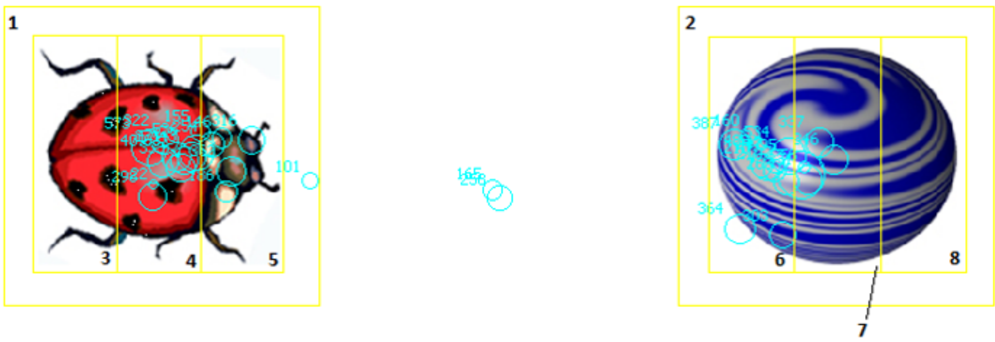


Figure 4.39. Interest areas of the reference and located object. IA 1 and 2 are the overall interest areas for respectively the reference and located object, fixations in these IAs are analysed for hypothesis 1. In this example, IA 3 and 5 are the front and back of the reference object, which were analysed for hypothesis 2.

GHz Pentium 4 PC, which controlled the experiment and logged the position of the eye throughout the experiment. Throughout the task, participants used a chinrest.

4.2.1.3 Procedure. At the beginning of the experiment, we used a nine-point sequence to calibrate and validate eye position. Instructions were presented to the participants both verbally (spoken by a native English narrator) and visually on the screen. Then we presented the objects used in the study so that participants would know the labels for each object, and could easily identify the objects when describing them. At the start of the experiment an example trial was presented, one of four combinations (intrinsic: in front of/ behind, relative: to the left/ right) of the two objects in Figure 4.399, followed by six probe trials. The order of trials was pseudo-randomized across 12 different lists. Participants were asked to look at each picture, wait for the question (Where is object A in relation to object B), and then answer out loud with a full sentence.

The design consisted of a single variable with 3 levels (example trial: intrinsic, relative, neutral), and it was manipulated between participants. The dependent variables were the IA participants looked at on the example and the probe trials, and the type of reference frame produced by participants in the probe trials.

4.2.2 Results.

We first examined whether the linguistic example affected which reference frame participants chose across the six probe trials. We categorized participants' responses for each probe trial as using either the intrinsic frame or the relative frame. We analysed the frequency of intrinsic and relative frame terms for all six probe trials separately. Table 4.18 shows the percentage of participants using each reference frame for each probe by condition. Note that all participants in this experiment used either an intrinsic or a relative description (two participants for which the microphone failed to record the response are eliminated from this analysis). There is a main (between participant) effect of condition, $F(1,79) = 16.775$, $p < .001$, $\eta p^2 = .298$, showing that participants in the intrinsic condition used the intrinsic reference frame more often that participants in both the relative and neutral condition. There was also a main effect of trial, but no interaction between trial and condition, $p = .26$, $\eta p^2 = .03$. The data shows that the trial effect is driven by the relative and neutral condition, in which participants start using the intrinsic condition more over time, although they never use it in an extent comparable to the intrinsic condition.

We next examined the eye movement data. First, we examined fixations on the example trial. There were more fixations on the reference object in the intrinsic condition following the onset of the prepositional phrase compared to the relative and neutral conditions (see Figure 4.40). In order to examine the time spent attending the reference and located objects, we summed the fixation durations for the reference and located objects separately in the example and probe trials. In the example trial we calculated from the onset of the question (*"Where is the marble in relation to the ladybird"*) until the end of the trial; in the probe trials recording started at the onset of the question until the beginning of the response (onset of spatial description). These data were analysed with a 7 (trial: primer trial and 6 probe trials) × 3 (condition: intrinsic, relative, neutral) × 2 (object: reference, located) mixed model ANOVA. This produced a main effect of object, $F(1,81) = 19.37$, $p < .001$, $\eta p^2 = .193$. Overall, more time was spent looking at the reference object (*M* = 51%, *SE* = 1.7%) compared to the located object (*M* = 37.5%, *SE* = 1.7%). The interaction between condition and object was also significant, $F(2,81) = 3.3$, $p = .042$, $\eta p^2 = .075$. More time was spent looking at the reference object in the intrinsic condition (*M* = 57.6%, *SE* = 3.3%) compared to both the neutral (*M* = 46.4%, *SE* = 2.2%) and relative (*M* = 49% *SE* = 3.4%) conditions. We ran a separate 7 (trials) × 3

Table 4.18. *Percentage of intrinsic (Int) and relative (Rel) frame use for each probe for each condition.*

	Probe 1		Probe 2		Probe 3		Probe 4		Probe 5		Probe 6	
Primer	Int*	Rel**	Int	Rel	Int	Rel	Int	Rel	Int	Rel	Int	Rel
Intrinsic (21)	76%	24%	86%	14%	86%	14%	86%	14%	86%	14%	86%	14%
Relative (19)	16%	84%	21%	79%	37%	63%	37%	63%	37%	63%	32%	68%
Neutral (42)[11]	19%	81%	19%	81%	24%	76%	31%	69%	31%	69%	38%	62%

* Int. = intrinsic; ** Rel = relative

[11] We planned to add more participants to the neutral condition in order to get meaningful group sizes of participants spontaneously choosing the intrinsic and the relative reference frame. However, the ratio of intrinsic : relative reference frame response was about 1:4, so that we got convinced that it was not feasible to add participants until a meaningful sample of spontaneous 'intrinsic' responses was reached.

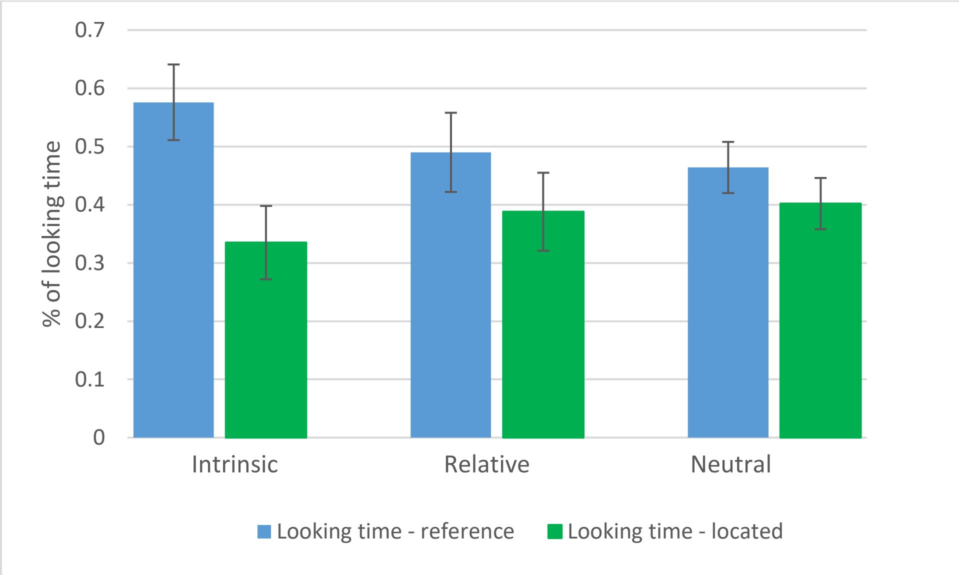


Figure 4.40. Interaction between condition and object with respect to looking time. Error bars represent 95% confidence intervals.

(condition) ANOVA over fixations on the reference object, to test this difference between conditions. There was a main effect of condition, $F(2,81) = 3.989$, $p = .022$, $\eta p^2 = .09$, but no main effect of trial $p = .657$, $\eta p^2 = .008$ and no interaction between trial and condition, $p = .794$, $\eta p^2 = .016$.

This analysis shows that looking behaviour across conditions was consistent across example and probe trials, consistent with the view that eye movement patterns are maintained from the linguistic example across all probe trials. As the data on the probe trials are prior to the description produced by participants on each of these trials, these findings suggest that eye movement patterns are "replayed" across different and future visual scenes, which then likely influences future spatial descriptions.

Table 4.19. Results of the ANOVA on eye-movement data

Source	***df* and *F* value**	**MS**	**Significance**	**ηp²**
Trial* (T)	F(6,486) = .730	0.022	0.606	0.009
T × condition*	F(612,486) = .214	0.006	0.996	0.005
Object (O)	F(1,81) = 19.370	4.666	<.001	0.193
O × condition	F(2,81) = 3.300	0.795	0.042	0.075
T × O	F(6,486) = .778	0.15	0.587	0.01
T × O × condition	F(12,486) = .664	0.128	0.786	0.016

*Greenhouse-Geisser

Finally, we analysed the amount of time spent looking at the parts of the reference object, separating areas that are diagnostic (front and back) vs. not diagnostic (centre) of object orientation. In order to do so, we summed the fixation durations at the front (one-third of the object) and back (one-third of the object) of the reference object and subtracted the time spent looking at the middle (one-third of the object). A 3 × 7 (condition: intrinsic, relative, neutral) × (trial: example, probes 1-6) mixed model ANOVA showed no main effect of trial, p = .241, ηp^2 = .016, no main effect of condition, p = .211, ηp^2 = .038, nor an interaction between trial and condition, p = .139, $\eta p2$ = .036. However, relatively more time was spent looking at the front and back of the reference object in the intrinsic condition (*M* = 53.2%, *SE* = 7.7%) than in either the relative (*M* = 36.5%, SE = 8.1%) or the neutral (*M* = 37.7%, SE = 5.3) conditions.

4.2.3 Discussion.

We found effects of linguistic example on reference frame choice as expressed in language, robust across probe trials. In the neutral condition, there was a strong preference to use the relative frame. After an intrinsic prime, 84% of probe trials had an intrinsic description, compared to 27% in the neutral condition, and 30% in the relative example condition. This represents a considerable influence of previous description on reference frame selection, persisting with impressive consistency across the extended duration of the experiment. More informative is the eye movement data regarding how participants looked at visual scenes following the example description and prior to their own descriptions across the probe trials.

With respect to looking times on the reference and located objects, the data provide the first evidence that reference frames, expressed in language, are associated with different visual attentional patterns. Across all the trials (example trial and six probes) more time was spent looking at the reference object than the located object following the offset of the prepositional phrase (diagnostic of reference frame) when the example trial contained a preposition denoting an intrinsic reference frame compared to a relative or neutral frame. This is consistent with earlier work showing more coarse grained effects of prepositions on object reference using the Visual World Paradigm (Chambers et al., 2002), and is also consistent more broadly with a range of studies showing the rapid interplay between language and visual attention (for a review, see Henderson & Ferreira, 2004). It is particularly striking that the looking time behaviour was consistent across the example trial and six probe trials, just as the verbal responses were consistent across trials.

The analyses of looking behaviour to parts of the reference objects did not produce reliable results. This may well be a result of the size of objects used. The mean size of objects on the screen was 5 cm × 5 cm, so participants could apprehend the entire object within para-foveal vision, and thus, eye movements to left and right sides of the object were unnecessary to orient front and back. Future studies could increase the size of the objects to investigate how participants establish object axes across multiple fixations.

4.3 General Discussion

We examined effects of linguistic descriptions on the types of reference frames people use when describing simple spatial scenes. Our data are consistent with the view that language draws attention to the visual world in specific ways, consistent with other visual world paradigm studies. We consider the mechanisms involved below, and implications for theories of language production more generally.

While sizeable increases in intrinsic frame use were following an intrinsic description of a spatial scene, it is notable that in the neutral condition there was a preference to use relative descriptions (71% across probes). There is considerable disagreement regarding general preference for reference frames in the literature. For example, some authors have argued that the ease of perceptual availability and reduction in computational effort makes the relative frame dominant/most preferred (e.g., Levelt, 1982; Linde & Labov, 1975), while others argued that the intrinsic frame is preferred (Carlson-Radvansky & Irwin, 1993; Miller & Johnson-Laird, 1976). Discourse and task context effects can have a strong influence (cf. Tenbrink, 2011). Indeed, reference frame choice has been shown to be affected by a range of situational influences, including the embedding of the objects in more complex and real-world scenes (Johannsen & De Ruiter, 2013a), and the communicative context in which the speaker is situated (Galati & Avraamides, 2013).

Our results provide the first empirical evidence that more time is spent looking at the reference object following an intrinsic description compared to when a neutral or relative description is given. Reference frames have received much attention in the field of spatial language and spatial cognition, but thus far, eye-tracking data regarding reference frames has not been available. Not only are reference frames theoretically differentiable (cf. Levinson, 1996), but it would seem that they are associated with differential allocation of attention to the reference and located objects when one looks at a spatial scene following the use of a reference frame in a spatial description by an interlocutor.

The novel and arguably most important finding in this study was that the same pattern of looking time differences occurred for each of the probe trials, prior to participants' descriptions. Moreover, there was no interaction between trial (example trial, probes 1-6) and object (reference, located), which indicates that the same pattern held between the example and probe trials. These results suggest effects of priming. When participants were primed in a specific linguistic frame of reference, they were more likely to align to that reference frame. The lack of an interaction effect shows that participants did not switch between reference frames during the probe trials. In terms of the interactive alignment or misalignment models, our participants performed eye-movements consistent with the reference frame that was used and re-played these eye-movements when they were asked to describe the object arrangements themselves. In the interactive misalignment account one might hypothesise that participants perform different eye-movements. If alignment is based on a mechanism of mis-alignment, participants would be expected to test other possible descriptions upon hearing an example trial. However, the interactive misalignment model is based on interaction. Since dialogue and monologue have been shown to be different skills, it is important to recognize this limitation (Pickering & Garrod, 2004; Schober, 1993; Watson et al., 2004).Therefore, the lack of explorative eye-movements to consider other frame of references could be due to the lack of an interlocutor. The misalignment might be useful in interaction, but when there is no interaction, there is no reason to test whether a hearer has the same situation model, as there is no collaboration or end goal. A potential follow-up study to tease the interactive alignment and interactive misalignment accounts apart should combine the strengths of both types of study, including a confederate for actual interaction and a between-participant design.

However, we know that language draws attention to the world in specific ways, and in turn, that looking at the world in a particular way affects choice of syntactic structure (e.g., Gleitman et al., 2007). Rather than regarding linguistic descriptions and visual cueing as independent and different parameters, one can argue that there is a close interplay between language and visual attention, such that they support each other to maximise alignment between interlocutors. This leads to an increased likelihood of the same type of description being used as that used on the example (prime) trial.

Finally, we can return to the range of influences on choice of language considered at the outset of this paper. Consistent with work on the influence of visual (cued) attention on syntactic structure (Gleitman et al., 2007), we have suggested that drawing attention to the intrinsic frame either through an intrinsic prior (example) description may direct visual attention to the reference object, increasing the likelihood of producing an intrinsic description compared to no-prime or relative-prime conditions. This provides a parsimonious approach to the effect of multiple influences on spatial description, starting with the assumption that language and visual changes can direct the attention of the speaker in similar ways. It also affords a means to test whether a possible strategic route to spatial description, where for example people deliberately choose to ignore past information (systematic misalignment in dialogue; Healey, 2008), results in overriding visual attentional patterns (akin to dual process models of semantic priming; see Mummery, Shallice, & Price, 1999), or inhibition of the influence of past spatial description in the first instance.

In summary, we have presented the first evidence (to our knowledge) for distinguishable looking behaviour patterns as a function of reference frames expressed in language. Using between-participants designs where past information can be systematically manipulated immediately prior to a probe trial description provides a clean way of testing how language choice is affected by multiple constraints.

SECTION 3

-

General Discussion

Chapter 5 – General Discussion and Conclusion

5.1 Chapter overview

The experimental work outlined in this thesis aimed to investigate the influence of language on cognition and perception, specifically the influence of spatial language on memory for object location, and the influence of language expressing reference frames on visual attention. In the present chapter, we will first summarize the main findings presented in this thesis, before considering how these findings advance theory within the field. Lastly, we discuss directions for future research.

5.2 Summary of Results

5.2.1 Chapter 2

In the first series of experiments (see Chapter 2), we set out to test whether language affects memory for object location and to elucidate the mechanism involved. In our studies, we teased apart three different models, each having different predictions of the mechanism via which language might influence spatial memory: the Expectation model, the Congruence model, and the Attention allocation model. To do so, we used spatial demonstratives as a vehicle to test spatial memory using the memory game procedure (Coventry et al., 2008).

We teased apart models explaning the mechanism underlying the influence of language on perception: Expectation, Congruence, and Attention. Our results favoured the Expectation model as an explanation as to how language influences spatial memory. The Expectation model is consistent with models of predictive coding, suggesting that people continuously predict future states in the world (cf., Clark, 2013). The Expectation model (Coventry et al., 2014) suggests that spatial language can elicit a specific expectation about a scene. *This* and *my* have a proximal predictive value, such that objects referred to with these words are linguistically encoded to be closer by compared to objects referred to with implicitly distal terms like *that* and *your* (e.g., Clark & Sengul, 1978; Coventry, Valdés, Castillo, & Guijarro-Fuentes, 2008; Diessel, 2006). The Expectation model takes these predictive values and predicts that, encoding objects with distal language will distort memory in a way that objects are misremembered to be further away, compared to objects encoded with proximal language terms.

In the Congruence model, it is not so much the language, but the congruency of language and space that influences memory. Research on congruency found that, for example, action performance is quicker and more accurate in tasks in which a spatial scene and spatial language are congruent compared to an incongruent situation (cf. Bonfiglioli et al., 2009; Stevens & Zhang, 2013). Following the Theory of Event coding (Hommel, Musseler, Aschersleben, & Prinz, 2001), suggesting that perception and action share an indistinguishable underlying representational medium, we extend these findings to spatial memory. If memory is more accurate when language describing a spatial situation is congruent with that situation, then memory for object location should be more accurate in trials in which an object is placed closer to a participant combined with *this* or *my* compared to memory for objects placed at the same locations combined with *that* or *your*. Similarly, memory should be more accurate for objects placed further away if the placement is verbally encoded using *that* and *your* compared to *this* and *my*. The crucial difference between the Expectation model and the Congruence model is that the latter predicts a cross-over interaction. Our results in Chapter 2 all showed a main effect, but no cross-over interaction, thereby supporting the predictions by the Expectation model. (In Experiment 3, we did find an interaction, but the interaction showed objects placed with *your* were misremembered less accurately when placed further away, an effect opposite to the prediction the congruence model would make). Language did have an effect on memory for object location, in which objects placed with *that* were systematically misremembered to be further away than objects placed with *this*, irrespective of distance.

The third model we tested was the Attention allocation model. Since language focuses attention (Allopenna et al., 1998; Tanenhaus et al., 1995) and longer looking times are associated with better memory performance (Huebner & Gegenfurtner, 2010), objects placed under language conditions eliciting a stronger focus of attention will be remembered more accurately. People may spend more time looking at objects in closer proximity (Garrido-Vásquez & Schubö, 2014), a system that could be activated by proximal language. This could be consistent with the schema of focus for demonstrative reference (Figure 2.25), in which *this* allocates a higher focus of attention to a referent compared to *that* (Strauss, 2002). Alternatively, the definiteness of demonstratives compared to, for example, determiners means that demonstratives focus attention of a conspecific more strongly to a referent (cf. Peeters & Özyürek, 2016). A model based on focussed attention, which we coined the Attention Allocation model, would have similar behavioural predictions as the Expectation model. If people spend more time looking at objects that are linguistically encoded to be closer by (e.g., *this, my)*, compared to objects linguistically encoded distal (e.g., *that, your)*, memory performance should be more accurate in the former and lead to a main effect of language. Furthermore, as participants tend to overestimate object location in memory (cf. Coventry et al., 2014), the attention effect would have the same direction as the Expectation model. The difference is that in the Expectation model, language directly affects memory, whereas in the Attention allocation model, memory performance is affected by language via the amount of attention they elicit. If demonstrative pronouns (*this/that)* have a stronger effect on focus of attention compared to an indefinite referential expression (*the*), the data would show a difference between demonstratives and the determiner. To test the attention allocation account, we had participants engage in the memory game while wearing eye-tracker glasses. We found no difference in fixation times between the language conditions (although we did not find evidence there is no attention effect based on language). This suggests that attention is not driving the main effects that we found across experiments.

In summary, results overall supported the Expectation Model. The driving mechanism of the Expectation model is consistent with accounts of predictive coding (Clark, 2013; Friston, 2005) and more broadly with the idea that prediction plays a major part in cognition. The theory of predictive coding suggests that people continuously predict possible future states based on current perceptual input combined with learned associations (Friston, 2009). The brain prepares a response based on this prediction and only needs to process the *error signal*, the difference between the prediction and the updated visual input, once the new state of the world emerges. Adjusting the prediction according to the error signal costs less energy than responding to surprises. Therefore, Friston (2009) coined this the free-energy principle. The associations that afford prediction might be learned via the principle of statistical learning (Newport & Aslin, 2004; Pulvermüller, 2012; Saffran et al., 1996) – by learning co-occurrences of words and meaning or words and situations (see also Andrews et al., 2009).

Our results are also consistent with a probabilistic inference model, as discussed previously (Cibelli et al., 2016; Huttenlocher et al., 1991, 2004; Regier &

Xu, 2017). Probabilistic inference models assume that memory systems are based on two tiers, the perceptual, ‘fine-grained’ representation of an object or event, and the more general category to which an object or event belongs. During memory retrieval (retrieval in itself resulting in lability of the memory trace), these two tiers merge, in which the strength of each tier is dependent on the amount of uncertainty in the ‘fine-grained’ representation. That is, one is more certain about the fine-grained representation, the more general category knowledge has a lower impact. However, the more uncertainty there is on the ‘fine-grained’ representation, the stronger the category effect gets. This uncertainty can be caused by fading memory, fatigue, or any other factor inducing uncertainty in the mental representation. In our study, the linguistic encoding can elicit a specific categorical-representation of an object. If the meaning of demonstratives (*this, that*), is based on the numerous examples of situations in which people use demonstratives - as statistical learning theory suggests (Pulvermüller, 2012) - their meaning and therefore their predictive value is based on this common use, in which *this* refers to objects close by a speaker, or in a setup with two or more objects, *this* refers to the object closer by the speaker than other objects. Upon retrieval, this category could influence the memorized location of an object.

Our results show that memory can be influenced by seemingly small differences in the spatial linguistic description of a situation. However, there is a myriad of ways in which the spatial world can potentially be carved up, and there are 6,000-8,000 languages used around the world that do so in different ways (Evans & Levinson, 2009; Landau & Jackendoff, 1993). For example, more than a third (37.4%) of the worlds’ languages employs a three-way demonstrative system (Diessel, 2005). This variation opens up the possibility that different demonstrative systems may pose different predictions about object location in memory. In Chapter 3, we tested the Expectation model in a language in which the use of the demonstrative system is more complex: Japanese.

5.2.2 Chapter 3

To explore potential limits of the Expectation model, we tested Japanese speakers, since Japanese employs a three-way demonstrative system (based on distance and/or position of interlocutors), opening the possibility of more complex demonstrative use and therefore a different influence of language on memory for object location. Over four experiments, our aim was twofold: we wanted to establish the parameters Japanese demonstratives use to encode space, and second, we wanted to find out if and how Japanese demonstratives affect memory for object location. In addition to the two Japanese studies we ran two English studies, to replicate earlier findings (Coventry et al., 2008; Gudde et al., 2016, Chapter 2) and to afford comparison of the English and Japanese data.

We identified four different theoretical models describing the Japanese demonstrative system. The first account is the distance model, suggesting that the three Japanese demonstratives (*kono, sono, ano*) encode egocentric distance of the referent from the speaker; *kono* for close to speaker, *sono* for medium distance, and *ano* for far distance (e.g., Ootsuki (1889), as cited in Nakamura, 2012). The second explanation is a person-centred account, in which demonstratives encode the interlocutor's territory in which the referent is located; *kono* for close to the speaker, *sono* for close to the hearer, and *ano* for far from both (Sakuma, 1936; Sakata , 1971, as cited by Hasegawa, 2012; Hattori, 1992; Niimura & Hayashi, 1994; Sakata, 1992). The two other possibilities suggest a combination of these two models. Model 3, the dual-system model suggests that demonstratives encode distance or territory as a function of the spatial configuration of interlocutors. When hearer and speaker share a territory, demonstratives encode distance from this shared territory. When hearer and speaker do not share territory, demonstratives encode territory (Kamio, 1994). A variation on this dual-system model is that *sono* can refer to a *interactional* territory instead of the hearers' territory (Yoshimoto, 1992, 1986). In this case, *sono* would not be used for objects close to the hearer, but for objects within the interactional dyad, and objects out of the interactional area, further from interlocutors, would be encoded with *ano.* The fourth model is the double-binary model in which the three demonstratives form a double-binary contrast. For example, *kono* and *sono* can contrast space, *kono* and *ano* can contrast specific territories, but *sono* and *ano* are not a contrasting pair (Hasegawa, 2012).

We found that Japanese demonstrative use is based on both *distance* and *territory*. This rejects the influence of either parameter on its own (Model 1 and 2). Furthermore, our data rejects the double binary model (Model 4) – most our participants used all three demonstratives, which is impossible in the double-binary model. Our results were consistent with the dual-system model (Model 3) with the idea of an interactional space. Participants used *kono* almost exclusively for near locations, and near locations were almost exclusively referred to with *kono*. *Ano* was the preferred demonstrative for objects out of reach when speaker and hearer shared territory, but when the hearer sat opposite the speaker, the use of *sono* and *ano* was about 50/50. This was irrespective of whether the object was within or out of interlocutors' reach. Therefore, the results were most consistent with the *interactional space* dual-system model: *sono* was used irrespective of the objects' location relative to the interlocutor, and was used for objects located at all distances out of participants' reach, but between participant and conspecific. Note that in our study, when participant and interlocutor sat opposite each other, all objects were within their interactional space. Therefore, using this specific setup, we cannot see whether encoding is different for objects within this interactional space and out of interactional space. The data also showed a gender effect, in which agent affected male participants, whereas agent did not affect female participants.

The English results replicated previous findings, in which there was a main effect of agent and of location. Interestingly, results revealed an interaction between location and position, in which *that* was used less frequently for objects placed in the near space of the conspecific. This is of note, since English does not explicitly encode the position of an interlocutor. However, these results are consistent with Coventry et al. (2014), who found that other parameters which are not explicitly encoded in English do nonetheless influence English demonstrative use (e.g., ownership, visibility, familiarity). These results are important because they are inconsistent with the perspective of cross-linguistic differences, but support the perspective of universal systems of demonstrative pronouns.

Turning to the influence of language on memory, data showed the increased complexity of Japanese demonstrative encoding. Whereas the English data replicated previous findings (both in the language and the memory experiment), showing a main effect of language but no interactions with space, the Japanese results were less clear. There were no main effects of language or space, but data showed an interaction of demonstrative and space, and there was an interaction of conspecific, demonstrative, space, and gender. The gender effect is interesting, as it has not been found before (we did not find a gender effect in the English memory studies in Chapter 2 nor 3).

In the analysis of male participant data, we found an interaction between demonstrative and space when participant and conspecific were seated side-by-side, indicating an effect of language. This effect disappeared in trials in which the conspecific was opposite the participant. These results are consistent with the dual-system model, in which demonstratives encode distance when hearer and speaker share territory, but territory when they do not. Female data showed an interaction between demonstrative and space, but no three-way interaction with position. An explanation for the difference between male and female participants in Japanese could be that evolutionarily, males (across species) are more territorial. The differences between English and Japanese could be explained by the finding that the Japanese society scores higher on masculinity (Hofstede, 1980, 1998). The potential for cross-cultural differences could be picked up in future research.

Overall, in Chapter 2 and 3 we have seen demonstratives encode distance from a speaker and can affect memory for object location: objects placed with *that* are misremembered to be further away than objects placed with *this.* Furthermore, we found that a parameter that is explicitly encoded in Japanese, but not in English - position of a conspecific – does influence English demonstrative use. This finding supports a universal view of demonstrative systems. It is important to note that the effect was stronger in Japanese than it was in English, which might suggest that the weighting of parameters varies across languages, so that parameters that are explicitly encoded have stronger effects.

5.2.3 Chapter 4

In Chapter 4, we explored the influence that reference frames have on attention allocation. We set out to determine whether there are differences in how people look at the world based on linguistic cueing - the use of different reference frames in the description of a specific spatial scene. To do this, we ran an eye-tracking study, presenting simple spatial arrays of two objects. We found that participants show different patterns of eye-movements when they are hear spatial descriptions employing different reference frames, and that these patterns persist throughout following trials. To our knowledge, it is the first time that specific patterns of eye-movements are found as a result of the use of different reference frames.

In our study, we cued participants with an auditory description of a simple visual array (e.g., “the marble is in front of the ladybird”). During the succeeding probe trials, we presented different object pairs, and recorded participants’ eye-movements and their verbal description of the scene. Results showed distinctive eye-movement patterns based on the reference frame that were used in the prime trial. When the intrinsic reference frame is used (e.g., *"the marble is in front of the ladybird")*, participants paid most attention to the reference object. This is consistent with our hypothesis that the reference object requires the most attention in order to confirm the intrinsic axis (front/back). In contrast, in the relative condition we found attention more equally divided between both objects. Perhaps the most important finding however is, that these fixation patterns occurred for each of the probe trials, "prior" to participants' descriptions. There was no interaction between trial (example trial, probes 1 – 6) and object (reference, located), indicating that the fixation pattern held between the example and probe trials. This offers a possible mechanism to explain the linguistic alignment effects: language draws attention to the world in specific ways, leading to looking at a new spatial scene in the same way. In turn, looking at the world in a specific way leads to the increased likelihood of the same type of description being used as was used on the example (prime) trial, consistent with the effects of endogenous and exogenous visual cueing on choice of syntactic structure (e.g., Gleitman et al., 2007). One can argue that there is a close interplay between language and visual attention, such that they support each other to maximize alignment between interlocutors.

Consistent with the work on the influence of visual (cued) attention on syntactic structure (Gleitman et al., 2007), we suggest that drawing attention to the intrinsic frame through an intrinsic prior (example) description may direct more visual attention to the reference object compared to the relative reference frame, which increases the likelihood of the alignment in reference frame use (Watson et al., 2004).

5.3 Methodological concerns

A common pitfall in psychological research is the trade-off between ecological validity and control: the more experimental control, the lower the ecological validity, but the higher ecological validity, the lower the experimental control. This may especially affect the language production experiments in Chapter 3. The dynamic environment in which language usually takes place is limited in a lab environment. Therefore, the ecological validity of these specific studies might have been traded for increased experimental control.

Previous research testing spatial demonstrative production or comprehension of demonstratives ranges between observational research with high ecological validity but low experimental control (e.g., Enfield, 2003), testing participants' intuitions about demonstratives by asking which words they would use in hypothetical situations (e.g., Wilkins, 1999), and studies with high experimental control but low ecological validity, such as designs in which congruence of demonstratives and pictures is tested (Peeters et al., 2015; Stevens & Zhang, 2013). As we know that there are many different possible parameters on which language use is based, it is impossible to acquire data in real life environments while controlling for all these variables. In our experiments, we tested participants' language use in a controlled lab-environment. This experimental control might however have influenced the ecological validity. When people have to choose their language, it is possible the controlled lab-environment limited the dynamic characteristics of language use that would occur in more naturalistic settings. For example, we cannot differentiate between distances on a scale bigger than our table using our lab setting. It has been suggested that the three-way distinction in Spanish demonstratives (este, ese, aquel) is difficult to capture in a lab setting, as aquel is used to refer to objects that are beyond a lab-environment. In that case, the use of aquel might be adapted to contrast between distances in a lab-environment, but the environment might limit the more natural use. On the other hand, the memory game procedure has shown to be a good procedure to tease apart other parameters (ownership, visibility, familiarity, language) under high experimental control.

Some specific limitations occur in Chapter 4. For example, the 'interaction' participants experienced is subject to the same ecological vs. experimental trade-off. The verbal cueing and questions were pre-programmed instead of uttered in an actual two-way interactional setting as one would experience in everyday life. This could affect the use and interpretation of language (Pickering & Garrod, 2004; Schober, 1993; Watson et al., 2004). This means that in terms of dialogue, the results might not be generalizable to free-dialogue. On the other hand, the between subject design does ensure that participants do not become aware of the aim of the study, which could increase the ecological validity. A procedure that could combine the strengths of both studies, for example an interactive design including a confederate and eye-tracker glasses, might solve the limitations of both types of studies.

A second potential issue in Chapter 4, as discussed previously, has to do with our hypothesis that participants would fixate more on the front/back of the reference object compared to the middle; since the front and back are the areas that provide information on the intrinsic axes. Even though we found evidence based on an arguably crude analysis of fixations on the reference object or the located object, we could not find a difference of fixations within the interest areas of the reference objects. A methodological issue that could explain this is that the figures were relatively small. That means a participant could see the entire reference object within peripheral view if s/he fixated on any point on the reference object. Therefore, s/he did not need to shift fixations across the object to determine the direction of the intrinsic axis. Future studies could be run using larger objects in which participants cannot see the entire object within one fixation.

More generally, the trade-off applies to the memory data in Chapter 2 and 3 as well, although one could argue that to test the specific influence of languages on spatial memory, the lab-environment made the results cleaner. Since participants were aware of the spatial memory task they were performing, it would be expected that memory performance is more accurate in a controlled environment compared to the ‘real world’. Following the suggestion that memory for object location is a combination of the actual location and contextual information (Coventry et al., 2014), the actual location memory should be stronger in an experimental setting, leaving less influence of contextual information. The fact that we are still able to find effects of language on memory could therefore imply that the effect is stronger in real-life environments. In real life, people do not necessarily focus on exact object location. Therefore, one would expect that real-life memory for object location suffers from more uncertainty, compared to when a participant knows s/he will be tested on fine-grained spatial memory.

Another issue is the use of eye-tracker glasses in Experiment 4. We did not find any indication that languages have a different influence on fixation, but we did not find evidence that it does not. It is possible that different demonstratives elicit different reference frames or searching strategies when one is encoding a spatial memory, but that the ability of the current generation of eye-tracker glasses does not provide the fine-grained measurements needed to observe these. Moreover, there are some basic problems with the use of the eye-tracker glasses. For example, in our study we looked at fixations on objects placed at different distances. This means that

that the angle of vision, and therefore the standard calibration error, are different for each distance. In the manual coding, we needed some leniency to account for this, which might have introduced noise obscuring a potential effect.

5.3.1 Chapter 2: Influence of Cognition on Perception Checklist (Firestone & Scholl, 2015)

In Chapter 1.4.4 we discussed a set of problems that can be frequently found in studies testing the influence of top-down cognition on perception. Firestone and Scholl (2015) proposed a checklist of 6 items that should be considered (see Table 1.1). In these subsections we will review these six considerations for each chapter.

Research designs need to have uniquely disconfirmatory predictions. In the experiments presented in Chapter 2, we teased apart 3 different models with different predictions and tested which of these predictions matched our results best. One of these models was specifically based on the effect attention could have. In Experiment 4, we tested whether the differences were caused by differences in attention time, but we found no evidence for this while replicating the behavioural memory results.

The final part of every tested participant was a structured interview, testing whether participants realised what the study was about and/ or whether they used unexpected strategies. Any participant who would have successfully answered this question would have been eliminated, but no participant did – all participants bought into the idea that the demonstrative recall task was to increase cognitive load. When participants do not realise what is being tested, they cannot adjust their responses to suit (or contradict) the hypothesis.

We did not make any claims as to actual differences in perception. Since all trials were identical – save for the demonstrative used in the instruction – and counterbalanced, it is unlikely that results are shaped by structural differences in stimuli salience or memory performance other than the influence the language condition had on memory.

5.3.2 Chapter 3: Influence of Cognition on Perception Checklist (Firestone & Scholl, 2015)

Studies in Chapter 3 were similar to those in Chapter 2. Although we did not test for attentional effects in Chapter 3, results from Experiment 4 show that it is unlikely that specific attentional differences between language caused structural effects. Apart from that, the same precautions discussed in section 5.3.1 hold for Chapter 3. We counterbalanced the same stimuli and locations between trials, minimizing the potential for random effects to have a structural influence on our results.

5.3.3 Chapter 4: Influence of Cognition on Perception Checklist (Firestone & Scholl, 2015)

In Experiment 9, we compared participants' eye-movements in two different conditions. Attentional effects were not peripheral, but the main aim of the study. Furthermore, the design allowed for results that could support and contradict our hypothesis (or be inconclusive). Next to that, we used a between-participants design, minimizing the chance of participants realising they were in a specific condition, which was confirmed upon debrief.

Participants were presented with counterbalanced but identical probe-trials, after only the linguistic description in the prime trial was manipulated, minimizing the influence of random low-level differences.

We tested whether linguistic cueing influenced eye-movement patterns and succeeding linguistic descriptions, we made no claim regarding actual perceptional differences. Furthermore, the task was an online verbal description task, in which we again made no specific claims about top down effects on actual perception, merely on top down effects on attention focussing..

5.4 Theoretical implications

The presented research advances the literature on memory, and the influence of language on non-linguistic processes. Our results support models of predictive coding while rejecting other models of cognition. Furthermore, we show for the first time that spatial prepositions elicit identifyable eye-movement patterns.

In Chapter 2, results showed a robust effect of language on memory for object location, specifically the influence of demonstratives and possessives on spatial memory. In our study, the expectation, elicited by contextual knowledge – or language – concatenates with the actual object location and drives the memory error. This expectation is consistent with models of predictive coding. However, it is important to note that the Expectation model is currently underspecified regarding whether its effect occurs at encoding or retrieval. Our effects are also consistent with an extended probabilistic inference model as proposed by Regier and colleagues

(2016; 2017), suggesting that memory (in their studies for colour) is based on a two tier system of the actual perception and the category a colour belongs to. The probabilistic inference model, in contrast to the Expectation model, clearly implicates effects at retrieval, as uncertainty is needed for the tier of object knowledge to have an effect. There cannot be uncertainty while a participant still sees an object. If we extend their model to the domain of spatial memory (Huttenlocher et al., 1991, 2004), it would predict that uncertainty of the actual memory location opens the potential for contextual information to affect memory for object location. The continuous nature of a location memory error provides this uncertainty upon recall. Object knowledge could account for the difference in memorised location and actual location. In effect, the probabilistic inference model makes the same predictions as the Expectation model, and it might be argued that the probabilistic inference model would make similar predictions if extended to memory.

The predictions of the Expectation model and the probabilistic inference model could be teased apart using fMRI. The difference between the two models is that the Probabilistic inference model takes effect at retrieval, whereas the Expectation model has no strong claim on whether the effect happens at encoding or retrieval. If we would find differences in activation between conditions (e.g., of ownership, familiarity, language) at encoding, this could not be due to uncertainty and might therefore help tease apart these two models.

In Chapter 3, we found that Japanese demonstratives encode both distance and position of an interlocutor. These results align with only one of the identified models explaining the Japanese demonstrative system. Since most of our participants used all three demonstratives, and we found an effect of both distance from a participant and the position of the conspecific, we can falsify three models: the double binary, and both models suggesting influence of either distance or position. Our participants used all three demonstratives available to them in all different situations, consistent with both the parameter of distance and position of a conspecific (the territory model). The two most basic models, claiming that Japanese demonstratives are used either to encode distance or territory are thereby falsified. The fourth model we identified, described a double binary, in which *kono* and *sono*, and *kono* and *ano* contrast respectively territory and distance. However, our participants mostly used all three demonstratives for objects placed under different conditions of distance and position of a conspecific. Thereby, our data are not consistent with the double binary model. Our results support the dual system model in which both distance and position are influencing demonstrative use. The effect of position was found in English as well – although the effect was stronger in Japanese. The finding that position of conspecific, not explicitly encoded in English demonstratives, does influence English demonstrative use supports the idea that despite seemingly explicit differences between demonstrative systems, demonstratives are based on a universal set of underlying non-linguistic parameters. The research makes advances in theoretical understanding.

Further, Chapter 3 adds to the debate between universals in language vs. linguistic relativity. Coventry et al. (2014) identified a number of parameters which, despite not being explicitly encoded in English, affect English demonstrative use nonetheless (and memory for object location in a consistent manner). We found that position of a conspecific, explicitly encoded in Japanese but not in English, does affect English demonstrative use. This pattern seems to be consistent with universals in language. We could speculate that there are underlying universal mechanisms important in the use and contrasting function of demonstratives across languages, although this leaves open the possibility that parameters might be weighted differently across languages. We do however appreciate that great claims require great evidence, and many more languages need to be tested before language universals about demonstratives could be claimed as a position. Furthermore, it is important to note that the strength of the position effect was different between languages – Japanese speakers were affected more strongly by the position of a conspecific than English speakers. The different weightings of different parameters do however provide an interesting question for linguistic typology research: Is the effect of a parameter that is more frequently explicitly encoded in languages stronger than the effect of a parameter that is less frequently encoded explicitly. For example, physical distance is a parameter that is more commonly explicitly encoded than other parameters (e.g., ownership), and so far the effect of physical distance has been found to be rather strong compared to other parameters. If universal underlying parameters exist, parameters with a stronger effect on cognition and demonstrative use should more commonly be encoded explicitly.

In Chapter 4, we saw that linguistic cueing influences the way people scan the visual world. To our knowledge, we showed for the first time that different descriptions of the visual world lead to specific eye-movement *signatures*. Our data extend effects of visual cueing from syntactic structure to the conceptual domain. The choice of spatial language reflects the choice of an underlying frame of reference, which is usually taken to be at a conceptual level of representation rather than a lexical level. In most situations, multiple reference frames are appropriate for use (Carlson-Radvansky & Logan, 1997), and the speaker needs to select between these reference frames for language production. Verbal cueing affects this reference frame selection – how one talks about the world conceptually – consistent with the influence of past linguistic information on conceptualisation (Garrod & Anderson, 1987; Watson et al., 2004).

5.5 Future directions

The mechanism via which language influences memory can be investigated further using different behavioural settings, neuro-imaging methodologies, and neural modelling.

5.5.1 Behavioural methods. In Chapter 3, our results showed differences between Japanese and English participants. While their respective languages may cause these results, it is important to determine whether this is actually the case. To do so, we can run a follow-up study in which we vary, between participants, different versions of the memory game: a language production task, the memory version without language, and a memory version where language is manipulated. The procedure would look as follows. First, we replicate the language production experiment we used in Chapter 3. Second, we can use a no-language version of the memory version of the memory game, in which the participants perform a verbal interference task (Coventry et al., 2014; Garden et al., 2002), to test the strong version of the linguistic relativity hypothesis. In this task, no language is used to instruct object placement – the experimenter just places an object at a specific location. During the encoding of object location, participants repeat a meaningless phonological sequence ‘Ba Be Bi Bo Bu’. If we find a difference between Japanese and English speakers, this would suggest that memory performance is different even when language can not be accessed. If the differences disappear, this would suggest that the differences found in the previous where language was manipulated, were a result of the language. Third, a replication of the language manipulation version (Experiment 5 and 6) can be run, to confirm the results we found in the experiments presented in Chapter 3. These three parts can be run with the same participants. This within participant design allows us to test whether a specific use of demonstratives correlates with (the strength of) influence of demonstratives. For example, whether a stronger effect of position of a conspecific in demonstrative use is correlated with a stronger influence of position of a conspecific in the memory task. Furthermore, this design would help replicate the findings of the four experiments we currently ran, potentially providing more evidence for the differences between language and gender groups.

Other studies regarding the universality of parameters underlying demonstrative systems could focus on other parameters that are encoded in specific languages but not in others. For example, it has been reported that some indigenous languages contrast demonstratives based on whether a referent is uphill or downhill (for example in Yupno, see Cooperrider et al., 2016). An experimental design testing this could for example make use of a table placed at an angle. English participants can be tested on their demonstrative use while referring to objects that are higher or lower on the table.

Other variations on the procedure could inform the relationship between actual object location and expectation. A way to test the influence of the expectation or contextual information is by making the actual object location cue less strong. In previous research, a probabilistic model was suggested to underly memory. In this model, memory consist of two tiers, the actual perceptual memory, and a more ‘general’ knowledge about an object or event. Upon recall, the memory is influenced by both tiers, as a function of uncertainty. When the actual representation is more certain, there is a lower influence of ‘general object knowledge’. When there is more uncertainty in the perceptual memory, there is more influence of general knowledge (Huttenlocher et al., 1991, 2004; Regier & Xu, 2017). This prediction is similar to Expectation models’ . Uncertainty can be caused by for example fading memory or fatigue. We can manipulate the perceptual memory by presenting the objects for a shorter time. Currently, all object presentations are 10 seconds, but if we, for example, cut this presentation time in half, the actual object location memory will be weaker, leaving more space for contextual influence.

Alternatively, we can introduce a longer time-delay between encoding and retrieval. A longer timedelay might weaken the actual memory, again increasing the influence that contextual information could have on remembered object location (i.e., Bae, Olkkonen, Allred, & Flombaum, 2015; Cibelli et al., 2016).

Another line of studies could use a slightly different spatial configuration in the procedure (see Figure 5.37). In our study, we found an effect consistent with a dual-system model of Japanese demonstratives. However, there are two different versions of this model. The original model (Hoji et al., 2003; Kamio, 1994) suggests that objects near a hearer are referred to with *sono* only if they are within reaching distance from a hearer. A revised version of the model (Yoshimoto, 1992, 1986), suggested that *sono* would refer to all objects that are within the interactional dyad. Our study showed an effect of position, in which participants referred to an object moreoften with *sono* when the object was within the interactional dyad. However, this was regardless of whether the object was within the reach of the hearer, suggesting our data supports the updated version of the dual-system model. We can confirm this in a study in which the participant sits at the centre of the long side of the table, with the experimenter next to the participant, or on one of the short sides of the table (right in the example), so we can the influence of an *interactional* space (the space within the interactional dyad) and 'out of either territory'. Objects can be placed within the participants' reach (space 1), between participant and experimenter (space 2), but also out of reach of both participant and experimenter (space 3). By adding two locations marked with a red square, we would have objects within reach

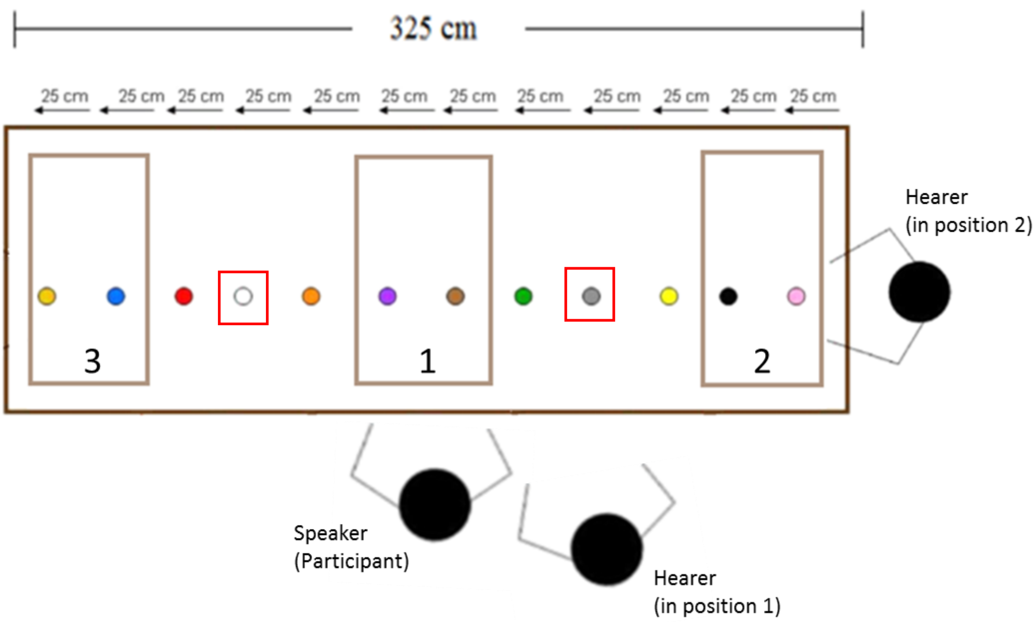


Figure 5.37. Procedural setup to tease apart the *interactional* space and *far-from-both* space.

of either interlocutor, but also out of reach of both interlocutors, both within interactional space and out of interactional space.

We could also adjust the spatial configuration of locations. In the current studies, we only used a horizontal plane (table) with distances along a midline from the participant to test memory. In Chapter 3, we added a conspecific, seated opposite or next to the participant. The orientation of locations could be varied; locations could for example be placed on a vertical instead of a horizontal plane. This way it could be tested how peri-personal space relates to the body. If peri-personal space is within arms' reach, then objects placed at the same distance from the body, but on the floor, should be considered extra-personal space. If peri-personal space is the space surrounding the body, it would be expected that it does not matter at which height an object is placed (as long as the height does not exceed participants' length. Furthermore, a three dimensional setup like this one could test peri-personal space, not only in front of the participant, but also to the sides or behind a participant.

In this work we present data supporting the Expectation model (Coventry et al., 2014), consistent with theories of predictive coding (Bar, 2009; Friston, 2009) and the probabilistic model proposed by Regier and colleagues (2016; 2017). Next to exploring the influence of expectation as a result of demonstrative use in a cross-linguistic study, it would be interesting to test the Expectation model across domains. As we saw in the domain of colour perception, there is a large body of research looking at the influence of top down cognition, suggesting a probabilistic inference model. However, the Expectation model can be tested in other domains. The Expectation model might work via a process of predictive coding. Accounts of predictive coding suggest that the brain continuously predicts future states, to avoid surprise and safe energy (Friston, 2009). To do this, only a baseline of perceptual information and deviations from prediction are encoded. If this is the case, one would expect that deviations are memorized more accurately, or in other words, that deviations have a memory enhancing effect. This means that, regardless of the domain tested, trials including an certain amount of surprise should be memorized better than trials that are in line with predictions. The enhanced fine-grained performance can be expected to lead to a reduced uncertainty, and therefore to a reduced influence of general knowledge.

5.5.2 Neuro imaging methods. We can extend on our current findings concerning and extend the limits of the Expectation model within the domain of spatial cognition, using brain imaging methodologies. As mentioned in the discussion comparing the Expectation model and the Probabilistic interference model (section 5.2.1), we cannot distinguish between language effects happening at encoding or retrieval in our current procedure. However, the theoretical implications of a potential effect at encoding would be strong. If the effect happens at encoding, this would show an online effect of higher-level processes on perception. Brain imaging methodologies, e.g., fMRI and TMS, could help elucidate the current findings and situate them in knowledge about neural correlates. fMRI scanning can identify the different brain areas involved. For example, literature suggests that the brain constructs multiple functionally distinct representations of space (di Pellegrino & Làdavas, 2015), which are centred around body parts in different primate species (e.g., hand-centred, head-centred, trunk-centred). Some areas that seem to be involved in the representation of visual peri-personal space in relation to the hand (peri-hand space) are the intraparietal sulcus (IPS) and lateral occipital complex (LOC). A specific interesting finding was that visual perception is dominantly displayed in the posterior IPS and LOC, whereas the anterior IPS was more sensitive to a proprioceptive representation of the hand, showing tactile hand-specific activation, which suggests that the aIPS uses multisensory information (Makin, Holmes, & Zohary, 2007). This neurological distinction between peri-personal space and extra-personal space (di Pellegrino & Làdavas, 2015; Makin et al., 2007; Rizzolatti, Fadiga, Fogassi, & Gallese, 1997) is of particular interest for this thesis and future research into the findings presented in this thesis. Findings of this distinction are also reported in a dissociation for peri-personal and extra-personal neglect (Lane, Ball, Smith, Schenk, & Ellison, 2013).

Activation in the SPOC is strongest when interacting within reaching distance, or peri-personal space (Gallivan et al., 2011), even in the absence of overt responses. With regard to our studies, it would be interesting to see if the SPOC area is involved in memory for proximal object location. It could then be contrasted whether proximal language has similar effects to memory for proximal location. For example, proximal language could activate the SPOC area, as proximal language often indicates an action response. Furthermore, using TMS it is possible to explore causal relations between activated or inhibited brain areas and the effects we found so far.

5.5.3 Modelling. In Chapter 2, we tested the predictions of three theoretical models regarding the underlying mechanism via which language influences memory. While these models can represent mechanisms and relations underlying different cognitive processes, they are not themselves neural models – they cannot perform computations (c.f., Lipinski, Schneegans, Sandamirskaya, Spencer, & Schöner, 2012; Logan & Sadler, 1996). Future research could focus on specifying neural models and teasing apart the models representing the interaction between memory and language.

In previous chapters, we already introduced the Attentional Vector-Sum model (Regier & Carlson, 2001), and the probabilistic inference model (i.e., Cibelli et al., 2016; Regier & Xu, 2017). Two examples of neural models that we did not yet discuss, but are an interesting future direction of research are the Category Adjustment Model (CAM) (Huttenlocher et al., 1991) and models based on the Dynamic Field Theory (DFT) (e.g., Lipinski et al., 2012; Lipinski, Simmering, Johnson, & Spencer, 2010).

The CAM is a class of probabilistic models ((i.e., Crawford, Regier, & Huttenlocher, 2000; Huttenlocher et al., 2000; Regier & Xu, 2017). The model suggests that memory is encoded at two levels, a fine-grained memory of the particular experience (e.g., an event or an object) and a general category in which the experienced object or event falls. A category is a bounded region that includes a range of fine-grained stimulus values, the end of a category is marked by a boundary value. The prototype value is the central value within a category (Huttenlocher et al., 1991). The category (or prototypical value) can bias a memory towards a more categorical representation. In Chapter 1 we discussed probabilistic inference models with regards to colour perception. When there is an amount of uncertainty of the particular memory, the memory of the experience is merged with the prototypical value of the colour, via a process of probabilistic inference. This integration is suggested to work via a Bayesian framework (Deneve & Pouget, 2004). The CAM has also been applied to spatial cognition, for example if participants are asked to remember the location of a dot on an object, memory is biased towards the centre of the object (Crawford, Huttenlocher, et al., 2000; Holden et al., 2013).

However, there are two problems for the CAM. First, as mentioned previously, the effect of a prototype of category would only be expected to work upon memory recall. After all, while you see an object there should be no uncertainty regarding the exact location of an object. But in Chapter 3, we presented language production data that are parallel to the memory data – if an object is placed under conditions in which participants preferentially refer to objects using *that*, they are also remembered to be further away. This could suggest that the effect causing the bias in memory is also apparent in an online production task, when participants refer to an object while looking at it. The second problem is the constitution of the prototypical value. It is unclear how these prototypes or categories are generated. In the CAM, categories seem to be fixed. Applied to the results presented in Chapter 2, an *a posteriori* explanation of the data using the CAM would be that the prototype of an object placed with *this* would be a location close to the participants, whereas the prototypical location of a *that*-object would be away from the participant. Therefore, an object placed with *that* has a bias away from the person (and an object placed with *this* has a bias towards the person). This mechanism would predict, as the Expectation model does, that objects placed with *that* are misremembered to be further away than objects placed with *this*.

A models that does include the potential for learning is based on the DFT. Dynamic neural fields represent neural population dynamics in which different levels of information are processed by the activation or inhibition within and between different neural levels. The dynamic nature of the model constitutes feedback mechanisms, enabling the model to integrate long-term memory traces with experience. For example in Lipinski et al. (2010), learning is presented to solve the problem of contrasting previous results found by Huttenlocher (1991, 2004) and Spencer and Hund (2002). Huttenlocher et al. (1991, 2004) showed that the use of spatial categories improves the overall accuracy of memory, as the use of spatial boundaries constrain the memory error within category boundaries. However, they argued that the distribution of the presented stimuli did not influence location estimation, whereas Spencer and Hund (2002) found that experience did bias performance. Lipinski et al. (2010) presented participants with multiple blocks of the same memory task, and found that in the first block of stimulus presentations performance (when participants did not have experience in the task), performance was unbiased by previous target presentation, as Huttenlocher et al (2004) stated. However, after additional experience biases shifted towards the previous target distributions as Spencer and Hund (2002) showed.

Based on DFT, an architecture of spatial language production can be modelled (Lipinski et al., 2012). Lipinski et al., presented a model that can perceive different objects (based on colour) and produce spatial language to express the relation of different object in relation to one another. The architecture of the model consists of different fields. There is a *target field* in which the location of the object (which itself is identified by *colour-space fields* and related *colour-term nodes*) is encoded. The location activation (which holds strong local excitation/ lateral inhibition, in order to end up with one single activity peak defining the specific location of the object) in the target field feeds back to other levels of representation, binding (for example) the colour and the location of an object. By mapping the location of multiple objects, the model is able to derive the spatial orientation of the objects in relation to each other via *spatial relation nodes*, and is able to assign spatial terms (e.g., above/below/left/right) describing this relation. Spatial relation nodes activate corresponding *spatial term nodes*, assigning a linguistic description to the spatial relation processed by the field. The heart of the model is the *reference frame transformation field*. The transformation field is a field in which all combinations of target and reference positions are defined, thereby creating a single combined representation of the information from the object-centred field (object location) and the reference field (spatial term). The coupling between target, reference, and object-centred fields is bidirectional, meaning that by having input of two of the three fields the accompanying representation the third field can be activated (Lipinski et al., 2012). This model could be consistent with our data, although there are some differences between the information the model uses and our procedure. In our procedure for example, only one object is present at any given time (although participants could be using the different locations that are visible that are between them and the object). Based on the object location and the location of the object relative to for example themselves, or the locations that are visible on the table, they might base their language production.

It is interesting to note that if this model contains similar fields as a potential DFT memory model would, it might potentially explain our results via the *reference field*. This reference field holds an activation based on the spatial term that is used to describe the object location, and has a bi-directional inhibiting relation with the target field, which holds the objects' location, while both fields feed into the transformation field. If memory for object location is derived from the activation and the relational excitation/ inhibition between these fields, this could be a mechanism via which language biases memory for object location. In our memory-version of the procedure, the object location is biased by the spatial term (*this/that*). If the language use has an inhibitory effect on the object location, it effectively creates the uncertainty of the perceptual memory that opens the door for a biasing function of language.

5.6 General conclusion

To summarize, the work carried out in this thesis explored the relationship between language and non-linguistic processes. English demonstrative use influences spatial memory, consistent with predictions of the Expectation model. The Expectation model fits with current theories of predictive coding. Our results therefore add to the field of cognitive psychology, memory, and cognition.

Furthermore, our results suggest that there might be universal mechanisms underlying demonstratives. Position of a conspecific influences both Japanese and English demonstrative use, whereas the position of a conspecific is not explicitly encoded in English. However, the level of explicitness of demonstratives can influence the strength of the parameters. In Japanese, the effect of the position of a conspecific was stronger than in English. These results add to the fields of psychology and linguistics. Finally, we found an association between distinctive eye-movement patterns and frames of reference. Spatial descriptions influence the way someone looks at the world, showing eye-movement *signatures* for the different reference frames. In other words, we showed that language has a robust effect on non-linguistic processes, focussing on spatial memory and visual attention. The work has shed light on a mechanism explaining *how* memory works, via a two-tier system, in which language and prediction can play a part, and has suggested a mutual influence between attention and spatial language.

References

Allopenna, P. D., Magnuson, J. S., & Tanenhaus, M. K. (1998). Tracking the Time Course of Spoken Word Recognition Using Eye Movements: Evidence for Continuous Mapping Models. *Journal of Memory and Language*, *38*(4), 419–439. https://doi.org/10.1006/jmla.1997.2558

Alogna, V., Attaya, M., Aucoin, P., Bahnik, S., Birch, S., Birt, A., & Zwaan, R. (2014). Registered Replication Report: Schooler and Engstler-Schooler (1990). *Perspectives on Psychological Science*, *9*(5), 556–578. https://doi.org/10.1177/1745691614545653

Altmann, G. T. M., & Kamide, Y. (2007). The real-time mediation of visual attention by language and world knowledge: Linking anticipatory (and other) eye movements to linguistic processing. *Journal of Memory and Language*, *57*(4), 502–518. https://doi.org/10.1016/j.jml.2006.12.004

Andrews, M., Vigliocco, G., & Vinson, D. (2009). Integrating experiential and distributional data to learn semantic representations. *Psychological Review*, *116*(3), 463–98. https://doi.org/10.1037/a0016261

Athanasopoulos, P., Bylund, E., Montero-Melis, G., Damjanovic, L., Schartner, A., Kibbe, A., … Thierry, G. (2015). Two Languages, Two Minds: Flexible Cognitive Processing Driven by Language of Operation. *Psychological Science*, 1–9. https://doi.org/10.1177/0956797614567509

Bae, G. Y., Olkkonen, M., Allred, S. R., & Flombaum, J. I. (2015). Durham Research Online EXCEPTIONALLY. *Journal of Experimental Psychology. General*, *144*(4), 744–763. https://doi.org/http://dx.doi.org/10.1037/xge0000076

Balcetis, E., & Dunning, D. (2010). Wishful seeing: more desired objects are seen as closer. *Psychological Science*, *21*(1), 147–52. https://doi.org/10.1177/0956797609356283

Bar, M. (2009). The proactive brain: memory for predictions. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*, *364*(1521), 1235–43. https://doi.org/10.1098/rstb.2008.0310

Barsalou, L. W. (1999). Perceptual symbol systems. *Behavioral and Brain Sciences*, *22*(4), 577–660. https://doi.org/10.1017/S0140525X99002149

Barsalou, L. W. (2008). Grounded cognition. *Annual Review of Psychology*, *59*, 617–45. https://doi.org/10.1146/annurev.psych.59.103006.093639

Baumeister, R. F., & Sommer, K. L. (1997). What do men want? Gender differences and two spheres of belongingness: comment on Cross and Madson (1997). *Psychological Bulletin*, *122*(1), 38-44-55. https://doi.org/10.1037/0033-2909.122.1.38

Belli, R. F. (1988). Color blend retrievals: Compromise memories or deliberate compromise responses? *Memory & Cognition*, *16*(4), 314–326. https://doi.org/10.3758/BF03197042

Bird, C. M., & Burgess, N. (2008). The hippocampus and memory: insights from spatial processing. *Nature Reviews. Neuroscience*, *9*(3), 182–194. https://doi.org/10.1038/nrn2335

Bock, K., Dell, G. S., Chang, F., & Onishi, K. H. (2007). Persistent structural priming from language comprehension to language production. *Cognition*, *104*(3), 437–458. https://doi.org/10.1016/j.cognition.2006.07.003

Bonfiglioli, C., Finocchiaro, C., Gesierich, B., Rositani, F., & Vescovi, M. (2009). A kinematic approach to the conceptual representations of this and that. *Cognition*, *111*(2), 270–4. https://doi.org/10.1016/j.cognition.2009.01.006

Boroditsky, L. (2011). How language shapes thought. *Scientific American*, (April), 63–65.

Bowden, J. (2014). Mental Deixis in Taba. *Deixis and Spatial Expressions in Languages of Indonesia*, *56*, 79–100. Retrieved from http://hdl.handle.net/10108/77647%5D

Bowerman, M. (1996). Learning how to structure space for language -- A crosslinguistic perspective. *Language and Space*.

Branigan, H. P., Pickering, M. J., & Cleland, A. A. (2000). Syntactic co-ordination in dialogue. *Cognition*, *75*(2), 13–25. https://doi.org/10.1016/S0010-0277(99)00081-5

Brennan, S. E., & Clark, H. H. (1996). Conceptual pacts and lexical choice in conversation. *Journal of Experimental Psychology. Learning, Memory, and Cognition*, *22*(6), 1482–1493. https://doi.org/10.1037/0278-7393.22.6.1482

Brown, A. M., Lindsey, D. T., & Guckes, K. M. (2011). Color names, color categories, and color-cued visual search: Sometimes, color perception is not categorical. *Journal of Vision*, *11*(12), 2–2. https://doi.org/10.1167/11.12.2.Introduction

Brown, P., & Levinson, S. C. (1993). Linguistic and nonlinguistic coding of spatial arrays: Explorations in Mayan cognition. *Cognitive Anthropology Research Group at the Max Planck Institute for Psycholinguistics*, *Working Pa*(24).

Brown, R. W., & Lenneberg, E. H. (1954). A study in language and cognition. *The Journal of Abnormal and Social Psychology*, *49*(3), 454–62. https://doi.org/10.1037/h0057814

Bruner, J. S. (1983). *Child's Talk: Learning to use language*. Oxford: Oxford University Press.

Bruner, J. S., Postman, L., & Rodrigues, J. (1951). Expectation and the Perception of Color. *The American Journal of Psychology*, *64*(2), 216. https://doi.org/10.2307/1418668

Bullier, J. (2001a). Feedback connections and conscious vision. *Trends in Cognitive Sciences*, *5*(9), 369–370. https://doi.org/10.1016/S1364-6613(00)01730-7

Bullier, J. (2001b). Integrated model of visual processing. *Brain Research Reviews*, *36*(2–3), 96–107. https://doi.org/10.1016/S0165-0173(01)00085-6

Burenhult, N. (2003). Attention, Accessibility, and the Assressee: the Case of the Jahai Demonstrative "Ton." *Pragmatics*, *379*(13:3), 363–379.

Burgess, N. (2006). Spatial memory: how egocentric and allocentric combine. *Trends in Cognitive Sciences*, *10*(12), 551–557. https://doi.org/10.1016/j.tics.2006.10.005

Burgess, N. (2008). Spatial cognition and the brain. *Annals of the New York Academy of Sciences*, *1124*, 77–97. https://doi.org/10.1196/annals.1440.002

Burigo, M., & Knoeferle, P. (2015). Visual Attention during Spatial Language Comprehension. *PloS One*, *10*(1), 1–21. https://doi.org/10.4119/unibi/2699034

Butterworth, B., & Reeve, R. (2008). Verbal counting and spatial strategies in numerical tasks: Evidence from indigenous Australia. *Philosophical Psychology*, *21*(4), 443–457. https://doi.org/10.1080/09515080802284597

Byrne, P., Becker, S., & Burgess, N. (2009). Remembering the past and imagining the future: a neural model of spatial memory and imagery, *114*(2), 340–375. https://doi.org/10.1037/0033-295X.114.2.340

Byron, D. K., & Stoia, L. (2005). An analysis of proximity markers in collaborative dialogs. *Proceedings from the Annual Meeting of the Chicago Linguistic Society*, *41*(2), 17–32. Retrieved from http://cls.metapress.com/index/7N1136131762U237.pdf

Capirci, O., Iverson, J. M., Pizzuto, E., & Volterra, V. (1996). Gestures and words during the transition to two-word speech. *Journal of Child Language*, *23*(3), 645–673. https://doi.org/10.1017/S0305000900008989

Carey, D. P., & Johnstone, L. T. (2014). Quantifying cerebral asymmetries for language in dextrals and adextrals with random-effects meta analysis. *Frontiers in Psychology*, *5*(NOV), 1–23. https://doi.org/10.3389/fpsyg.2014.1128

Carlson-Radvansky, L. a., & Irwin, D. E. (1994). Reference Frame Activation during Spatial Term Assignment. *Journal of Memory and Language*. https://doi.org/10.1006/jmla.1994.1031

Carlson-Radvansky, L. A., & Irwin, D. E. (1993). Frames of reference in vision and language: Where is above? *Cognition*, *46*(3), 223–244. https://doi.org/10.1016/0010-0277(93)90011-J

Carlson-Radvansky, L. A., & Jiang, Y. (1998). Inhibition accompanies reference-frame selection. *Psychological Science*, *9*(5), 386–391. https://doi.org/10.1038/sj/gt/3301593

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. https://doi.org/10.1006/jmla.1997.2519

Carlson, L. a., & Van Deman, S. R. (2008). Inhibition within a reference frame during the interpretation of spatial language. *Cognition*, *106*(1), 384–407. https://doi.org/10.1016/j.cognition.2007.03.009

Casarotti, M., Michielin, M., Zorzi, M., & Umiltà, C. (2007). Temporal order judgment reveals how number magnitude affects visuospatial attention. *Cognition*, *102*(1), 101–17. https://doi.org/10.1016/j.cognition.2006.09.001

Casasanto, D., & Boroditsky, L. (2008). Time in the mind: using space to think about time. *Cognition*, *106*(2), 579–93. https://doi.org/10.1016/j.cognition.2007.03.004

Chambers, C., Tanenhaus, M. K., Eberhard, K. M., Filip, H., & Carlson, G. N. (2002). Circumscribing Referential Domains during Real-Time Language Comprehension. *Journal of Memory and Language*, *47*(1), 30–49. https://doi.org/10.1006/jmla.2001.2832

Choi, S., & Bowerman, M. (1991). Learning to express motion events in English and Korean: The influence of language-specific lexicalization patterns. *Cognition*, *41*(1–3), 83–121. https://doi.org/10.1016/0010-0277(91)90033-Z

Choi, S., McDonough, L., Bowerman, M., & Mandler, J. M. (1999). Early sensitivity to language-specific spatial categories in English and Korean. *Cognitive Development*, *14*(1999), 241–268. https://doi.org/10.1016/S0885-2014(99)00004-0

Chomsky, N., & Foucault, M. (2006). *The Chomsky-Foucault Debate*. New York, NY: The New Press.

Cibelli, E., Xu, Y., Austerweil, J. L., Griffiths, T. L., & Regier, T. (2016). The Sapir-Whorf hypothesis and probabilistic inference: Evidence from the domain of color. *PLoS ONE*, *11*(7), 1–28. https://doi.org/10.1371/journal.pone.0158725

Clark, A. (2013). Whatever next? Predictive brains, situated agents, and the future of cognitive science. *The Behavioral and Brain Sciences*, *36*(3), 181–204. https://doi.org/10.1017/S0140525X12000477

Clark, E. V. (1978). From gesture to word: on the natural history of deixis in language acquisition. In J. S. Bruner & A. Garton (Eds.), *Human growth and development* (pp. 85–120). London: Oxford University Press.

Clark, E. V., & Sengul, C. J. (1978). Strategies in the acquisition of deixis. *Journal of Child Language*, *5*(3), 457–475. https://doi.org/10.1017/S0305000900002099

Clark, H. H. (1996). *Using Language*. Cambridge: Cambridge University Press.

Constable, M. D., Kritikos, A., & Bayliss, A. P. (2011). Grasping the concept of personal property. *Cognition*, *119*(3), 430–7. https://doi.org/10.1016/j.cognition.2011.02.007

Cooperrider, K. (2016). The Co-Organization of Demonstratives and Pointing Gestures. *Discourse Processes*, *53*(April), 632–656. https://doi.org/10.1080/0163853X.2015.1094280

Cooperrider, K., Slotta, J., & Núñez, R. (2016). Uphill and Downhill in a Flat World: The Conceptual Topography of the Yupno House. *Cognitive Science*, 1–32. https://doi.org/10.1111/cogs.12357

Coventry, K. R., Christophel, T. B., Fehr, T., Valdés-Conroy, B., & Herrmann, M. (2013). Multiple routes to mental animation: language and functional relations drive motion processing for static images. *Psychological Science*, *24*, 1379–88. https://doi.org/10.1177/0956797612469209

Coventry, K. R., Griffiths, D., & Hamilton, C. J. (2014). Spatial demonstratives and perceptual space: describing and remembering object location. *Cognitive Psychology*, *69*, 46–70. https://doi.org/10.1016/j.cogpsych.2013.12.001

Coventry, K. R., Valdés, B., Castillo, A., & Guijarro-Fuentes, P. (2008). Language within your reach: near-far perceptual space and spatial demonstratives. *Cognition*, *108*(3), 889–95. https://doi.org/10.1016/j.cognition.2008.06.010

Crawford, L. E., Huttenlocher, J., & Engebretson, P. H. (2000). Category effects on estimates of stimuli: perception or reconstruction? *Psychological Science*, *11*(4), 280–284. https://doi.org/10.1111/1467-9280.00256

Crawford, L. E., Regier, T., & Huttenlocher, J. (2000). Linguistic and non-linguistic spatial categorization. *Cognition*, *75*(3), 209–235. https://doi.org/10.1016/S0010-0277(00)00064-0

Cross, S. E., & Madson, L. (1997). Models of the self: self-construals and gender. *Psychological Bulletin*, *122*(1), 5–37. https://doi.org/10.1037/0033-2909.122.1.5

Cunningham, S. J., Vergunst, F., Macrae, C. N., & Turk, D. J. (2013). Exploring early self-referential memory effects through ownership. *British Journal of Developmental Psychology*, *31*(3), 289–301. https://doi.org/10.1111/bjdp.12005

Cunningham, S., Turk, D. J., Macdonald, L. M., & Macrae, N. C. (2008). Yours or mine? Ownership and memory. *Consciousness and Cognition*, *17*(1), 312–8. https://doi.org/10.1016/j.concog.2007.04.003

Dehaene, S. (1992). Varieties of numerical abilities. *Cognition*, *44*(1–2), 1–42. https://doi.org/10.1016/0010-0277(92)90049-N

Dehaene, S., Bossini, S., & Giraux, P. (1993). The mental representation of parity and number magnitude. *Journal of Experimental Psychology: General*, *122*(3), 371–396. https://doi.org/10.1037/0096-3445.122.3.371

Delk, J. L., & Fillenbaum, S. (1965). Differences in Perceived Color as a Function of Characteristic Color. *The American Journal of Psychology*, *78*(2), 290. https://doi.org/10.2307/1420503

Deneve, S., & Pouget, A. (2004). Bayesian multisensory integration and cross-modal spatial links. *Journal of Physiology Paris*, *98*(1–3 SPEC. ISS.), 249–258. https://doi.org/10.1016/j.jphysparis.2004.03.011

Dessalegn, B., & Landau, B. (2008). More Than Meets the Eye: The Role of Language in Binding and Maintaining Feature Conjunctions. *Psychological Science*, *19*(2), 189–195. https://doi.org/10.1111/j.1467-9280.2008.02066.x

Dessalegn, B., & Landau, B. (2013). Interaction between language and vision: it's momentary, abstract, and it develops. *Cognition*, *127*(3), 331–44.

https://doi.org/10.1016/j.cognition.2013.02.003

Deutscher, G. (2005). *The Unfolding of Language: the evolution of mankinds greatest invention*. London: The Random House Group.

di Pellegrino, G., & Làdavas, E. (2015). Peripersonal space in the brain. *Neuropsychologia*, *66*, 126–133. https://doi.org/10.1016/j.neuropsychologia.2014.11.011

Diessel. (2005). Distance contrasts in demonstratives. In M. Haspelmath, M. Dryer, D. Gil, & B. Comrie (Eds.), *World atlas of language structures* (pp. 170–173). Oxford, UK: Oxford University Press.

Diessel, H. (1999). *Demonstratives: form, function and grammaticalization*. Amsterdam/Philadephia: John Benjamins Publishing Company.

Diessel, H. (2006). Demonstratives, joint attention, and the emergence of grammar. *Cognitive Linguistics*, *17*(4), 463–489. https://doi.org/10.1515/COG.2006.015

Diessel, H. (2014). Demonstratives, Frames of Reference, and Semantic Universals of Space. *Language and Linguistics Compass*, *8*(3), 116–132. https://doi.org/10.1111/lnc3.12066

Dobnik, S., Kelleher, J. D., & Koniaris, C. (2014). Priming and Alignment of Frame of Reference in Situated Conversation. In *Proceedings of SemDial 2014 (DialWatt): The 18th Workshop on the Semantics and Pragmatics of Dialogue* (pp. 43–52). Edinburgh.

Dolscheid, S., Shayan, S., Majid, A., & Casasanto, D. (2013). The thickness of musical pitch: psychophysical evidence for linguistic relativity. *Psychological Science*, *24*(5), 613–21. https://doi.org/10.1177/0956797612457374

Drivonikou, G. V, Kay, P., Regier, T., Ivry, R. B., Gilbert, A. L., Franklin, A., & Davies, I. R. L. (2007). Further evidence that Whorfian effects are stronger in the right visual field than the left. *Proceedings of the National Academy of Sciences*, *104*(3), 1097–1102. https://doi.org/10.1073/pnas.0610132104

Eberhard, K. M., Spivey-Knowlton, M. J., Sedivy, J. C., & Tanenhaus, M. K. (1995). Eye movements as a window into real-time spoken language comprehension in natural contexts. *Journal of Psycholinguistic Research*, *24*(6), 409–436. https://doi.org/10.1007/BF02143160

Enfield, N. J. (2003). Demonstratives in Space and Interaction: Data from Lao Speakers and Implications for Semantic Analysis. *Language*, *79*(1), 82–117. Retrieved from http://www.jstor.org.proxy.lib.pdx.edu/stable/4489386

Ernst, M. O., & Banks, M. S. (2002). Humans integrate visual and haptic information in a statistically optimal fashion. *Nature*, *415*(6870), 429–433. https://doi.org/10.1038/415429a

Evans, N., & Levinson, S. C. (2009). The myth of language universals: language diversity and its importance for cognitive science. *The Behavioral and Brain Sciences*, *32*(5), 429-48-494. https://doi.org/10.1017/S0140525X0999094X

Everett, D. L. (2005). Cultural Constraints on Grammar and Cognition in Pirah・ Another Look at the Design Features of Human Language. *Current Anthropology*, *46*(4), 621–646. Retrieved from http://www.jstor.org/stable/10.1086/431525

Farran, E. K., & O'Leary, B. (2015). Children's ability to bind and maintain colour–location conjunctions: the effect of spatial language cues. *Journal of Cognitive Psychology*, *5911*(October), 1–8. https://doi.org/10.1080/20445911.2015.1092980

Feist, M. I., & Gentner, D. (2007). Spatial language influences. *Memory & Cognition*, *35*(2), 283–296.

Firestone, C., & Scholl, B. J. (2015). Cognition does not affect perception: Evaluating the evidence for "top-down" effects. *Behavioral and Brain Sciences*, *4629*(2016), 1–77. https://doi.org/10.1017/S0140525X15000965

Fischer, M. H., Castel, A. D., Dodd, M. D., & Pratt, J. (2003). Perceiving numbers causes spatial shifts of attention. *Nature Neuroscience*, *6*(6), 555–6. https://doi.org/10.1038/nn1066

Fodor, J. A. (1989). The modularity of mind : an essay on faculty psychology. *Reasoning: Studies of Human Inference and Its Foundations*, (1983), 878–914. https://doi.org/10.2307/413815

Frank, M. C., Everett, D. L., Fedorenko, E., & Gibson, E. (2008). Number as a cognitive technology: Evidence from Pirahã language and cognition. *Cognition*, *108*(3), 819–824. https://doi.org/10.1016/j.cognition.2008.04.007

Frank, M. C., & Goodman, N. D. (2014). Inferring word meanings by assuming that speakers are informative. *Cognitive Psychology*, *75*, 80–96. https://doi.org/10.1016/j.cogpsych.2014.08.002

Friston, K. (2003). Learning and inference in the brain. *Neural Networks : The Official Journal of the International Neural Network Society*, *16*(9), 1325–52.

https://doi.org/10.1016/j.neunet.2003.06.005

Friston, K. (2005). A theory of cortical responses. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*, *360*(1456), 815–36. https://doi.org/10.1098/rstb.2005.1622

Friston, K. (2009). The free-energy principle: a rough guide to the brain? *Trends in Cognitive Sciences*, *13*(7), 293–301. https://doi.org/10.1016/j.tics.2009.04.005

Galati, A., & Avraamides, M. N. (2013). Flexible spatial perspective-taking: conversational partners weigh multiple cues in collaborative tasks. *Frontiers in Human Neuroscience*, *7*(September), 618. https://doi.org/10.3389/fnhum.2013.00618

Gallivan, J. P., McLean, A., & Culham, J. C. (2011). Neuroimaging reveals enhanced activation in a reach-selective brain area for objects located within participants' typical hand workspaces. *Neuropsychologia*, *49*(13), 3710–3721. https://doi.org/10.1016/j.neuropsychologia.2011.09.027

Garden, S., Cornoldi, C., & Logie, R. H. (2002). Visuo-spatial working memory in navigation. *Applied Cognitive Psychology*, *16*(1), 35–50. https://doi.org/10.1002/acp.746

Garrido-Vásquez, P., & Schubö, A. (2014). Modulation of visual attention by object affordance. *Frontiers in Psychology*, *5*(FEB), 1–11. https://doi.org/10.3389/fpsyg.2014.00059

Garrod, S., & Anderson, A. (1987). Saying what you mean in dialogue: A study in conceptual and semantic co-ordination. *Cognition*, *27*(2), 181–218. https://doi.org/10.1016/0010-0277(87)90018-7

Garrod, S., & Pickering, M. J. (2004). Why is conversation so easy? *Trends in Cognitive Sciences*, *8*(1), 8–11. https://doi.org/10.1016/j.tics.2003.10.016

Gennari, S. P., Sloman, S. A., Malt, B. C., & Fitch, W. T. (2002). Motion events in language and cognition. *Cognition*, *83*(1), 49–79. https://doi.org/10.1016/S0010-0277(01)00166-4

Gentner, D., & Loftus, E. (1979). Integration of verbal and visual information as evidenced by distortions in picture memory. *The American Journal of Psychology*, *92*(2), 363–375. Retrieved from http://www.jstor.org/stable/1421930

Gentner, Özyürek, A., Gürcanli, Ö., & Goldin-Meadow, S. (2013). Spatial language facilitates spatial cognition: Evidence from children who lack language input.

Cognition, *127*(3), 318–330. https://doi.org/10.1016/j.cognition.2013.01.003

Gilbert, A. L., Regier, T., Kay, P., & Ivry, R. B. (2006). Whorf hypothesis is supported in the right visual field but not the left. *Proceedings of the National Academy of Sciences of the United States of America*, *103*(2), 489–494. https://doi.org/10.1073/pnas.0509868103

Gleitman, L. R., January, D., Nappa, R., & Trueswell, J. C. (2007). On the give and take between event apprehension and utterance formulation. *Journal of Memory and Language*, *57*(4), 544–569. https://doi.org/10.1016/j.jml.2007.01.007

Göbel, S. M., Shaki, S., & Fischer, M. H. (2011). The Cultural Number Line : A Review of Cultural and Linguistic Influences on the Development of Number Processing. *Journal of Cross-Cultural Psychology*, *42*(4), 543–565. https://doi.org/10.1177/0022022111406251

Goldstone, R. L. (1995). Effects of Categorization on Color Perception. *Psychological Science*, *6*(5), 298–304. https://doi.org/10.1111/j.1467-9280.1995.tb00514.x

Gordon, P. (2004). Numerical cognition without words: evidence from Amazonia. *Science (New York, N.Y.)*, *306*(5695), 496–499. https://doi.org/10.1126/science.1094492

Gudde, H. B., Coventry, K. R., & Engelhardt, P. E. (2015). The Influence of Language on Memory for Object Location. In R. Dale, C. Jennings, T. Matlock, D. Noelle, A. Warlaumont, & J. Yoshimi (Eds.), *Proceedings of the 37th Annual Conference of the Cognitive Science Society* (pp. 1–6). Austin, TX: Cognitive Science Society.

Gudde, H. B., Coventry, K. R., & Engelhardt, P. E. (2016). Language and memory for object location. *Cognition*, *153*, 99–107. Retrieved from http://www.sciencedirect.com/science/article/pii/S0010027716301044

Gumperz, J. J., & Levinson, S. C. (1991). Rethinking Linguistic Relativity. *Current Anthropology*, *32*(5), 613–622. https://doi.org/10.1086/204009

Gumperz, J. J., & Levinson, S. C. (1996). *Rethinking Linguistic Relativity*. Cambridge, UK: Cambridge University Press. https://doi.org/10.1086/204009

Hansen, T., Olkkonen, M., Walter, S., & Gegenfurtner, K. R. (2006). Memory modulates color appearance. *Nature Neuroscience*, *9*(11), 1367–8. https://doi.org/10.1038/nn1794

Hasegawa, Y. (2012). Deictic and anaphoric uses of the Japanese demonstratives ko-so-a. *Journal of Japanese Linguistics*, *28*.

Hattori. (1992). Demonstratives. In K. & Takabo (Ed.), *Demonstratives* (p. 47).

Hayward, W. G., & Tarr, M. J. (1995). Spatial language and spatial representation. *Cognition*, *55*(1), 39–84. https://doi.org/10.1016/0010-0277(94)00643-Y

Healey, P. G. T. (2008). Interactive Misalignment: The Role of Repair in the Development of Group Sub-languages. In R. Cooper & R. Kempson (Eds.), *Language in Flux* (pp. 13–39). College Publications.

Healey, P. G. T., Purver, M., & Howes, C. (2014). Divergence in dialogue. *PloS One*, *9*(2), e98598. https://doi.org/10.1371/journal.pone.0098598

Heider, E. R. (1972a). Probabilities , Sampling , and Ethnographic Method : The Case of Dani Colour Names. *Royal Anthropological Institute of Great Britain and Ireland*, *7*(3), 448–466. Retrieved from http://www.jstor.org/stable/2800917

Heider, E. R. (1972b). Universals in color naming and memory. *Journal of Experimental Psychology*, *93*(1), 10–20. https://doi.org/10.1037/h0032606

Heine, B., & Kutteva, T. (2002). *World Lexicon of grammaticalization*. Cambridge, UK: Cambridge University Press.

Hemmer, P., & Persaud, K. (2014). Interaction between categorical knowledge and episodic memory across domains. *Frontiers in Psychology*, *5*(June), 584. https://doi.org/10.3389/fpsyg.2014.00584

Henderson, J., & Ferreira, F. (2004). *The Interface of Language, Vision, and Action: Eye Movements and the Visual World*. Psychology Press.

Hermer-Vazquez, L., Moffet, A., & Munkholm, P. (2001). Language, space, and the development of cognitive flexibility in humans: The case of two spatial memory tasks. *Cognition*, *79*(3), 263–299. https://doi.org/10.1016/S0010-0277(00)00120-7

Hermer-Vazquez, L., Spelke, E., & Katsnelson, A. (1999). Sources of flexibility in human cognition: Dual-task studies of space and language. *Cognitive Psychology*, *39*(1), 3–36. https://doi.org/10.1006/cogp.1998.0713

Hofstede, G. (1980). Motivation, leadership, and organization: Do American theories apply abroad? *Organizational Dynamics*, *9*(1), 42–63. https://doi.org/10.1016/0090-2616(80)90013-3

Hofstede, G. (1998). *Masculinity and Femininity: The Taboo Dimension of National Cultures*. London: SAGE Publications.

Hoji, H., Kinsui, S., Takubo, Y., & Ueyama, A. (2003). The demonstratives in modern Japanese. *RoutledgeCurzon Asian Linguistics Series*, *97*, 1–21.

Holden, M. P., Newcombe, N. S., & Shipley, T. F. (2013). Location memory in the real world: Category adjustment effects in 3-dimensional space. *Cognition*, *128*(1), 45–55. https://doi.org/10.1016/j.cognition.2013.02.016

Hollingworth, A., & Rasmussen, I. (2010). NIH Public Access. *Journal Experimental Psychology: Human Perception and Performance*, *36*(3), 1–37. https://doi.org/10.1037/a0017836.Binding

Hollingworth, H. L. (1910). The Central Tendency of Judgment. *Journal of Philosophy*, *7*(17), 461–469.

Hommel, B., Musseler, J., Aschersleben, G., & Prinz, W. (2001). The theory of event-coding (TEC): A framework for perception and action planning. *Behavioural and Brain Sciences*, *24*, 849–937. https://doi.org/10.1017/S0140525X01000103

Huebner, G. M., & Gegenfurtner, K. R. (2010). Effects of viewing time, fixations, and viewing strategies on visual memory for briefly presented natural objects. *Quarterly Journal of Experimental Psychology (2006)*, *63*(7), 1398–413. https://doi.org/10.1080/17470210903398139

Huttenlocher, J., Hedges, L. V., & Duncan, S. (1991). Categories and particulars: Prototype effects in estimating spatial location. *Psychological Review*, *98*(3), 352–376. https://doi.org/10.1037//0033-295X.98.3.352

Huttenlocher, J., Hedges, L. V., & Vevea, J. L. (2000). Why do categories affect stimulus judgment? *Journal of Experimental Psychology: General*, *129*(2), 220–241. https://doi.org/10.1037/0096-3445.129.2.220

Huttenlocher, J., Hedges, L. V, Corrigan, B., & Crawford, L. E. (2004). Spatial categories and the estimation of location. *Cognition*, *93*, 75–97. https://doi.org/10.1016/j.cognition.2003.10.006

Jarbou, S. O. (2010). Accessibility vs. physical proximity: An analysis of exophoric demonstrative practice in Spoken Jordanian Arabic. *Journal of Pragmatics*, *42*(11), 3078–3097. https://doi.org/10.1016/j.pragma.2010.04.014

Jiang, Y., Olson, I., & Chun, M. (2000). Organization of visual short-term memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *26*(3), 683–702. https://doi.org/10.1037/0278-7393.26.3.683

Johannsen, K., & De Ruiter, J. P. (2013a). Reference frame selection in dialog:

priming or preference? *Frontiers in Human Neuroscience*, *7*(October), 667. https://doi.org/10.3389/fnhum.2013.00667

Johannsen, K., & De Ruiter, J. P. (2013b). The role of scene type and priming in the processing and selection of a spatial frame of reference. *Frontiers in Psychology*, *4*(APR), 1–7. https://doi.org/10.3389/fpsyg.2013.00182

Kahneman, D., & Treisman, A. M. (1984). Changing Views of Attention and Automaticity. In R. Parasuraman & R. Davies (Eds.), *Varieties of Attention* (pp. 29–61). New York: Academic Press.

Kamio, A. (1994). The theory of territory of information: The case of Japanese. *Journal of Pragmatics*, *21*(1), 67–100.

Kay, P., & McDaniel, C. K. (1978). The linguistic significance of the meanings of basic color terms. *Language*, *54*(3), 610–646. https://doi.org/10.1353/lan.1978.0035

Kay, P., & Regier, T. (2003). Resolving the question of color naming universals. *PNAS*, *100*(15), 9085–9089. Retrieved from www.pnas.org?cgi?doi?10.1073?pnas.1532837100

Kemmerer, D. (1999). “Near” and “far” in language and perception. *Cognition*, *73*(1), 35–63. https://doi.org/10.1016/S0010-0277(99)00040-2

Kemmerer, D., & Tranel, D. (2000). A double dissociation between linguistic and perceptual representations of spatial relationships. *Cognitive Neuropsychology*, *17*(5), 393–414. Retrieved from http://www.tandfonline.com/doi/abs/10.1080/026432900410766

King, J. a, Trinkler, I., Hartley, T., Vargha-Khadem, F., & Burgess, N. (2004). The hippocampal role in spatial memory and the familiarity--recollection distinction: a case study. *Neuropsychology*, *18*(3), 405–417. https://doi.org/10.1037/0894-4105.18.3.405

Knoeferle, P., & Crocker, M. W. (2006). The coordinated interplay of scene, utterance, and world knowledge: evidence from eye tracking. *Cognitive Science*, *30*, 481–529. https://doi.org/10.1207/s15516709cog0000_65

Kveraga, K., Ghuman, A. S., & Bar, M. (2007). Top-down predictions in the cognitive brain. *Brain and Cognition*, *65*(2), 145–168. https://doi.org/10.1016/j.bandc.2007.06.007

Lakoff, G., & Johnson, M. (2008). *Metaphors we live by*. Univesity of Chicago press.

Lamme, V. A. F., & Roelfsema, P. R. (2000). The distinct modes of vision offered by feedforward and recurrent processing. *Trends in Neurosciences*, *23*(11), 571–579. https://doi.org/10.1016/S0166-2236(00)01657-X

Landau, B., & Jackendoff, R. (1993). “What” and “where” in spatial language and spatial cognition. *Behavioral and Brain Sciences*, *16*(2), 217–265. https://doi.org/10.1017/S0140525X00029733

Lane, A. R., Ball, K., Smith, D. T., Schenk, T., & Ellison, A. (2013). Near and far space: Understanding the neural mechanisms of spatial attention. *Human Brain Mapping*, *34*(2), 356–366. https://doi.org/10.1002/hbm.21433

Levelt, W. J. M. (1982). Cognitive styles in the use of spatial direction terms. *Speech, Place, and Action: Studies in Deixis and Related Topics*, 251–268. Retrieved from http://pubman.mpdl.mpg.de/pubman/item/escidoc:67007:4/component/escidoc:468509/Levelt_Cognitive_1982.pdf

Levelt, W. J. M., & Kelter, S. (1982). Surface Form and Memory in Question Answering. *Cognitive Psychology*, *14*, 78–106.

Levinson, S. C. (1996). Frames of Reference and Molyneux’s Question: Crosslinguistic Evidence. In *Language and Space* (pp. 109–169).

Levinson, S. C. (2004). *Space in Language and Cognition: Explorations in Cognitive Diversity*. Cambridge, UK: Cambridge University Press.

Levinson, S. C., Kita, S., Haun, D. B. M., & Rasch, B. H. (2002). Returning the tables: Language affects spatial reasoning. *Cognition*, *84*(2), 155–188. https://doi.org/10.1016/S0010-0277(02)00045-8

Li, P., Abarbanell, L., Gleitman, L., & Papafragou, A. (2011). Spatial reasoning in Tenejapan Mayans. *Cognition*, *120*(1), 33–53. https://doi.org/10.1016/j.cognition.2011.02.012

Li, P., & Gleitman, L. (2002). Turning the tables: language and spatial reasoning. *Cognition*, *83*(3), 265–94. https://doi.org/10.1016/S0010-0277(02)00009-4

Linde, C., & Labov, W. (1975). Spatial networks as a site for the study of langauge and thought. *Language and Cognitive Processes*, *51*, 924–939. Retrieved from http://www.jstor.org/stable/412701 Accessed : 29-03-2

Lipinski, J., Schneegans, S., Sandamirskaya, Y., Spencer, J. P., & Schöner, G. (2012). A neurobehavioral model of flexible spatial language behaviors. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *38*(6), 1490–1511. https://doi.org/10.1037/a0022643

Lipinski, J., Simmering, V. R., Johnson, J. S., & Spencer, J. P. (2010). The role of experience in location estimation: Target distributions shift location memory biases. *Cognition*, *115*(1), 147–153. https://doi.org/10.1016/j.cognition.2009.12.008

Loewenstein, J., & Gentner, D. (2005). Relational language and the development of relational mapping. *Cognitive Psychology*, *50*(4), 315–53. https://doi.org/10.1016/j.cogpsych.2004.09.004

Loftus, E. F. (1975). Leading questions and the eyewitness report. *Cognitive Psychology*, *7*(4), 560–572. https://doi.org/10.1016/0010-0285(75)90023-7

Loftus, E., Miller, D. G., & Burns, H. J. (1978). Semantic integration of verbal information into a visual memory. *Journal of Experimental Psychology: Human Learning & Memory*, *4*(1), 19–31. https://doi.org/10.1037/0278-7393.4.1.19

Loftus, E., & Palmer, J. (1974). Reconstruction of Automobile Destruction : An Example of the Interaction Between Language and Memory. *Journal of Verbal Learning and Verbal Behavior*, *589*, 585–589.

Logan, G. D., & Sadler, D. D. (1996). A computational analysis of the apprehension of spatial relations. *Language and Space*.

Longo, M. R., & Lourenco, S. F. (2006). On the nature of near space: effects of tool use and the transition to far space. *Neuropsychologia*, *44*(6), 977–81. https://doi.org/10.1016/j.neuropsychologia.2005.09.003

Longo, M. R., & Lourenco, S. F. (2007). Spatial attention and the mental number line: Evidence for characteristic biases and compression. *Neuropsychologia*, *45*(7), 1400–1407. https://doi.org/10.1016/j.neuropsychologia.2006.11.002

Lupyan, G., & Clark, A. (2015). Words and the world: Predictive coding and the language-perception-cognition interface. *Current Directions in Psychology*, 1–10. https://doi.org/10.1177/0963721415570732

Lupyan, G., Thompson-Schill, S. L., & Swingley, D. (2010). Conceptual Penetration of Visual Processing. *Psychological Science*, *21*(5), 682–691. https://doi.org/10.1177/0956797610366099

Luz, S., & van der Sluis, I. (2008). Production of Demonstratives in Dutch, English and Portuguese Dialogues. *Proceedings of the 13th European Workshop on Natural Langauge Generation (ENLG)*, (September), 181–186.

Lynott, D., & Coventry, K. (2014). On the ups and downs of emotion: testing between conceptual-metaphor and polarity accounts of emotional valence-spatial location interactions. *Psychonomic Bulletin & Review*, *21*(1), 218–26. https://doi.org/10.3758/s13423-013-0481-5

Maes, A., & de Rooij, C. (2007). (How) Do Demonstratives Code Distance. In A. Branco, T. McEnery, R. Mitkov, & F. Silva (Eds.), *Proceedings of the 6th Discourse Anaphora and Anaphor Resolution Colluquium Daarc 2007* (pp. 83–89). Lagos: Centro Linguistica da Universidade do Porto.

Majid, A., Bowerman, M., Kita, S., Haun, D. B. M., & Levinson, S. C. (2004). Can language restructure cognition? The case for space. *Trends in Cognitive Sciences*, *8*(3), 108–114. https://doi.org/10.1016/j.tics.2004.01.003

Makin, T. R., Holmes, N. P., & Zohary, E. (2007). Is that near my hand? Multisensory representation of peripersonal space in human intraparietal sulcus. *The Journal of Neuroscience : The Official Journal of the Society for Neuroscience*, *27*(4), 731–40. https://doi.org/10.1523/JNEUROSCI.3653-06.2007

Marr, D. (1982). The Philosophy and the Approach. In *Vision: A computational investigation into the human representation and processing of visual information* (Vol. 79, pp. 213–219). New York: Freeman.

Mazoyer, B., Zago, L., Jobard, G., Crivello, F., Joliot, M., Perchey, G., … Tzourio-Mazoyer, N. (2014). Gaussian mixture modeling of hemispheric lateralization for language in a large sample of healthy individuals balanced for handedness. *PLoS ONE*, *9*(6), 9–14. https://doi.org/10.1371/journal.pone.0101165

McCloskey, M., & Zaragoza, M. (1985). Misleading Postevent Information and Memory for Events: arguments and evidence against memory impairment hypotheses. *Journal of Experimental Psychology: General*, *114*(1).

Meier, B. P., & Robinson, M. D. (2004). Why the sunny side is up: association between affect and vertical position. *Psychological Science*, *15*(4), 243–7. https://doi.org/10.1111/j.0956-7976.2004.00659.x

Meteyard, L., Bahrami, B., & Vigliocco, G. (2007). Motion detection and motion verbs: Language affects low-level visual perception. *Psychological Science*, *18*(11), 1007–1013. https://doi.org/10.1111/j.1467-9280.2007.02016.x

Meteyard, L., Cuadrado, S. R., Bahrami, B., & Vigliocco, G. (2012). Coming of age: A review of embodiment and the neuroscience of semantics. *Cortex*, *48*(7), 788–804. https://doi.org/10.1016/j.cortex.2010.11.002

Meyer, K., & Damasio, A. (2009). Convergence and divergence in a neural architecture for recognition and memory. *Trends in Neurosciences*, *32*(7), 376–382. https://doi.org/10.1016/j.tins.2009.04.002

Miller, G. A., & Johnson-Laird, P. N. (1976). *Language and Perception* (Vol. 23). Belknap Press.

Mills, G., & Healey, P. (2008). Semantic negotiation in dialogue: the mechanisms of alignment. *SigDial Workshop on Discourse and Dialogue*, (June), 46–53. Retrieved from http://www.eecs.qmul.ac.uk/~gj/MillsHealey2008sigdial.pdf

Mummery, C. J., Shallice, T., & Price, C. J. (1999). Dual-process model in semantic priming: A functional imaging perspective. *NeuroImage*, *9*(5), 516–525. https://doi.org/10.1006/nimg.1999.0434

Munnich, E., Landau, B., & Dosher, B. A. (2001). Spatial language and spatial representation: A cross-linguistic comparison. *Cognition*, *81*(3), 171–207. https://doi.org/10.1016/S0010-0277(01)00127-5

Nakamura, J. (2012). *Cognitive semantics of the qualitative suffix in the Japanese spatial demonstratives koko , soko , and asoko : SPACE , THING , and TIME*.

Naveh-Benjamin, M. (1987). Coding of spatial location information: An automatic process? *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *13*(4), 595–605. https://doi.org/10.1037//0278-7393.13.4.595

Naveh-Benjamin, M. (1988). Recognition memory of spatial location information: another failure to support automaticity. *Memory & Cognition*, *16*(5), 437–445. https://doi.org/10.3758/BF03214224

Nelson, T. O., & Chaiklin, S. (1980). Immediate memory for spatial location. *Journal of Experimental Psychology: Human Learning & Memory*, *6*(5), 529–545. https://doi.org/10.1037//0278-7393.6.5.529

Newport, E. L., & Aslin, R. N. (2004). Learning at a distance I. Statistical learning of non-adjacent dependencies. *Cognitive Psychology*, *48*(2), 127–162. https://doi.org/10.1016/S0010-0285(03)00128-2

Niimura, T., & Hayashi, B. (1994). English and Japanese Demonstratives: A contrastive analysis of second language acquisition. *Issues in Applied Linguistics*, *5*(2), 217–220. https://doi.org/10.5811/westjem.2011.5.6700

O'Keefe, J., & Nadel, L. (1978). *The Hippocampus as a cognitive map*. *Hippocampus*. Oxford, UK: Oxford University Press.

Oh, S.-Y. (2001). A Focus-Based Study of English Demonstrative Reference.

Journal of English Linguistics, *29*(2), 124–148.

Okazaki, T. (2011). The domains of the deictic uses on the demonstrative pronouns, *40*, 29–48.

Ozturk, O., Shayan, S., Liszkowski, U., & Majid, A. (2013). Language is not necessary for color categories. *Developmental Science*, *16*(1), 111–115. https://doi.org/10.1111/desc.12008

Özyürek, A. (1998). An Analysis of the Basic Meaning of Tukish Demonstratives in Face-to-Face Conversational Interactin. In *Oralité et gestualité: Communication multimodale, interaction* (pp. 604–614). Paris: L'Harmattan.

Papafragou, A., Hulbert, J., & Trueswell, J. (2008). Does language guide event perception? Evidence from eye movements. *Cognition*, *108*(1), 155–184. https://doi.org/10.1016/j.cognition.2008.02.007

Papafragou, A., Massey, C., & Gleitman, L. (2002). Shake, rattle, "n" roll: The representation of motion in language and cognition. *Cognition*, *84*(2), 189–219. https://doi.org/10.1016/S0010-0277(02)00046-X

Pederson, E., Danziger, E., Wilkins, D., Levinson, S. C., Kita, S., & Senft, G. (1998). Semantic Typology and Spatial Conceptualization. *Linguistic Society of America*, *74*(3), 557–589. Retrieved from http://www.jstor.org/stable/417793

Pederson, E., & Wilkins, D. (1996). A cross-linguistic questionnaire on "demonstratives." In S. C. Levinson & A. Majid (Eds.), *Manual for the 1996 Field Season* (pp. 1–11). Nijmegen, Max Planck Institute for Psycholinguistics.

Peeters, D., Hagoort, P., & Ozyürek, A. (2014). Electrophysiological evidence for the role of shared space in online comprehension of spatial demonstratives. *Cognition*, *136*, 64–84. https://doi.org/10.1016/j.cognition.2014.10.010

Peeters, D., Hagoort, P., & Özyürek, A. (2015). Electrophysiological evidence for the role of shared space in online comprehension of spatial demonstratives. *Cognition*, *136*, 64–84. https://doi.org/10.1016/j.cognition.2014.10.010

Peeters, D., & Özyürek, A. (2016). This and That Revisited: A Social and Multimodal Approach to Spatial Demonstratives. *Frontiers in Psychology*, *7*(February), 2014–2017. https://doi.org/10.3389/fpsyg.2016.00222

Pertzov, Y., Dong, M. Y., Peich, M.-C., & Husain, M. (2012). Forgetting what was where: the fragility of object-location binding. *PloS One*, *7*(10), e48214. https://doi.org/10.1371/journal.pone.0048214

Pertzov, Y., & Husain, M. (2014). The privileged role of location in visual working memory. *Attention, Perception & Psychophysics*, *76*(7), 1914–24. https://doi.org/10.3758/s13414-013-0541-y

Pica, P., Lemer, C., Izard, V., & Dehaene, S. (2004). Exact and Approximate Arithmetic in an Amazonian IndigeneGroup.. *Science*, *306*(5695), 499–503. https://doi.org/10.1126/science.1102085

Pickering, M. J., & Garrod, S. (2004). Toward a mechanistic psychology of dialogue. *The Behavioral and Brain Sciences*, *27*(2), 169-190-226. https://doi.org/10.1017/S0140525X04000056

Pickering, M. J., & Garrod, S. (2006). Alignment as the basis for successful communication. *Research on Language and Computation*, *4*(2–3), 203–228. https://doi.org/10.1007/s11168-006-9004-0

Pilling, M., Wiggett, A., Ozgen, E., & Davies, I. R. L. (2003). Is color "categorical perception" really perceptual? *Memory & Cognition*, *31*(4), 538–551. https://doi.org/10.3758/BF03196095

Pinto, Y., Gaal, S. Van, Lange, F. P. De, Lamme, V. a F., & Seth, A. K. (2015). Expectations accelerate entry of visual stimuli into awareness. *J. Vision*, *15*(8), 1–15. https://doi.org/10.1167/15.8.13.doi

Postma, A., & Haan, E. H. F. De. (1996). What Was Where? Memory for Object Locations. *Quarterly Journal of Experimental Psychology*, (1), 178–200. Retrieved from http://web.b.ebscohost.com/ehost/pdfviewer/pdfviewer?sid=16c9d2a7-2cbb-40b7-a936-feb0a9328cd8%2540sessionmgr198&vid=1&hid=109

Postma, A., Kessels, R. P. C., & van Asselen, M. (2008). How the brain remembers and forgets where things are: the neurocognition of object-location memory. *Neuroscience and Biobehavioral Reviews*, *32*(8), 1339–45. https://doi.org/10.1016/j.neubiorev.2008.05.001

Pulvermüller, F. (2012). Meaning and the brain: The neurosemantics of referential, interactive, and combinatorial knowledge. *Journal of Neurolinguistics*, *25*(5), 423–459. https://doi.org/10.1016/j.jneuroling.2011.03.004

Pylyshyn, Z. (1999). Is vision continuous with cognition? The case for cognitive impenetrability of visual perception. *The Behavioral and Brain Sciences*, *22*(3), 341-365-423. https://doi.org/10.1017/S0140525X99002022

Rasmussen, T., & Milner, B. (1977). the Role of Early Left-Brain Injury in Determining Lateralization of Cerebral Speech Functions. *Annals of the New*

York Academy of Sciences, *299*(1), 355–369. https://doi.org/10.1111/j.1749-6632.1977.tb41921.x

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. https://doi.org/10.1037//0096-3445.130.2.273

Regier, T., & Xu, Y. (2017). The Sapir-Whorf hypothesis and inference under uncertainty. *Wiley Interdisciplinary Reviews: Cognitive Science*, 1–15.

Restle, F. (1970). Speed of adding and comparing numbers. *Journal of Experimental Psychology*, *83*(2, Pt.1), 274–278. https://doi.org/10.1037/h0028573

Rizzolatti, G., Fadiga, L., Fogassi, L., & Gallese, V. (1997). The space around us. *Science*, *277*(5323), 190–191.

Roberson, D., & Davidoff, J. (2000). The categorical perception of colors and facial expressions: the effect of verbal interference. *Memory & Cognition*, *28*(6), 977–986. https://doi.org/10.3758/BF03209345

Roberson, D., Davies, I., & Davidoff, J. (2000). Color categories are not universal: replications and new evidence from a stone-age culture. *Journal of Experimental Psychology: General*, *129*(3), 369–398. https://doi.org/10.1037/0096-3445.129.3.369

Saffran, J. R. (2002). Constraints on Statistical Language Learning. *Journal of Memory and Language*, *47*(1), 172–196. https://doi.org/10.1006/jmla.2001.2839

Saffran, J. R., Aslin, R. N., & Newport, E. L. (1996). Statistical learning by 8-month-old infants. *Science*. Retrieved from http://www.sciencemag.org/content/274/5294/1926.short

Sakata, Y. (1992). Chapter 6. In S. Kinsui & Y. Takubo (Eds.), *Demonstratives (Shiji-shi)* (2nd ed., p. 54). Tokyo: Hituzi Shobo.

Sapir, E. (1929). The Status of Linguistics as a Science. *Language*, *5*, 207–214.

Schober, M. F. (1993). Spatial perspective-taking in conversation. *Cognition*, *47*, 1–24.

Schooler, J. W., & Engstler-Schooler, T. Y. (1990). Verbal overshadowing of visual memories: Some things are better left unsaid. *Cognitive Psychology*, *22*(1), 36–71. https://doi.org/10.1016/0010-0285(90)90003-M

Shaki, S., & Fischer, M. H. (2008). Reading space into numbers - a cross-linguistic comparison of the SNARC effect. *Cognition*, *108*(2), 590–599. https://doi.org/10.1016/j.cognition.2008.04.001

Shaki, S., Fischer, M. H., & Petrusic, W. M. (2009). Reading habits for both words and numbers contribute to the SNARC effect. *Psychonomic Bulletin & Review*, *16*(2), 328–31. https://doi.org/10.3758/PBR.16.2.328

Shi, Z., Zhou, A., Han, W., & Liu, P. (2011). Effects of ownership expressed by the first-person possessive pronoun. *Consciousness and Cognition*, *20*(3), 951–5. https://doi.org/10.1016/j.concog.2010.12.008

Slobin, D. I. (1996). From Thought and Language To Thinking Fo Speaking. In J. J. Gumperz & S. C. Levinson (Eds.), *Rethinking linguistic relativity* (pp. 70–96). Cambridge, UK: Cambridge University Press.

Slobin, D. I. (2003). Language and Thought Online : Cognitive Consequences of Linguistic Relativity. In *Language in Mind: Advances in the study of language and thought* (pp. 159–192). Cambridge, MA: MIT Press.

Smith, L. B., & Samuelson, L. K. (2010). The Spatial Foundations of Language and Cognition. In K. S. Mix, L. B. Smith, & M. Gasser (Eds.), *The Spatial Foundations of Language and Cognition* (pp. 1–328). https://doi.org/10.1093/acprof:oso/9780199553242.001.0001

Spencer, J. P., & Hund, A. M. (2002). Prototypes and particulars: Geometric and experience-dependent spatial categories. *Journal of Experimental Psychology: General*, *131*(1), 16–37. https://doi.org/10.1037/0096-3445.131.1.16

Spivey, M. J., Tyler, M. J., Eberhard, K. M., & Tanenhaus, M. K. (2001). Linguistically mediated visual search. *Psychological Science*, *12*(4), 282–286. https://doi.org/10.7897/2277-4343.04135

Stevens, J., & Zhang, Y. (2013). Relative distance and gaze in the use of entity-referring spatial demonstratives: An event-related potential study. *Journal of Neurolinguistics*, *26*(1), 31–45. https://doi.org/10.1016/j.jneuroling.2012.02.005

Stevens, J., & Zhang, Y. (2014). Brain mechanisms for processing co-speech gesture: A cross-language study of spatial demonstratives. *Journal of Neurolinguistics*, *30*(1), 27–47. https://doi.org/10.1016/j.jneuroling.2014.03.003

Strauss, S. (2002). This, that, and it in spoken American English: A demonstrative system of gradient focus. *Language Sciences*, *24*(2), 131–152. https://doi.org/10.1016/S0388-0001(01)00012-2

Takahashi, T. (1992). Chapter 3. In *Shiji-shi (demonstratives)* (2nd ed., pp. 38–46).

Tokyo: Hituzi Shobo.

Talmy, L. (1975). Figure and Ground in Complex Sentences. *Proceedings of the First Annual Meeting of the Berkeley Linguistics*, 419–430.

Tanaka, J. W., & Presnell, L. M. (1999). Color diagnosticity in object recognition. *Perception & Psychophysics*, *61*(6), 1140–1153. https://doi.org/10.3758/BF03207619

Tanenhaus, M. K., Spivey-Knowlton, M. J., Eberhard, K. M., & Sedivy, J. C. (1995). Integration of visual and linguistic information in spoken language comprehension. *Science*, *268*(5217), 1632–1634. https://doi.org/10.1126/science.7777863

Tenbrink, T. (2011). Reference frames of space and time in language. *Journal of Pragmatics*, *43*(3), 704–722. https://doi.org/10.1016/j.pragma.2010.06.020

ter Horst, A. C., van Lier, R., & Steenbergen, B. (2011). Spatial dependency of action simulation. *Experimental Brain Research*, *212*(4), 635–44. https://doi.org/10.1007/s00221-011-2748-2

Thelen, E. (2005). Dynamic Systems Theory and the Complexity of Change. *Psychoanalytic Dialogues: The International Journal of Relational Perspectives*, *15*(2), 255–283. https://doi.org/10.1080/10481881509348831

Thierry, G., Athanasopoulos, P., Wiggett, A., Dering, B., & Kuipers, J.-R. (2009). Unconscious effects of language-specific terminology on preattentive color perception. *Proceedings of the National Academy of Sciences of the United States of America*, *106*(11), 4567–70. https://doi.org/10.1073/pnas.0811155106

Tomasello, M. (1999). The Human Adaptation for Culture, *28*(1999), 509–529. https://doi.org/10.1146/annurev.soc.28.110601.140936

Tomasello, M., Carpenter, M., Call, J., Behne, T., & Moll, H. (2005). Understanding and Sharing intentions: the origins of cultural cognition. *Behavioral and Brain Sciences*, 675–735.

Treisman, A.M., Gelade, G. (1980). Treisman & Gelade (1980) feature integration theory of attention, Cogn. Cognitive Psychology.

Treisman, A., & Zhang, W. (2006). Location and binding in visual working memory. *Memory & Cognition*, *34*(8), 1704–1719. https://doi.org/10.3758/BF03195932

Trueswell, J. C., & Papafragou, A. (2010). Perceiving and remembering events cross-linguistically: Evidence from dual-task paradigms. *Journal of Memory and Language*, *63*(1), 64–82. https://doi.org/10.1016/j.jml.2010.02.006

Turk, D. J., Gillespie-Smith, K., Krigolson, O. E., Havard, C., Conway, M. a., & Cunningham, S. J. (2015). Selfish learning: The impact of self-referential encoding on children’s literacy attainment. *Learning and Instruction*, *40*, 54–60. https://doi.org/10.1016/j.learninstruc.2015.08.001

Turk, D. J., van Bussel, K., Waiter, G. D., & Macrae, C. N. (2011). Mine and me: exploring the neural basis of object ownership. *Journal of Cognitive Neuroscience*, *23*(11), 3657–68. https://doi.org/10.1162/jocn_a_00042

Umiltà, C., Priftis, K., & Zorzi, M. (2009). The spatial representation of numbers: Evidence from neglect and pseudoneglect. *Experimental Brain Research*, *192*(3), 561–569. https://doi.org/10.1007/s00221-008-1623-2

Vinson, D. W., Engelen, J., Zwaan, R. A., Matlock, T., & Dale, R. (2017). Implied motion language can influence visual spatial memory. *Memory & Cognition*. https://doi.org/10.3758/s13421-017-0699-y

Voyer, D., Postma, A., Brake, B., & Imperato-McGinley, J. (2007). Gender differences in object location memory: A meta-analysis. *Psychonomic Bulletin & Review*, *14*(1), 23–38. https://doi.org/10.3758/BF03194024

Wang, R. F., & Spelke, E. S. (2002). Human spatial representation: insights from animals. *Trends in Cognitive Sciences*, *6*(9), 376–382. https://doi.org/10.1016/S1364-6613(02)01961-7

Watson, M. E., Pickering, M. J., & Branigan, H. P. (2004). Alignment of Reference Frames in Dialogue. *Proceedings of the 26th Annual Conference of the Cognitive Science Society*.

Whorf, B. L. (1956). Language, Thought, and Reality: selected writings of benjamin lee whorf. In J. B. Carroll (Ed.), *Language, Thought, and Reality: selected writings of benjamin lee whorf*. Cambridge, MA: The M.I.T. Press.

Wilkins, D. P. (1999). The 1999 demonstrative questionnaire: “this” and “that” in comparative perspective. In D. Wilkins (Ed.), *Manual for the 1999 Field Season* (pp. 1–24). Nijmegen, Max Planck Institute for Psycholinguistics.

Winawer, J., Witthoft, N., Frank, M. C., Wu, L., Wade, A. R., & Boroditsky, L. (2007). Russian blues reveal effects of language on color discrimination. *Proceedings of the National Academy of Sciences of the United States of America*, *104*(19), 7780–5. https://doi.org/10.1073/pnas.0701644104

Wolff, P., & Holmes, K. J. (2011). Linguistic relativity. *Wiley Interdisciplinary Reviews: Cognitive Science*, *2*(3), 253–265. https://doi.org/10.1002/wcs.104

Yoshimoto. (1992). Demonstratives in Japanese. In Takabo & Kinsui (Eds.), *Demonstratives* (1992nd ed., p. 105).

Yoshimoto, K. (1986). On Demonstratives KO/SO/A in Japanese. *言語研究 (Gengo Kenkyu)*, *90*, 48–72.

Zorzi, M., Priftis, K., & Umiltà, C. (2002). Brain damage: neglect disrupts the mental number line. *Nature*, *417*(6885), 138–139. https://doi.org/10.1038/417138a

Zwaan, R. a., & Radvansky, G. a. (1998). Situation models in language comprehension and memory. *Psychological Bulletin*, *123*(2), 162–185. https://doi.org/10.1037//0033-2909.123.2.162

Appendices

Appendix A

Gudde, H. B., Coventry, K. R., & Engelhardt, P. E. (2016). Language and memory for object location. *Cognition*, *153*, 99–107. Retrieved from http://www.sciencedirect.com/science/article/pii/S0010027716301044

Language and Memory for Object Location

Harmen B. Gudde*[12*], Kenny R. Coventry*, and Paul E. Engelhardt

University of East Anglia, UK

Word count: 6993

[*] *Corresponding authors at the School of Psychology, University of East Anglia, Norwich, NR4 7TJ, UK.

Email address: h.gudde@uea.ac.uk or k.coventry@uea.ac.uk

Abstract

In three experiments, we investigated the influence of two types of language on memory for object location: demonstratives (*this*, *that*) and possessives (*my*, *your*). Participants first read instructions containing demonstratives/possessives to place objects at different locations, and then had to recall those object locations (following object removal). Experiments 1 and 2 tested contrasting predictions of two possible accounts of language on object location memory: the *Expectation Model* (Coventry, Griffiths, & Hamilton, 2014) and the *congruence account* (Bonfiglioli et al., 2009). In Experiment 3, the role of attention allocation as a possible mechanism was investigated. Results across all three experiments show striking effects of language on object location memory, with the pattern of data supporting the Expectation Model. In this model, the expected location cued by language and the actual location are concatenated leading to (mis)memory for object location, consistent with models of predictive coding (Bar, 2009; Friston, 2003).

Keywords: memory; object location; spatial demonstratives; possessives; peripersonal/ extrapersonal space

Language and Memory for Object Location

1. Introduction

The relationship between language and non-linguistic representations is a fundamental topic in the cognitive sciences. Often this relationship is approached from the standpoint of the extent to which non-linguistic representations are necessary for language comprehension (e.g. within the framework of 'embodied' cognition; cf. Barsalou, 1999). However, equally important is the extent to which language can influence non-linguistic processes (Coventry, Christophel, Fehr, Valdés-Conroy, & Herrmann, 2013). Language can direct the attention of a conspecific to the spatial world; spatial expressions, such as *these coins* or *the cup is on the table* serve to direct the attention of a hearer to regions of space (Miller & Johnson-Laird, 1976). And the pairing of language with visual events and images also affects what is recalled about the spatial world. For example, Loewenstein and Gentner (2005) found that children performed better in a mapping task when spatial relations were paired with spatial language at encoding (e.g. "I'm putting the book on the shelf"). They argue that relational language fosters the development of representational structures that facilitate cognitive processing (see also Hermer-vazquez, Spelke, & Katsnelson, 1999).

Language can facilitate the binding and maintenance of color-location conjunctions (Dessalegn & Landau, 2008, 2013; Farran & O'Leary, 2015). For example, in a memory experiment, four-year olds performed a task in which a target (e.g. a square split in half by two different colors) was presented which they then had to find in an array. Performance was enhanced if the target was accompanied by spatial cues (e.g. "yellow is on top"). There was no additional benefit for children verbalizing the linguistic cue themselves over just hearing the cue, as long as they had a stable understanding of the spatial terms (Farran & O'Leary, 2015).

As well as facilitating memory, language presented with a spatial scene can also lead to memory errors (Feist & Gentner, 2007; Gentner & Loftus, 1979). For instance, Feist and Gentner (2007) showed that recognition memory for spatial scenes was shifted in the direction of the spatial relational language (spatial prepositions) presented with scenes at encoding. In their study they presented participants with ambiguous pictures depicting spatial relations accompanied with or without spatial sentences. When participants responded in a later yes-no recognition task, spatial language at encoding was associated with more false positives (in cases where the spatial language at encoding was associated with a more prototypical version of the spatial relation than the relation actually shown). Feist and Gentner (2007) suggest this is a result of an interactive encoding of language and visual memory, in which language influences the way people encode visual scenes. More broadly language can be used as a tool in a task to aid memory and/or processing of spatial information (see for example Frank, Everett, Fedorenko, & Gibson, 2008; Li, Abarbanell, Gleitman, & Papafragou, 2011) consonant with some weaker variants of so-called 'linguistic relativity' (see Wolff & Holmes, 2011 for review).

The effects of language on memory are not limited to spatial cognition. It has also been found that presenting possessive pronouns in combination with a memory task enhances response times and memory for objects (Shi et al., 2011). Shi et al. presented Chinese nouns preceded by a pronoun (*my/ his*). Participants had to scale the presented nouns for likeability and were given a surprise memory test. In the 'my' condition, participants responded faster and showed a better memory performance for the nouns than in the 'his' condition.

Although it has been shown that language can influence memory, it has yet to be demonstrated *how* it does so. In this paper, our focus is on the (possible) influence of spatial demonstratives and possessives on memory for object location. The continuous nature of object location memory errors affords testing directly between a number of possible mechanisms regarding how language affects memory for object location.

Spatial demonstratives (e.g. *this/ that)* are among the earliest words children learn (Diessel, 2006) and have been shown to be associated with discrete zones of peri-personal (near) and extra-personal (far) perceptual space (Coventry, Valdés, Castillo, & Guijarro-Fuentes, 2008; Diessel, 2006; Maes & de Rooij, 2007; Stevens & Zhang, 2013; cf. Peeters, Hagoort, & Ozyürek, 2014). However, this distinction is flexible and graded. Near space can be extended or contracted by tool or weight use (Longo & Lourenco, 2006), and the use of *this* is similarly extended when participants use a stick to point at objects (Coventry et al., 2008). In addition to distance, demonstrative choice is also affected by other variables. Coventry et al. (2014) explored the relationship between object knowledge and distance on both demonstrative choice in English and memory for object location. Across seven experiments they found that object familiarity (i.e. familiar versus unfamiliar colored shapes), object ownership (whether the participant owned the object or not) and object visibility (whether the object was covered with an opaque cover or not) all affected demonstrative choice to describe object location and (non-linguistic) memory for object location. For example, unfamiliar objects (low frequency color-shape combinations, such as a viridian nonagon) were misremembered as being further away than they actually were relative to familiar objects (e.g. red square). In order to account for both the demonstrative choice data and the memory data, Coventry et al. (2014) proposed a model of the influence of object knowledge on both measures. In their *Expectation Model*, memory for object location is a combination of where an object is located and where an object is expected to be located (see Figure 1a). The expectation of the objects' location is combined with the actual object location (with an associated estimation error) in memory, as follows:

$$M_D = f(D_a, D_{exp}, D_{err})$$

where M = signed memory error, D = distance, $_a$ = actual, $_{exp}$ = expected and $_{err}$ = estimation error.

Coventry et al. (2014) acknowledge that the model may operate at encoding of object location or at retrieval. If the former is the case, it is assumed that an object expected to be in peripersonal space (such as an object owned by the participant), activates peripersonal space as the participant encodes object location, and therefore that the actual representation of location at encoding, and later memory is a concatenation of expectation of where an object is most likely to be located and where it is actually located. The alternative possibility is that the location errors emerge only at retrieval, consistent with effects found in the verbal overshadowing (Alogna et al., 2014; Schooler & Engstler-Schooler, 1990) and eye-witness testimony literatures (E. Loftus et al., 1978; E. Loftus & Palmer, 1974; McCloskey & Zaragoza, 1985).

Coventry et al. (2014) did not examine the influence of language on memory for object location, but by extension the expectation model makes predictions regarding how language might impact upon memory for location. As *this* is associated with near space and *that* with far space, one can assume that the expected distance value associated with *that* would be greater than the expected value distance associated with *this*. Combined with the actual distance, the expectation model therefore predicts a main effect of language on memory for object location, with *that* associated with (mis)memory for objects further away than they actually were compared to *this* (Figure 1a). In addition the Expectation Model predicts an effect of location, in which memory for objects further away is worse than for objects closer by.

In contrast to the expectation model, there is a considerable body of work within an 'embodied cognition' framework providing evidence for the importance of congruence/incongruence effects between language and space that makes different predictions from the expectation model. A growing body of literature suggests that participants' performance is affected by congruence/incongruence between language or concepts and space. For example, it has been shown that participants respond more quickly to positively valenced stimuli in a congruent high location than an incongruent low location, and vice versa for negative stimuli (e.g. Barsalou, 2008; Meier & Robinson, 2004; cf. Lynott & Coventry, 2014). What one might term a 'congruence account' has been extended to movement planning, whereby movements are prepared based on given language (Bonfiglioli et al., 2009; see also Stevens & Zhang, 2013). For example, Bonfiglioli et al. (2009) required participants to grip an object after listening to an instruction that indicated whether the object was near or far. A significant interaction was found in which performance was better when the descriptive language and space were congruent compared to incongruent situations - reaction times were significantly longer when language was incongruent with space compared to when language and space were congruent. Bonfiglioli et al., (2009) therefore concluded that they found interference effects from the used pronouns at the level of movement planning. In line with the Theory of Event-Coding (TEC) (Hommel et al., 2001), we extend these findings in the action literature to memory. In the TEC, it is suggested that perception and action share an indistinguishable underlying representational medium. This would entail that, for example, memory and action are based on the same cognitive codes. Therefore, if an effect of interference due to incongruence is found in action planning, it should be found in memory (Hommel et al., 2001). When we extend the effects of congruence on action to memory for object location, we would therefore predict a similar interaction. Congruence in language and space would be expected to enhance the accuracy of memory for location, with greater errors (without specification of direction) when there is a mismatch between the demonstrative and location, as follows:

$$M_D = f(D_a, C, D_{err})$$

where M = signed memory error, D = distance, $_a$ = actual, C = congruence of language with location and $_{err}$ = estimation error (Figure 1b). This means that when a congruent demonstrative is used to describe an object's location (e.g. 'this' for an object close by, or 'that' for an object further away), memory for object location is expected to be more accurate than when language and situation are incongruent (e.g. 'that' for an object close by, 'this' for an object further away).

Distinct from both the Expectation and Congruence models, the possible effect of language on memory should also be considered in relation to the allocation of attention. A large literature shows that language affects where one looks in a visual scene, for example in terms of fixating particular objects when they are mentioned (e.g., Allopenna, Magnuson, & Tanenhaus, 1998; Tanenhaus, Spivey-Knowlton, Eberhard, & Sedivy, 1995). Building on this earlier work, it is possible that language also affects the amount of time one spends looking at an object. Given that *this* is associated with proximity to a speaker, one might speculate that participants might look longer at an object at a location when preceded with *this* compared with *that*, as visual attention is allocated preferentially to near objects compared to objects further away (Garrido-Vásquez & Schubö, 2014). Following evidence that longer looking times are associated with better memory performance (e.g. Huebner & Gegenfurtner, 2010), one might then predict better accuracy of recall for trials preceded with this compared with that. In Experiment 3 we used eye tracking during the encoding phase, to investigate whether differences are driven by attention allocation (Figure 1c). The goal of the present studies was to test whether language affects memory for object location, and to elucidate the mechanism involved. Specifically, we aimed to tease apart these three accounts by examining the effects of demonstrative and possessives on memory for object location. The first experiment tested whether spatial demonstratives affected memory for object location with contrasting predictions from two possible models of how language affects memory: congruence vs. expectation. Experiment 2 tests whether the effects found for demonstratives also occur for possessives (*my/your*) – terms which have also been associated with the peripersonal/extrapersonal space distinction. Experiment 3 tests predictions from the attention allocation model using eye tracking.

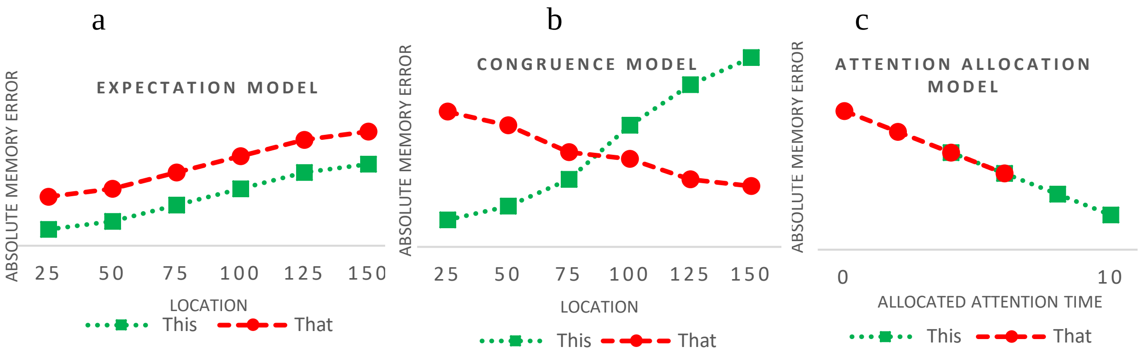


Figure 1. Predictions by the different models, from left to right: a. Expectation Model, b. Congruence Model, c. Attention Allocation model. On the y-axis the difference between the actual location and the remembered location is presented, a higher value on the y-axis means an object is remembered as being further away than it actually was. In 1a and 1b, the six distances from the participant (in cm) used in Experiment 1 and 2 are plotted. In 1c, the x-axis represents the total possible fixation time (10 seconds) for participants in Experiment 3. The lines represent the influence of demonstratives (*this/that*). In Figure 1c more attention leads to a smaller memory error, and '*this'* is predicted to elicit more attention than '*that'*.

2. Experiment 1: The influence of demonstratives on spatial memory

This experiment tested whether spatial demonstratives paired with an object at encoding affected memory for object location, with objects placed at varying distances in front of participants on a table and then removed. The main goal was to test between the expectation and congruence models. Critically, the expectation model predicts a main effect of demonstrative on object location memory and the congruence account predicts an interaction between demonstrative and distance.

2.1 Method

2.1.1 Participants.

Thirty-six native English speaking students[13] were tested, receiving either course credit or payment for their participation. Stereoacuity was measured using the Randot Stereotest (Stereo Optical Inc. Chicago, USA). Two participants did not have a threshold of at least 40 arcseconds and therefore were excluded. Two additional participants were excluded because they had more than 10% incorrect answers in the

[13] Sample size is based on Coventry et al., 2014

memory task. This left 32 participants, 9 males and 23 females, with an age range of 18 – 31 years old (M = 20.78, SD = 3.14).

2.1.2 Materials

Six distinguishable, different colored shapes on plastic discs (e.g. yellow triangle/blue heart), 6.5 cm in diameter, were placed on six different locations. The locations were spaced equidistantly along a midline from the participants' edge of a large conference table (L = 320cm, W = 90cm), starting at 25cm from the participant up to 150cm. The three dots that were closest to the participants were located within peripersonal space, while the remaining three dots were within extra-personal space (confirmed for each participant). The table was covered with a black cloth so that no spatial cues were present.

2.1.3 Procedure and Design

Participants were asked to sit as close to the table as was comfortable, to ensure that all participants were approximately the same distance from the objects. Then, they played a 'memory game' as used previously by Coventry et al. (2014); participants were told the experiment was testing memory for object location. On each trial, the participant read out an instruction card indicating which object had to be placed on which location. The instructions all had the form: "Place DEMONSTRATIVE, OBJECT COLOR, OBJECT NAME, on the COLOR dot" (e.g. "Place this red triangle on the blue dot"). Following the instruction, participants closed their eyes while the experimenter placed the object as instructed. The participant was then given 10 seconds to view the object and to memorize the object location before the object and the dots were removed and the experimenter went behind a curtain to present an indication stick. Next, the participant verbally instructed the experimenter to match the near edge of the indication stick to the remembered near edge of the object location. Participants were then required to verbally indicate the demonstrative used on the instruction card to ensure they had attended to the instructions (see Figure 2).

There were two demonstratives (*this/ that)* and a neutral determiner (*the),* six locations and six objects. Participants were presented with six practice trials, after which 54 experimental trials were conducted (consisting of 3 trials of every term on every location: 3×3×6). The indication stick was presented at a distance of 10cm (counterbalanced to be further or nearer) from the actual location. Within the first 10 trials, there were three filler trials in which the indication stick was presented at a distance of 20cm from the object location, to prevent the initial placement of the stick becoming a cue for the object location. Every trial in which a participant could not remember the demonstrative was repeated at the end of the experiment (if a participant couldn’t remember >10% of the trials s/he was excluded). Also trials in which a participants’ estimate of the object location was >25cm from the original location were repeated. If the criterion was not met after repeating the trial was eliminated. At the end of the experiment, reaching distance was measured for each

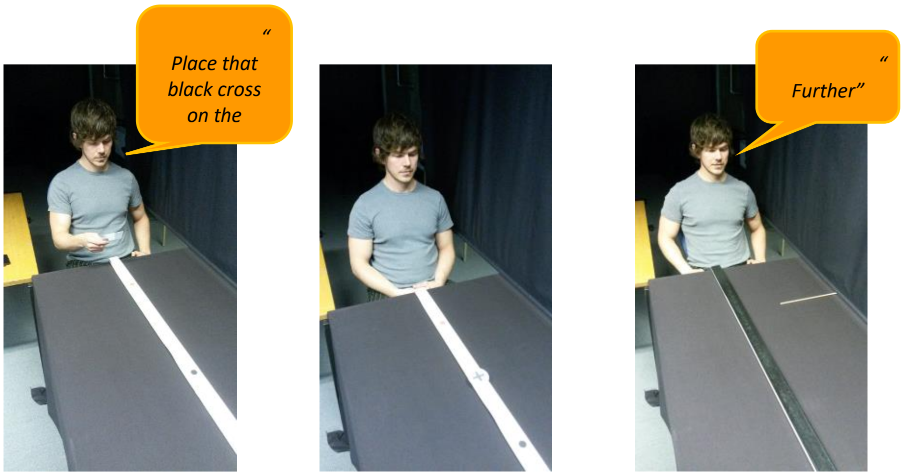
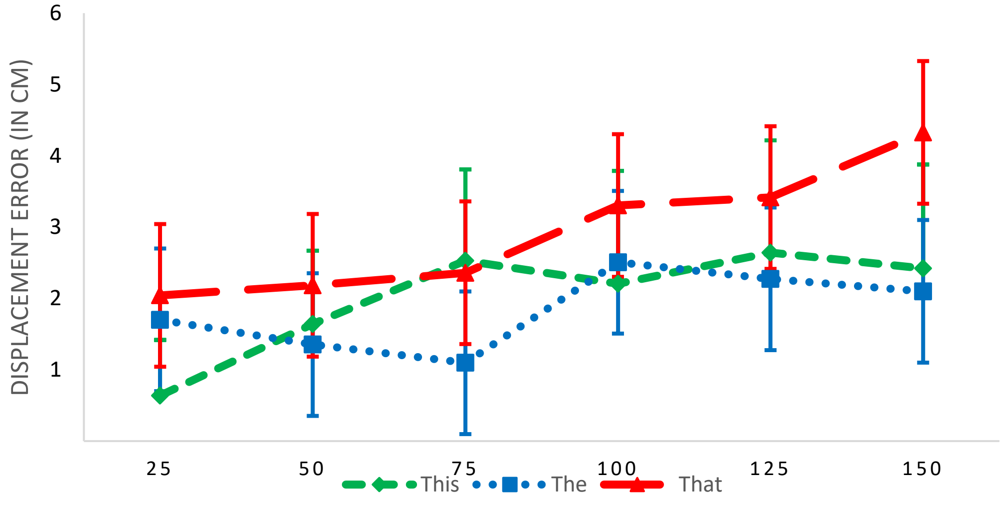


Figure 2. The participant reads out the instruction card, then memorizes the object location and finally instructs the experimenter

participant. Every participant could reach only the first three dots. The ‘memory game' cover meant that participants were not aware that we were interested in the differences between demonstratives (confirmed during debrief).

2.2 Results and Discussion

The memory displacement data – that is, the difference between the recalled distance and the actual distance between the recalled distance and the actual distance measured in centimeters – are displayed in Table 1 (see supplementary data). Note that a positive value indicates that an object was (mis)remembered as further away than it actually was. A 3 × 6 (demonstrative × location) ANOVA was performed on the memory displacements. The assumption of sphericity was violated in both the location and the demonstrative × location analysis. We therefore used the Greenhouse-Geisser correction for these analyses. There was a main effect of demonstrative $F(2,62) = 6.68$, $MSE = 10.04$, $p < .01$, $\eta p^2 = .18$, showing an effect of language on memory for object location: follow up (LSD) tests showed significant differences between locations accompanied by the *that* ($M = 2.94$, $SE = .42$) compared to both the *this* ($M = 2.01$, $SE = .41$) and the *the* ($M = 1.84$, $SE = .47$) conditions (both *p's* < .01; see Figure 3). There was a marginal effect of location $F(5,155) = 2.33$, $MSE = 25.49$, $p = .08$, $\eta p^2 = .07$, revealing that memory for object location deteriorated with distance, consistent with the previous studies. Importantly, there was no interaction between demonstrative and location $F(10,310) = 1.4$, $MSE = 9.13$, $p = .21$, $\eta p^2 = .04$. The results therefore support the expectation model rather than the congruence model; *this* leads to more accurate object location memory than *that,* irrespective of the congruence between the specific demonstrative and location. We next considered whether the same pattern of results might emerge with a different language manipulation involving possessives.



Locations (in cm) from participants

Figure 3. Results of Experiment 1, error bars are 95% confidence intervals.

3. Experiment 2: The influence of possessives on spatial memory

Some studies have shown that ownership improves memory for objects (S. Cunningham et al., 2008; Shi et al., 2011; Turk et al., 2015) and influences how people physically interact with objects (Constable et al., 2011). For example, Cunningham et al. (2008) had a participant and a confederate sort cards with pictures of shopping items into their own basket or the other person's basket. At the end of the trials participants completed a surprise memory test for the objects depicted on the cards. Participants had more accurate memories for self-owned objects than objects owned by a conspecific.

In another study, specifically targeting memory for object location, Coventry et al. (2014), found that object ownership affected memory for object location (and demonstrative choice). Using the memory game paradigm participants were given a set of coins in payment at the start of the experiment, and the coins placed at different to-be-remembered locations were either those coins or coins owned by the experimenter of the same denominations. Participants misremembered the conspecific's coins as being further away than their own coins.

One of the problems with the ownership studies described above is that they cannot easily distinguish between an effect of the abstract concept of ownership and an effect of the possessives (*my/your*) used to indicate ownership during task instruction. For example, in the study of Coventry et al. (2014), coins were given to participants as participant payment at the start of the task to confer ownership, but language during the task itself involved by necessity the use of possessives (e.g. "Place your coin on the red dot") in order to disambiguate which coin was to be placed during the task. It is therefore unclear whether the effect of ownership is driven by the language indicating ownership (possessives in all cases), the conceptual representation of ownership itself, or a combination of the two. Here we investigated whether possessives have the same influence on memory for object location as did the demonstratives in Experiment 1, whether personal possessives *alone* are able to drive memory effects, and again, whether the expectation vs. congruence models offer a better account as to how possessives affect memory for object location.

3.1 Method

3.1.1 Participants

Thirty-nine native English speaking participants were tested, as in Experiment 1. Five participants were excluded as they did not score above the threshold of 40 arcseconds (N=2), had more than 10% mistakes in the memory task (N=2) or could not reach the 50cm point (N=1). This left 34 participants; 14 male and 20 female, with an age range of 18 – 44 years old (M = 23.76, SD = 4.87).

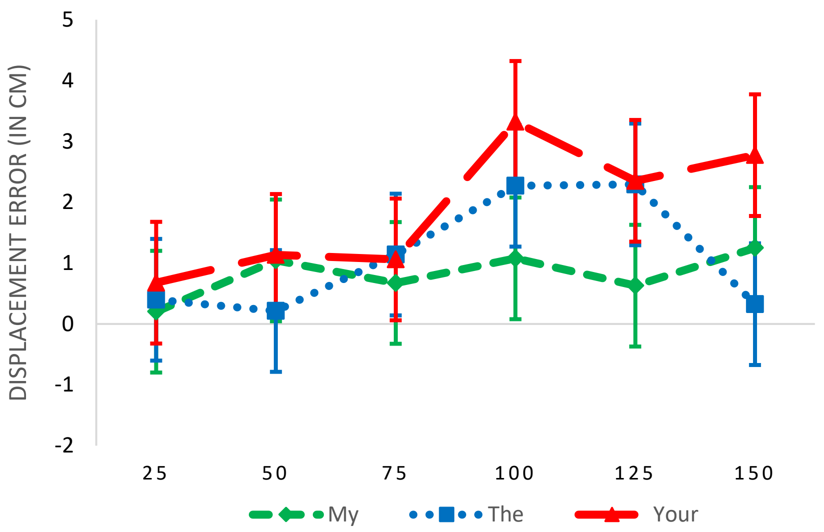
3.1.2 Procedure and Design

The procedure was similar to Experiment 1, with the exception that the demonstratives were replaced with possessives (*my*, *your;* the *the* condition was retained). In order to able to distinguish between an actual ownership effect and a language effect of possessives, participants did not own any of the objects, and all objects were used in all language conditions.

3.2 Results and Discussion

The memory displacement data are displayed in Table 1 (see supplementary data). A 3 × 6 (possessive × location) ANOVA was performed on the difference (in centimeters) between the actual position of an object and the memorized position. There was a main effect of possessive $F(2,66) = 8.25$, $MSE = 7.62$, $p = .001$, $\eta p^2 = .2$, showing that objects in the *your* condition ($M = 1.89$, $SE = .43$) were remembered as being significantly further away than objects in both the *my* condition ($M = .81$, $SE = .34$) and the *the* condition ($M = 1.11$, $SE = .34$), both *p's* < .01; see Figure 4). A significant effect of location was also found $F(5,165) = 3.47$, $MSE = 18.07$, $p = .01$, $\eta p^2 = .1$, showing that accuracy deteriorated as the objects were placed further away. These results are compatible with earlier studies on ownership. However, as all objects were used in all language conditions, there was no actual sense of ownership over any of the objects; the ownership was only marked by the use of possessives. This shows that possessives on their own affect memory for object location.

Additionally there was an interaction between possessive and location $F(10,330) = 2.25$, $MSE = 10.37$, $p = .03$, $\eta p^2 = .06$. As can be seen in Figure 4, the interaction pattern is consistent with the expectation model and is not consistent with the congruence account: there is no cross-over between peripersonal and extrapersonal space as would be expected in the congruence account. However, it is the case that the effect of distance does seem to vary as a function of language. To further unpack this, we ran three one-way ANOVAs to test location effects by term, revealing that there was only a reliable peri-personal/extra-personal effect in the "*your*" and "*the*" conditions ($p < .05$). This effect was absent in the "*my*" condition ($p > .05$; see Figure 4). This suggests that memory for possessed objects maybe particularly enhanced, overriding any effect of peripersonal versus extrapersonal space.



Locations (in cm) from participants

Figure 4. Results of Experiment 2, error-bars are 95% confidence intervals.

4. Experiment 3: The influence of attention on spatial memory

So far the results are consistent the expectation account. However, it is important to also consider the possibility that the results might be driven by the allocation of attention. Visual attention is allocated preferentially to objects nearby, compared to objects further away (Garrido-Vásquez & Schubö, 2014) and longer fixation times lead to better memory performance (Huebner & Gegenfurtner, 2010). Therefore, the predictions of memory error in the Expectation Model and what we have coined an "Attention Allocation Model" are similar, but differ in underlying mechanism. The Expectation Model predicts that memory for object location is a function of the language used to refer to the object (and the expectation of location associated with that language) combined with actual object location. The Attention Allocation model suggests memory for object location is a function of the fixation time and the object location. The results of Experiment 1 could therefore be alternatively explained by the Attention Allocation Model as driven by longer fixation times to objects paired with "this" versus "that" rather than differences in expectation values. In this experiment, we used eye tracking to measure participants' looking time during encoding. That allowed us to measure the time a participant is focused on the object in each language condition to see whether attention might account for the main effect of language reported above.

A second aim of Experiment 3 is to explore the connection between demonstratives and reference frames. As peri-personal space is the area within our grasp, this can be seen as an 'action space' in which objects are mapped onto an egocentric reference frame, compared to extra-personal space which may be mapped onto an allocentric reference frame (ter Horst et al., 2011). If the language effect that was found in the first two experiments is driven by the expectation raised by the specific use of language, then this expectation may result in different use of reference frames. We explored whether encoding object location onto an egocentric reference frame resulted in more searching behavior along the sagittal line, to encode

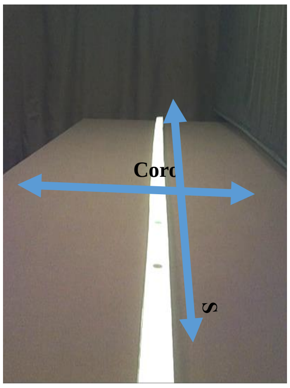


Figure 5. Coronal and Sagittal plane

distance from the participant, compared to encoding onto an allocentric reference frame which could result in more searching behavior along the coronal line (see Figure 5). Results could help distinguish between models that predict solely an influence of egocentric representations on spatial memory versus ‘two-system’ models that predict a parallel egocentric and allocentric representations in object location memory (see Burgess, 2006)

4.1 Method

4.1.1 Participants

Nineteen participants were tested as in Experiment 1. Three participants were excluded from the analysis as the eye-tracker could not be calibrated. This left 16 suitable participants for the analyses, 5 male and 11 female, with an age range of 18 – 22 (*M* = 19.19, *SD* = 1.17).

4.1.2 Procedure and Design

The procedure was based on Experiments 1 and 2, but in this experiment, participants wore SMI eye-tracker glasses (30Hz binocular eye tracking glasses). For this reason, only 4 positions were used – two locations in peripersonal space and two in extrapersonal space. (The first location was too close for the eye-tracker and the furthest location was not useable because the area of interest was too distorted). Before the experiment started, the glasses were calibrated using marks on the wall. After that, calibration was validated four times throughout the experiment by having participants look at the four different locations on the table. The eye-tracking data were coded using semantic gaze mapping[14]. As the angle from the participant to the object was different for every location, the standard error in calibration of the eye-tracker image was slightly different per location. These distortions had to be

[14] This involves the manual coding of video-based eye-tracking data, by which fixations are coded on a gaze map.

accounted for in the semantic gaze mapping. Therefore, the coding was slightly less stringent for further locations compared to closer locations. For the furthest location any fixation within an area of 6.5 cm (equivalent to the diameter of the object discs) around the object was marked as a fixation on the object. In the nearest location any fixation within an area of 3.25 cm (half an objects' diameter) was marked as a fixation on the object (see Figure 6)[15]. The gaze mapping data were used in a 3 × 4 (demonstrative × location) design, investigating the differences in total fixation time (ms) on the object.

4.2 Results and Discussion

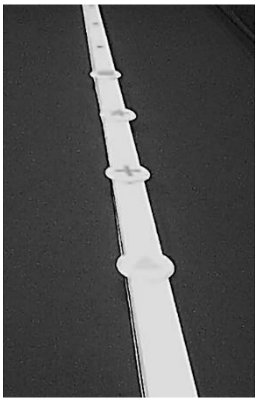
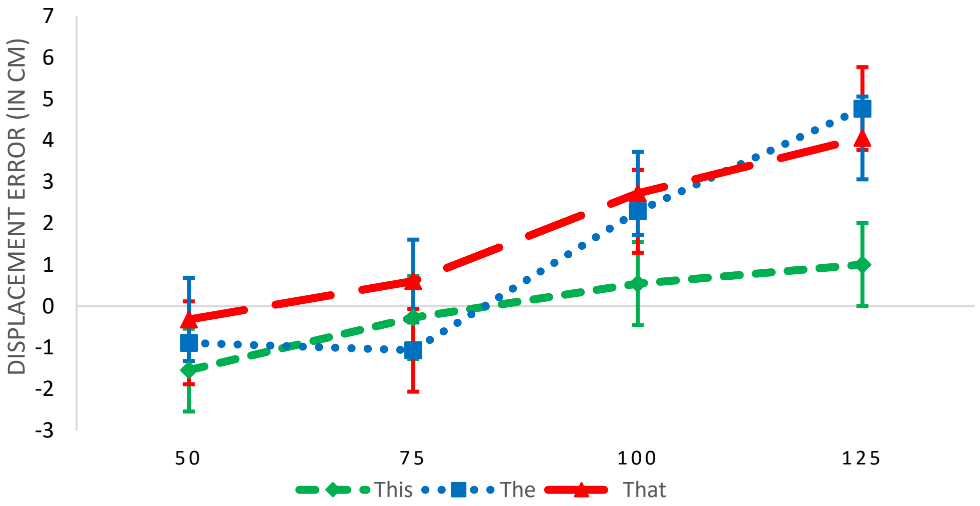


Figure 6. Object

The memory displacement data and the fixation times are displayed in Table 1 (see supplementary data). The memory data were analyzed in a 3 × 4 (demonstrative × location) ANOVA. A main effect of demonstrative was found $F(2,30) = 5.77$, $MSE = 10.02$, $p < .01$, $\eta p^2 = .28$, in which recalled distances for object location in the *that* condition ($M = 1.77$, $SE = .68$) were significantly further away than those in the *this* condition ($M = -.07$, $SE = .79$), $p < .05$. The *this* condition distances were also significantly closer than in the *the* condition ($M = 1.3$, $SE = .59$),

[15] Although the coding was adjusted for the different distances, this does not detract from the results as the adjustments were conducted across the different language conditions.

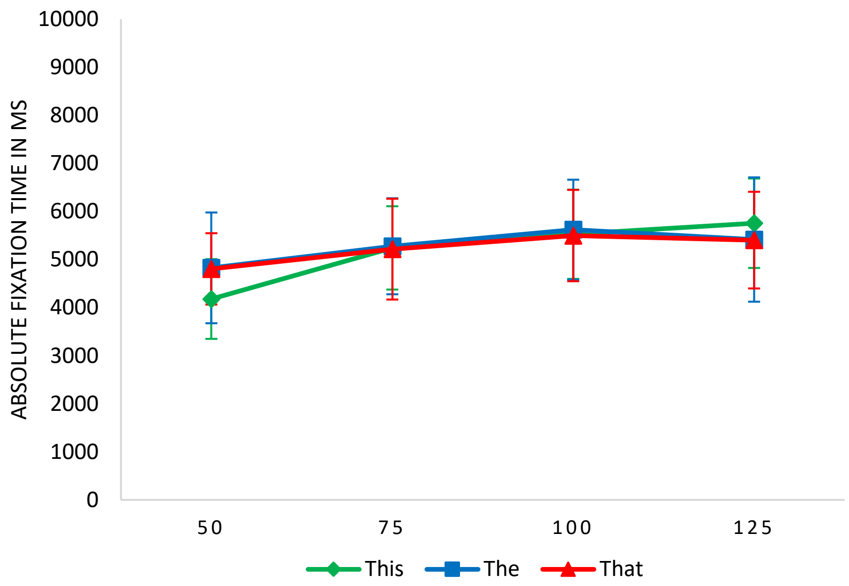
p < .05 (see Figure 7). This replicates the result of Experiment 1. There was also a main effect of location *F*(3,45) = 9.69, *MSE* = 29.77, *p* = .001, ηp^2 = .39, in which participants' accuracy deteriorated as locations were further away. There was no interaction effect between demonstrative and location *F*(6,90) = 1.61, *MSE* = 9.26, *p* = .15, ηp^2 = .1, which means that the effect of language was the same across locations.



Locations (in cm) from participants

Figure 7. Behavioral data of Experiment 3, error-bars are 95% confidence intervals

To see whether the language effects found were driven by a mechanism as hypothesized by the Expectation Model or the Attention Allocation Model, we next examined the gaze data collected during encoding. A 3 × 4 (demonstrative × location) analysis of object fixation time showed no effect of language, *F*(2,30) = .13, *MSE* = 1974647.31, *p* = .81, ηp^2 = .009 ("*this*" *M* = 5175.70, *SE* = 345.44; "*that*" *M* = 5230.97, *SE* = 257.65; "*the*" *M* = 5285.14, *SE* = 416.76) suggesting that the language effect is not driven by differences in attention (See Figure 8). There was a location effect *F*(3,45) = 4.66, *MSE* = 1997163.36, *p* < .01, ηp^2 = .24, showing that participants fixated longer on locations further away. However, this location effect could be due to the differences in coding caused by distance, as explained in Figure 6. There was no interaction effect between demonstrative × location $F(6,90) = .62$, $MSE = 1442394.41$, $p = .71$, $\eta p^2 = .04$.



Location (in cm) from participants

Figure 8. Gaze data from Experiment 3, based on summed fixation time on the objects in the respective language condition at the respective locations averaged per trial (in ms). Error-bars are 95% confidence intervals.

In a second analysis, we explored the connection between demonstratives and reference frames, and specifically to test whether people use different coordinate systems to remember object locations, based on spatial language. Fixations were coded as sagittal searching behavior, if a sequence of two or more fixations fell within a range on either side (left/ right) of the white location stick, the range being 3.25cm from the sides for the closest location and 6.5cm from the furthest location. These distances were based on the size of an object on the respective location as represented on the screen (the actual objects had a diameter of 6.5cm). Fixations were coded as searching behavior along the coronal line if a sequence of two or more fixations fell within a range above or below the object location. The range was half an objects' size for the closest location and one objects' size for the furthest location along the coronal plane. Fixations coded as fixations on the actual object were excluded from this analysis, so no fixation was used twice. After this coding, a ratio of fixations was calculated (coronal / (sagittal + coronal) (see Figure 7). A 3 × 4 (demonstrative × location) ANOVA was performed. There was no main effect for demonstrative $F(2,30) = .15$, $MSE = .05$, $p = .86$, $\eta p^2 = .01$; "*this*" $M = .42$, $SE = .07$; "*that*" $M = .4$, $SE = .09$; "*the*" $M = .42$, $SE = .08$), nor location $F(3,45) = .2.25$, $MSE = .16$, $p = .13$, $\eta p^2 = .13$, nor an interaction with distance, $F(6,90) = .78$, $MSE = .05$, $p = .59$, $\eta p^2 = .05$, suggesting that the language effect was not caused by differences in search-behavior. Based on these data, we cannot distinguish between different models for the use of reference frames in memory for object location.

5. General Discussion

The results of all three experiments show that language affects memory for object location, with main effects of language in all three studies. The use of both demonstratives (Experiment 1 and 3) and possessives (Experiment 2) affects memory for object location. These results are consistent with previous studies showing an influence of language on memory for spatial relations (Feist & Gentner, 2007; Loewenstein & Gentner, 2005), and also with the effects of object knowledge on object location memory reported by Coventry et al. (2014). We have found robust effects of language on object location memory together with an effect of distance on memory for object location. We first consider explanations for these results prior to implications for theories of language and memory more generally.

Three possible accounts of the influence of language on object location memory were set out prior to designing the present series of experiments: the congruence account, the expectation account, and an attention allocation model. The difference between the expectation and congruence models is the prediction of an interaction in the latter, and a main effect of language without an interaction in the former. The expectation model, proposed by Coventry et al. (2014), to explain object knowledge effects on memory, maintains that language elicits an expectation about an objects' location which is concatenated with actual object location, leading to the prediction that the language effect should be the same for objects in near space and far space. In contrast, the congruence account predicts that memory should be better for trials in which language is congruent with the object location, predicting an interaction between language and location; congruent trials (where *this/ that* are respectively combined with *near/ far* space) should be remembered better than incongruent trials (in which *this/ that* are respectively combined with *far/near* space). In Experiments 1 and 3 there was no interaction, supporting the expectation account. In Experiment 2 (possessives) there was an interaction, but this effect was driven by the absence of a location effect for the "*my"* condition and not by congruence/incongruence contrasts. Thus, as a whole, results of the current experiments all support the expectation model.

Experiment 3 tested the third possibility that different types of language might result in different amounts of attention being paid to objects/locations, with associated differences in memory performance. Put simply, the longer one spends looking at an object, the better one's memory for object location. The eye tracking data from this experiment revealed no differences in viewing time as a function of demonstrative. Also participants did not present different searching behavior based on different demonstratives, allowing us to rule out the attention allocation model.

Given that the results support the expectation model, there are three keys issues that merit discussion. First, we can consider the relationship between the expectation model and memory models more broadly. Memory for object location is often taken to involve memory for the location in which an object is positioned, memory for the object itself, and a binding between object location and object (see for example Postma & Haan,1996). Previously, Coventry et al. (2014), finding effects of object knowledge on memory for object location, argued against memory models that prioritize object location over object knowledge (e.g. the model of Jiang et al., 2000, who argued that location may act as an anchor to which object properties are attached). However, the effects of language on memory for object location and the previous effects of object knowledge reported are consistent with variants of *object file* theory (Kahneman & Treisman, 1984), in which object location is one of the features integrated in the file. Location features appear not to be bound to an absolute location but are defined relative to an abstract representation, which leads to memory errors (see Hollingworth & Rasmussen, 2010; Pertzov, Dong, Peich, & Husain, 2012). This focus on relative location can explain how spatial language can cue memory for object location, via the expectation of the objects' location relative to the speaker or another object. Wang and Spelke (2002) suggested that the human representational system depends in some way on language, by which humans can go beyond the limits of orientation systems as are found in animals. This influence of language skills may facilitate more flexible problem solving (Hermer-Vazquez et al., 2001). One can speculate that the advantage is that such a relative, dynamic system enables us to mentally process arrays in different contexts (e.g. a desk or the universe), using the same language and concepts.

Second, one needs to unpack in more detail how the expectation model works, and in particular, how the expectation values form and how they combine with the actual distance information available. Coventry et al. (2014) do not offer detail regarding this, but they assume that the expectation model works via the prediction of object location as a product of the history and context of past bindings between language, objects and location. For example, objects owned by people are more likely to be near people than equivalent objects owned by someone else. This likelihood is then used to predict future encounters with objects: if an object is owned, one would anticipate the object is nearer than if one does not own the object. This anticipation works similarly for the visibility and familiarity parameters Coventry et al. (2014) identified. Respectively, visible objects are usually closer than objects one cannot see, and familiar objects are more likely to be near us than unfamiliar objects. This anticipation-mechanism could be accommodated by correlational learning (see Pulvermüller, 2012) - the process in which neurons that fire together strengthen their connections and become more tightly associated (also known as potentiation), and such a mechanism has been implicated not only in mapping language to perception (Coventry et al., 2013), but also how one learns how words co-occur to form meaningful language structures during language learning (Saffran, Aslin, & Newport, 1996; Saffran, 2002).

The Expectation Model can also be extended outside spatial language, both in cases where language is explicit during a task (as in our studies), but also in cases where language may not be explicit, but may nevertheless affect non-linguistic performance. For example evidence from color perception (Jerome S. Bruner et al., 1951; Delk & Fillenbaum, 1965) has shown similar effects of the influence of object knowledge on memory. Object knowledge influences categorization of objects, so that participants judge objects within a category to have a more similar hue than objects between categories. For example, in an array of letters and numbers, participants judged symbols within a respective category to be more similar in color than between categories, even if the two, between categories, target symbols were identically colored (Goldstone, 1995). In another series of studies, it was shown that color memory and color perception judgements are influenced by the characteristic color of an object, these object knowledge effects were stronger in objects with a high color diagnosticity (e.g. yellow for a banana) than in objects with low color diagnosticity (yellow for a lamp) (Belli, 1988; Jerome S. Bruner et al., 1951; Hansen et al., 2006; Tanaka & Presnell, 1999). These ‘top down’ effects on color perception are consistent with the idea that knowledge of expected hue combines with actual hue information leading to categorization errors. Such an account merits further testing in this domain.

More broadly the Expectation Model is consistent with models of predictive coding (A. Clark, 2013; Friston, 2005). Clark (2013) suggests that people use prediction to minimize energy costs (the free-energy principle, Friston, 2009). In his model, the brain receives input from the perceptual system and uses existing knowledge to *predict* or anticipate the new state of the world based on that perceptual input (A. Clark, 2013; Friston, 2003). The brain prepares a response based on this prediction and only needs to process the *error signal*, the difference between the prediction and the updated visual input, once the new state of the world emerges. This means that instead of processing or ‘creating’ a full response, the brain only needs to adjust the predicted response to be appropriate to the actual input. In the Expectation Model, the prediction is based upon learned associations between language, objects and locations (for example via statistical/ correlational learning), and it is these associations that can then reduce the work needed to process continually changing object location bindings on a moment to moment basis.

A third issue that needs discussion is whether the effects of language operate at the level of encoding or retrieval. One possibility is that *this*, for example, actually activates peripersonal space more when looking at an object than *that*, and therefore that the memory differences are a direct result of differences in peripersonal space activation during encoding. Such a view is consistent with recent models of perception (e.g. Bar, 2009) that incorporate top-down predictions from memory as a mechanism during the act of perceiving. Alternatively, it is possible that the influence of language only occurs at retrieval, with remembered distances migrating in the direction of the remembered demonstrative/possessive. In order to test between these alternatives, it is possible to run neuroimaging studies to measure the degree of peripersonal space activation while viewing objects under different object knowledge and/or language conditions (see Coventry et al., 2014 for discussion). A second way one can get at this issue relates to memory decay: if the influence of language operates at retrieval, then the longer the time interval the greater the effects of language there should be. We are currently exploring these possibilities.

In summary, we found a main effect of language (demonstratives and possessives) on memory for object location across experiments. We teased apart the predictions of three different models explaining this mechanism: the Expectation model, the Congruence model, and the Attention allocation model. Overall, results favored the Expectation model, suggesting that the expected location of an object – cued by language (e.g., *this* for referents close by; *that* for referents further away) – and the actual location are combined, leading to (mis)memory for object location.

Acknowledgments

All authors have no competing interest nor conflict of interest to declare. This project has received funding from the European Union's Seventh Framework Programme for research, technological development, and demonstration under grant agreement no 316748.

Appendix B

The Influence of Language on Memory for Object Location

Harmen B. Gudde (h.gudde@uea.ac.uk)
School of Psychology, University of East Anglia
Norwich, NR4 7TJ, UK

Kenny R. Coventry (k.coventry@uea.ac.uk)
School of Psychology, University of East Anglia
Norwich, NR4 7TJ, UK

Paul E. Engelhardt (p.engelhardt@uea.ac.uk)
School of Psychology, University of East Anglia
Norwich, NR4 7TJ, UK

Abstract

In this study, the influence of two types of language on memory for object location was investigated: demonstratives (*this, that)* and possessives (*my, your*). Participants read instructions (containing *this/that/my/your/the*) to place objects at different locations. They then had to recall those object locations. Experiments 1 and 2 tested the contrasting predictions of two possible accounts of language on memory: the *expectation* model (Coventry et al., 2014) and the *congruence* account (Bonfiglioli et al., 2009). In Experiment 3, the role of attention as a possible mechanism was investigated. The results across all three experiments show striking effects of language on object location memory; objects in the "*that"* and "*your"* condition were misremembered to be further away than objects in the "*this"* and "*my"* condition. The data favored the expectation model: expected location cued by language and actual location are concatenated leading to (mis)memory for object location.

Keywords: memory for object location; spatial demonstratives; possessives

Introduction

Language is often used to direct the attention of a conspecific to the spatial world, and the pairing of language with visual images affects what is recalled about those images. For example, Loewenstein and Gentner (2005) found that children performed better in a mapping task when spatial relations were paired with spatial language at encoding. Relational language fosters the development of representational structures that facilitate mental processing (see also Hermer-Vazquez, Spelke, & Katsnelson, 1999). But language presented with a spatial scene can also lead to memory errors. For instance, Feist and Gentner (2007) showed that recognition memory for spatial scenes was shifted in the direction of the spatial relational language (spatial prepositions) presented with scenes at encoding.

Although these studies show an effect of language on memory, it is not yet known how language affects memory. Here we focus on spatial demonstratives and possessives, and the possible effect these terms might have on memory for object location. In doing so, the continuous nature of object location memory errors allows us to contrast different possible mechanisms regarding how language affects memory for object location.

Demonstratives (this, that) have been shown to be associated with discrete zones of peri-personal and extra-personal (near and far) space (Diessel, 2006). However, this distinction is flexible. Near space can be contracted or extended by weight or tool use (Coventry et al., 2008; Longo & Lourenco, 2006), and the use of "this" is similarly extended when participants use a stick to point at objects (Coventry et al., 2008). Object knowledge also affects both perception and memory for object location. For example, Balcetis and Dunning (2010) showed that participants perceived desirable objects as being closer to themselves than less desirable objects. Previous research has also shown that several object properties, including ownership, visibility, and familiarity, influence the use of spatial demonstratives in English and memory for object location (Coventry et al., 2014).

In order to account for the influence of object knowledge on memory for object location (and by extrapolation, language), Coventry et al. (2014) proposed an *expectation model.* In this model, memory for object location is a function of the actual object location concatenated with where an object is expected to be (with a constant estimation error). For example, an object owned by the participant is expected to be nearer than an object owned by someone else. This results in the participant misremembering an object owned by someone else as further away than an object they owned at the same location. This expectation model also makes a prediction (tested below) regarding the direction of memory errors when demonstratives occur with objects at encoding. Specifically, it might be expected that "*this"* activates near space and "*that"* activates far space, and when conjoined with actual location, should lead to objects paired with "*that"* being (mis)remembered as being further away than they actually were (relative to objects paired with "*this"*).

In contrast to the expectation model proposed by Coventry et al. (2014), there is a considerable body of work within an embodied cognition framework providing evidence for an alternative "*congruence account*" between language and space that makes different predictions from the expectation model. For example, it has been shown that participants respond more quickly to positive stimuli in a congruent high location than an incongruent low location, and vice versa for negative stimuli (e.g. Barsalou, 2008; Meier & Robinson, 2004; cf. Lynott & Coventry, 2014). Moreover, this congruence account has been extended to movement planning where movements were prepared based on language (Bonfiglioli et al., 2009). Participants were required to grip an object after listening to an instruction that indicated whether the object was near or far; RTs were significantly longer when language was incongruent with space compared to when language and space were congruent. In extrapolating this congruence account to memory, congruence in language and space would be predicted to enhance the accuracy of memory for location.

The goal of the present study was to test whether language affects memory for object location, and to unpack the mechanism involved. Specifically, we aimed to tease apart these two accounts by examining the effects of demonstratives and possessives on memory for object location. The first experiment tested whether spatial demonstratives affect memory for object location with contrasting predictions from the two different accounts of how language affects memory: congruence v. expectation. Experiment 2 tests whether similar effects occur for possessives (*my*, *your*) – terms which have also been associated with the peripersonal/ extrapersonal space distinction. Experiment 3 tests whether the effects found in Experiment 1 might be a result of a third variable – i.e. language affecting the amount of attention paid to an object at a given location.

Experiment 1

This experiment tested whether spatial demonstratives paired with an object at encoding affected memory for object location. The main goal was to test between the expectation and congruence models. Critically the expectation model predicts a <u>main effect</u> of demonstrative on object location memory and the congruence account predicts <u>an interaction</u> between demonstrative and distance, such that memory should be more accurate when language and object location are congruent.

Participants

In this study, 36 participants were tested. All were native English speakers receiving either course credit or payment for their participation. Stereoacuity was measured using the Randot Stereotest (Stereo Optical Inc. Chicago, USA) and participants who did not have a threshold of at least 40 arcseconds were excluded (N=2). Two additional participants were excluded because they produced more than 10% incorrect answers in the demonstrative memory task (see below). This left 32 participants for the analyses, 9 male and 23 female, with an age range of 18 – 31 years (*M* = 20.8, *SD* = 3.1).

Materials

Six distinguishable, different colored shapes on plastic discs (e.g. yellow triangle, blue heart), 6.5 cm in diameter, were placed on six different locations. The locations were spaced equidistantly along a midline from the participants' edge of a large conference table (L = 320cm, W = 90cm), starting at 25cm from the participant up to 150cm (Coventry, Valdés, Castillo, & Guijarro-Fuentes, 2008). The three dots that were closest to the participants were located within their peri-personal space, while the remaining three dots were within their extra-personal space (this was confirmed for each participant). The table was covered with a black cloth so that no spatial cues where present on the table.

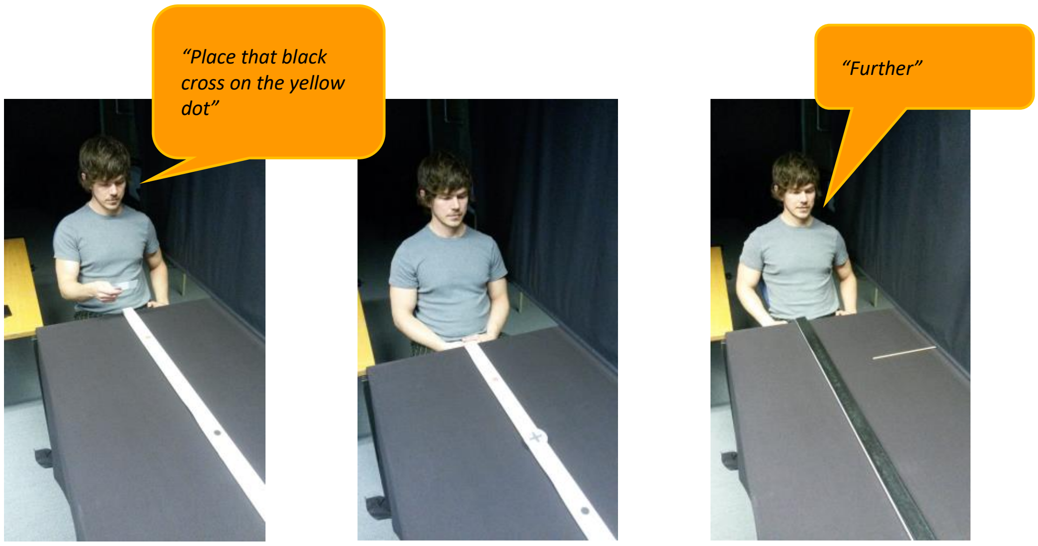


Figure 1. The participant reads out the instruction card, then memorizes the object location and finally instructs the experimenter verbally to match the indication stick to the object location.

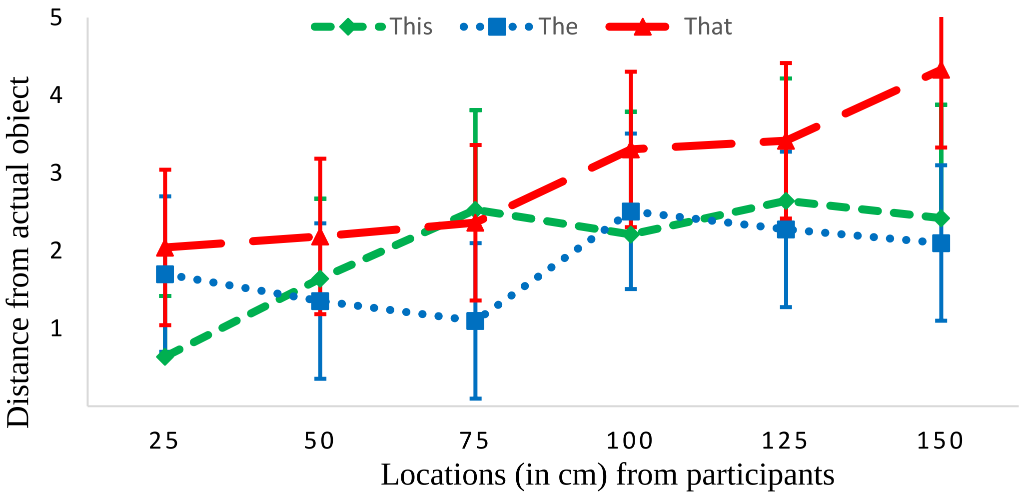
Procedure and Design

Participants were asked to sit as close to the table as was comfortable, to ensure that all participants were approximately the same distance from the objects. Participants were told the experiment was testing memory for object location. On each trial, the participant read out an instruction card indicating the placement of an object on a location. The instructions all had the form: "Place DEMONSTRATIVE, OBJECT, on the COLOR dot" (e.g. Place this red triangle on the blue dot). Following the instruction participants closed their eyes while the experimenter placed the object as instructed. The participant was then given 10 seconds to memorize the object location, before the object and the dots were removed and the experimenter went behind a curtain to present an indication stick (to prevent the experimenter from cueing the participant). The participant verbally instructed the experimenter to match the near edge of the indication stick to the remembered near edge of the object location (thus ensuring that participants didn't move and kept the same distance from the table throughout the experiment). At the end of each trial, participants were required to verbally indicate the demonstrative used on the instruction card to ensure they had attended to the instructions (see Figure 1).

There were two demonstratives (*this, that)* and a neutral determiner (*the)*, six locations and six objects. Participants were presented with six practice trials, after which 54 experimental trials were conducted (consisting of 3 trials of every term on every location: 3×3×6). The indication stick was presented at a distance of 10cm (counterbalanced to be further or nearer) from the actual object location. To prevent the initial placement of the indication stick becoming a cue for object location there were three filler trials within the first 10 trials, in which the indication stick was presented at 20cm from the object location. Remembered distance was measured in millimeters. When a participant couldn't remember the demonstrative the trial was repeated at the end of the experiment (unless a participant couldn't remember >10%, in that case s/he was excluded). At the end of the experiment reaching distance was measured to check that every participant could reach only the first three dots but not the furthest three dots. The "memory game" cover meant that participants were not aware that we were interested in the differences between demonstratives (confirmed during debrief).

Results and Discussion

A 3 × 6 (demonstrative × location) ANOVA was performed on the difference (in millimeters) between the remembered position of the object and the actual position. There was a main effect of demonstrative $F(2,62) = 6.68$, $p < .01$, $\eta p^2 = .18$, showing a direct effect of language on memory for object location: follow up t-tests showed significant differences between locations accompanied by the demonstrative "*that*" ($M = 2.94$, $SD = .42$) compared to both the "*this*" ($M = 2.01$, $SD = .41$) and "*the*" ($M = 1.84$, $SD = .47$) conditions (both *p's* < .05; see Figure 2[16]). There was a marginal effect of location $F(5,155) = 2.33$, $p = .08$, $\eta p^2 = .07$, revealing that memory for object location deteriorated with distance. More importantly, there was no significant interaction between demonstrative and location $F(10,310) = 1.4$, $p = .21$, $\eta p^2 = .04$. The results therefore support the expectation model rather than the congruence model; "*that*" leads to objects being misremembered as further away compared to "*this*", irrespective of the congruence between the specific demonstrative and location.



Experiment 2

Some studies have shown that ownership influences memory for objects (Cunningham et al., 2008) and how people interact with objects (Constable, Kritikos,

[16] On the Y-axis, the absolute difference is presented (cm). A positive value means that objects were remembered as being further away than they were.

& Bayliss, 2011). Coventry et al. (2014), using the memory game, found that object ownership affected memory for object location and demonstrative choice. In Experiment 2, we investigated whether possessives (*my*, *your*) have the same influence on memory for object location as demonstratives, with the prediction that "*your"* objects would be associated with misremembered distances further away compared to "*my"* objects.

Participants

In this study 39 participants were tested. All participants were native English speakers receiving either course credit or payment for their participation. Stereoacuity was measured as in Experiment 1. Two participants did not score a threshold of 40 arcseconds, two participants had more than 10% errors in the possessive memory task and one participant could not reach the second (50cm) dot. These participants were excluded from analysis, leaving 34 participants; 14 male and 20 female, with an age range of 18 – 44 years (M = 23.8, SD = 4.9).

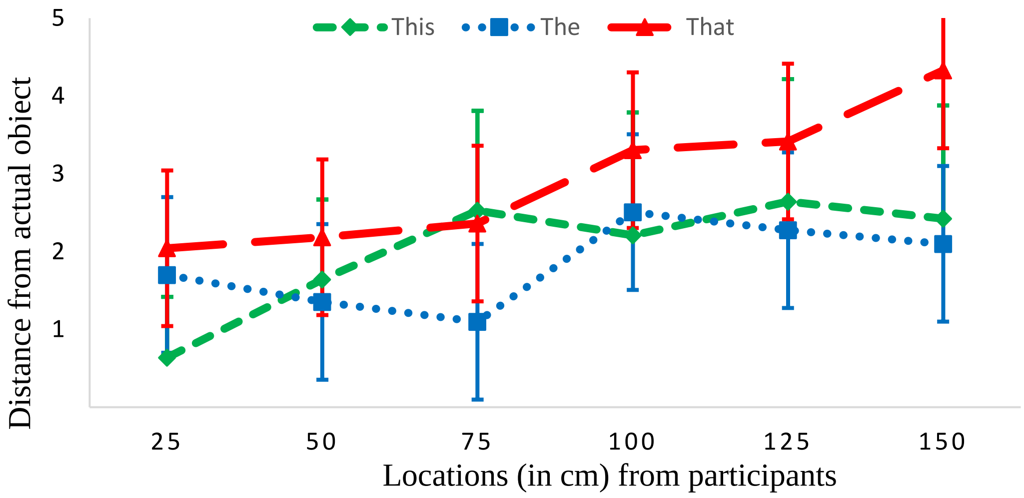
Procedure and Design

The procedure was similar to Experiment 1, with the exception that the demonstratives were replaced with possessives (*my*, *your*).

Results and Discussion

A 3 × 6 (possessive: *my, your, the* × location) ANOVA was performed on the difference (in millimeters) between the actual position of an object and the memorized position. There was a main effect of possessive $F(2,66) = 8.25$, $p = .001$, $\eta p^2 = .2$, showing that objects in the "*your"* condition (M = 1.89, SD = .43) were remembered significantly further away than objects in both the "*my"* condition (M = .81, SD = .34) and the *"the"* condition (M = 1.11, SD = .34), both $p < .05$; See Figure 3). A significant effect of location was found $F(5,165) = 3.47$, $p = .01$, $\eta p^2 = .1$, showing that accuracy deteriorated as the objects were placed further away. Additionally there was an interaction between possessive and location $F(10,330) = 2.25$, $p = .03$, $\eta p^2 = .06$. This effect indicates that the influence of language is different at different locations. To unpack this interaction three one-way ANOVAs were performed, to test location effects per term. These showed that there was only a reliable peri-personal/extra-personal effect in the "*your"* and "*the"* conditions (p <

.05). This effect was absent in the "*my"* condition (*p* > .05; see Figure 3). This suggests that memory for possessed objects maybe particularly enhanced, overriding any effect of peri-personal versus extra-personal space. Note that this interaction effect is not as predicted by the congruency account (congruence between language and space should lead to more accuracy; this is not what was found).



Experiment 3

So far the results support the expectation account. However, there is a third, alternative account that we have thus far not considered. It could be the case that "*this"* causes participants to look at an object and object location for longer than "*that"*, leading to better accuracy of recall. In this experiment, we used eye tracking to investigate this alternative hypothesis.

Participants

In this experiment, 19 participants were tested. All participants were native English speakers receiving either course credit or payment for their participation. Stereoacuity was measured as in Experiment 1; all participants had appropriate depth perception and their reach stretched between the 75cm and 100cm location. For three participants the eye-tracker could not be calibrated. These participants were excluded from analysis. This left 16 suitable participants for the analyses, 5 male and 11 female, with an age range of 18 – 22 years (*M* = 19.2, *SD* = 1.2).

Procedure and Design

The procedure was based on Experiments 1 and 2, but in this experiment, participants wore SMI eye-tracker glasses (30Hz binocular eye tracking glasses). For this reason, only 4 positions were used (the first location was too close for the eye-tracker and the furthest location was not useable because the area of interest was too distorted). Before the experiment started the glasses were calibrated using marks on the wall. After that, calibration was validated four times throughout the experiment by having participants look at the four different locations on the table. The eye-tracking data were coded using semantic gaze mapping[17]. As the visual angle from the participant to the object was different for every location, the standard error in calibration was slightly different per location. These differences in error had to be accounted for in the semantic gaze mapping. Therefore the coding was slightly less stringent for further locations compared to closer locations. For the furthest location any fixation within 6.5 cm of the object was marked as a fixation on the object. In the nearest location any fixation within half an object's size of the object was marked as a fixation on the object (see Figure 5). The gaze mapping data were used in a 3 × 4 (demonstrative × location) design, investigating the differences in total fixation time (ms) on the object.

Results and Discussion

The memory data was first analyzed in a 3 × 4 (demonstrative × location) ANOVA. A main effect of demonstrative was found $F(2,30) = 5.77$, $p < .01$, $\eta_p^2 = .28$, in

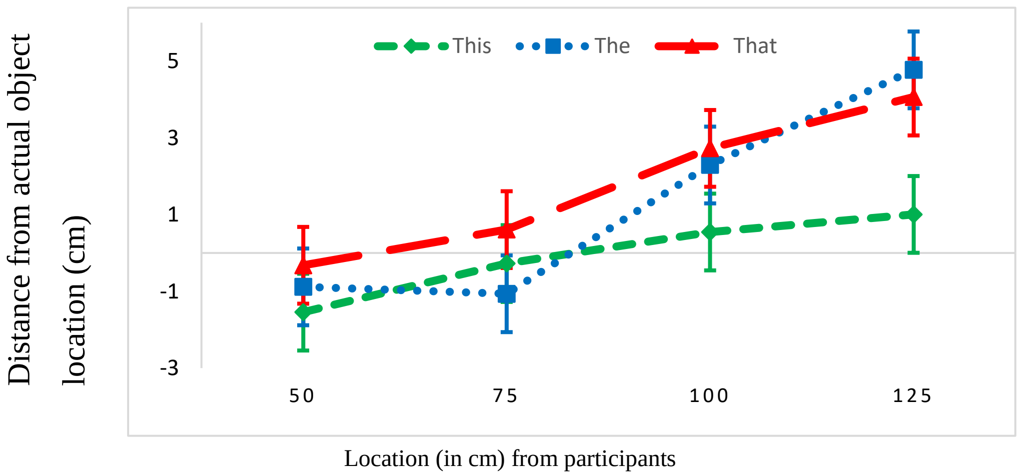


Figure 4. Results of Experiment 3. Error-bars are 95% confidence intervals.

[17] This involves the manual coding of video-based eye-tracking data, by which fixations are coded on a gaze map.

which recalled distances for object location in the "*this"* condition (*M* = -.07, *SD* = .79) were closer than in the "*that"* condition (*M* = 1.77, *SD* = .68) and the "*the"* condition (*M* = 1.2, *SD* = .59), both $p < .05$ (See Figure 4). The significant difference between the *"this"* and *"that"* condition is consistent with the results of Experiment 1. There was also a main effect of location $F(3,45) = 9.69$, $p = .001$, $\eta p^2 = .39$, in which participants' accuracy deteriorated as locations were further away. There was no interaction effect between demonstrative and location $F(6,90) = 1.61$, $p = .15$, $\eta p^2 = .1$, which means that the effect of language was the same across locations.

Regarding the gaze data; there was no significant difference in the amount of time objects were fixated as a result of language condition $F(2,30) = .13$, $p = .81$, $\eta p^2 = .009$, suggesting that language doesn't change the amount of time participants attended to a specific object/location. There was a location effect $F(3,45) = 4.66$, $p < .01$, $\eta p^2 = .24$, showing that participants fixated longer on locations further away, although the lack of measurement accuracy in the far locations may have influenced this effect. There was no interaction effect between demonstrative × location $F(6,90) = .62$, $p = .71$, $\eta p^2 = .04$, meaning that the influence of language was similar across different location conditions.

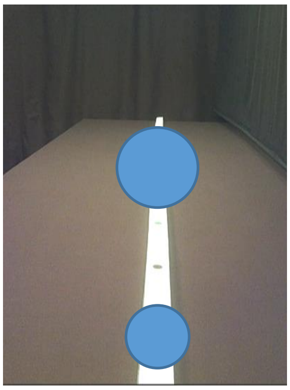


Figure 5. Object area in semantic gaze mapping

General Discussion

The results of all three experiments show that language affects memory for object location, with main effects of language in all cases. The use of both demonstratives (Experiment 1 and 3) and possessives (Experiment 2) affects memory for object location. These results are consistent with previous studies showing an influence of language on memory for spatial relations (Feist & Gentner, 2007; Loewenstein & Gentner, 2005), but our results show the first evidence of the influence of language on memory for object location. The results are also consistent with the manipulations of object knowledge on object location memory reported by Coventry et al. (2014).

In addition to the influence of language on object location memory, the experiments also revealed effects of location (the effect was marginal in Experiment 1), suggesting that participants' memory for object location deteriorates as the object is further away. These results again replicate effects of distance found in Coventry et al. (2014), and provide further evidence of the mapping between perceptual space and language and memory.

In these experiments, we have also been able to test between rival accounts of the influence of language on object location memory - the congruence account and the expectation account. We also considered a further possibility that language might affect the amount of time participants fixate an object. The difference between the expectation and congruence models is the prediction of an interaction effect in the latter, and the absence of an interaction effect in the former. The expectation account maintains that language elicits an expectation about an objects' location which is concatenated with actual object location, leading to the prediction that the language effect should be the same for objects in near space and far space. In contrast, the extended congruence account predicts that memory should be better for trials in which language is congruent with the situation, predicting an interaction effect between language and location; congruent trials (where *this/ that* are respectively combined with *near/ far* space) should be remembered better than incongruent trials (in which *this/ that* are respectively combined with *far/near* space). In Experiments 1 and 3 there was no interaction, supporting the expectation account. In Experiment 2 there was an interaction. However, this effect was driven by the absence of a location effect for the "*my"* condition and not by congruence/incongruence contrasts. Thus, as a whole, results of the current experiments all support the expectation account.

Experiment 3 tested the third possibility that different types of language might result in different amounts of attention being paid to objects/locations, with associated differences in memory performance. Put simply, the longer one spends looking at an object, the better one's memory for object location. The eye tracking data from this experiment revealed no differences in fixations on objects as a function of demonstrative, allowing us to discount this third possibility.

Overall the results support the expectation model. However, it remains to be established if this model operates at the level of encoding or at retrieval. One possibility is that "*this*", for example, actually activates peripersonal space more when looking at an object than "*that*", and therefore that the memory differences are a direct result of differences in peripersonal space activation during encoding. Such a view is consistent with recent models of perception (e.g. Bar, 2009) that incorporate top down predictions from memory as a mechanism during the act of perceiving. Alternatively, it is possible that the influence of language only occurs at retrieval, with remembered distances migrating in the direction of the remembered demonstrative/possessive. In order to test between these alternatives, it is possible to run neuroimaging studies to measure the degree of peripersonal space activation while viewing objects under different object knowledge and/or language conditions (see Coventry et al., 2014 for discussion). We are currently exploring this possibility.

Acknowledgments

This project has received funding from the European Union's Seventh Framework Programme for research, technological development and demonstration under grant agreement no 316748.

Appendix C

Abstract published for the 39th annual conference of the Cognitive Science Society (Gudde & Coventry, in press):

Demonstratives are among the most frequent words in all languages, but demonstrative systems vary considerably between languages. In two experiments, we tested demonstrative use and the influence of demonstratives on spatial memory in Japanese and English – languages with purportedly very different demonstrative systems. Participants engaged in a ‘memory game’, tapping their use of demonstratives to describe objects located on the sagittal plane (Experiment 1) and the influence of demonstratives on memory for object location (Experiment 2). In addition to distance from speaker, the experiments also manipulated the position of a conspecific (next to or opposite participants). Distance and position of conspecific both affected demonstrative choice and memory in Japanese, with similar effects in English even though English does not explicitly encode the position of a conspecific. We discuss possible universals underlying demonstrative systems and the influence of language on memory.

Appendix D

Information sheets for the memory game.

School of Psychology

Memory for Object Location

<u>Participant Information Sheet</u>

Thank you for your interest in this study. Before you decide whether to take part, please read the following information carefully (this sheet is for you to keep). You may ask me any questions if you would like more information.

What is this research looking at?
The aim of this study is to look at memory for objects and their location when they are no longer in view.

Do I have to take part?
It is up to you to decide to join the study. We will describe the study and go through this information sheet. If you agree to take part, we will then ask you to sign a consent form. You are free to withdraw at any time, without giving a reason. This would not affect you in any way.

What will happen if I agree to take part? The task involves watching objects being placed at various locations on a table. You will be asked to look at and remember the object's location for a few seconds. At the end of this part of the study you will be asked a couple of questions about the objects you have seen.

The study should take approximately one hour.

Are there any problems with taking part? There should be no problems that occur during the study, however you will be asked to sit at a table for the duration of the study.

Will it help me if I take part? There will no direct benefit to you in taking part, you will however be contributing to the research here at the University and to our knowledge of how people remember objects and object locations.

How will you store the information that I give you? All information which you provide during the study will be stored in accordance with the 1998 Data Protection Act and kept strictly confidential. The chief investigator will be the custodian of the anonymous research data. All data will be kept for 5 years, and will then be securely disposed of. All the data collected from the study will be stored anonymously on a password protected computer. Your consent form will be kept separately in a locked filing cabinet. Only the research team will have access to your data.

How will the data be used? Your anonymous data will be collated together with other data collected during the study. This collated date may be presented in scientific journal articles or presented to conferences. Neither you nor your individual data will be identifiable able amongst this group data.

What happens if I agree to take part, but change my mind later? If you decide to withdraw during the study please let the investigator know. Any personal information or data that you have provided will be destroyed immediately. After you have completed the study you can still withdraw your personal information and data be contacting the investigator (Harmen Gudde: h.gudde@uea.ac.uk) within one month of the study. After this time the data will have been collated and may have been submitted for publication.

How do I know that this research is safe for me to take part in? All research within the University is looked at by an independent group of people, called a Research Ethics Committee, to protect your safety, rights, wellbeing and dignity. This research was approved by the Psychology Research Ethics Committee at the University of East Anglia.

You are under no obligation to agree to take part in this research.
If you do agree you can **withdraw at any time without giving a reason.**

Contacts
You are welcome to discuss your participation with the experimenter, Harmen Gudde: H.Gudde@uea.ac.uk; Phone 01603 591638 or the Principal Investigator, Head of School Professor Kenny Coventry: K.Coventry@uea.ac.uk; Phone 01603 597145.
If you have any worries or concerns about this research you want to share with someone not involved in the study you may contact: Chair of the Ethics Committee Dr. Piers Fleming, P.Fleming@uea.ac.uk; Phone 01603 59 3386 .

School of Psychology
Consent Form

Memory for Object Location

Name of Researcher: Harmen Gudde

Please initial all boxes

1. I have read and understand the information sheet Memory for Object Location and have had the opportunity to questions and have had these answered satisfactorily. ☐

2. My participation is voluntary and I know that I am free to withdraw at any time, without giving any reason and without it affecting me at all ☐

3. I know that no personal information (such as my name) will be shared outside of the research team or published in the final report(s) from this research ☐

4. I agree to take part in the above study ☐

Participant's signature………………………………………………Date……………………….

Researcher Contact details:
Harmen Gudde, email: h.gudde@uea.ac.uk

Do also contact us if you have any worries or concerns about this research.
School of Psychology Ethics Committee:
ethics.psychology@uea.ac.uk; Phone 01603 597146
Head of School Professor Kenny Coventry:
k.coventry@uea.ac.uk; Phone 01603 597145

"Memory for Object Location"
Participant Debrief Sheet

Thank you for participating in this study. Your help with our work is greatly appreciated. Research in Psychology is only possible with the time and effort that participants generously give up. Your assistance today has made a genuine contribution to the Psychology research in the School, and the field in general.

In this research we are investigating the way language influences peoples memory for object location. At every presentation we altered the use of language to prime your memory. Using the real position of the object compared to the position you directed the researcher to we can see whether the language that was used actually changed your memory. This research can be useful for integrating natural language interfaces in computer systems.

NB Please *refrain* from sharing any of the information and hypotheses you have received during or after the experiment with individuals who may also take part in this study.

The data we have received through your participation will be analysed anonymously with the data we have gathered from other participants. Note that you may request your data to be removed from the sample at any time.

All information which you provided during the study will be stored in accordance with the 1998 Data Protection Act and kept strictly confidential. The chief investigator will be the custodian of the anonymous research data. All data will be kept for 5 years, and will then be securely disposed of. All data will be secured by a password, but your contacts will be kept separate from your results. No data will be shared with anyone that is not involved in the study.

Contacts

This study has been cleared in accordance with the ethical review processes of the University of East Anglia and within the guidelines of the British Psychology Society.

You are welcome to discuss your participation with the experimenter, Harmen Gudde: H.Gudde@uea.ac.uk; Phone 01603 591638 or the Principal Investigator, Head of School Professor Kenny Coventry: K.Coventry@uea.ac.uk; Phone 01603 597145.

If you have any worries or concerns about this research you want to share with someone not involved in the study you may contact: Chair of the Ethics Committee Dr. Piers Fleming, P.Fleming@uea.ac.uk; Phone 01603 59 3386 .

Appendix E

Information sheets for Experiment 9.

Participant Information Sheet

Thank you for your interest in this study. Before you decide whether to take part, please read the following information carefully (this sheet is for you to keep). You may ask any questions if you would like more information.

What will I do in this research?

In this experiment you will see some simple spatial relations that you have to describe while your eye-movements are measured with the eye-tracker.

Do I have to take part?

It is up to you to decide to join the study. We will describe the study and go through this information sheet. If you agree to take part, we will then ask you to sign a consent form. You are free to withdraw at any time, without giving a reason. This would not affect you in any way.

What will happen if I agree to take part?

If you agree you will be asked to describe some spatial situations that are presented on the screen. The experiment will take about 5 minutes.

Will it help me if I take part?

Taking part in this study will probably not help you personally, but by participating you will benefit the programme of research at the School of Psychology at UEA. It is therefore very much appreciated if you would take part.

How will you store the information that I give you?

All information you provided during the study will be stored in accordance with the 1998 Data Protection Act and kept strictly confidential. The chief investigator will be the custodian of the anonymous research data. All data will be secured by a password and your contacts will be kept separate from your results. No data will be shared with anyone that is not involved in the study. Anonymized data will be stored in line with open practice.

How will the data be used?

All data will be anonymised, processed and might be analysed. Data may be presented in journals and at conferences.

What happens if I agree to take part, but change my mind later?

You have the right to withdraw during the experiment and up until you leave the lab. After this time it will no longer be possible to identify your data for withdrawal since your experimental data is not paired with your name.

How do I know that this research is safe for me to take part in?

This study has been cleared in accordance with the ethical review processes of the University of East Anglia and within the guidelines of the British Psychology Society. For any questions please contact the Experimenter, Principal Investigator or Chair of the Ethics Committee, contacts are shown below.

You are under no obligation to agree to take part in this research.
If you do agree you can **withdraw at any time without giving a reason.**

Contacts

You are welcome to discuss your participation with the experimenter, Harmen Gudde: H.Gudde@uea.ac.uk;

Phone 01603 591638 or the Principal Investigator, Head of School Professor Kenny Coventry: K.Coventry@uea.ac.uk; Phone 01603 597145.
If you have any worries or concerns about this research you want to share with someone not involved in the study you may contact: Chair of the Ethics Committee Dr. Piers Flemming, P.Fleming@uea.ac.uk; Phone 01603 59 3386 .

Consent Form
Object location in static visual scenes
Name of Researcher: Harmen Gudde

Please initial all boxes

5. I have read and understand the information sheet (Object location in static visual scenes) and have had the opportunity to questions and have had these answered satisfactorily. ☐

6. My participation is voluntary and I know that I am free to withdraw at any time, without giving any reason and without it affecting me at all ☐

7. I know that no personal information (such as my name) will be shared outside of the research team or published in the final report(s) from this research ☐

8. I agree to take part in the above study ☐

Participant's signature…………………………………………Date………………………..

Researcher Contact details:

Harmen Gudde
H.Gudde@uea.ac.uk
Tel: +44 (0)1603 59 1638

Do also contact us if you have any worries or concerns about this research.

School of Psychology Ethics Committee:
ethics.psychology@uea.ac.uk; Phone 01603 597146
Head of School Professor Kenny Coventry:
K.Coventry@uea.ac.uk; Phone 01603 597145

“Object Location in Static Visual Scenes”
Participant Debrief Sheet

Thank you for participating in this study. Your help with our work is greatly appreciated. Research in Psychology is only possible with the time and effort that participants generously give up. Your assistance today has made a genuine contribution to the Psychology research in the School, and the field in general.

In this research we are investigating the way people describe spatial situations. Using the eye-tracker we measured your eye movements to see which areas of interest you have checked when you analysed and described each of the situations we have presented. This research can be useful for integrating natural language interfaces in computer systems.

NB Please refrain from sharing any of the information and hypotheses you have received during or after the experiment with individuals who may also take part in this study.

The data we have received through your participation will be analysed anonymously with the data we have gathered from other participants. Note that you may request your data to be removed from the sample at any time.

All information which you provided during the study will be stored in accordance with the 1998 Data Protection Act and kept strictly confidential. The chief investigator will be the custodian of the anonymous research data. All data will be kept for 5 years, and will then be securely disposed of. All data will be secured by a password, but your contacts will be kept separate from your results. No data will be shared with anyone that is not involved in the study.

Contacts

This study has been cleared in accordance with the ethical review processes of the University of East Anglia and within the guidelines of the British Psychology Society.

You are welcome to discuss your participation with the experimenter, Harmen Gudde: H.Gudde@uea.ac.uk; Phone 01603 591638 or the Principal Investigator, Head of School Professor Kenny Coventry: K.Coventry@uea.ac.uk; Phone 01603 597145.

If you have any worries or concerns about this research you want to share with someone not involved in the study you may contact: Chair of the Ethics Committee Dr. Piers Flemming, P.Fleming@uea.ac.uk; Phone 01603 59 3386 .

Appendix F

Object-pairs used in the probe trials of study 4.1