University of Glasgow

Dunn, Benjamin Mark (2016) *Which way is up? Grounded mental representations of space.* PhD thesis.

http://theses.gla.ac.uk/7460/

Copyright and moral rights for this thesis are retained by the author

A copy can be downloaded for personal non-commercial research or study

This thesis cannot be reproduced or quoted extensively from without first obtaining permission in writing from the Author

The content must not be changed in any way or sold commercially in any format or medium without the formal permission of the Author

When referring to this work, full bibliographic details including the author, title, awarding institution and date of the thesis must be given

Glasgow Theses Service
http://theses.gla.ac.uk/
theses@gla.ac.uk

Which way is up?

Grounded Mental Representations of Space

Benjamin Mark Dunn, BA, PGDip, MRes.

Submitted in fulfilment of the requirements for the Degree of PhD

Institute of Neuroscience and Psychology
College of Medical, Veterinary and Life Sciences
University of Glasgow

January 2016

© Benjamin Mark Dunn. 29th January 2016

Abstract

Processing language is postulated to involve a mental simulation, or re-enactment of perceptual, motor, and introspective states that were acquired experientially (Barsalou, 1999, 2008). One such aspect that is mentally simulated during processing of certain concepts is spatial location. For example, upon processing the word “moon” the prominent spatial location of the concept (e.g. ‘upward’) is mentally simulated.

In six eye-tracking experiments, we investigate how mental simulations of spatial location affect processing. We first address a conflict in previous literature whereby processing is shown to be impacted in both a facilitatory and inhibitory way. Two of our experiments showed that mental simulations of spatial association facilitate saccades launched toward compatible locations; however, a third experiment showed an inhibitory effect on saccades launched towards incompatible locations. We investigated these differences with further experiments, which led us to conclude that the nature of the effect (facilitatory or inhibitory) is dependent on the demands of the task and, in fitting with the theory of *Grounded Cognition* (Barsalou, 2008), that mental simulations impact processing in a dynamic way.

Three further experiments explored the nature of verticality – specifically, whether ‘up’ is perceived as away from gravity, or above our head. Using similar eye-tracking methods, and by manipulating the position of participants, we were able to dissociate these two possible standpoints. The results showed that mental simulations of spatial location facilitated saccades to compatible locations, but only when verticality was dissociated from gravity (i.e. ‘up’ was above the participant’s head). We conclude that this is not due to an ‘embodied’ mental simulation, but rather a result of heavily ingrained visuo-motor association between vertical space and eye movements.

List of Tables

List of Figures

Acknowledgements

I would like to take this opportunity to express my appreciation and gratitude to my supervisor Dr Christoph Scheepers who, in the past three years, has educated and supported me above and beyond what is presented in this thesis. Under your supervision, I have grown to become an independent, freethinking academic.

My thanks are extended to Dr Yuki Kamide who, along with Dr Scheepers, gave me this opportunity. Furthermore, I am thankful to the ESRC (Economic and Social Research Council) for funding my doctoral work, as well as the wider project led by Dr Kamide and Dr Scheepers.

I am also grateful to Dr Sara Sereno who has helped and advised me on numerous occasions. Similar thanks go to my officemates Dr Dominic Thompson and Kieran O'Shea for allowing me to offload my concerns.

Finally, I would like to thank my family for the unconditional love and support I receive from them every day. In particular, Laura and Elliot have taught me that there is much more to life than just academia.

Author’s declaration

I declare that this thesis is my own work carried out under the normal terms of supervision.

Data reported throughout this thesis has been presented at the following academic conferences by the author:

Conference Talks:

Dunn B.M., Kamide Y. & Scheepers, C. (2015) Processing Spatial Language. *Talk presented at the 30th Annual Conference of the British Psychological Society Postgraduate Affairs Group.* Glasgow, UK.

Dunn B.M., Kamide Y. & Scheepers, C. (2014) Hearing "moon" and looking up: Word-related spatial associations facilitate saccades to congruent locations. *Paper presented at the 36th Annual Conference of the Cognitive Science Society.* Quebec City, Canada.

Conference Posters:

Dunn B.M., Kamide Y. & Scheepers, C. (2015) Processing Spatial Language: Facilitatory and Inhibitory Effects Depend on Task Demands. *Poster presented at the 21st Annual Conference for Architectures and Mechanisms for Language Processing Conference (AMLaP).* Valletta, Malta.

Dunn B.M., Kamide Y. & Scheepers, C. (2014) Which way is up? Are word-related spatial associations affected by body position?: Evidence from eye movements. *Poster presented at the 7th Annual Conference on Embodied and Situated Language Processing (ESLP).* Rotterdam, The Netherlands.

Chapter 1 Introduction

Dante Alighieri wrote, “Heaven wheels above you, displaying to you her eternal glories, and still your eyes are on the ground” (Sisson, 1993, p. 261). If we think of heaven as above us, should we not look up towards it?

Descriptions of spatial locations pervade our everyday lives; yet understanding a spatial location implied by language requires a mental representation. In order to comprehend language, it is thought that people construct a mental model (Garnham, 1987; Johnson-Laird, 1983) or situation model (Van Dijk & Kintsch, 1983; Zwaan & Radvansky, 1998) allowing them to form an internal representation of the described event. For instance, when we hear the word “moon”, we know that the moon itself is in the sky and in order to see it we must look upwards. Yet, even when we are sitting inside and unable to see the sky, we can still imagine the moon in its usual location. This internal mental representation of the moon’s spatial location is partly made up of our experiential knowledge – we have experience of looking up to the moon in the sky.

The aim of the present research is twofold. Firstly, to explore how mental representations of spatial locations implied by language influence our processing. And secondly, whether our body position influences the way in which we mentally represent spatial language. In a series of experiments outlined below, I will show that the way in which we process spatial language is dynamic, and influenced by the task at hand. Furthermore, our mental representations of space are influenced by our body position.

Here I provide a brief overview of the work contained in this thesis. Further discussion of each topic is provided in the relevant chapters. Chapter 2 starts by providing a review of the relevant literature. The remainder of the thesis is divided into two main parts: Part I (Chapters 2-5) investigates how our mental representations of space affect eye-movement behaviour. Chapter 4 begins by investigating a conflict in the previous literature whereby mental representations of space have been shown to both facilitate and inhibit processing. I present three separate experiments showing the differing influence that vertical associations of space have on eye movements in both a lexical decision and target detection task. These are followed by two further experiments investigating how manipulating different aspects of a task impacts on processing and eye-movement behaviour. Following on from this, Chapter 5 focuses on how our mental representations of space can affect our eye-movement trajectories. I present an experiment investigating whether the vertical spatial associations of a concept can influence attention in such a way that saccade trajectories differ depending on perceived directionality.

Part II furthers the research from Part I. In Chapter 6, I present an experiment manipulating the body position of participants in an attempt to understand how our mental representations of space are affected by our body position. The final two experiments investigate this in more detail. These experiments use a similar methodology to those in Part 1 to measure eye-movement behaviour, but also incorporate the effects of manipulating body position explored in the first experiment of Part 2.

Finally, Chapter 7 concludes the thesis and presents a summary of the findings, a discussion of the implications, and directions for future work.

Part I. Eye-movement Behaviour

Chapter 2 Literature Review

2.1 Mental Representations Involved in Language Comprehension

It is proposed that language comprehension involves forming an internal representation of the described objects, situations, and events, known as a mental or situation model (Garnham, 1996; Glenberg & Langston, 1992; Johnson-Laird, 1983; Van Dijk & Kintsch, 1983). However, more than just understanding the semantic aspect of language, Garnham (1987) proposed that comprehenders use their existing knowledge to mentally represent the information given in the text as an imagined visual scene. For example, whilst comprehending the sentence "Lisa was at her desk writing her thesis" we use our existing knowledge to bridge the gap between the linguistic information supplied by the text and our imagined mental model. The linguistic information actually tells us very little here: A female called Lisa is within close proximity to a desk and is writing a thesis that belongs to her. Yet, we can use our own knowledge and experience to enrich our representation of the sentence beyond that of just the language itself (Glenberg, Meyer, & Lindem, 1987; Zwaan, 1999; Zwaan & Radvansky, 1998). You might imagine that Lisa is a student, she is sat down on a swivel chair, and she is typing at a computer. Importantly, none of these assumptions were explicitly mentioned in the text, but using our experiential knowledge of desks and thesis writing we create a more complete mental image. Did you imagine a swivel chair or perhaps one with four legs? Either way, the very fact that we can ask such a question shows that my mental representation of the situation most certainly differs to your mental representation in some aspects as we have a wealth of different experiences. Similarly, there will be many parts of our mental representations that overlap with one another's. Furthermore, our experiential knowledge of concepts can be combined in order to create mental representations of described events of which we have no experience. The well-known idiom "pigs might fly" is used to imply the impossibility of an event, and yet by combining our experiential knowledge of 'pigs' and 'flying', a mental representation of a pig flying above the Glaswegian skyline is no more difficult (depending on your knowledge of Glasgow) to visualise than imagining Lisa writing her thesis.

Mental representations of language are not limited to sentential events, but also occur at word level (Dudschig, Lachmair, de la Vega, De Filippis, & Kaup, 2012; Dudschig, Souman, Lachmair, de la Vega, & Kaup, 2013; Estes, Verges, & Barsalou, 2008; Gozli, Chasteen, & Pratt, 2013; Lachmair, Dudschig, De Filippis, de la Vega, & Kaup, 2011; Zwaan & Yaxley, 2003). For example, if our earlier protagonist Dante were presented with the word “heaven”, it would activate his existing semantic and experiential knowledge of the concept (God, clouds, up, angels, high, death, and so on) leading to a mental representation of heaven. It becomes clear then, that mental representations are a crucial part of how we understand language. Indeed, one of the questions considered in this thesis is the way in which word level mental representations of vertical spatial locations affect online processing.

2.2 The Symbol Grounding Problem

One problem that mental models commonly encounter is the symbol grounding problem. Broadly, the problem is how an abstract representation (or symbol) derives meaning from anything other than the other abstract symbols we associate with it. If Dante’s representation of ‘heaven’ is made up of ‘God’, ‘clouds’, ‘angels’, and so on, then what is his representation of say, ‘God’ made up of? Ultimately, the problem is that none of these symbols have any true meaning. Indeed, in John Searle’s now famous *Chinese Room* thought experiment he proposes that one could be *perceived* as understanding Chinese symbols without actually comprehending the meaning of them (Searle, 1980). Searle argues that he could receive Chinese symbols (uninterpretable to him) and by using a set of rules given to him in English, he could return different Chinese symbols that indicate a coherent response despite not understanding any of the given symbols. Hence, to the Chinese reader, it looks as if Searle has understood the incoming symbols and produced a valid response. This, argues Searle, is how a computer could be perceived as understanding language by following a set of given rules and yet never actually comprehend the meanings of the input or output. In terms of the symbol grounding problem, Searle shows how a language can consist of symbols and rules without ever having any meaning. Using symbols to provide meaning to other symbols is why Harnad (1990) likens the symbol grounding problem to learning Chinese from a Chinese-Chinese dictionary whereby continuously ‘grounding’ symbols with other symbols would be never-ending. For instance, in Dante’s case, one symbol contributing to his understanding of “heaven” is “clouds”, which in turn he may understand as “vapour”, which in turn he understands as “gas” and so on. Thus, the symbol grounding problem asks how our understanding of a concept is made up of anything other than a collection of yet more abstract symbols.

More recently, Barsalou (1999) proposed that the meaning of conceptual symbols is understood perceptually. Barsalou states that perceptual components are grounded in our sensory experiences (including introspection and proprioception). Ultimately, we understand the meaning of a concept by mentally simulating our experience of it in a limitless number of ways.

2.3 Grounded language comprehension

The theory of *Grounded Cognition* (Barsalou, 2008) proposes that the brain's modal systems underlie cognitive processes. It rejects the more traditional idea that amodal symbols represent knowledge in semantic memory. A main focus of grounded cognition is the role that mental simulations (i.e. partial re-enactments of perceptual, motor and introspective states acquired during experience with the world, body, and mind) play in cognitive processing (Barsalou, 1999, 2008; Decety & Grèzes, 2006; Goldman, 2006). This approach argues that we mentally simulate perceptual, motor and introspective states experienced upon interaction with a concept in order for later retrieval. For instance, on Dante's ascent to heaven, he must climb to the top of the mountain *Purgatory*. Upon later retrieval of the concept "mountain", Dante's mental simulation comprises of his experiences and his knowledge of mountains (e.g. how a mountain looks and feels, the action of climbing a mountain, his fear of heights). However, mental simulations are not just mental imagery. Mental imagery requires a more conscious effort to construct and represent a concept, whereas mental simulations seem to occur automatically.

According to Barsalou (2008), this is what happens when we comprehend language – and most of it occurs at a nonconscious level. There is much research to support the idea of language comprehension as mental simulations. For example, Intraub and Hoffman (1992) presented participants with a mixture of paragraphs and pictures and found that, upon later presentation of the paragraphs (some of which described the previously presented pictures), participants were confident that they had seen a picture described by the paragraph even when they had not. These results suggest that upon the initial encoding of the paragraphs, participants experienced a mental representation allowing them to internally visualise the described scene. Hence, during the recognition phase participants confused their mental representation of a scene with images they had actually seen. In a similar vein, Potter and colleagues have shown that including pictures during text did not significantly interfere with sentence processing (Potter, Kroll, Yachzel, Carpenter, & Sherman, 1986).

More specific to the work presented in this thesis, is how grounded cognition explains perceptual simulations of space.

2.4 Perceptual simulation

A large amount of research has focused on how perceptual simulations are affected by language comprehension. In this manner, Stanfield and Zwaan (2001) presented participants with a sentence implying a horizontal or vertical orientation of an object (e.g., He hammered the nail into the [wall / floor]) followed by a congruent or incongruent depiction of the object (a nail laying flat / standing up). When asked if the visual target had been mentioned in the sentence, participants responded faster when the cued orientation matched that of the target. Moreover, Zwaan, Stanfield, and Yaxley (2002) presented cue sentences implying the shape of an object (e.g. The ranger saw the eagle [in the sky / in its nest]) and showed that participants were quicker to identify a visual target (e.g. picture of an eagle) when its shape (wings outstretched, compared with wings folded) was compatible with the shape implied by the sentence. Zwaan and Yaxley (2003) found facilitatory effects of perceptual-spatial representations using a semantic relatedness task. Specifically, when two words were presented on the computer screen in line with the real-world vertical arrangement of the objects they referred to (e.g. the word ‘branch’ displayed above the word ‘root’), participants were faster to provide a relatedness judgement than when the order was reversed (‘root’ displayed above ‘branch’). The same was found with abstract concepts such as ‘master’ and ‘slave’ (Schubert, 2005). Taken together, these findings suggest that modality-specific perceptual features (i.e. size, shape, location) acquired during previous experiences are mentally simulated during language comprehension.

Glenberg and Kaschak (2002) introduce the ‘action-sentence compatibility effect’ (ACE) whereby providing sensibility judgements on sentences implying a ‘towards’ or ‘away’ motion (e.g. ‘[open / close] the drawer’) took longer when the response location was in an incompatible position. That is, after reading “open the drawer”, participants took longer to respond when a ‘yes’ response involved moving their hand away from them to press a button (see also Borghi, Glenberg, & Kaschak, 2004; Tucker & Ellis, 2004). Furthermore, Zwaan and Taylor (2006) showed that linguistically implied rotation (e.g. ‘Liza opened the pickle jar’) led to faster sensibility judgements when the motor response involved turning a knob in a direction compatible with the linguistic stimuli (e.g. anticlockwise). These findings show that mentally simulating linguistically implied motor movements directly impacts on actual motor movements.

Kaschak et al. (2005) presented participants with sentences describing motion either towards or away from the participant (e.g. ‘the car approached you’ vs. ‘the car left you in the dust') concurrently with a dynamic visual stimulus implying 'towards' or 'away' motion. However, their results show that participants took longer to make judgements (i.e. did the sentence make sense?) on the sentences when the described motion matched that of the dynamic visual stimulus, thus suggesting that the visual display interfered with a perceptual simulation evoked by the linguistic stimuli. The conflicting patterns of results (specifically between the ACE and Kaschak et al., 2005) are explained in terms of temporal overlap and integratability. Notably, the perceptual stimulus (a rather abstract visual display) used by Kaschak et al. (2005) is not easily integrated into a simulation about cars moving towards or away from you and hence interferes with responses. With regards to the ACE, the ease at which one integrates the 'towards' or 'away' movement with the mental simulation activated by the sentence leads to there being a compatibility, as oppose to a mismatch, effect. This idea is also supported by work conducted by Zwaan, Madden, Yaxley, and Aveyard (2004), outlined below.

Together, these studies indicate that perceptual simulations are automatically activated whilst comprehending language, as none of the tasks required explicit comprehension or judgements of shape/direction/orientation, yet responses were still affected. Furthermore, these simulations combine experiential knowledge with contextual information given by the linguistic stimuli such that, for instance, a perceptual simulation of an eagle differs depending on the retrieval cues.

However, we live in a dynamic world and the studies above investigate merely a snapshot of everyday life. Numerous studies have shown that mental representations are dynamic (e.g. Freyd, 1983, 1987; Freyd & Finke, 1984). For example, participants viewing photographs implying action movements (e.g. jumping from a wall) took longer to classify the two images as different from one another when the implied motion was moving forward in time compared to backward in time (Freyd, 1983). Thus, it is suggested that participants formed a dynamic mental representation and 'filled in the gap' between the two photographs. Hence, seeing a photograph of the person jumping from the wall followed by a photograph taken slightly later was harder to distinguish than seeing them in the reverse order (in which the person went from mid-air, back to the wall).

There are also a number of studies supporting the view that dynamic mental representations are evoked by language. Whilst much of the previously presented research suggests that a mental representation of a specific linguistically described object, situation or event affects our processing of a static visual scene, not all language describes a static event. Zwaan et al. (2004) presented participants with sentences implying differing directions of movement; participants either heard sentences implying a 'towards' movement (e.g. ‘The shortstop hurled the softball at you’) or an ‘away’ movement (e.g. ‘You hurled the softball at the shortstop’). After each sentence, participants saw a sequence of two separate images showing the previously mentioned object (e.g. the softball). Crucially, the images either depicted a ‘towards’ movement by showing a smaller softball followed by a larger version of the same softball, or an ‘away’ movement by reversing the presentation order. Participants were quicker to classify the two images as the same object when the apparent motion matched the motion implied by the earlier presented sentence.

More recently, Lindsay, Scheepers, and Kamide (2013) presented visual scenes depicting an agent (i.e. a student) at one end of a path and a goal object (i.e. picnic basket) at the other. Eye-movement analyses showed that participants were quicker to fixate the goal object after hearing sentences with a fast verb compared with a slow verb (e.g. ‘The student will [run / stagger] along the trail to the picnic basket’). Furthermore, participants fixated the path for a longer period (and the goal object for a shorter period) in the slow verb condition compared to the fast verb condition. Analogous results were shown by Speed and Vigliocco (2013). These results suggest that participants were mentally simulating the described movement and in particular its implied speed of motion, which in turn affected the decision of where, and in what space of time, to deploy their visual attention (i.e. the ‘path’ region vs. the ‘goal’). Collectively, these studies show how we are able to create a dynamic representation of a described event, which can be updated as more information becomes available. Moreover, perceptual simulation of a motion event affects the processing of non-linguistic stimuli.

Overall, the studies presented here show that language automatically activates a mental simulation by combining the available linguistic information and our experiential knowledge. Furthermore, mental simulations are dynamic in their nature and affect our processing, attention and behaviour.

2.5 Perceptual Simulations of space

When Dante heard “heaven”, part of his existing knowledge that became active was the spatial location – he knows (or believes) that heaven is up. A number of common words activate perceptual associations of space. ‘Push’ and ‘pull’ seem to be mentally represented on a horizontal axis, whereas ‘float’ and ‘sink’ are more likely to activate vertical mental representations (Richardson, Spivey, Barsalou, & McRae, 2003). In addition, ‘bird’ and ‘head’ entail upward spatial associations, whereas ‘foot’ and ‘snake’ are more likely to invoke downward spatial associations (Estes et al., 2008; Meteyard, Bahrami, & Vigliocco, 2007; Šetić & Domijan, 2007; Zwaan & Yaxley, 2003). Mental representations of vertical and horizontal spatial axes have been shown to occur with various different concepts. Indeed, these perceptual representations of space are similar to the SNARC effect (spatial-numerical association of response codes) whereby mental representations of numbers seem to occur along a mental number line with smaller numbers being represented as closer to the perceiver than larger numbers (Dehaene, Bossini, & Giraux, 1993; Fischer, Castel, Dodd, & Pratt, 2003). Interestingly, the direction by which the numbers increase along the mental number line seems to depend on the individual's established reading direction. Whilst readers of left-to-right languages (i.e. English) show evidence for a number line increasing from left to right, readers of right-to-left languages (i.e. Arabic) show the opposite (Maass & Russo, 2003; Tversky, Kugelmass, & Winter, 1991; Zebian, 2005).

Meier and Robinson (2004) showed how valence is represented on a vertical axis. Positive words were categorised quicker when they appeared at the top of the screen, whereas negative words were categorised quicker when they appeared at the bottom of the screen. These results support the theory that abstract concepts are grounded by physical metaphors, such that 'happy' and 'sad' are upward and downward concepts respectively (Lakoff & Johnson, 1999; Piaget & Inhelder, 1969). Meier and Robinson furthered these results by showing that people displaying higher levels of neuroticism or depressive symptoms were more likely to detect lower versus higher targets (Meier & Robinson, 2006). Furthermore, divine concepts seem to be represented vertically – with 'God' as up and 'Devil' as down (Chasteen, Burdzy, & Pratt, 2010; Meier, Hauser, Robinson, Friesen, & Schjeldahl, 2007). In contrast, there is also a body of research suggesting that valence is represented on the horizontal axis. Recently, Casasanto (2009) showed that handedness affects horizontal representations of emotional valence. Left-handed individuals were more likely to associate positive concepts with leftward space and negative concepts with rightward space whereas right-handed individuals showed the opposite pattern.

Finally, there is evidence to suggest that the abstract concept of power is represented vertically. As outlined earlier, Schubert (2005) found that people were quicker to make a semantic-relatedness judgement when words associated with powerfulness such as "master" were presented above words associated with powerlessness, such as "slave" (see also: Giessner & Schubert, 2007; Zanolie et al., 2012). Nevertheless, this seems to be a function of the relationship between two concepts. Lakens and colleagues showed that powerful words (e.g. 'king', 'boss') were rated as higher in vertical space when presentation was interspersed with powerless words (e.g. 'slave', 'defendant') compared with when they were rated independently. Hence, a master is not so powerful without their slave (Lakens, Semin, & Foroni, 2015). Similar vertical space representations have been found using Chinese honorifics (Lu, Zhang, He, Zheng, & Hodges, 2013). Participants were quicker to judge the vertical orientation of an arrow when it was preceded by a compatible versus incompatible honorific (i.e. an elevating word followed by an arrow pointing upward, or denigrating word followed by an arrow pointing downward). Furthermore, Taylor, Lam, Chasteen, and Pratt (2015) showed that participants were quicker to detect targets appearing at the top (vs. bottom) of the screen after reading high self-esteem words (e.g. ‘brave’); the reverse was true for low self-esteem words (e.g. ‘timid’).

In summary, we have seen that vertical (and in some cases horizontal) representations of space encompass a wide range of concepts. Power, emotional valence, and self-esteem all seem to be represented vertically. However, the work presented here will focus on vertical representations of spatial location, and how these can modulate our attention.

2.6 Perceptual representations of spatial location modulate attention

A number of studies have shown how perceptual representations of spatial location can direct attention (Bergen, Lindsay, Matlock, & Narayanan, 2007; Dudschig et al., 2013; Gozli et al., 2013).

Bergen et al. (2007) showed how sentences implying dynamic upward or downward movement interfered with identification of a letter at the top or bottom of the screen respectively. For example, reading “The cellar flooded” [downward noun] or “The lizard ascended” [upward verb] hindered identification of an ‘X’ at the bottom or top of the screen respectively. Similarly, Estes et al. (2008) showed that presenting participants with upward (e.g. “hat”) or downward (e.g. “boot”) associated words interfered with detection of a target in a compatible on-screen location.

More recently, there is evidence to show that perceptual representations of spatial location can facilitate, rather than inhibit processing. Dudschig et al. (2013) visually presented participants with a series of German words with upward or downward spatial associations (e.g. “Sonne” and “Maus”, which translate to ‘sun’ and ‘mouse’) along with a number of non-words; the participant was required to make a lexical decision. This decision was made by looking towards a target at either the top or bottom of the screen. Results showed a significant interaction between spatial association and response location, with saccades to targets at the top of the screen being launched more quickly after reading an upward word vs. a downward word. Conversely, saccades were launched more quickly to the bottom of the screen after reading a downward word vs. an upward word. Despite not being an explicit part of the task, perceptual representations of space were activated by the words, affected attention, and therefore influenced eye-movement behaviour. However at this stage, it is impossible to conclude whether compatible word-target spatial associations (i.e. 'sun' and looking to the top of the screen) facilitated processing, or whether incompatible word-target spatial associations (i.e. 'sun' and looking to the bottom of the screen) interfered with processing, as the researchers did not use a comparable baseline condition. The lack of baseline condition is a recurrent theme throughout research in the current field and one aim of this doctoral thesis is to show why including such a condition is a necessity in any future research.

The research outlined above shows that perceptual representations of space activated by our experiential knowledge of a concept seem to modulate our attention. Yet, there is disparity in the results, which means it remains unclear as to whether attending to a location compatible with the perceived spatial location of a concept is facilitated or inhibited. Indeed, there are numerous studies supporting both a facilitatory and inhibitory effect of processing.

For instance, when participants were required to make a lexical decision on nouns associated with a upward or downward location (e.g. 'roof' vs. 'root'), they were quicker to respond in the direction compatible with the word (Lachmair et al., 2011). That is, participants were quicker to press a button located higher up after reading upward nouns (e.g. 'roof') compared to downward nouns; the reverse was true for a response button located lower down. The results show by Lachmair and colleagues seem to fit with the aforementioned Action-Sentence Compatibility Effect (ACE), albeit using nouns with perceived spatial locations as opposed to sentences implying action (Glenberg & Kaschak, 2002). However, the discrepancy between inhibitory and facilitatory effects remains. That is, Lachmair et al. (2011) do not include a baseline condition therefore making the results inconclusive as to whether compatible cue-target trials facilitated processing or incompatible cue-target trials inhibited processing.

Three previously discussed experiments show how perceptual representations of spatial location interfere with detection of a target in a compatible on-screen location (Bergen et al., 2007; Estes et al., 2008; Richardson et al., 2003). Perhaps then, target detection tasks show inhibitory effects on processing, and tasks that require a deeper level of processing show facilitatory effects – i.e. lexical decision (Dudschig et al., 2013; Lachmair et al., 2011) or semantic relatedness (Schubert, 2005; Zwaan & Yaxley, 2003). This potential explanation is inconsistent with the results of a recent target detection experiment (Dudschig et al., 2012). Similar to the aforementioned lexical decision task

(Dudschig et al., 2013), participants read words implying an upward or downward spatial location. Following the word presentation, a small target appeared at either the top or bottom of the screen – participants pressed the space bar once they had detected the target. Response times showed a compatibility effect with faster detection of a target appearing at the top of the screen following an upward word vs. downward word, and the reverse pattern of results for a target appearing at the bottom of the screen. Whilst these results seem to show a facilitatory effect on processing exhibited by a target detection task – something that does not fit with the above explanation – they once again lack a neutral baseline condition. Therefore it remains inconclusive as to whether compatible word-target associations facilitated processing or incompatible word-target associations inhibited processing.

Estes et al. (2008) hypothesised that processing would be hindered if the representation of a cued concept did not have overlapping features with the target object. Alternatively, if the perceptual simulation of the cued object shares features with the target object then processing is facilitated. This goes some way to explaining the conflicting results in the previously reviewed literature. For example, perceptual simulations of an eagle in the sky/nest shared overlapping features with the visual form of the target (Zwaan et al., 2002); or perceptual simulations of vertical/horizontal concepts shared few features with a target (a small square/circle) and therefore hindered detection (Richardson et al., 2003). Of course, since Estes and colleagues proposed the perceptual-featural overlap hypothesis in 2008, there have been a number of studies concluding facilitatory effects on processing despite the target object (in most cases, a small circle or square) sharing few features with the cued concept (Dudschig et al., 2012; Dudschig et al., 2013; Lachmair et al., 2011).

One final, but important point is the difference between target detection and target discrimination. In a target discrimination paradigm more than one target appears (e.g. an X at the top and an O at the bottom of a display) and the participant must locate a particular one. Such tasks have shown inhibitory effects on processing (e.g. Estes et al., 2008). Conversely, those using a target detection paradigm – locating one target as soon as it appears – seem to show facilitatory effects on processing (e.g. Dudschig et al., 2012; Lachmair et al., 2011)[1]. Yet, in a direct comparison between target detection and target discrimination, Gozli et al. (2013) showed that spatial word cues facilitated detection of a

[1] This is not always the case (e.g. Bergen et al., 2007; Gozli et al., 2013; Richardson et al., 2003).
[2] The University of Glasgow subject pool consists of over 7500 active members ranging from age 16-75. Each subject is vetted via self-report questionnaires concerning a number of filters. All participants involved in the

target in a compatible location (e.g. 'God' followed by a target at the top of the display); and whilst the response times for the detection task were significantly faster than for the discrimination task, the compatibility effect was the same. Taking into account the results shown by Gozli and colleagues (2013), we consider the aforementioned target discrimination procedures in the same vein as target detection procedures throughout the discussions in this thesis. This issue will, however, be revisited in further detail in Chapter 7.

In summary, there have been a number of studies published in the last 15 years focusing on how perceptual simulations of space modulate attention and affect behaviour. There is an ongoing debate about whether processing spatial concepts facilitates or inhibits behavioural responses to compatible or incompatible locations. A consistent flow of research is being published and arguing in favour of facilitatory and inhibitory effects on processing. Yet ultimately, the results of the majority of studies outlined above remain inconclusive, as facilitatory or inhibitory effects do not become clear unless compared to a neutral baseline condition. In a number of experiments presented in *Part 1* of this thesis, I will investigate whether perceptual simulations of spatial concepts facilitate or inhibit responses to (in)compatible locations and whether these effects are dependent upon the task at hand. Comparing the results to a neutral baseline condition will allow for a reliable conclusion to be drawn about the facilitatory or inhibitory effects on processing.

Chapter 3 Norming of Stimuli

3.1 Linguistic Stimuli

Eight of the nine experiments presented here used the same 120 word stimuli. Prior to any experimental data collection, a spatial association norming study and principal component analysis were performed (detailed below). Where appropriate, the analyses conducted throughout this thesis use the principal components as covariates in an attempt to control for confounding effects the linguistic stimuli have on the dependent variable.

Spatial Association Norming

An internet-based rating study was conducted to verify the intended spatial associations per condition. Participants from the University of Glasgow subject pool[2] rated 402 English candidate words for vertical association on a Likert scale ranging from −5 to +5 (detailed below). The words were split into 15 lists (each seen by at least 21 participants), with 25-30 items per list. Underneath each printed candidate word, there was an 11-point bipolar scale on which participants had to provide their spatial association ratings. The leftmost point on the scale (scored as −5) was labelled "down" (for downward association), the rightmost point (scored as +5) was labelled "up" (for upward association), and the midpoint (scored as 0) was labelled "neutral" (for no vertical association). Participants also marked a word as 'known' if they were familiar with the word or 'unsure' if they were not. Eleven cases (0.1%) with 'unsure' ratings were removed from analysis. The mean rating for the final selection of 'up' words was +3.65 (N = 40; min. +3.00; max. +4.36); the 'neutral' words scored an average of +0.03 on the scale (N = 40; min. −0.27; max. +0.43); finally, the 'down' words had an average rating of −3.48 (N = 40; min. −4.48; max. −2.82).

One hundred and twenty words were chosen as linguistic stimuli. There were 40 'up' (e.g. *moon*), 40 'down' (e.g. *sewer*) and 40 'neutral' (e.g. *letter*) words. The 'up' and 'down' stimuli each consisted of 20 verbs, 12 nouns and 8 adjectives whilst the 'neutral' condition had 20 verbs, 11 nouns and 9 adjectives. All words are listed in the Appendix.

Unlike much past research (e.g. Dudschig et al., 2013), we chose to norm a mixture of word classes, and ultimately, this is reflected in our chosen linguistic stimuli. Language contains a number of different word classes and we aimed to make our linguistic stimuli as

[2] The University of Glasgow subject pool consists of over 7500 active members ranging from age 16-75. Each subject is vetted via self-report questionnaires concerning a number of filters. All participants involved in the experiments in this thesis: were British English, had normal vision and hearing, had no learning disabilities, were native English monolingual speakers, were not synesthetes, and did not have autism spectrum disorder. Participants also submitted a screening questionnaire to the experimenter confirming their status as a monolingual English speaker.

representative as possible. In this instance, we intended to collect 20 adjectives, 20 verbs and 20 nouns for each direction category (180 words in total); however the ratings of the candidate words did not allow this (e.g. the 20th highest ranked adjective was *aloft*, with a mean rating of +2.36, which was not reliably associated with 'upwards' and therefore could not be included as part of the linguistic stimuli). To this end, we took the decision to use 20 verbs and a combination of 20 nouns and adjectives as these reflected the spatial association ratings more reliably. Whilst it would have been possible to increase the amount of normed candidate words to potentially meet the initial aim of 180 linguistic stimuli, we felt that the chosen 120 words was a representative reflection of language as a whole. It could be argued that the use of different word classes may mask any possible effect that perceptual simulations of space may have on motor responses. For example, most previous research uses only nouns (e.g. Dudschig et al., 2012; Dudschig et al., 2013; Lachmair et al., 2011). However, we feel that in order to truly investigate whether language processing is grounded and hence affects motor responses, one must use a sample of words that reflects different word classes. Indeed, Bergen et al. (2007) have shown how processing sentenced with 'up' and 'down' verbs impacts upon response behaviour. The results to our norming study seem to suggest that perceptual simulations of space are just as prevalent in verbs and adjectives as they are in nouns – some of the highest (and lowest) ranked words support this (e.g. mean ranking for 'descending' was -4.22, and for 'elevating' was +4.24).

Across the three word direction conditions, one-way ANOVAs showed no significant differences in lexical frequency (Baayen, Piepenbrock, & Van Rijn, 1993), number of syllables, and number of phonemes, lexical decision response times (Balota et al., 2007; Keuleers, Lacey, Rastle, & Brysbaert, 2012)[3], arousal (Warriner, Kuperman, & Brysbaert, 2013) or concreteness. To gather concreteness ratings, we asked 38 new participants to rate each of the 120 words on a scale of 1 (very abstract) to 7 (very concrete).

The ANOVAs also included valence and dominance ratings for each word (Warriner et al., 2013). Valence ratings of 'down' words (M = 4.15, SD = 1.07) differed significantly to both 'neutral' (M = 5.16, SD = 1.31) and 'up' words (M = 6.07, SD = 0.97). 'Neutral' and 'up' words did not differ significantly (note however, that $p = .053$). 95% confidence intervals (CIs) showed that 'down' words showed significantly lower

[3] Given the high correlation ($r = .77$) between the British Lexicon Project (BLP) and the English Lexicon Project we chose to use the more recent BLP lexical decision response times in the principal component analysis detailed below.

valence ratings compared with 'neutral' words (-1.02 ± 0.76) and 'up' words (-1.92 ± 0.73). Similarly, dominance ratings of 'down' words (M = 4.50, SD = 0.84) differed significantly to both 'neutral' (M = 5.52, SD = 0.61) and 'up' words (M = 5.55, SD = 0.69). 95% CIs showed that 'down' words had significantly lower dominance ratings compared with 'neutral' words (-1.03 ± 0.49) and 'up' words (-1.05 ± 0.47).

Due to these differences in valence and dominance ratings, and the fact that repeated presentation of linguistic stimuli has been shown to inflate traditional 'matched by item' differences to a significant level and therefore increase the chance of Type I error (Scheepers, 2014), we included these (and other) factors as covariates in relevant analyses by using a principal component analysis (outlined below). We chose to use principal components rather than include all variables as covariates in our analyses for two reasons. Firstly, the principal component analysis allows for fewer variables to be considered as covariates in the experimental analyses and secondly, it makes the estimation of their influence easier as each component is orthogonal to one another, thus avoiding colinearity.

Finally, all words as well as 120 non-word fillers, were recorded as separate sound files using a computer generated male British-English voice ('Brian', implemented in IVONA Reader software; http://www.ivona.com/en/reader). We chose to use an artificial voice for two reasons. Firstly, *IVONA Reader* allows for control over variables that are otherwise difficult to control in human voices (i.e. word stress and intonation), thus each word recording had a steady tone with no rising or falling intonation and avoided any potentially confounding effects arising from the use of a human voice. Secondly, we are interested in whether an artificial voice can impact on cognitive processing in a similar way to a human voice.

Spoken durations for 'up' words (M = 708 ms, SD = 129 ms), 'down' words (M = 721 ms, SD = 160 ms), and 'neutral' words (M = 710 ms, SD = 156 ms) did not differ reliably from one another (*p*s > .6). The 120 non-word fillers were pronounceable psuedowords constructed from novel composites of existing English phonemes (e.g. *asteng*). Each sound file had the volume normalised to -6dB (peak level) using Sound Studio (Felt Tip Software).

Principal Component Analysis

To further control for the confounding effects the linguistic stimuli may have on the dependent variable, we performed a principal component analysis (PCA) on all 120 word stimuli. Even though we matched the linguistic aspects of the stimuli by items, we also decided to include the factors as covariates in relevant analyses. Using a principal component analysis insures the resulting factor loadings are orthogonal, and hence make it easier for complex designs (like those presented in this thesis) to converge.

For each word, the length, BLP lexical decision response time (Keuleers et al., 2012), CELEX lemma frequency (Baayen et al., 1993), concreteness (our own ratings – see *Chosen Stimuli* Section above) and valence, arousal and dominance ratings (Warriner et al., 2013) were recorded. These were used in a PCA employing a *varimax with Kaiser normalisation* rotation method. Table 3.1 shows the factor loadings after varimax rotation for each variable and the three principal components.

Factors	**Principal Components**		
	Lexical Access	**Valence-Dominance**	**Arousal**
Frequency	.791*	.132	.042
Length	-.690*	-.007	.326
Lexical Decision RTs	-.684*	-.003	.051
Concreteness	.720*	.327	.104
Valence	.153	.882*	-.068
Dominance	.072	.892*	-.028
Arousal	-.050	-.076	.960*

Table 3.1. Results of the PCA showing the factor loadings after varimax rotation of the seven factors and three principal components. The factors assigned to each principal component is marked with an asterisk.

To this end, the three principal components were classified as *Lexical Access*, which consists of positive frequency, negative length, negative lexical decision response time, and positive concreteness loadings; *Valence-Dominance*, which consists of positive valence and dominance loadings; and *Arousal*, which consists of positive arousal loadings. The three principal component scores for each word were, where appropriate, entered as covariates in the experimental analyses reported throughout this thesis – more detail can be found in the Methods section of each experiment.

Chapter 4 Saccade Latency Effects

4.1 Facilitation vs. Inhibition: Baseline problems

A number of studies have shown conflicting results in terms of whether a perceptual simulation of a concept with prominent, automatically activated spatial associations facilitates or inhibits our ability to attend to a compatible or incompatible location (Bergen et al., 2007; Dudschig et al., 2012; Dudschig et al., 2013; Estes et al., 2008; Lachmair et al., 2011; Richardson et al., 2003). As outlined above (see Chapter 2.6), none of these studies included a neutral baseline condition; therefore reliably concluding facilitatory or inhibitory effects on processing is not possible. As an example, we shall review the results shown by Dudschig et al. (2013). To recap, participants were presented with 39 ‘up’ associated German nouns (e.g. ‘sun’), 39 ‘down’ related nouns (e.g. ‘mouse’) and 78 non-words. Participants’ task was to perform a lexical decision by looking to a response target (a small circle) at the top or bottom of the screen confirming whether they had seen a word or non-word – target location was counterbalanced between experimental blocks. From the results, reproduced in Figure 4.1, Dudschig et al. (2013) conclude that a concept’s typical location in the world facilitates the launching of a saccade to a compatible location.

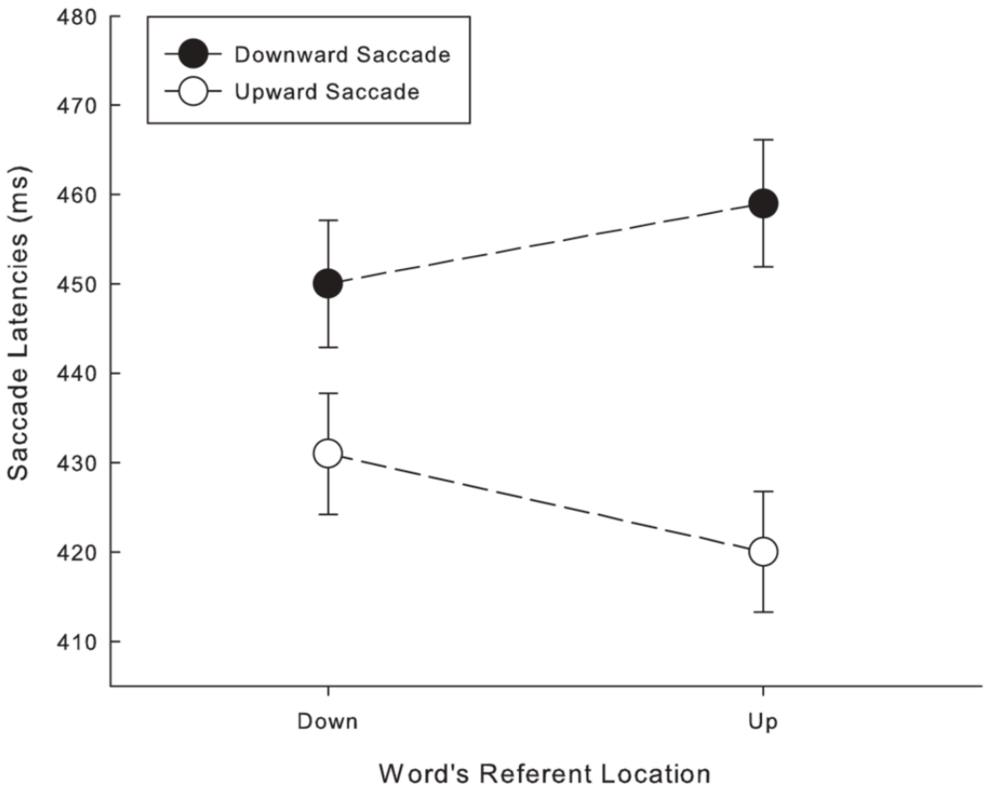


Figure 4.1. Results from Dudschig et al. (2013). Without a baseline condition, facilitatory or inhibitory effects cannot be reliably concluded.

However, if we assume that baseline responses – that is, a lexical decision to a concept not associated with vertical space – is on par with the 'compatible' responses, it becomes clear that the results could show an inhibitory effect on processing. For instance, reading the word "sun" interferes with the mechanisms involved in directing attention downwards, resulting in longer response times.

4.2 Experiment 1

In order to clearly identify whether the perceptual simulation of a concept facilitates or inhibits attentional responses to (in)compatible locations, a neutral baseline condition was introduced. Experiment 1 uses an eye-movement activated lexical decision task similar to that used by Dudschig et al. (2013); however along with the 40 upward associated and 40 downward associated words, the design includes 40 neutral words – these are words that have no vertical spatial association (see Chapter 3 for a review of all linguistic materials). Further differences between Experiment 1 and Dudschig et al. (2013) include using English words rather than German, counterbalancing response location between participants to avoid demand characteristics (Dudschig and colleagues counterbalanced within participants, but between experimental blocks), only presenting the 120 words and 120 non-words once (as opposed to four times) and presenting words auditorily rather than visually in order to avoid any reading confounds on eye movements (i.e. reading left to right could affect vertical saccadic responses depending on word length, frequency and other linguistic confounds). Finally, the norming procedure (see Chapter 3) showed that adjectives and verbs imply vertical spatial locations in a similar vein to nouns hence Experiment 1 uses adjectives, verbs and nouns.

4.2.1 Method

Participants

Thirty individuals (22 Female; M=24.2 years) from the University of Glasgow participated in the study, each receiving £4 or course credits; none of these participated in the previously described norming study. All participants had either normal or corrected-to-normal vision and were monolingual native English speakers.

Materials

The 120 critical words and 120 matched non-words outlined in the *Linguistic Stimuli* section in Chapter 3 were used in this experiment.

Apparatus

The stimuli were presented on a 21-inch CRT monitor (19.9-inch viewable display) of a DELL Optiplex GX 720 desktop computer with a display refresh rate of 150 Hz. Chin and forehead rests, positioned at a distance of 70 cm from the screen, were used to minimise head movements. Participants' eye movements were continuously monitored using a desk-mounted SR Research EyeLink 1000 eye-tracker, sampling at 1000 Hz. Although viewing was binocular, only the dominant eye was tracked, as established by a variation of the Miles test (Miles, 1930; Roth, Lora, & Heilman, 2002). Stimulus presentation and data collection were controlled using Experiment Builder software (SR Research).

Procedure

Each participant was presented with 240 auditory stimuli (120 words and 120 non-words, outlined in Section 3.1) in an individually determined random order. Prior to the experimental trials, participants undertook 5 practice trials to ensure they understood the task. As shown in Figure 4.2, each trial began with the presentation of a central fixation cross for drift correction. While the participant kept looking at the cross and 150 ms after drift correction, the sound file was played via headphones; at the onset of the sound file, a green and a red square appeared on the screen. Each square measured 10 × 10 screen pixels and appeared 200 pixels (4.70º of visual angle) above and below the central fixation cross, respectively. The participant's task was to decide, as quickly and accurately as possible, whether what they had just heard was an actual English word or not by looking at either the green square (if they thought they heard a word) or the red square (if they thought they heard a non-word). The location of the red and green square was counterbalanced across participants; 15 participants had the red square at the top and the green square at the bottom (and vice versa for the remaining 15 participants) for all 240 trials. This between-subject manipulation lowered the chances of participants figuring out the purpose of the study. Each trial terminated when a fixation was detected in one of the target areas (dashed rectangles in Figure 4.2), or after a timeout of 3000 ms, respectively. The target areas for the trial-terminating gaze trigger were defined as the inside edge of the coloured square to the top or bottom edge of the screen (190 pixels, 4.45º) and were 800 pixels (25.59º) wide. Before the first trial and after every 40 subsequent trials, the eye-tracker was recalibrated and validated using a 9-point fixation procedure – only 'Good' validations were accepted ('Good' validations are determined as having an average gaze error <0.5º and a maximum error <1.0º). An experimental session lasted approximately 40 minutes.

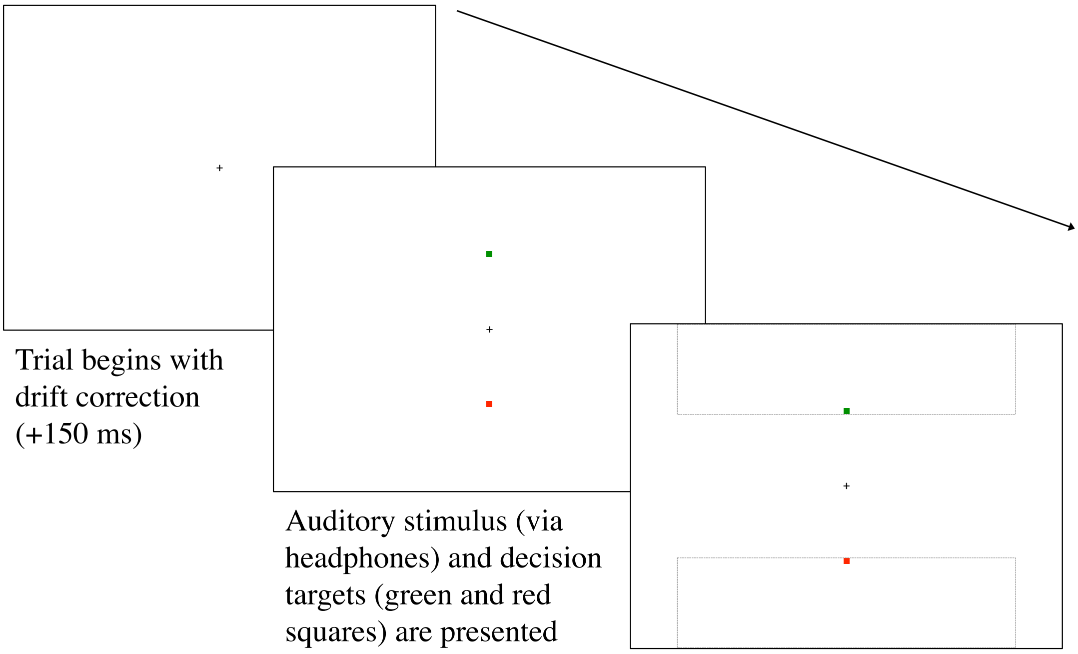


Figure 4.2. Schematic of a single trial. The green (for 'yes') and red (for 'no') squares represent the targets for lexical decision; their vertical positioning was counterbalanced across participants.

Data Analysis

Lexical decision accuracy was greater than 95% in each condition. Only the critical word trials were considered for analysis. The dependent variable of interest was the *saccade launch latency for correct 'word' decisions* (saccade towards green square), measured from the *offset* of the auditory word presentation until the eye started moving away from the central fixation cross (as determined by saccadic acceleration and velocity thresholds). Trials that contained eye-blinks, multiple saccades, off-target saccades (not landing within 100 pixels of the green square), or terminating after the 3000 ms timeout, were excluded from analysis (affecting less than 10% of the critical trials). Saccade launch latency outliers of more than 2.5 *SD*s away from the mean of a given subject × condition combination were also removed (affecting less than 3% of the data).

Inferential analyses were based on Generalised Estimating Equations (GEE; Hanley, Negassa, & Forrester, 2003; Hardin & Hilbe, 2003). GEE was favoured over Generalised Linear Mixed-Effects Models (GLMEM) due to the increased likelihood of convergence in the complex designs. Indeed, test runs with GLMEM suggested that the latter was likely to encounter convergence problems when a design-appropriate 'maximal' random effect structure was used (Barr, Levy, Scheepers, & Tily, 2013). Initial analyses conducted using GLMEM had not converged after a 36-hour period iterating, hence for practical purposes we used GEE. The major limitation of GEE (compared with GLMEM) is the inability to simultaneously compute by-subject and by-item analyses. In this sense, the separate analyses of by-subject and by-item effects are more comparable to F1 and F2 values in ANOVA analyses. However, GEE also takes into account the random effects of participants and items.

Since saccade launch latencies (or RT distributions in general) tend to be positively skewed, we implemented a Gamma regression approach by using a *Gamma* distribution and *Log* link function in the GEE model specifications. The *Log* link required all data points to be positive, which was not always the case because perceivers occasionally started to move their eyes before the end of the word (recall that the latencies were determined relative to word offset). To deal with this, we added a 700 ms constant to each data point before performing the inferential analyses. In the descriptive means reported below, this constant has been removed.

Two types of analyses were performed. In the *by-subject* analysis, word direction (‘up’, ‘neutral’, ‘down’) was entered as within-subjects factor and saccade direction (‘upwards’, ‘downwards’) as between-subjects factor. In the *by-item* analysis, word direction was between- and saccade direction within-items. To account for imbalances in item-specific control variables, the three principal components identified in the principal component analysis (Section 3.1, above) were entered as additional within-subjects / between-items covariates[4]. Analyses without these covariates yielded nearly identical results. The latter will not be reported in detail but see the grey lines in Figure 4.3 for comparison. All analyses assumed an exchangeable covariance matrix for repeated measurements.

4.2.2 Results

Table 4.1 shows Generalized Score Chi-Square statistics from the Gamma regression analyses in GEE. A first point to note is that two of the control covariates (principal components) had a significant overall influence on saccade launch latency: Lexical access was associated with a positive coefficient (+.045 ± .003 *SE* by subjects; +.043 ± .012 *SE* by items) and arousal with a negative coefficient (−.033 ± .003 *SE* by subjects; −.033 ± .010 *SE* by items), meaning that greater lexical difficulty led to an increase, and greater arousal to a decrease in saccade launch latency.

[4] In this experiment, and throughout this thesis, the covariates were entered as main effect terms only. The purpose was to neutralise by-item imbalances in these control predictors across word direction conditions.

Effect		**By Subjects**		**By Items**	
	df	**GS χ^2**	***P***	**GS χ^2**	***P***
Word Direction (W)	2	13.70	< .01	3.19	0.20
Saccade Direction (S)	1	2.03	0.15	48.33	< .001
W × S Interaction	2	12.53	< .01	26.90	< .001
Lexical Access*	1	26.46	< .001	9.84	< .01
Valence-Dominance*	1	.75	0.39	.020	0.89
Arousal*	1	26.17	< .001	8.74	< .01

Table 4.1. Inferential results (Generalized Score Chi-Squares, degrees of freedom, p-values) from the Gamma regression analyses in GEE. Control predictors (covariates) are marked with an asterisk.

With regards to our experimental manipulations, the main effect of word direction (ca. 25 ms higher saccade launch latencies in the ‘neutral’ condition compared to the other word direction conditions) was significant within-subjects but not between-items (presumably due to reduced power in the latter case). Likewise, the main effect of saccade direction (ca. 50 ms higher saccade launch latencies for downward than for upward saccades overall) was significant within-items but not between-subjects (again, suggesting reduced power for the *between* factor). Note that previous research (including Dudschig et al., 2013; Goldring & Fischer, 1997; Miles, 1936) has shown a similar general disadvantage for downward saccades.

Most crucially, there was a significant word direction × saccade direction interaction in both the by-subject and the by-item analysis. Figure 4.3. shows the covariate-adjusted by-subject means and *SE*s per condition (for comparison, non-adjusted means and *SE*s are shown in grey). 95% CIs for simple effects showed that upward saccades were launched more quickly upon hearing ‘up’ words like *moon* than upon hearing ‘neutral’ words like *letter* (by-subject contrast: 57 ± 10 ms; by-items: 61 ± 23 ms); the comparison between ‘down’ words like *sewer* and ‘neutral’ words like *letter* was not significant (by-subjects: 2 ± 8 ms; by-items: 6 ± 26 ms). Conversely, downward saccades were launched quicker after ‘down’ words like *sewer* compared to ‘neutral’ words like *letter* (by-subjects: 35 ± 17 ms; by-items: 45 ± 26 ms), whereas the contrast between ‘up’ words like *moon* and ‘neutral’ words like *letter* was not significant (by-subjects: 15 ± 10 ms; by-items: 17 ± 22 ms).

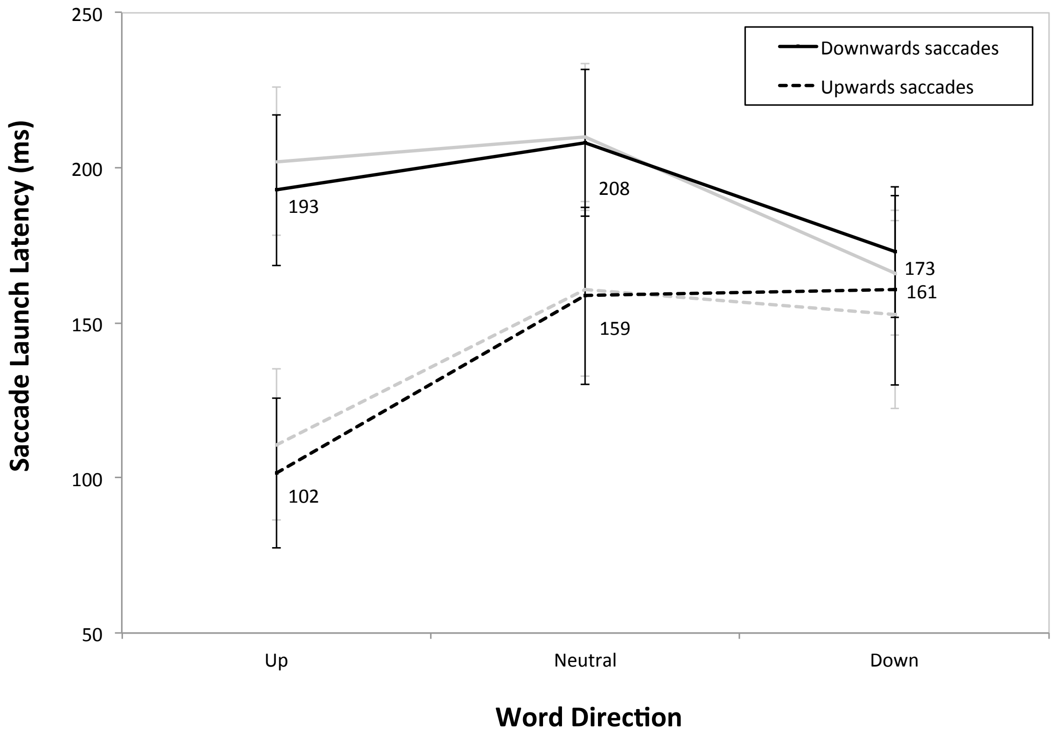


Figure 4.3. By-subject means and SEs per condition. The black lines show covariate-adjusted data and the grey lines show data from an analysis that did not include the covariates.

4.2.3 Discussion

Using an eye-movement activated lexical decision task, the present experiment investigated how perceptual simulations of spatial location activated by words affect launch latencies for saccades towards compatible or incompatible spatial locations. The first point of note is the main effect of saccade direction, with saccades being launched to targets at the top of the display significantly quicker than those at the bottom of the display. Since the 1930s, eye-tracking literature has shown that upward saccades are launched consistently quicker than downward saccades (e.g. Dudschig et al., 2013; Goldring & Fischer, 1997; Miles, 1936); and recently it has been shown that pre-saccade fixations are significantly shorter before upward (vs. downward) saccades (Greene, Brown, & Dauphin, 2014). To this end, the upward advantage shown in our results seems to reflect the general bias of the visuo-motor system rather than a direct impact of our experimental manipulations. Indeed, this effect is observed throughout the different experiments presented in this thesis. However, a disadvantage for downward responses is also apparent in hand movement tasks. For example Lachmair et al. (2011), outlined above (Section 2.6), used a lexical decision task requiring upward or downward hand movement responses and showed significantly faster responses for upward compared with downward movements. A similar upward response advantage for hand movements in a linguistic task employing a

Stroop-like methodology is shown in Dudschig and Kaup (in press). Furthermore, an advantage for upward (vs. downward) responses is present in literature investigating how reading direction affects mental number line representations, otherwise referred to as the 'vertical SNARC effect' (e.g. Hung, Hung, Tzeng, & Wu, 2008; Ito & Hatta, 2004). Taken together, it seems that there is a general motoric bias for upward motor movements that is apparent across a number of different paradigms.

With regards to our manipulations, the results clearly showed that 'direction' words facilitate saccades towards compatible locations in the vertical dimension, but crucially, do not inhibit saccades towards incompatible locations. Importantly, saccadic response latencies to words in the neutral (baseline) condition did not significantly differ from those in the spatially incompatible conditions.

The results from Experiment 1 build on those presented by Dudschig et al. (2013) by confirming the facilitatory role that perceptual simulations of linguistic stimuli have on visuo-spatial processing. However, the results do not support the perceptual-featural overlap hypothesis proposed by Estes et al. (2008). To recap, the perceptual-featural overlap hypothesis states that a lack of featural overlap between the perceptual simulation of the cued concept and the target object would lead to an interference effect. Thus in Experiment 1, one would expect that the perceptual simulations evoked by the 'up' and 'down' direction words would lead to slower identification of a target in a compatible location. Apparently, this was not the case – indeed, despite the target (a small green square) having no featural overlap with the cued concepts the results still clearly show a facilitatory effect on processing. This is discussed further in the General Discussion following Experiment 2 (Section 4.4).

By way of further exploring the pattern of results shown in Experiment 1, Experiment 2 set out to investigate whether a different pattern of results would be seen with targets presented in the horizontal dimension. If the facilitatory pattern of results shown in Experiment 1 was due to perceptual simulations of spatial location affecting visuo-motor response mechanisms, then saccading to a horizontal location incompatible with the activated upward or downward perceptual simulation should result in response times no different to the neutral baseline. That is, hearing an 'up' word like "moon" would activate an upward perceptual simulation and hence saccading to a target to the left or right of the display would not be facilitated.

4.3 Experiment 2

The second experiment was identical to the first except that the target locations were now presented in the horizontal dimension. If the pattern of results shown in Experiment 1 was due to perceptual simulations of the cued concepts facilitating responses to compatible locations, it should be expected that there would be no significant differences in saccadic launch latency times to horizontal targets.

4.3.1 Method

Participants

Thirty individuals (22 Female; M=22.8 years) from the University of Glasgow participated in the study, each receiving £4 or course credits; none of these participated in Experiment 1. All participants had either normal or corrected-to-normal vision and were native monolingual English speakers.

Materials

The 120 direction words and 120 matched non-words outlined in the *Linguistic Stimuli* section above were used in this experiment.

Apparatus and Procedure

The apparatus and procedure was identical to that of Experiment 1, except the location of the target squares, which were presented in the horizontal dimension. Each square appeared 200 pixels (6.50º of visual angle) left and right of the central fixation cross, respectively. The target areas for the trial-terminating gaze trigger were defined as the inside edge of the coloured square to the left or right edge of the screen (184 pixels, 5.98º) and spanned the entire height of the screen (768 pixels, 17.86º).

Data Analysis

Lexical decision accuracy was greater than 95% in each condition. Data were analysed in an identical way to Experiment 1. To recap, only the critical word trials were considered for analysis. The dependent variable of interest was the *saccade launch latency for correct 'word' decisions* (saccade towards green square), measured from the *offset* of the auditory word presentation until the eye started moving away from the central fixation cross (as determined by saccadic acceleration and velocity thresholds). Trials that contained eye-blinks, multiple saccades, off-target saccades (not landing within 100 pixels of the green square), or terminating after the 3000 ms timeout, were excluded from analysis (affecting less than 8% of the critical trials). Saccade launch latency outliers of more than 2.5 *SD*s away from the mean of a given subject × condition combination were also removed (affecting ca. 2.5% of the data).

Inferential analyses were based on GEE, which was again favoured over GLMEM. In order to accurately model skewed RT data (see Experiment 1 for details), we implemented a Gamma regression approach by using a *Gamma* distribution and *Log* link function in the GEE model specifications. The *Log* link required all data points to be positive, which was not always the case because perceivers occasionally started to move their eyes before the end of the word (recall that the latencies were determined relative to word offset). To deal with this, we added a 500 ms constant to each data point before performing the inferential analyses. In the descriptive means reported below, this constant has been removed.

Two types of analyses were performed. In the *by-subject* analysis, word direction (‘up’, ‘neutral’, ‘down’) was entered as within-subjects factor and saccade direction (‘leftwards’, ‘rightwards’) as between-subjects factor. In the *by-item* analysis, word direction was between- and saccade direction within-items. To account for imbalances in item-specific control variables, the three principal components identified in the principal component analysis section above were entered as additional within-subjects / between-items covariates. Analyses without these covariates yielded nearly identical results and will not be reported in detail, but see the grey lines in Figure 4.4 for comparison. All analyses assumed an exchangeable covariance matrix for repeated measurements.

4.3.2 Results

Table 4.2. shows Generalized Score Chi-Square statistics from the Gamma regression analyses in GEE. A first point to note is that, similar to the vertical dimension (Experiment 1) two of the control covariates (principal components) had a significant overall influence on saccade launch latency: Lexical access was associated with a positive coefficient (+.047 ± .005 *SE* by subjects; +.048 ± .014 *SE* by items) and arousal with a negative coefficient (−.038 ± .005 *SE* by subjects; −.036 ± .014 *SE* by items), meaning that greater lexical difficulty led to an increase, and greater arousal to a decrease in saccade launch latency.

Effect		By Subjects		By Items	
	df	GS χ^2	*P*	GS χ^2	*P*
Word Direction (W)	2	11.93	< .01	1.97	0.37
Saccade Direction (S)	1	0.91	0.34	10.62	< .01
W × S Interaction	2	1.12	0.57	0.53	0.77
Lexical Access*	1	21.52	< .001	8.66	< .01
Valence-Dominance*	1	0.51	0.48	0.096	0.76
Arousal*	1	19.46	< .001	6.38	< .05

Table 4.2. Inferential results (Generalized Score Chi-Squares, degrees of freedom, p-values) from the Gamma regression analyses in GEE. Control predictors (covariates) are marked with an asterisk.

With regards to our experimental manipulations, the main effect of word direction (ca. 19 ms higher saccade launch latencies in the 'neutral' condition compared to the other word direction conditions) was significant within-subjects but not between-items (presumably due to reduced power in the latter case). Likewise, the main effect of saccade direction (ca. 24 ms higher saccade launch latencies for leftward than for rightward saccades overall) was significant within-items but not between-subjects (again, suggesting reduced power for the *between* factor).

Unlike Experiment 1, there was no significant word direction × saccade direction interaction in either the by-subject or the by-item analysis. Figure 4.4 shows the covariate-adjusted by-subject means and *SE*s per condition (for comparison, non-adjusted means and *SE*s are shown in grey).

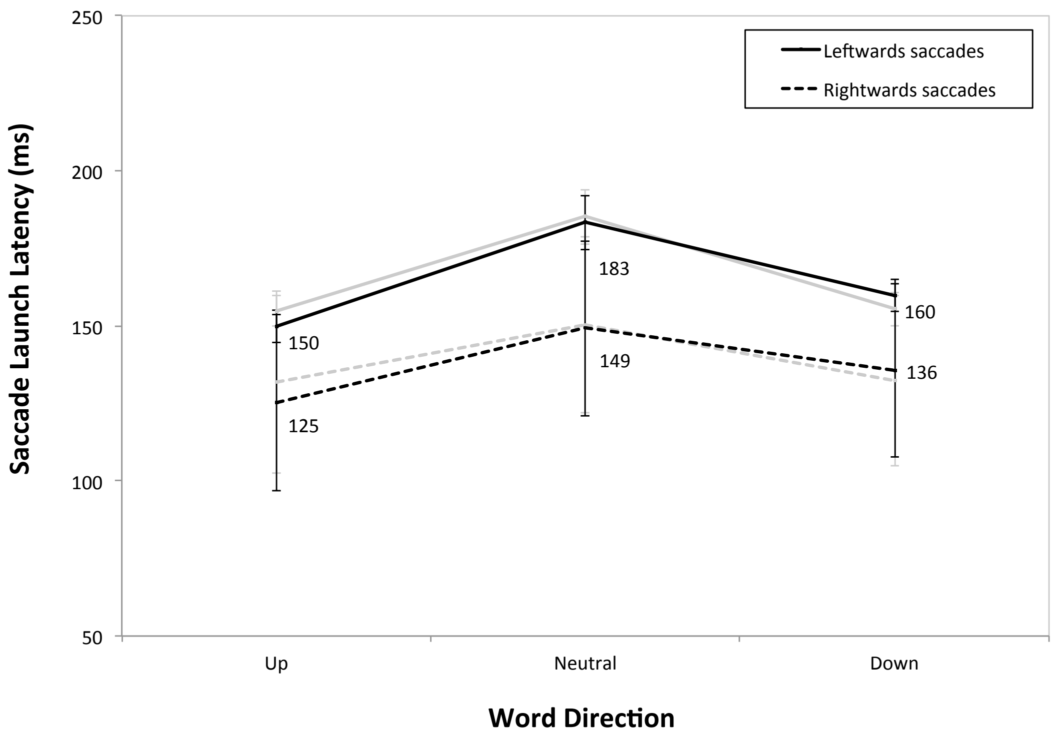


Figure 4.4. By-subject means and SEs per condition. The black lines show covariate-adjusted data and the grey lines show data from an analysis that did not include the covariates.

4.3.3 Discussion

Experiment 2 replicated the procedure of Experiment 1, but used horizontal targets rather than vertical targets. The aim of Experiment 2 was to act as a control experiment by reinforcing the facilitatory perceptual simulation effects shown in Experiment 1. If the facilitatory results seen in Experiment 1 were due to the perceptual simulation of a concept activating a spatial representation, then responding to a location that does not overlap with the active spatial representation should show no difference compared to baseline responses. The results from Experiment 2 essentially confirmed this prediction.

However, the present results show that responses to a horizontal target location were significantly slower if they were preceded by a 'neutral' word compared to an 'up' or 'down' direction word. This was true for both left and right target locations. So the question here is why does responding to a neutral word take longer than responding to a direction word? This is potentially due to the norming procedure employed to select the direction words. Recall that during norming participants were required to rate words on their verticality (see Chapter 3); those falling in the middle of the scale were deemed to be neither 'up' words nor 'down' words and were chosen as our 'neutral' words. Upon closer inspection, the neutral word category contains words that could be deemed as horizontally associated (e.g. *equator*, *wide*, *belt*, *waistline*). As such, perceptual simulations of the neutral words may have affected responses to left and right locations. However this does not explain the results from Experiment 1. Recall that those results also showed a significant main effect of word direction, with saccades launched after 'neutral' words being significantly longer (ca. 25 ms) than those launched after 'up' or 'down' words. Therefore, we conclude that there is a negative bias towards the 'neutral' words not controlled for by the three covariates (principal components). Given this main effect of word direction, we should consider the implications it has on determining the pattern of results. That is, as the neutral words act as a baseline condition, could the negative bias lead to a misinterpretation regarding whether results show a facilitatory or inhibitory effect? In Experiment 1, one can explain the significant main effect of word direction (a 25 ms disadvantage for neutral words) as a result of the facilitatory responses. Both 'up' and 'down' words are followed by two possible response locations – one is compatible and the other incompatible; contrast this to 'neutral' words which, due to their lack of spatial association, are followed by neither compatible nor incompatible response locations. As approximately half of all 'up' and 'down' word responses were compatible, the facilitatory effect shown in RTs shortened the overall mean RT for the relevant word condition. Given that there is no 'compatible' neutral location, there is no opportunity for the overall mean RT to be shortened in a similar way. Due to this, the overall mean RTs for 'up' and 'down' words are significantly shorter than the overall mean RT for 'neutral' words, which manifests itself as a main effect of word direction. In contrast, Experiment 2 has no compatible or incompatible response locations for either 'up', 'down' or 'neutral' words, thus one would not expect the main effect of word condition yet, it still occurs. In this case, one must consider that the 'neutral' words may perhaps induce a negative attentional bias such that an aspect not controlled for by our principal components leads to longer response times. In hindsight, we could have undertaken an experiment to investigate whether such an attentional bias could affect eye-movement responses; for example, a go/no-go lexical decision task using all 240 word and nonword stimuli in which participants responded to the 120 words (i.e. 'go' trials) and did not respond to the 120 nonwords (i.e. 'no go' trials). This would allow for a simple response time measure providing insight on whether the neutral words do indeed show a negative attentional bias. If present, we would need to be sure that the results to Experiment 1 (and further experiments) do indeed show a facilitatory effect by accounting for the aforementioned attentional bias during the analyses.

Another point to consider is the main effect of saccade direction (ca. 24 ms advantage for right saccades over left saccades). Whilst it is only speculative, this rightwards advantage could potentially be due to a motoric bias in English monolinguals such that eye movements from left to right reflect reading direction. Alternatively, the difference in saccade launch latency times may be a positivity effect. Research outlined above (e.g. Casasanto, 2009) shows that rightwards is often associated with positive valence and correctness, hence responding to the 'yes' or 'word' target when making a lexical decision would be quicker when it appears on the right side of the screen. However, this seems to be related to dominant sides, or handedness with left-handers showing the opposite pattern (Casasanto & Chrysikou, 2011). Given that we did not collect data on handedness, this point remains unclear. Further investigation beyond the scope of the current thesis is required to investigate this rightward advantage, therefore a comprehensive conclusion is not possible at present.

What is clear however is that there were no significant differences between saccadic launch latencies to 'up' and 'down' words. This suggests that overlapping spatial associations between the mentally simulated concept and the response location explain the differences shown in Experiment 1. This is discussed in further detail below.

4.4 General Discussion of Experiments 1 and 2

Two experiments aimed to investigate further the conflict in the current literature whereby perceptual simulations of spatial concepts have been shown to both facilitate and inhibit response behaviour. The basis of both experiments was to test the perceptual-featural overlap hypothesis proposed by Estes et al. (2008), and to confirm whether the effects concluded by Dudschig et al. (2013) were truly facilitatory or inhibitory, in nature.

Using an eye-movement activated lexical decision task, Experiment 1 showed that responding to an on-screen location is facilitated when preceded by a word implying a compatible spatial location. In a similar task, Experiment 2 showed no difference in response times to ‘up’ and ‘down’ words when responding to horizontal on-screen locations. The question is why do compatible direction words facilitate saccades? Previously, facilitation was explained as the result of featural overlap between the cue and target – hence, a lack of featural overlap led to inhibitory effects (Estes et al., 2008). Unlike previous experiments showing facilitation (e.g. Zwaan et al., 2004; Zwaan et al., 2002), the visual form and shape of the targets in the current experiments had no overlapping features with the cue; however, there is at least one aspect of the target that might overlap with the cue word, namely the vertical direction that leads to the target location. From a *Grounded Cognition* point of view, experiential traces associated with a concept become reactivated upon later presentation (Barsalou, 1999, 2008). Hearing the word “moon” would reactivate all experiential traces of the related concept (including perhaps, the motor action of looking up to see the moon), and therefore, compatible visuo-motor responses (saccading upwards) should be facilitated. By contrast, if a given word’s vertical association is incompatible with the direction of the required saccadic response, then its influence on saccade launch latency is no different from that of a vertically ‘neutral’ word like “letter”. This suggests that experiential traces associated with words would not interfere with, or inhibit, incompatible saccadic responses.

What about the inhibitory effects shown by some of the studies discussed earlier (see Section 2.6: Estes et al., 2008; Richardson et al., 2003)? Note that these studies did not record eye movements, making it difficult to compare them with the present results. However, these studies also used a different paradigm; rather than asking participants to make a lexical decision, they were required to identify a target in a vertical on-screen location after hearing or reading ‘up’ or ‘down’ words. If the spatial location of a concept is a prominent enough feature to become reactivated by an abstract target in a compatible location, the aforementioned studies should also have shown a facilitatory effect. With this in mind, we aimed to investigate the inhibitory effects shown by studies using a target detection paradigm (without lexical decision), but with a similar procedure to Experiments 1 and 2.

4.5 Experiment 3

In Experiment 3 we used a method that has previously shown inhibitory effects when researching spatial locations of words – target detection (Estes et al., 2008; Richardson et al., 2003). A different set of participants than those from Experiments 1 and 2 were presented with the same spoken words (but not the non-words) and 120 matched filler words (see Materials, Section 4.5.1). Each word was followed by a green square presented either above, below (in the same locations as Experiment 1), left or right (in the same locations as Experiment 2) of the centre of the screen. Participants were required to listen to the word and fixate the target once it appeared. Once again participants responded using eye movements, however unlike the previous experiments, only one target was present per trial. Crucially, by using the same critical words as Experiments 1 and 2 we formed a baseline condition in a similar way, allowing facilitatory or inhibitory effects of perceptual-spatial word associations to manifest themselves in a way comparable to the previous results. If the vertical association of a concept is a prominent enough feature to overlap with the vertical location of the target, we would expect facilitatory visuo-motor responses (quicker saccades launched towards a compatible target vs. baseline condition) similar to those found in Experiment 1. Contrary to this, if the vertical association of a concept and the target's location are not prominent enough overlapping features (or, at least not enough to affect visuo-motor responses) then we would expect no difference in saccade launch times compared to the baseline condition. Finally, the inhibitory results discussed earlier may simply be due to the differences between the experimental paradigms (lexical decision vs. target detection) rather than perceptual-featural overlap between the cue and target. In this case, we might expect inhibitory effects (i.e. slower saccade launch times to compatible locations vs. baseline condition).

4.5.1 Method

Participants

Sixty individuals (38 Female, M=24.1 years) from the University of Glasgow participated in the study, each receiving £4 or course credits; none of these participated in any of the previous experiments. All participants had either normal or corrected-to-normal vision and were native monolingual English speakers.

Materials and Apparatus

The same 120 words from Experiment 1 were used alongside 120 filler words. The filler words consisted of 120 English words matched in length and frequency (Baayen et al., 1993) to the 120 critical words. All filler words are listed in the Appendix. The 240 words were presented using the same apparatus and setup as Experiments 1 and 2.

Procedure

Each participant was presented with 240 auditory stimuli (120 critical and 120 fillers) in an individually determined random order. Prior to the experimental trials, participants undertook 4 practice trials to ensure they understood the task. As shown in Figure 4.5, each trial began with the presentation of a central fixation cross for drift correction. While the participant kept looking at the cross, and 150 ms after drift correction, the sound file was played via headphones. At the offset of the word, a green square appeared on the screen. Each square measured 10 x 10 screen pixels and appeared 200 pixels either above, below, left or right of the central fixation cross; these equate to 4.70º of vertical and 6.50º of horizontal visual angle. The participant's task was to look towards the green square as quickly as possible. The vertical location of the green square was counterbalanced across critical words, and participants. Therefore, each participant saw the green square in each vertical location 60 times. The horizontal target locations were only ever preceded by a filler word. Running different experimental lists between participants ensured that each critical word was followed by a square in each vertical position. The vertical target locations for each word were randomized across 12 different lists thus minimizing intra-list effects. Crucially, the position of the green square could not be predicted by the word therefore lowering the chances of participants figuring out the purpose of the study. Each trial terminated when a fixation was detected in the target area surrounding the green square (dashed rectangle in Figure 4.5), or after 3000ms, respectively. Each target area for the trial-terminating gaze trigger was defined as the inside edge of the green square to the outer edge of the screen and were 800 pixels (25.59º) wide for up and down targets, and 768 pixels (17.86º, the entire screen height) high for left and right targets.

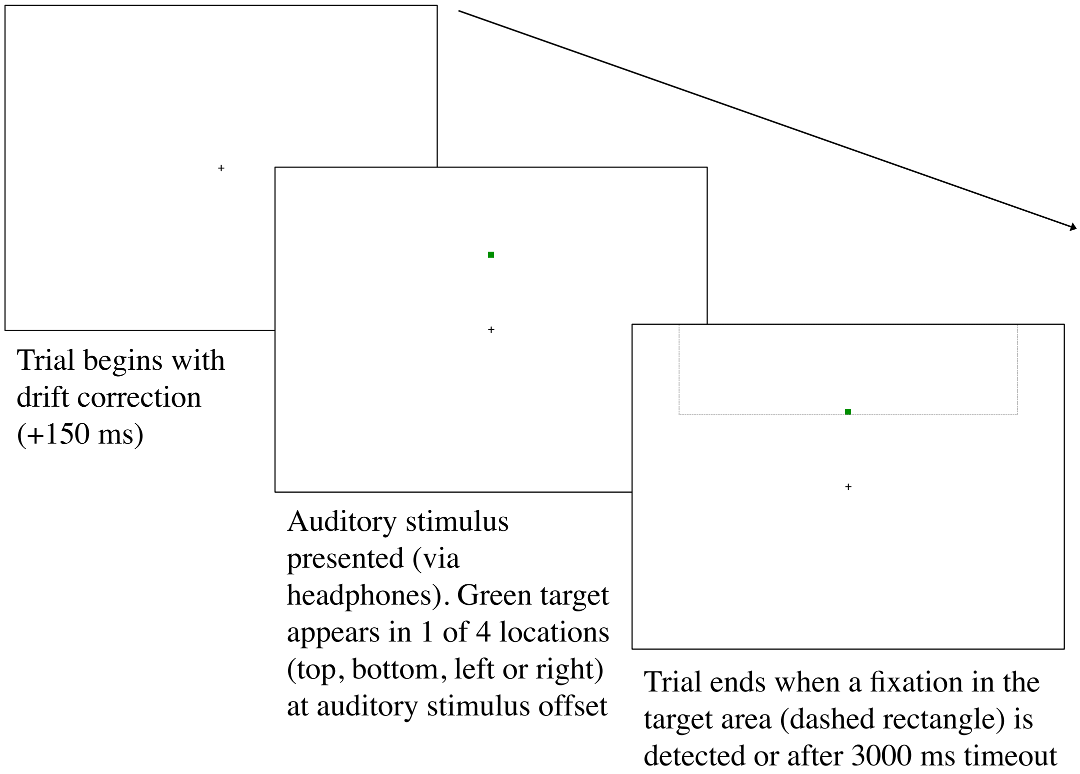


Figure 4.5. Schematic of a single trial. The green square appeared at either the top (as shown), bottom, left or right of the screen. The position was counterbalanced across trials and participants.

Before the first trial and after every 40 subsequent trials, the eye-tracker was calibrated and validated using a 9-point fixation procedure (only 'Good' validations were acceptable, see Experiment 1 for details). After every 40 trials, participants were presented with a word recognition questionnaire consisting of 10 words, some of which had been in the preceding 40 trials, and some which hadn't. The participant's task was to mark whether they thought the word had or hadn't appeared. This secondary task was used to further disguise the purpose of the study and ensured participants listened to (and processed) the words, rather than just launching saccades towards the targets.

Data Analysis

Accuracy in the secondary word recognition task was consistently greater than 72% (M = 83%) suggesting that participants attended to the word stimuli rather than just launched saccades towards the targets.

Target detection accuracy was greater than 99% in each condition. Only critical word trials (vertical target trials) were considered for analysis. The dependent variable of interest was the *saccade launch latency for looks to the target square*, measured from the *offset* of the auditory word presentation (which was also the onset of the target presentation) until the eye started moving away from the central fixation cross (as determined by saccadic acceleration and velocity thresholds). Trials that contained eye-blinks, multiple saccades, off-target saccades (not landing within 100 pixels of the green square), or terminating after the 3000 ms timeout, were excluded from analysis (affecting less than 5% of the critical trials). Saccade launch latency outliers of more than 2.5 *SD*s away from the mean of a given subject × condition combination were also removed (affecting less than 0.3% of the data).

Inferential analyses were based on GEE (see Experiment 1 for a description); implementing a Gamma regression approach by using a *Gamma* distribution and *Log* link function in the GEE model specifications.

Two types of analyses were performed. In the *by-subject* analysis, word direction ('up', 'neutral', 'down') and saccade direction ('upwards', 'downwards') were entered as within-subjects factors. In the *by-item* analysis, word direction was between- and saccade direction within-items. As the items were the same as the previous two experiments, the same three principal components identified in the principal component analysis (Section 3.1, above) were entered as additional within-subjects / between-items covariates. Analyses without these covariates yielded nearly identical results and will not be reported in detail, but see the grey lines in Figure 4.6 for comparison. All analyses assumed an exchangeable covariance matrix for repeated measurements.

4.5.2 Results

Table 4.3 shows Generalized Score Chi-Square statistics from the Gamma regression analyses in GEE. A first point to note is that, unlike for the previously presented experiments, none of the control covariates (principal components) had a significant overall influence on saccade launch latency.

Effect		**By Subjects**		**By Items**	
	df	**GS χ^2**	***P***	**GS χ^2**	***P***
Word Direction (W)	2	3.39	0.18	1.53	0.47
Saccade Direction (S)	1	29.46	< .001	68.86	< .001
W × S Interaction	2	14.49	< .01	12.66	< .01
Lexical Access*	1	0.35	0.56	0.12	0.73
Valence-Dominance*	1	0.08	0.79	1.09	0.30
Arousal*	1	1.32	0.25	1.83	0.18

Table 4.3. Inferential results (Generalized Score Chi-Squares, degrees of freedom, p-values) from the Gamma regression analyses in GEE. Control predictors (covariates) are marked with an asterisk.

With regards to our experimental manipulations, there was no main effect of word direction either within-subjects or between-items. However, the main effect of saccade direction (ca. 30 ms higher saccade launch latencies for downward than for upward saccades overall) was significant within-items and within-subjects. This disadvantage for downward saccades (presumably due to a general visuo-motor bias) is similar to that shown in Experiment 1 and previous research (e.g. Dudschig et al., 2013; Goldring & Fischer, 1997; Miles, 1936).

Most crucially, there was a significant word direction × saccade direction interaction in both the by-subject and the by-item analysis. Figure 4.6 shows the covariate-adjusted by-subject means and *SE*s per condition (for comparison, non-adjusted means and *SE*s are shown in grey). 95% CIs for simple effects showed that upward saccades were launched more *slowly* upon hearing 'down' words like *sewer* than upon hearing 'neutral' words like *letter* (by-subject contrast: 9 ± 5 ms; by-items: 7 ± 6 ms); the comparison between 'up' words like *moon* and 'neutral' words like *letter* was not significant (by-subjects: 1 ± 5 ms; by-items: 1 ± 7 ms). Downward saccades were launched more slowly after 'up' words like *moon* compared to 'neutral' words like *letter* (by-subjects: 6 ± 7 ms; by-items: 7 ± 9 ms, the difference only reached a significance level of $p = .06$), whereas the contrast between 'down' words like *sewer* and 'neutral' words like *letter* was not significant (by-subjects: 4 ± 8 ms; by-items: 4 ± 8 ms).

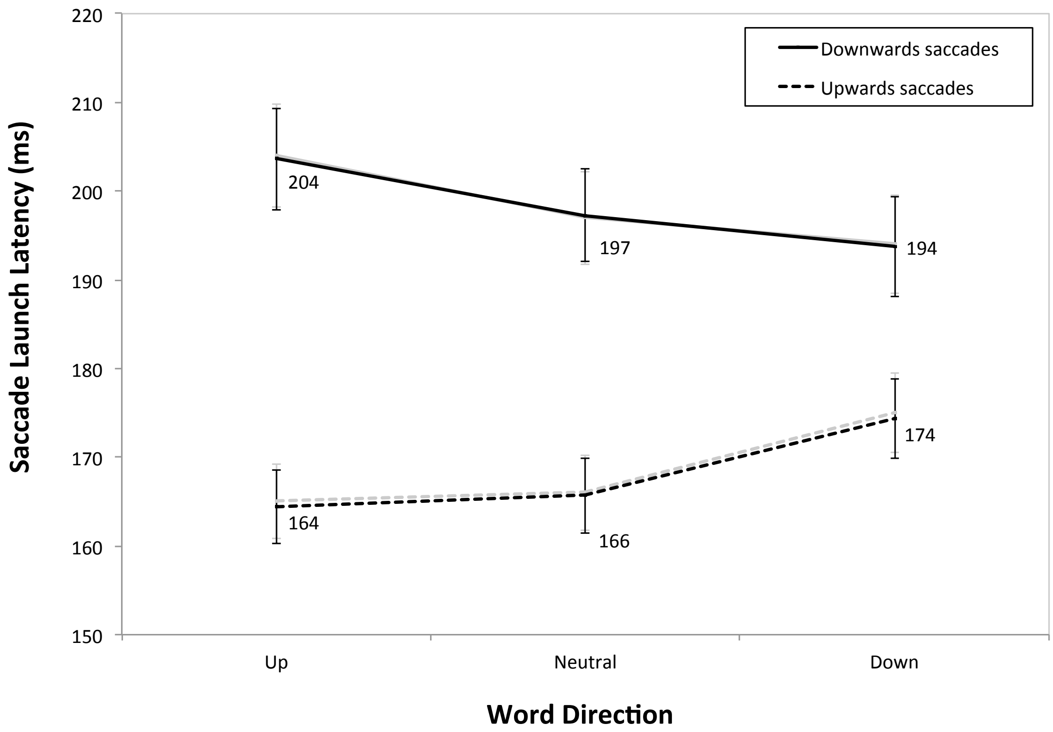


Figure 4.6. By-subject means and SEs per condition. The black lines show covariate-adjusted data and the (barely different) grey lines show data from an analysis that did not include the covariates.

4.5.3 Discussion

Experiment 3 used an eye movement activated target detection task to investigate whether a target's vertical location is a prominent enough feature to reactivate the perceptual-spatial representation previously activated by a word target and hence facilitate processing. The present results show that 'direction' words inhibit saccades to incompatible locations and, in contrast to Experiment 1, do not facilitate saccades to compatible locations. Note though, that these results are unlike previous target detection experiments that show an inhibitory effect of locating a target in a *compatible* location (Estes et al., 2008; Richardson et al., 2003). It is however, worth reiterating that the aforementioned studies did not record eye movements, making them difficult to compare to our results.

One notable outcome from the present experiment is the effective use of the neutral baseline condition. Consider the pattern of results if a baseline condition was not included (as is the case throughout the previously mentioned literature); saccades toward compatible target locations were launched significantly quicker than those to incompatible target locations. Thus without the baseline condition, the results would seemingly show a compatibility effect, which most likely (given our hypothesis) would have been incorrectly interpreted as a facilitatory effect. Given this, a key finding from the present experiment is that experimental designs without a baseline condition cannot reliably conclude a facilitatory or inhibitory effect on processing. This point becomes especially prominent in the current field whereby an ongoing debate surrounds the conflicting patterns of results (see Section 4.1, above). Experiment 3 shows (perhaps more so that Experiments 1 and 2) the necessity for a neutral baseline condition, and it is therefore imperative that future research in this field uses such a design as without it, results remain inconclusive and the 'facilitation versus inhibition' debate will never be concluded.

We are now posed with two questions. Firstly, why do the results from this experiment show inhibitory rather than facilitatory responses? And secondly, why are these results so different from the results of the vertical lexical decision task in Experiment 1? In response to the first question, we consider the perceptual-featural overlap account proposed by Estes et al. (2008).

It has been widely reported that the processing of a linguistic concept requires[5] a mental representation. The perceptual-featural overlap hypothesis reported by Estes and

[5] It is debated as to whether a mental representation is a requirement for understanding or is a peripheral event occurring as part of the understanding process. It is beyond the scope of the present thesis to discuss this in further detail; however we work on the assumption that representation occurs, regardless of whether it is required or not.

colleagues (2008), states that perception of a target is delayed if it shares few features with the mentally represented cue. In their results, the mental representation of the cue (e.g. *cowboy hat*) had few features that overlapped with the target (the letter *X*); hence the upward association activated by the cue interfered with detection of a target in the compatible location (i.e. the top of the screen). In contrast, one may expect the mental representation of 'cowboy hat' to facilitate the detection of a picture of a hat in a compatible location due to their overlapping features (Stanfield & Zwaan, 2001; Zwaan et al., 2004). Unfortunately there is no clarification as to how much of an overlap, or which features are required to facilitate detection of a target. Nonetheless, this theory does not explain the pattern of the present results. The facilitatory results shown in Experiment 1 could be accounted for if 'spatial location' were enough of an overlapping feature to facilitate responses to targets in compatible locations, however one would expect this pattern of results to be replicated in Experiment 3. Similarly, one would have expected a different pattern of results in the aforementioned research such that the upward association of 'cowboy hat' would have facilitated detection of a target in a compatible spatial location. Given that neither of these happened, it seems that the perceptual-featural overlap hypothesis does not provide a satisfactory explanation of our results. For that reason, we turn our discussion to the differences between our two tasks, and why they produce such contrasting results.

One major methodological difference between the lexical decision (Experiment 1) and target detection tasks (Experiment 3) is the time at which the target appears. The lexical decision experiments have both targets showing on the screen at the onset of the auditory stimuli such that a participant is able to (and should) respond as soon as is realistically possible after the lexical decision has been made. On the other hand, during the target detection procedure used in in Experiment 3 the target only appears at the offset of the auditory stimuli. Theoretically, in Experiment 1 an individual is unable to make a lexical decision until the word is heard and recognised; however a small number of saccades in both lexical decision tasks (Experiments 1 and 2) occurred before the auditory offset[6] – this is obviously not possible in the target detection task as the target does not appear until auditory offset. If a perceptual simulation of a concept occurs early, it could be the case that this facilitates earlier responses whereas a delay between the perceptual

[6] Further investigation showed that verbs with '-ing' endings occasionally showed negative latency times, suggesting that a lexical decision was made after the word stem. For example, "towering" led to a lexical decision after "tower-". Supplementary analyses removed all negative latency responses with the same pattern of results found, for this reason the negative latency results are included here rather than attempting to define an arbitrary cut-off point.

simulation occurring and a response being required may interfere with response behaviour. In order to test this assumption, Experiment 4 was designed with the intention of manipulating the stimulus onset asynchrony (SOA) such that earlier target presentation may lead to a facilitatory pattern of results whereas later target presentation may replicate the inhibitory pattern shown in Experiment 3.

In research investigating how a mental number line affects shifts in attention, stimulus onset time has been shown to be an important factor (Dodd, Van der Stigchel, Leghari, Fung, & Kingstone, 2008; Fischer et al., 2003). Recall that numbers seem to be represented on a left-to-right (for English speakers) mental number line such that smaller numbers are represented in leftward space, and larger numbers in rightward space. The SNARC effect, as it is known (see Section 2.5, above), shows that responding to a small (or large) number facilitates responses to the left (or right) of the participant. However, results show that this compatibility effect is not present at short SOAs (around 50-200 ms); it only becomes apparent after a short delay between the number cue and the target (300-500 ms) and fades with a longer delay to the extent that it is again no longer apparent after a delay of greater than 700 ms (Dodd et al., 2008; Fischer et al., 2003).

Whilst the spatial orienting results shown for the SNARC effect does not explain the pattern of results shown in Experiments 1 and 3, it does show that the time at which a compatible or incompatible spatial orientation response is required affects whether processing is facilitated or not, and hence it may be an important aspect of the task.

4.6 Experiment 4

In order to investigate whether the previously presented facilitatory and inhibitory effects shown in Experiments 1 and 3 are due to target onset times, Experiment 4 replicated the target detection procedure in Experiment 3 but manipulated the time at which the target appeared on the screen. The results from the lexical decision experiments (1 and 2) showed that participants occasionally responded before the offset of the auditory stimulus, thus it was hypothesised that early target presentation would lead to a facilitatory pattern of results with later target presentation leading to an inhibitory pattern of results. In order to allow participants to respond in a time similar to those shown in Experiment 1, targets should be presented on the screen prior to the offset of the auditory stimulus. Furthermore, later target presentation should allow for a replication of Experiment 3 (target presented at auditory stimulus offset), a target presented shortly after word offset and finally a target presented much later after word offset. To this end, four different SOA categories were defined (see below: Section 4.7.1 – Procedure).

4.6.1 Method

Participants

Sixty-one individuals (45 Female, M=21.5 years) from the University of Glasgow participated in the study, each receiving £4 or course credits; none of these participated in any of the previous experiments. All participants had either normal or corrected-to-normal vision and were native monolingual English speakers.

SOA Level	Mean (ms)	SD (ms)
70% of word duration	442	99
100% of word duration	632	98
130% of word duration	821	104
400 ms after offset	1032	107

Table 4.4. Mean and Standard Deviation of stimulus onset asynchrony times relative to word onset for each level of the SOA condition used in Experiment 4.

Materials and Apparatus

The same 120 words and 120 filler words from Experiment 3 were used. The 240 words were presented using the same apparatus and setup as Experiment 3.

Procedure

The procedure was identical to Experiment 3 apart from a difference in the time at which the target appeared on the screen. Four different stimulus onset asynchronies (SOAs) were employed, whereby each target appeared on the screen at 70%, 100% (as with Experiment 3) or 130% of the respective word's sound file duration or 400ms after word offset. Mean SOA times relative to word onset for each level are shown in Table 4.4. Of the 120 critical words, 30 appeared at each SOA, counterbalanced between subjects such that each word appeared after every SOA (this was also true of the filler words). SOAs were also randomised within each experimental list in order to create a further 8 lists (12 in total) thus minimising intra-list effects.

Data Analysis

One participant was removed from further analysis due to poor accuracy data (only saccading towards the target 67% of the time). Without them, target detection accuracy was greater than 98% in each condition with no significant differences between conditions.

Only critical word trials (vertical target trials) were considered for analysis. The dependent variable of interest was the *saccade launch latency for looks to the target square*, measured from the onset of the target presentation until the eye started moving away from the central fixation cross (as determined by saccadic acceleration and velocity thresholds). Trials that contained eye-blinks, multiple saccades, off-target saccades (not landing within 100 pixels of the green square), or terminating after the 3000 ms timeout, were excluded from analysis (affecting less than 5% of the critical trials). Saccade launch latency outliers of more than 2.5 *SD*s away from the mean of a given subject × condition combination were also removed (affecting less than 0.3% of the data).

Inferential analyses were based on GEE (see Experiment 1 for a description); implementing a Gamma regression approach by using a *Gamma* distribution and *Log* link function in the GEE model specifications.

Two types of analyses were performed. In the *by-subject* analysis, word direction (‘up’, ‘neutral’, ‘down’), saccade direction (‘upwards’, ‘downwards’) and SOA level (‘70%’, ‘100%’, ‘130%’, ‘400ms’) were entered as within-subjects factors. In the *by-item* analysis, word direction was between-items and saccade direction and SOA level were within-items. As the items were the same as the previously presented experiments (1, 2 and 3), the same three principal components identified in section principal component analysis section above were entered as additional within-subjects / between-items covariates. Analyses without these covariates yielded nearly identical results and will not be reported in detail, but see the grey lines in Figures 4.7 and 4.8 for comparison. All analyses assumed an exchangeable covariance matrix for repeated measurements.

4.6.2 Results

Table 4.5 shows Generalized Score Chi-Square statistics from the Gamma regression analyses in GEE. A first point to note is that none of the control covariates (principal components) had a significant overall influence on saccade launch latency.

With regards to our experimental manipulations, there was no main effect of word direction either within-subjects or between-items. However, the main effect of saccade direction (ca. 30 ms higher saccade launch latencies for downward than for upward saccades overall) was significant within-items and within-subjects. This disadvantage for downward saccades is similar to that shown in the previous experiments presented here and previous research (including Dudschig et al., 2013).

Effect		By Subjects		By Items	
	df	GS χ^2	***P***	GS χ^2	***P***
Word Direction (W)	2	4.31	0.12	3.12	0.21
Saccade Direction (S)	1	37.57	< .001	105.01	< .001
SOA Level (SOA)	3	39.90	< .001	77.38	< .001
W × S Interaction	2	12.72	< .01	13.78	< .01
W x SOA Interaction	6	4.47	0.61	2.53	0.87
S x SOA Interaction	3	4.27	0.23	3.28	0.35
W x S x SOA Interaction	6	9.68	0.14	6.14	0.41
Lexical Access*	1	1.15	0.28	2.44	0.12
Valence-Dominance*	1	1.86	0.17	2.29	0.13
Arousal*	1	0.54	0.46	0.60	0.44

Table 4.5. Inferential results (Generalized Score Chi-Squares, degrees of freedom, p-values) from the Gamma regression analyses in GEE. Control predictors (covariates) are marked with an asterisk.

There was also a main effect of SOA level, which was significant within-subjects and within-items. Rather than list all contrasts, we focus on contrasts to the 100% SOA condition, which is a replication of the target onset times used in Experiment 3. 95% CIs for simple effects showed that saccades were launched significantly slower after a 70% SOA compared to a 100% SOA (by-subject contrast: 13 ± 4 ms, by-items: 13 ± 3 ms). Similarly, after a 100% SOA saccades were launched significantly slower compared to a 130% SOA (by-subject contrast: 9 ± 3 ms, by items: 9 ± 3 ms), and a 400 ms SOA (by-subject contrast: 5 ± 4 ms, by-items: 5 ± 3 ms). Figure 4.7. shows the covariate-adjusted by-subject means and *SE*s per condition each of the four graphs represents a different SOA level. For comparison, non-adjusted means and *SE*s are shown in grey. These differ only slightly as the three covariates (principal components) did not significantly effect saccade launch latency times, hence the grey lines are very close to the black lines. Most interestingly, the overall word direction × saccade direction interaction was significant in both the by-subject and the by-item analysis. 95% CIs for simple effects showed that upward saccades were launched more *quickly* upon hearing ‘up’ words like *moon* than upon hearing ‘neutral’ words like *letter* (by-subject contrast: 3 ± 3 ms; by-items: 3 ± 3 ms); the comparison between ‘down’ words like *sewer* and ‘neutral’ words like *letter* was not significant (by-subjects: 1 ± 3 ms; by-items: 1 ± 3 ms). Downward saccades were launched more quickly after ‘down’ words like *sewer* compared to ‘neutral’ words like *letter* (by-subjects: 6 ± 4 ms; by-items: 6 ± 4 ms), whereas the contrast between ‘up’ words like *moon* and ‘neutral’ words like *letter* was not significant (by-subjects: 0 ± 4 ms; by-items: 0 ± 4 ms).

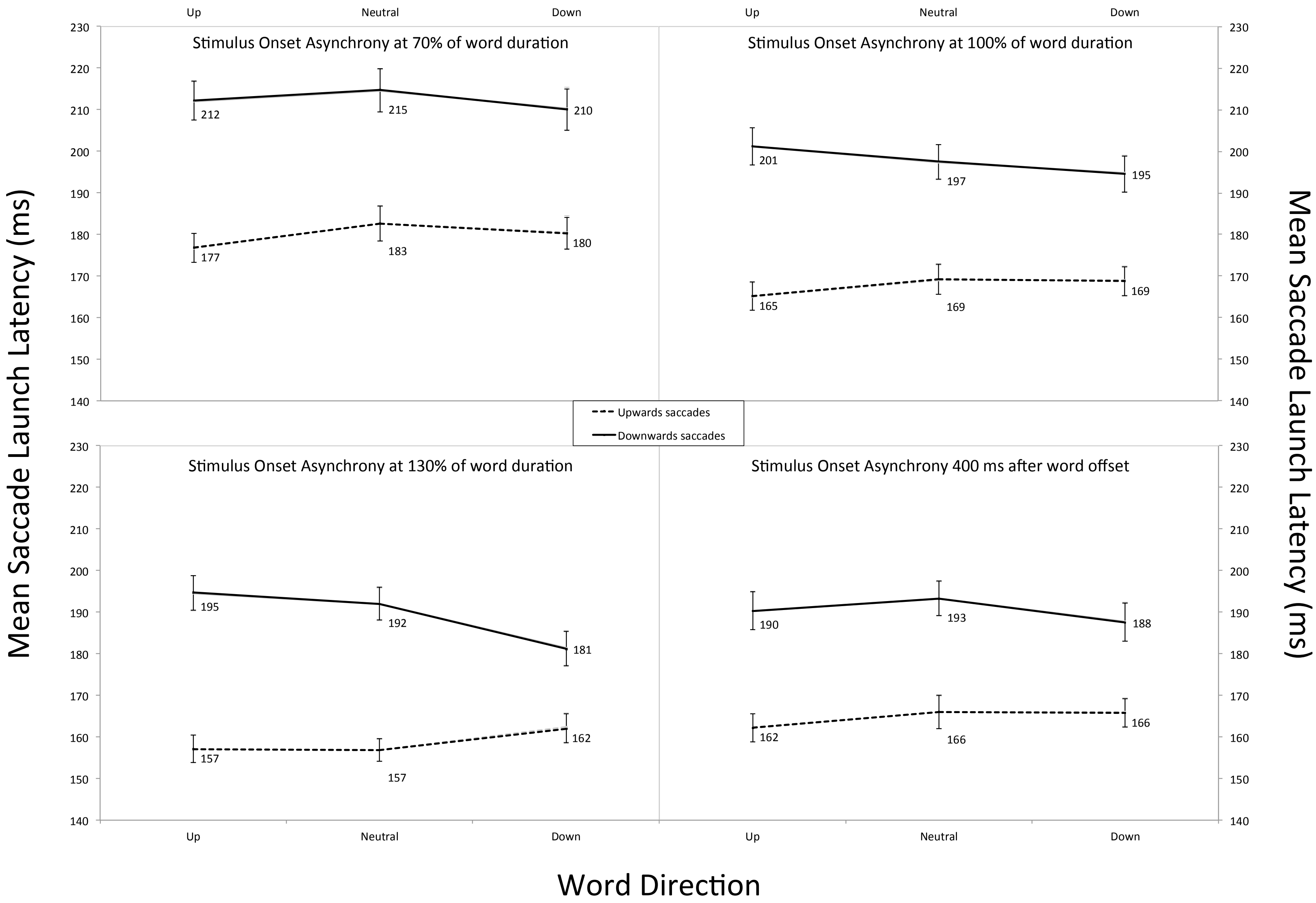


Figure 4.7. By-subject means and SEs for all SOA levels. The black lines show covariate-adjusted data and the grey lines show data from an analysis that did not include the covariates

However, given the small sizes of these contrasts we approach this interaction cautiously; this is discussed in more detail below. Figure 4.8. shows the covariate-adjusted by-subject means and *SE*s averaged across all SOA levels. Once again, non-adjusted means and *SEs* are shown in grey but differ only slightly due to the covariates not having a significant effect on the dependent variable.

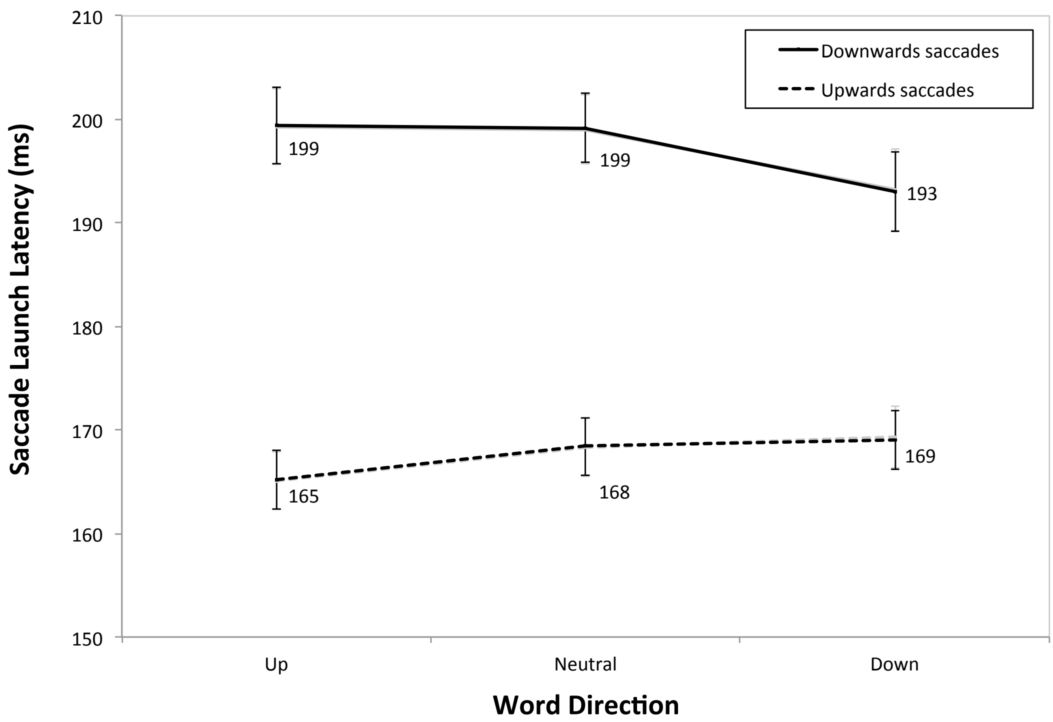


Figure 4.8. By-subject means and SEs combined across all SOA levels. The black lines show covariate-adjusted data and the (barely different) grey lines show data from an analysis that did not include the covariates.

4.6.3 Discussion

Experiment 4 investigated whether the delay between a cue and target is responsible for the previously documented differences in processing. Namely, whether processing is facilitated when a target is presented soon after the onset of a spoken word stimulus and inhibited when a target is presented at or after word offset. Based on the differing patterns of results shown in Experiments 1 (facilitatory effect) and 3 (inhibitory effect), it was hypothesised that detecting a target in a cue-compatible location would be facilitated if the visuo-motor processes overlapped with the perceptual simulation of the cued concept. Conversely, if the target appeared later (i.e. at, or after word offset) it was expected that saccades to an incompatible target location would be inhibited, similar to the results shown in Experiment 3.

The results of Experiment 4 show a significant main effect for SOA level; recall that launch latency is calculated from target onset. The general pattern was that the earlier a target appeared, the longer it took to launch saccades towards it (although this was not true for the 130% and 400 ms levels). Specifically, saccade launch latencies were longest after a 70% SOA, followed by 100% SOA, then 400 ms SOA, with the quickest launch latencies occurring after the 130% SOA. This pattern of results could be due to cognitive demands, such that early target presentation conflicts with the active conceptual processing mechanisms and hence visual search and saccade planning mechanisms are delayed until the cognitive load decreases. In short, doing two things at once causes a delay in responding; however this does not explain the full pattern of results (for example, saccade launch latency times were significantly longer following a 400 ms SOA compared with a 130% SOA). It is interesting to note that, a similar pattern of results was found in research on the SNARC effect – with early SOAs showing the slowest response times, later SOAs showing steadily decreasing response times and finally, SOAs greater than 750 ms showing progressively slower responses (Fischer et al., 2003). It is possible then, that the main effect of SOA shown in the results of Experiment 4 is a function of cognitive load similar to that shown when investigating spatial shifts in number perception. In conclusion, the SOA main effect suggests that when cognitive resources are used to process the word stimulus, fewer resources are allocated to the visual task of detecting a target therefore leading to longer response times.

Crucially, Experiment 4 also showed a significant two-way interaction between word direction (‘up’, ‘neutral’, ‘down’) and saccade direction (‘upward’, ‘downward’). Indeed, the overall pattern reveals a facilitatory effect on processing with saccades following ‘direction’ words being launched quicker to compatible locations compared to the neutral baseline. That is, within the time course of all the SOAs used, perceptual simulations of spatial concepts facilitated attentional shifts to a compatible location. This pattern of results is similar to that shown in Experiment 1, and was predicted here for early target presentation (i.e. 70% SOA). It should firstly be noted that the effect sizes for this interaction are extremely small (e.g. 3 ± 3 ms for ‘upward’ saccades) and, whilst statistically significant, the differences are bordering on being outside of the detection of the eye-tracking equipment[7].

[7] Technical specifications from SR Research state that the end to end sample delay for an EyeLink 1000 measuring at 1000Hz is: M < 1.8 ms and SD < 0.6 ms, hence the differences of 3 ms shown in Experiment 4 may not be reliable.

Nevertheless, it is worth asking why the facilitatory pattern shown in this interaction is no longer significant in the three-way word direction x saccade direction x SOA interaction. Unfortunately, the SOA levels used do not accurately reflect the time at which processing is facilitated. This is potentially due to the way SOA levels were defined. Using a percentage measure of the duration of each word allowed us to control the onset of a target relative to the perceptual simulation activated by the word. This dynamic measure was favoured over the traditional fixed SOAs used in much past research. If a fixed SOA were employed in Experiment 4, it would potentially have measured vastly differing points in the time course of the perceptual simulation. For example, if a target consistently appeared 200 ms prior to word offset, it would be vastly different for the spoken stimuli 'deep', which is 405 ms long compared to 'depreciating', which is 1045 ms long. In this example, only half of the information would be available to the listener in the trial using 'deep', compared to around 80% of the word being available in the trial using 'depreciating'. The percentage SOA design employed in Experiment 4 aimed to overcome this problem by making a target appear at a relative point in each word rather than at an arbitrary point.

Interestingly, the 100% SOA condition failed to replicate the effects shown in Experiment 3. In Experiment 3, saccades to incompatible locations were launched slower compared to the neutral baseline (significant for upward saccades, marginal for downward saccades). In the current experiment, the 100% SOA condition showed a similar trend for downward saccades and an opposing trend for upward saccades, however neither of the differences were significant. The reason why the previous result was not replicated remains unclear. However, it may be the case that manipulating SOA within-subjects may have impacted on responses as a whole and hence affected saccade launch latencies in the 100% SOA condition differently to those shown in Experiment 3. Realistically, the reason why the 100% SOA condition did not replicate the results of Experiment 3 cannot be identified from the present results and thus, further research is required.

The remaining question is why Experiment 3 showed an inhibitory effect and Experiment 1 show a facilitatory effect on processing? Aside from target onset, another major difference between Experiments 1 and 3 is the number of target locations. The targets for the lexical decision task always appear in the same location (for any single participant), whereas the target detection task has targets appearing randomly in one of four locations. The experimental procedures require the participant to respond by looking to a known location (the lexical decision task) or looking to an unpredictable location (target detection task). However prior to planning this saccade the participant must process the auditory stimulus, thus (for critical stimuli) activating the concept's vertical association. In the lexical decision task, this vertical association facilitates saccades if the 'word' target is in a compatible location compared to having no effect (relative to the baseline condition) if the target is in an incompatible location. In the target detection experiment, the uncertainty over the target's location means the initial vertical association activated by the auditory stimulus is potentially only useful in 25% of trials; that is, the activated upward association after hearing "moon" would only be useful for saccade planning if the target appeared at the top of the screen (vs. bottom, left or right). In this case, it is more cognitively efficient for the vertical association to be suppressed in preparation for planning a saccade to an incompatible location (which, including horizontal targets, is 75% of trial responses). This suppression of the initial vertical association may be enough to affect incompatible responses. At the point of response, the vertical association of the concept is being suppressed but is still active enough to interfere with the planning of a saccade to an incompatible target location. That is, when hearing "moon" and before planning the upcoming saccade the participant begins to suppress the upward association activated by the stimulus as it is unlikely (only 25% of the time) to be useful for the upcoming response. If an incompatible target appears (at the bottom of the screen), the upward association is still active enough to interfere with a downward response. However, if a compatible target appears (in this case, at the top of the screen), the participant is no quicker (or slower) to respond (relative to the baseline condition)[8]. Note that it could also be the case that critical responses to horizontal targets would be inhibited as these could be considered as incompatible locations. However, due to the experimental design (horizontal targets only appeared after filler words) we are unable to investigate this further.

The mechanisms involved here can be best described using a cost-benefit model of orienting attention. This idea is by no means new; Posner and his colleagues (Posner, 1980; Posner, Nissen, & Ogden, 1978; Posner, Snyder, & Davidson, 1980) have written extensively on behavioural outcomes when the proportion of valid cues is manipulated. Indeed, in a target detection task, participants were slower to detect the target after an invalid cue (e.g. an arrow pointing right followed by a target appearing on the left of the screen) versus a neutral cue (a fixation cross followed by a target on either side). This

[8] One may also consider the effects of Inhibition of Return (IOR; Posner & Cohen, 1984) upon compatible target detection. IOR suggests that orienting attention to a location (e.g. an upwards target), suppressing that orientation and then having to respond to that same location inhibits responses. However, considering our compatible responses (e.g. up word then up target) did not differ from our baseline responses (e.g. neutral word then up target) we are unable to conclude that IOR occurs here.

pattern of results is similar to those presented here (Experiment 3), however invalid cues only accounted for 20% of trials in Posner's study (Posner, 1980). Much research has shown how manipulating cue validity affects responses to spatial locations (e.g. Kingstone, 1992; Posner, 1980; Posner et al., 1978; Posner et al., 1980); however, manipulating valid cues to a level below chance has not been previously investigated.

Why the activated spatial orientation interferes with the planning of a saccade to an incompatible location, but does not facilitate saccades to a compatible location is still unknown. However, a recent series of experiments by Girardi and colleagues led to the conclusion that the cognitive effort involved in shifting spatial attention following a cue depends on how advantageous the shift is perceived to be for the upcoming response (Girardi, Antonucci, & Nico, 2013). This presents us with a new research question. If initial attentional orientation is suppressed when it is deemed not useful for the upcoming response task, would processing (and thus response times) be facilitated if the initial attentional orientation was deemed potentially useful for an upcoming task? Given the results of the two previous experiments, it is presumed that when attention is directed to a location via linguistic stimuli, it is either potentially useful for the upcoming task and therefore remains active (Experiment 1) or potentially not useful and begins to be suppressed (Experiment 3). In the first case, the spatial orientation activated by the word is either useful (or not) 50% of the time – that is, half of the time the spatial orientation activated by the word is compatible with the target location (however, recall that in Experiment 1 the targets were on the screen from word *onset* and the participant already knew the location of the targets therefore one would not need to predict a target location). In the second case, the "usefulness" reduces to 25%, as there are four possible response locations.

4.7 Experiment 5

Experiment 5 investigates whether the number of possible target locations modulates the previously reported cue-target compatibility effects. The experiment uses a similar design to Experiment 3, but increases the usefulness of activated spatial orientations to 50% by using only two possible response locations similar to Experiment 1. It is predicted that at this level it is no longer cognitively efficient to begin suppressing the attentional orientation, as it is just as likely to be useful when responding, as it is to not be useful. If the attentional orientation remains active, we should see responses to compatible locations facilitated relative to the baseline responses.

4.7.1 Method

Participants

Thirty-two individuals (16 Male, M=21.4 years) from the University of Glasgow participated in the study, each receiving £3 or course credits; none of these participated in any of the previous experiments. All participants had either normal or corrected-to-normal vision and were native monolingual English speakers. One participant did not complete the experiment fully due to a dislodged contact lens. Their part-dataset was discarded and another participant was recruited; hence there was a total of thirty-two complete datasets for analysis.

Materials and Apparatus

The same 120 critical words and 120 filler words from Experiment 3 were presented using the same apparatus and setup as Experiment 3.

Procedure

The procedure was similar to Experiment 3, however there were only 2 target locations. The 2 horizontal target locations from Experiment 3 were removed, thus participants responded to a target either at the top or bottom of the screen. Target location was counterbalanced between participants and items with each participant seeing 120 up and 120 down targets (60 of each were preceded by critical words and 60 by filler words). The 2 target locations for each word were randomized across 6 different lists thus minimizing intra-list effects; crucially, the position of the target could not be predicted by the word. All other aspects of the procedure were identical to Experiment 3, including target presentation at word offset.

Data Analysis

Target detection accuracy was greater than 99.5% in each condition. Unlike Experiment 3, all trials required vertical responses (saccading to a target at the top or bottom of the screen) however only those preceded by critical stimuli were considered for analysis. Once again, the dependent variable of interest was the *saccade launch latency for looks to the target square*, measured from the *offset* of the auditory word presentation (which was also the onset of the target presentation) until the eye started moving away from the central fixation cross and towards to target square (as determined by saccadic acceleration and velocity thresholds). Trials that contained eye-blinks, multiple saccades, off-target saccades (not landing within 100 pixels of the green square), or terminating after the 3000 ms timeout, were excluded from analysis (affecting less than 10% of the critical trials). Saccade launch latency outliers of more than 2.5 *SD*s away from the mean of a given subject × condition combination were also removed (affecting less than 0.1% of the data).

Inferential analyses were based on GEE (see Experiment 1 for a description); we implemented a Gamma regression approach by using a *Gamma* distribution and *Log* link function in the GEE model specifications. A similar procedure to Experiment 3 was used to define *by-subject* and *by-item* analyses. We again used the same principal components outlined in Chapter 3 as additional within-subjects / between-items covariates. All analyses assumed an exchangeable covariance matrix for repeated measurements.

4.7.2 Results

Table 4.6 shows Generalized Score Chi-Square statistics from the Gamma regression analyses in GEE. A first point to note is that, similar to Experiment 3, none of the control covariates (principal components) had a significant overall influence on saccade launch latency.

Effect		**By Subjects**		**By Items**	
	df	**GS χ^2**	***P***	**GS χ^2**	***P***
Word Direction (W)	2	6.85	< .05	5.45	0.065
Saccade Direction (S)	1	9.82	< .01	42.13	< .001
W × S Interaction	2	8.57	< .05	9.39	< .01
Lexical Access*	1	2.41	0.12	2.53	0.11
Valence-Dominance*	1	0.047	0.83	0.20	0.65
Arousal*	1	2.07	0.15	2.68	0.10

Table 4.6. Inferential results (Generalized Score Chi-Squares, degrees of freedom, p-values) from the Gamma regression analyses in GEE. Control predictors (covariates) are marked with an asterisk.

With regards to our experimental manipulations, there was a main effect of word direction within-subjects and marginally between-items (p = .065); saccade launch latencies to 'neutral' words (M = 168 ms) were significantly longer than those to 'up' words (M = 162 ms), but not 'down' words (M = 165 ms). The main effect of saccade direction (ca. 16 ms higher saccade launch latencies for downward than for upward saccades overall) was significant within-items and within-subjects. This disadvantage for downward saccades is similar to that shown in our previous experiments (1, 3 & 4) and previous research (e.g. Dudschig et al., 2013; Goldring & Fischer, 1997; Miles, 1936).

Most crucially, there was a significant word direction × saccade direction interaction in both the by-subject and the by-item analysis. Figure 4.9 shows the covariate-adjusted by-subject means and *SE*s per condition (for comparison, non-adjusted means and

*SE*s are shown in grey). 95% CIs for simple effects showed that upward saccades were launched more *quickly* upon hearing 'up' words like *moon* than upon hearing 'neutral' words like *letter* (by-subject contrast: 9 ± 5 ms; by-items: 9 ± 7 ms); the comparison between 'down' words like *sewer* and 'neutral' words like *letter* was not significant (by-subjects: 1 ± 5 ms; by-items: 0 ± 7 ms). Downward saccades were launched more quickly after 'down' words like *sewer* compared to 'neutral' words like *letter* (by-subjects: 7 ± 6 ms; by-items: 7 ± 6 ms), whereas the contrast between 'up' words like *moon* and 'neutral' words like *letter* was not significant (by-subjects: 2 ± 7 ms; by-items: 2 ± 8 ms).

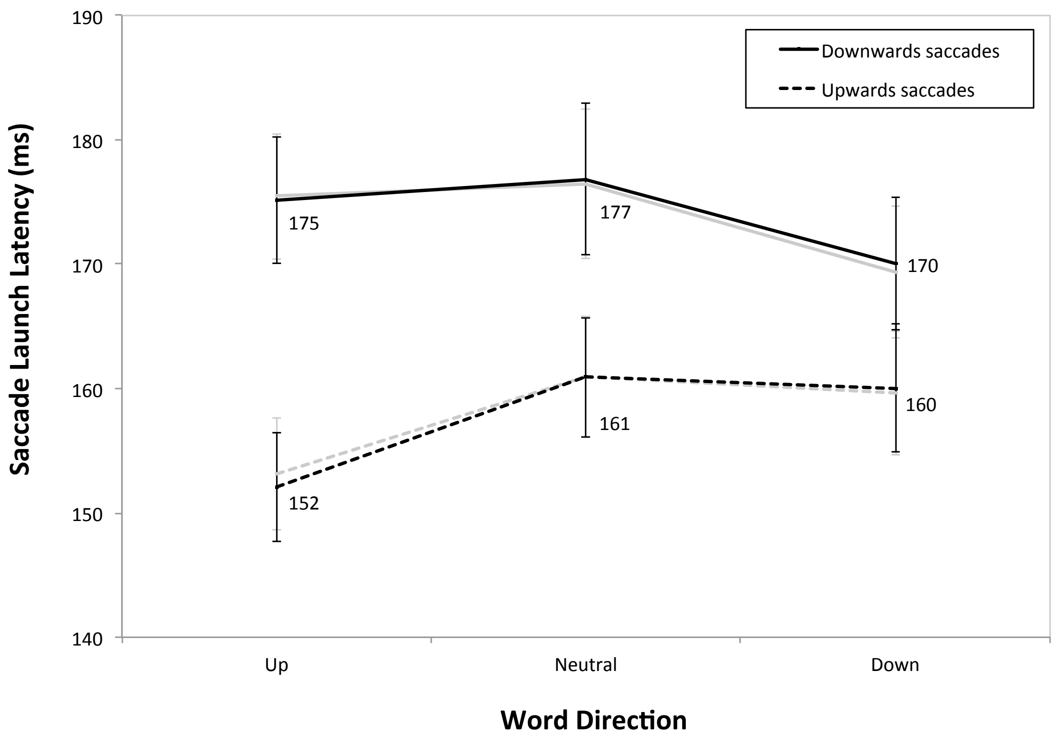


Figure 4.9. By-subject means and SEs per condition. The black lines show covariate-adjusted data and the (barely different) grey lines show data from an analysis that did not include the covariates.

4.7.3 Discussion

Experiment 5 investigated whether the proportion of valid cues, which was determined by the number of cue locations, mediated response times to cued spatial locations. The results clearly showed that 'direction' words facilitate saccades towards compatible locations in the vertical dimension but unlike Experiment 3, do not inhibit saccades towards incompatible locations.

In isolation, the results of the present experiment add to a growing body of literature showing that perceptual-spatial representations activated by linguistic stimuli can facilitate responses to congruent locations (relative to the baseline condition). Below, we discuss the implications of these results in context with those from Experiments 1 and 3.

Experiment 5 used a similar design to Experiment 3; participants were required to saccade to a target, which appeared at the offset of a word stimulus. The difference between the two designs was the number of possible target locations (2 vs. 4). Reducing the number of possible target locations from 4 to 2 has changed the pattern of results from inhibiting incompatible responses to facilitating compatible responses. The results to Experiment 5 are in line with the prediction that decreasing the amount of possible target locations (and hence, increasing the validity of the spatial orientation cued by word stimuli from 25% to 50%) would result in faster saccade launch latencies to compatible target locations and no difference (relative to the baseline condition) in saccadic responses to incompatible target locations. The results suggest that the initial attentional orientation that occurs upon presentation of a direction word stimulus remains active whilst planning the response, as it is now cognitively inefficient to be supressed. That is, after hearing "moon" attention is oriented upwards; this upwards orientation is potentially useful (or not) for 50% of responses hence the most cognitively efficient approach to responding would be for it to remain active rather than waste cognitive resources supressing it. This is unlike Experiment 3, in which attentional orientation is only useful for 25% of responses and is therefore better off being supressed.

The results from the present experiment seem to confirm the prediction outlined above (see section 4.7.3, Experiment 4 Discussion) that perceptual-spatial representations activated by linguistic stimuli only facilitate responses when they are congruent *and* useful (in terms of cognitive efficiency) rather than just congruent as was previously thought. This is discussed in further detail below.

It is worth considering the role that internal labelling may have upon the differing patterns of results shown across our experiments. Dudschig and Kaup (in press) suggest that compatibility effects, such as the facilitatory pattern of results shown above (Experiment 1), could be a result of motor responses being coded verbally. In Experiment 1 presented above, the two possible target locations remained consistent throughout the duration of an experimental session; hence each participant had the opportunity to create a response code (e.g. word-to-up or word-to-down). In this case, it is possible that the motor response is coded verbally (e.g. "look up") and therefore overlaps with the auditory stimulus code (i.e. an 'up' word). This would cause a compatibility effect, something that is subsequently reflected in faster response times to a compatible target location. Indeed, it would be expected that incompatible stimulus-response codes (e.g. 'down' word → word-to-up → "look up") would show no compatibility effect. In applying this idea to Experiment 3, in which such an internal response code would not be possible due to the unpredictable nature of the target (i.e. it appears in 1 of 4 possible locations), one would perhaps expect no compatibility effect. Whilst this idea goes someway to explaining the pattern of results in this thesis, there are still two major issues that remain unresolved. Firstly, why would a task in which no internal labelling occurs (Experiment 3) show an inhibitory effect on incompatible trials? And secondly, in a similar task where response locations are again unpredictable (Experiment 5), how would (a lack of) internal labelling explain the facilitatory effect shown in our results?

In answer to the first question, one could consider a dual process theory such as proposed by Neely (1977). For example, it is possible that processing word meanings is an ‘implicit’ process, whereas planning the motor response is a relatively ‘explicit’ process. It is generally proposed that implicit systems (sometimes called *System 1*) provide fast, automatic, low effort processing, whereas explicit systems (or *System 2*) provide slow, controlled, high effort processing (for a recent review, see Evans & Stanovich, 2013).

In Experiment 1, these two processes would seemingly interact harmoniously. In this case, word meanings (and their subsequently activated verticality) are processed in an implicit, bottom-up manner. Such a process would be fast and require little effort. In contrast, motor responses (which could be internally labelled) are processed in an explicit, top-down manner, which takes longer and requires more effort. If a word’s verticality is compatible with a response code it leads to a quicker motor response as the implicit and explicit processes overlap. Moreover, given that the response locations remain consistent throughout the experiment, a motor response requires less effort than, for example, Experiment 3 in which the response locations are unpredictable.

In Experiment 3, target location is unpredictable and motor responses cannot be labelled, thus increasing the effort required by the explicit systems. To this end, the explicit systems seem to take a strategic approach. Processing word meaning (as undertaken by implicit systems) is not required by the task, and along with the activated verticality associations, becomes supressed or inhibited. In the instances when the implicit and explicit processes are compatible (when a word’s verticality is compatible with a response location), motor responses are not facilitated as the explicit systems have attempted to inhibit the word’s verticality association in order to devote resources to the more complex response task. However, when the processes are incompatible, motor responses are inhibited. Presumably incompatible verticality associations require more effort to inhibit, and hence delay motor responses.

In Experiment 5 motor responses again cannot be labelled, however as there are only two possible locations, less effort (compared to Experiment 3) is required in order for explicit systems to prepare a motor response. Indeed, the results show a facilitatory effect – which brings us to the second question posed above. Perhaps the vertical saliency of the possible response locations becomes more apparent when there are only 2, compared with 4. With this, implicit processes (i.e. processing word meaning and subsequent verticality) may interact with explicit processes (i.e. vertical motor response) with verticality as a common factor. Furthermore, the implicit activation of a word's verticality no longer needs to be supressed or inhibited in order for explicit processes to prepare a motor response.

Of course, if a dual-processing theory such as proposed above is correct, it has far-reaching consequences within the domain of the present research. It seems to be assumed that if language is grounded, then the effects should be present in sensory-motor responses regardless of the paradigm. However, the experiments presented throughout this thesis show that motor responses to the same word stimuli can differ dependent on the task, rather than whether perceptual simulations of spatial association occur (or even exist). Crucially, it seems that an interaction between implicit processes, such as processing word meaning and explicit processes such as planning a motor response has a significant impact on findings. Indeed, without such an interaction, one may conclude that language may not be grounded at all. Yet, in reality the dependencies of the critical task have affected (or interacted with) the way the language has been processed (i.e. is it a necessary part of the procedure) such that motor responses may seem unaffected. As shown here, the dependencies of the critical task (e.g. complexity of task, response modality, stimulus/response saliency) are crucial when investigating language processing. With this in mind, further research is needed to investigate how language processing interacts with the dependencies of the critical task in order to fully understand which tasks are appropriate for investigating language processing, and more broadly grounded cognition.

4.8 Summary of Findings

A series of experiments presented in Chapter 4 have investigated how perceptual simulations of space are affected by different task demands. The chapter set out to explore why there is an unresolved conflict in the published literature surrounding the affect concepts associated with spatial locations have on processing, namely why some studies show a facilitatory effect and others show an inhibitory effect (e.g. Dudschig et al., 2012; Dudschig et al., 2013; Estes et al., 2008; Lachmair et al., 2011; Richardson et al., 2003).

In an attempt to further explore this conflict, we first identified a problem in the conclusions drawn from many previous studies – the lack of a baseline condition. In order to clarify whether faster (or slower) responses show evidence of facilitation (or inhibition), one must compare the response times to a baseline response. Without a baseline, the results remain inconclusive as to whether faster compatible (vs. incompatible) cue-target responses show evidence for a facilitatory effect on processing, or whether slower incompatible (vs. compatible) cue-target responses show evidence for an inhibitory effect on processing. It is somewhat surprising that none of the aforementioned studies used such a baseline setup and hence they remain inconclusive. Perhaps even more surprising given the conflict in the literature, is that conflicting results are continuing to be published without comparable baseline conditions (e.g. Dudschig, de la Vega, & Kaup, 2015; Estes, Verges, & Adelman, 2015). Experiment 1 presented here conclusively showed that perceptual simulations of spatial location activated by 'direction' words facilitate eye-movement responses to a compatible location relative to a neutral baseline. Importantly, the results showed no sign of an inhibitory effect. In an attempt to investigate the perceptual-featural overlap hypothesis (Estes et al., 2008), we hypothesised that the spatial location of a concept could be an overlapping feature (between the cue and target) and hence provide an explanation for the facilitatory effect seen in Experiment 1. However, using a target detection task (Experiment 3) to manipulate the spatial location overlap between cue and target showed an inhibitory effect on processing, such that participants responded more slowly to targets in a location incompatible with the cued location compared to the neutral baseline. Contrary to Experiment 1, there were no signs of a facilitatory effect. Perhaps more than any other experiment presented here, these results show the necessity for a baseline condition. Without the 'neutral' word condition, the results would only show that compatible responses are faster than incompatible responses and thus it would look as if there is a compatibility effect, or in other words compatible cue-target responses were facilitated.

Experiments 4 and 5 investigated why the previous experiments (and previous literature) had shown differing patterns of results – i.e. why Experiment 1 showed a facilitatory effect on processing and Experiment 3 showed an inhibitory effect on processing. Firstly, we explored how the stimulus onset time between cue and target affected processing. The findings were inconclusive as no significant interaction was found between word direction, saccade direction, and stimulus onset time. It is possible that the time at which a target is presented relative to the offset of the cue word may explain the differing patterns of results shown in previous experiments, however the results to Experiment 4 did not provide an explanation. The two-way interaction between word and saccade direction in Experiment 4 did however show the same facilitatory pattern of results as Experiment 1, albeit with very small differences in response times.

The final experiment in this chapter showed that the proportion of compatible cue-target responses seems to provide some explanation for the conflicting patterns of results shown in Experiments 1 and 3 (and previous literature). Experiment 5 replicated the procedure used in the target detection task (Experiment 3), however this time there were only two possible target locations – the top and bottom of the screen. In doing so, the results showed that ‘direction’ words facilitated eye-movement responses to compatible target locations relative to the neutral baseline. Crucially, there was no sign of an inhibitory effect. The key question is why a previous study using only two target locations concluded an inhibitory effect on processing (Estes et al., 2008). Before jumping to any conclusion, it is worth reiterating that Estes and colleagues did not track eye movements and hence the results do not allow for direct comparison to ours. The results in Experiment 3 showed that locating a target in a location *incompatible* with the cued location (e.g. a ‘down’ word followed by an ‘up’ target) was slower than the baseline condition; whereas the results from Estes et al. (2008) showed longer response times to targets in a location *compatible* with the cued location. In both cases, perceptual simulations occur whilst processing the cue word and hence attention is directed to a location compatible with the simulation. Without further research, it is difficult to speculate on why the differences between these two studies occur, however it could be a combination of the number of response locations (recall that the use of four possible response locations has previously shown inhibitory effects: Richardson et al., 2003) and the use of eye-movement measures as opposed to button pressing responses (note also that Gozli et al., 2013, Exp. 3A, replicated Estes and colleagues' results).

Why do the cue-target compatibility effects shown here seem to rely on how useful a cue is at predicting target location? Firstly, it is worth asking whether this is really the case. Consider an alternative; the mechanisms involved in visual search and attention are covertly attending many locations in anticipation of a target appearing at one of them, hence the resources available upon comprehending the cue word may be somewhat depleted compared to when there are less potential target locations. Due to the depleted resources, comprehending the word perhaps suffers and therefore does not facilitate the behavioural response (in this case, detection of a target in a compatible location). However, whilst this account explains why the use of 2 target locations (vs. 4 target locations) facilitates processing, it fails to explain the inhibitory effect shown in Experiment 3. If more cognitive resources are being used by visual search mechanisms, why do only incompatible cue-target trials show an inhibitory effect? On the other hand, if the pattern of results shown with 2 and 4 target locations is due to the ‘usefulness’ of the cue, then surely this assumes a strategic use of the activated spatial association, or does the 'suppression' of a less useful spatial association occur automatically?

In this case, perhaps a better explanation relates to task overlap, rather than featural overlap as suggested by Estes et al. (2008). If processing the meaning of a word overlaps with the critical task (e.g. lexical decision), then the response is facilitated. However, if processing word meaning interferes with the critical task, then responses are inhibited. Indeed, this theory also depends on the complexity of the critical task. In Experiment 1, the critical task is rather simple – look to a predictable location (the red and green targets remained consistent) depending on your lexical decision. Processing the meaning on the word was an integral part of the procedure, and together with a simple, predictable response task we found large facilitatory effects. Contrast this with Experiment 3: firstly, processing the meaning of the word was not as important (recall, participants had a recognition questionnaire after every 40 trials). Secondly, the critical task required more cognitive resources (attending 4 locations, responding to 1 unpredictable location); therefore if critical words induced spatial associations, these were subsequently inhibited, which is shown in the results. Finally, the procedure for Experiment 5 once again did not require deep processing of word meanings, but had a critical task more complex that Experiment 1 (as the target location was still unpredictable) yet easier than Experiment 3 (as there were less potential target locations). To this end, any spatial associations induced by the stimulus words did not require inhibition and in fact facilitated responses. Crucially, the effect was smaller than that shown in Experiment 1 for two reasons – word processing was not an integral part of the procedure, and the critical task was more complex. Moreover, we can apply this explanation to the results shown by Estes et al. (2008). Recall that the experiment used both a context (e.g. 'cowboy') and a cue word (e.g. 'hat' or 'boot') to denote spatial locations, and used a discrimination task (i.e. identify an X vs. O) rather than a detection task with only one target. Processing the two-word stimuli likely required more cognitive resources than processing a single word; similarly, discriminating between two targets is more complex than detecting one target. Due to these complexities – and similar to our Experiment 3 – any spatial association induced by the stimulus word pairs was inhibited in order to maximise performance on the critical target discrimination task. Unfortunately, only a limited amount can be concluded from the present experiments and in this case we are aware of the dangers of forming post hoc conclusions. Further research would be necessary to confirm the above theory. For example, an experimental procedure in which word processing is integral (e.g. lexical decision) in conjunction with a complex critical task (e.g. unpredictable 'word' and 'nonword' target locations) would provide some insight. In this scenario, one might predict an inhibitory effect as the critical task is similar to the target discrimination procedure used by Estes et al. (2008).

Finally, one other aspect of the target detection experiments presented here (Experiments 3, 4 and 5) that differed from the lexical decision experiments (Experiments 1 and 2) was the use of filler words rather than nonwords. By definition a lexical decision task requires a set of nonwords on which participants must make a judgement. However, in the target detection experiments we chose to include 120 filler words as a replacement for the nonwords. This made the experimental design have a similar structure to the lexical decision tasks, such that each participant was subjected to 240 trials. As outlined in Section 4.5.1, the 120 filler words were chosen on the basis that each word matched a critical 'direction' word in both length and frequency. Furthermore, the filler words contained no obvious spatial association (although, they were not subjected to a norming procedure). It is possible that presenting 120 words without spatial association weakened the 'global' verticality associations of the 120 direction words. That is, in Experiment 1 participants made a lexical decision on 120 direction words (80 of which were associated with verticality, and 40 neutral). Contrast this to the target detection experiments in which participants responded to 240 words. Again, 80 with vertical associations and 40 neutral, but this time with an additional 120 unrelated words. Indeed, this dilution of vertical associations may have led to a smaller, or weaker perceptual simulation in the cases where an upward or downward associated word was presented. If this is the case, one might expect this effect to manifest itself in the form of a smaller difference in response times between the direction and neutral word conditions. Indeed, the target detection experiments do show such a decrease; however the dilution of any effect does not explain the differing patterns of results between Experiments 3 and 5.

In summary, perceptual simulations of spatial location can facilitate and inhibit processing of a target in a compatible and incompatible location depending on the demands of the task. However, in order to draw valid conclusions from any future research, the use of a baseline condition is essential. Similarly, it seems that aspects of the task have a greater impact on processing and responses than previously literature had perhaps credited (i.e. number of response locations, eye-movement measures vs. button pressing, and target onset times). To this end, future research should accept that both facilitatory and inhibitory effects of processing occur and consider firstly, under which conditions does a particular effect occur, and secondly, why do different effects occur.

Chapter 5 Saccade Trajectory Effects

The previous experiments explored how perceptual simulations evoked by spatial concepts affected saccadic launch latency times; however there is evidence to suggest that perceptual simulations of space may also affect saccade trajectories. Using a signal detection task, Kamide (2007) showed that the shape of motion trajectory implied by a sentence influences saccade trajectory. Participants heard sentences with verbs implying an upper or lower trajectory (e.g. 'The ball will be [thrown / rolled] into the bin') alongside a concurrently presented visual scene depicting a boy, a ball and a bin. At verb offset, a small grey square appeared on the screen for 50 ms in a location either compatible or incompatible with the implied trajectory; the results suggested that detection of this signal was quicker when it appeared in a compatible rather than an incompatible location.

More recently, Kamide and colleagues presented participants with a visual scene depicting an agent (e.g. an 'alien') and a target object (e.g. a sofa) alongside a concurrently presented auditory sentence such as '[the agent] will [crawl / jump] onto the sofa.' The results showed that when participants fixated the path region between the agent and the target, the eyes were biased towards a higher location after hearing an upward verb (e.g. "jump") compared with a downward verb (e.g. "crawl"). This effect became more prominent when the depicted agent had to navigate around an obstruction (e.g. a television) appearing in the path region (Kamide, Lindsay, Scheepers, & Kukona, in press). Kamide et al. (in press) also tracked mouse movements and showed significantly higher mouse paths for upward vs. downward verbs.

These results suggest that the dynamic perceptual simulation evoked by the verb becomes active almost immediately after the verb has been comprehended, to the extent that it influenced visual attention and response behaviour.

The trajectory of a saccade is often assumed to be a straight, pre-programmed eye movement from one part of our surroundings to another. However, for nearly 50 years saccades have been shown to be curved and dynamic in their nature (e.g. Becker, 1989; Robinson, 1975; Sheliga, Riggio, & Rizzolatti, 1994; Yarbus, 1967). Saccade trajectory measures fall into two broad categories – curvature or deviation (Van der Stigchel, Meeter, & Theeuwes, 2006). Curvature measures look for differences over the course of a saccade trajectory from the start to the endpoint, whereas deviation measures show how external factors influence a trajectory (often, but not always, the external factors are mid-saccade). A number of studies have shown how saccade trajectories deviate both towards and away from visual distractors (for reviews, see Van der Stigchel, 2010; Van der Stigchel et al., 2006), however those particularly relevant to this thesis (e.g. Sheliga et al., 1994; outlined below) show that saccade trajectories deviate away from a cued location. When investigating the influence spatial attention has on saccade trajectories, it seems that directing attention to a spatial location causes a saccadic deviation away from this location when saccading towards a target. For instance, Sheliga et al. (1994) presented participants with four cue squares in a horizontal line at the top of a display – two in the left, and two in the right hemifield. Participants were required to fixate a central fixation cross, and without breaking their fixation, covertly attend to one of the four cue squares in which a cue (a cross) had appeared. Following this, participants made a saccade to a fifth square located at the bottom of the screen directly below the central fixation cross. The results showed that saccade trajectories (from central fixation to the response square) curved in the opposing direction from the hemifield in which the cue had appeared. Furthermore, saccade deviation was greater when preceded by cues in the two outermost locations compared with the two innermost locations. These results, along with several others, show that directing spatial attention to one location directly impacts saccade trajectories (Sheliga, Riggio, Craighero, & Rizzolatti, 1995; Sheliga, Riggio, & Rizzolatti, 1995; Van der Stigchel & Theeuwes, 2007).

Chapter 4 shows that perceptual simulations of space can directly impact on the time it takes to launch a saccade to a compatible or incompatible location. For the purposes of the present research, the question is whether the perceptual simulations of space activated by spatial concepts direct attention in a manner that affects saccade trajectories similar to those deviations outlined above. This issue is explored in Experiment 6.

5.1 Experiment 6

Given the rise in prominence of research investigating the interaction between vision and language, it is perhaps surprising that there is a lack of psycholinguistic research using saccade trajectory methodologies. Experiment 6 will investigate the effect that perceptual simulations of space have on saccade trajectories. It has already been shown throughout Chapter 4 that eye-movement behaviour is influenced by linguistically evoked perceptions of spatial location; hence the current experiment aims to see if saccade trajectories are affected in a similar way.

In order to investigate these effects, a design similar to the experiments presented in Chapter 4 is used. Participants will undertake an eye-movement activated lexical decision task requiring answers ‘in the horizontal dimension’ similar to Experiment 2 (see

Section 4.3). Assuming saccade trajectories are affected, the difference between conditions is difficult to predict. Studies of spatial attention such as those outlined above by Sheliga and colleagues (Sheliga, Riggio, Craighero, et al., 1995; Sheliga et al., 1994; Sheliga, Riggio, & Rizzolatti, 1995) would suggest a deviation in saccade trajectories away from a cued location. However, recall that Kamide (2007) showed faster detection of a signal in a location compatible with an implied motion trajectory, and Kamide et al. (in press) showed that fixations seem to follow the implied upward or downward trajectory of the verb. Due to the relative novelty of the methodological approach used in Experiment 6, the analysis will take an exploratory approach using a number of different trajectory measures outlined in recent reviews (Doyle & Walker, 2001, 2002; Ludwig & Gilchrist, 2002; Van der Stigchel et al., 2006). Each measure is presented in more detail in the Results section below (Section 5.1.2).

5.1.1 Method

Participants

Thirty-eight participants from the University of Glasgow participated in the study, each receiving £3; none of these had participated in any of the previous experiments. All participants had either normal or corrected-to-normal vision and were native monolingual English speakers.

Materials and Apparatus

The 120 direction words and 120 matched non-words outlined in the *Linguistic Stimuli* section above were used in this experiment. The linguistic stimuli were presented using the same apparatus and setup as Experiments 1-5. However, the response targets in the visual display did differ. Response targets were a green ‘Y’ that appeared in the right-centre of the screen (word) or a red ‘N’ appearing in the left-centre (non-word). Both visual targets were located 300 screen pixels horizontally (approximately 9.70º of visual angle) from the central fixation and measured 18 x 18 screen pixels (see Figure 5.1 below).

Procedure

Each participant was presented with 240 auditory stimuli in an individually determined random order. For each trial, participants had to make a lexical decision as to whether what they heard was a real English word or a made-up, non-word. Each trial began with the presentation of a centrally located ellipse for drift correction. Once fixated, the ellipse changed to a central fixation cross with a red ‘N’ on the left centre and green ‘Y’ on the right centre of the screen respectively (see Figure 5.1 below). At the same time, a sound file stimulus was played via headphones. The participant's task was to fixate either the green 'Y' or the red 'N' depending on whether they thought they heard a word or non-word. Each trial terminated after a fixation was recorded on either the 'Y' or 'N', or after a 5000ms timeout. A repeated-measures design was used; stimuli were randomised across participants with each participant providing lexical decisions on all 240 stimuli. Before the first trial and after every 40 subsequent trials, the eye-tracker was recalibrated and validated using a 9-point fixation stimulus; only 'Good' validations were acceptable ('Good' validations were determined as having an average gaze error <0.5º and a maximum error <1.0º). The entire experiment lasted approximately 40 minutes. Participants were then debriefed and any questions were answered.

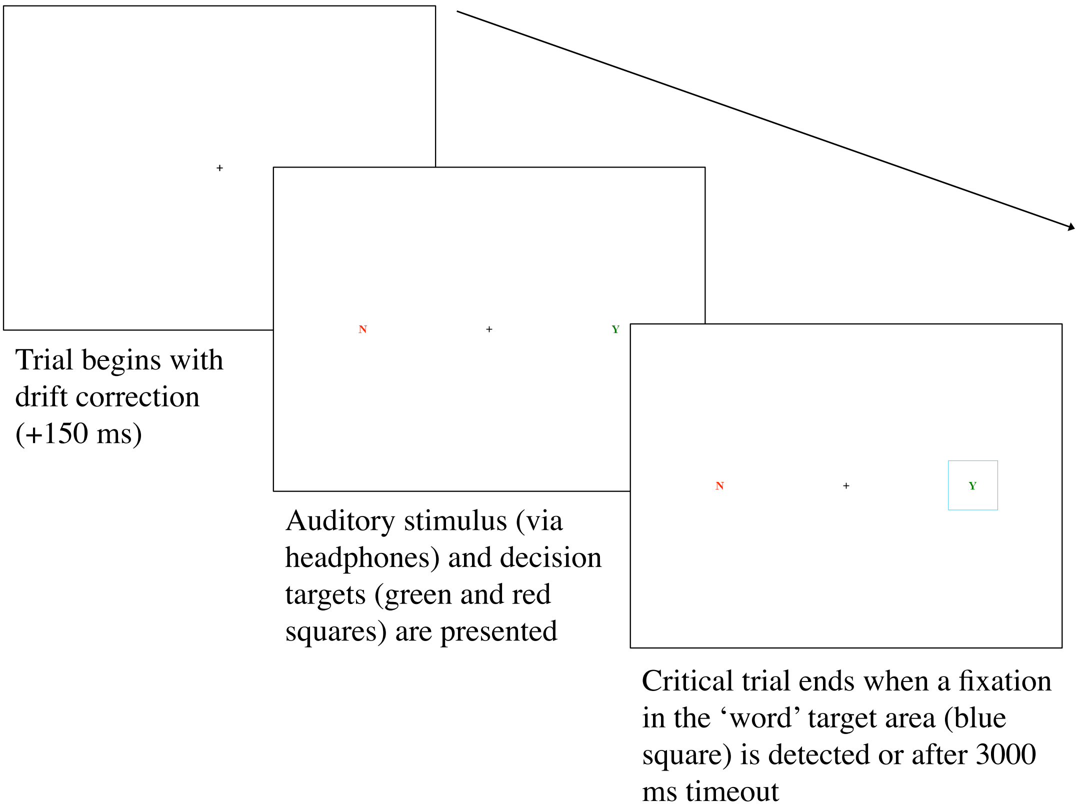


Figure 5.1. Schematic of a single trial. The green 'Y' (for 'word') and red 'N' (for 'non-word') letters represent the targets for lexical decision.

Data Analysis

Lexical decision accuracy for participants and items was generally over 80% (M = 92%). However, 1 participant and 2 items showed poor accuracy scores. The participant showed an accuracy score of 75% and was subsequently removed from further analyses. The two items were 'Loping' (neutral) and 'Butting' (neutral), which showed accuracy scores of 54% and 59% respectively. Given that these scores are only slightly greater than chance they were also removed from further analyses[9]. Furthermore, trials that contained eye-blinks, multiple saccades, off-target saccades (not landing within 50 pixels of the 'Y' target), or terminating after the 5000ms timeout, were excluded from any further analyses (affecting ca. 3% of trials). Outliers for each measure were determined as more than 2.5 *SDs* away from the mean of a given subject x condition combination and were also removed.

Data were analysed to see if there were significant differences in the saccade trajectories between the central fixation cross and the 'Y' target when participants heard the up, down or neutral target words. Three measures were used to analyse saccade trajectory – Initial Direction, Maximum Curvature and Area Curvature (see Figure 5.2 below). Saccade landing position was also analysed. All four of our chosen measures use screen pixels as their unit of measurement. Each measure is outlined in more detail below in the relevant section of the Results (Section 5.1.2).

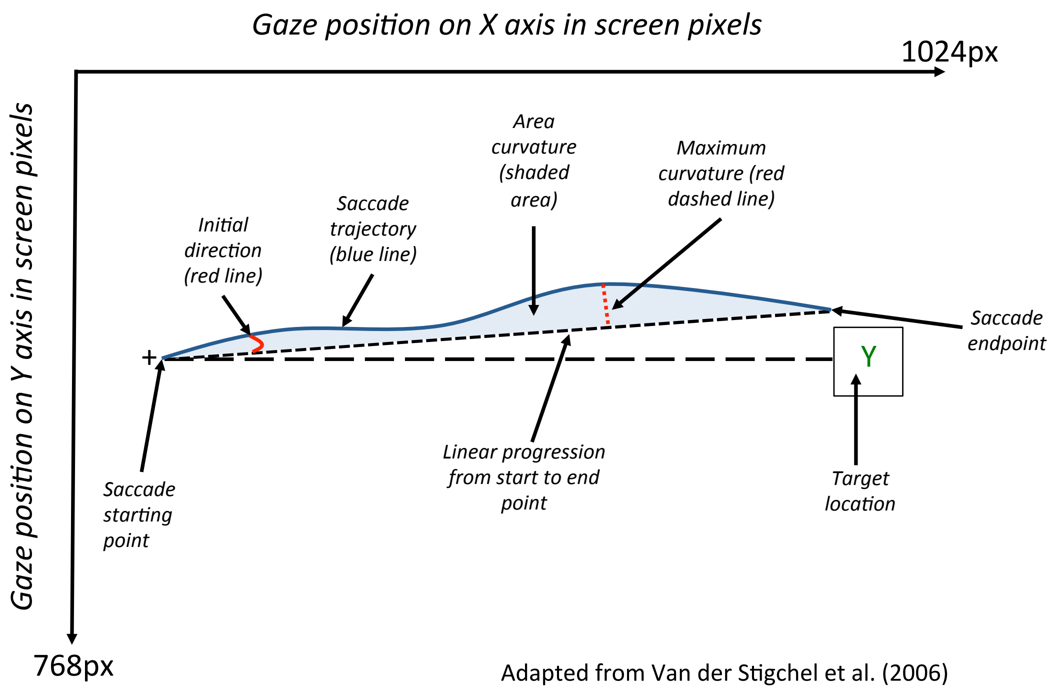


Figure 5.2. A visualisation of the Initial Direction, Maximum Curvature and Area Curvature saccade trajectory measures.

Inferential analyses were based on Generalised Estimating Equations (GEE; see Experiment 1 for description). For each measure, two types of analyses were performed. In the *by-subject* analysis, word direction ('up, 'neutral', 'down') was entered as a within-subjects factor. In the *by-item* analysis, word direction was entered as a between-item

9 Supplementary analyses conducted with the participant and items present did not show a different pattern of results.

factor. All analyses assumed an exchangeable covariance matrix for repeated measurements.

5.1.2 Results

Initial Direction

Initial direction provides a measure of the difference between the initial and overall directions of the saccade (Findlay & Harris, 1984). Initial direction was calculated 20ms from saccade onset (as recommended by Van Gisbergen, Van Opstal, Roebroek, & Noord, 1987). Results show the amount in screen pixels (px) by which the actual saccade trajectory deviates from the linear saccade progression (a straight line between saccade start and end; see Figure 5.2).

Table 5.1 shows the Generalized Score Chi-Square statistics from the linear regression analyses in GEE. None of the principal components showed a significant main effect on initial direction. There was no main effect of word direction with 'neutral' words (M = 1.23 px, SE = 1.00 px) showing no significant differences in initial direction to 'up' (M = 1.03 px, SE = .98 px) or 'down' words (M = 1.09 px, SE = .99 px). Similarly, 'up' and 'down' words did not differ significantly.

Effect		**By Subjects**		**By Items**	
	df	**GS χ^2**	***P***	**GS χ^2**	***P***
Word Direction (W)	2	1.22	0.54	1.68	0.43
Lexical Access*	1	0.25	0.62	1.94	0.16
Valence-Dominance*	1	1.50	0.22	1.45	0.23
Arousal*	1	0.14	0.71	0.33	0.57

Table 5.1. Inferential results (Generalized Score Chi-Squares, degrees of freedom, p-values) from the linear regression analyses in GEE. Control predictors (covariates) are marked with an asterisk.

Maximum Curvature

Maximum curvature is the largest absolute perpendicular deviation of the saccade trajectory from a straight line running from start to end point of the saccade (Smit & Van Gisbergen, 1990). A maximum curvature value for each critical saccade was calculated for analyses; reported curvature values are in screen pixels (px).

Table 5.2 shows the Generalized Score Chi-Square statistics from the linear regression analyses in GEE. Firstly, the principal component for valence-dominance showed a significant main effect ($p = .05$ between-items) suggesting that maximum saccade curvature is affected by the valence-dominance ratings of linguistic stimuli: Valence-Dominance was associated with a positive coefficient (+.37 ± .16 *SE* by subjects; +.37 ± .17 *SE* by items), meaning that greater valence and dominance ratings led to an increase in maximum curvature.

There was no main effect of word direction with 'neutral' words (M = -1.29 px, SE = 1.46 px) showing no significant differences in maximum curvature to 'up' (M = -1.28 px, SE = 1.40 px) or 'down' words (M = -.95 px, SE = 1.44 px). Similarly, 'up' and 'down' words did not differ significantly.

Effect		**By Subjects**		**By Items**	
	df	**GS χ^2**	***P***	**GS χ^2**	***P***
Word Direction (W)	2	0.65	0.72	1.08	0.58
Lexical Access*	1	0.025	0.88	0.35	0.55
Valence-Dominance*	1	4.51	< .05	3.85	0.05
Arousal*	1	1.04	0.31	1.27	0.26

Table 5.2. Inferential results (Generalized Score Chi-Squares, degrees of freedom, p-values) from the linear regression analyses in GEE. Control predictors (covariates) are marked with an asterisk.

Area Curvature

Area curvature was calculated according to the description provided by Ludwig and Gilchrist (2002) whereby the distance along a straight line for each millisecond sample is multiplied by the perpendicular deviation (measured in screen pixels). The sum of these deviations gives the total area curvature (in screen pixels squared; px^2) for a given saccade.

Table 5.3 shows the Generalized Score Chi-Square statistics from the linear regression analyses in GEE. None of the principal components showed a significant main effect on area curvature. There was no main effect of word direction with 'neutral' words (M = 26.57 px^2, SE = 232.20 px^2) showing no significant differences in area curvature to 'up' (M = 59.83 px^2, SE = 217.68 px^2) or 'down' words (M = 55.77 px^2, SE = 217.15 px^2). Similarly, 'up' and 'down' words did not differ significantly.

Effect		**By Subjects**		**By Items**	
	df	**GS χ^2**	***P***	**GS χ^2**	***P***
Word Direction (W)	2	0.72	0.70	0.54	0.76
Lexical Access*	1	0.097	0.76	0.28	0.60
Valence-Dominance*	1	1.47	0.23	1.21	0.27
Arousal*	1	0.016	0.90	0.063	0.80

Table 5.3. Inferential results (Generalized Score Chi-Squares, degrees of freedom, p-values) from the linear regression analyses in GEE. Control predictors (covariates) are marked with an asterisk.

Saccade Landing Position

Saccade landing position was calculated by normalising each saccade such that it begins at the centre of the display on the vertical axis. Saccade landing position is the normalised saccade endpoint on the Y-axis shown in Figure 5.2 (see above). The values are reported in screen pixels and represent the amount of deviation from the horizontal midpoint of the target area such that positive values represent a saccade endpoint lower (as '0 px' is the top left of the display) than the horizontal midpoint.

Table 5.4 shows the Generalized Score Chi-Square statistics from the linear regression analyses in GEE. None of the principal components showed a significant main effect on saccade landing position. There was no main effect of word direction with 'neutral' words (M = 3.83 px, SE = 0.29 px) showing no significant differences in saccade landing position to 'up' (M = 3.86 px, SE = 0.29 px) or 'down' words (M = 3.43 px, SE = 0.30 px). Similarly, 'up' and 'down' words did not differ significantly.

Effect		**By Subjects**		**By Items**	
	df	**GS χ^2**	***P***	**GS χ^2**	***P***
Word Direction (W)	2	1.03	0.60	1.27	0.53
Lexical Access*	1	0.49	0.49	0.27	0.61
Valence-Dominance*	1	2.88	0.09	2.54	0.11
Arousal*	1	1.82	0.18	0.58	0.45

Table 5.4. Inferential results (Generalized Score Chi-Squares, degrees of freedom, p-values) from the linear regression analyses in GEE. Control predictors (covariates) are marked with an asterisk.

5.1.3 Discussion

Experiment 6 investigated how perceptual simulations of space affect saccade trajectories. Previously, linguistically induced perceptual simulations of motion trajectory have been shown to facilitate signal detection in a plausible location (Kamide, 2007) and to impact upon fixation locations in vertical space (Kamide et al., in press). Furthermore, saccade trajectory has been shown to deviate away from a spatial location previously attended to (Sheliga, Riggio, Craighero, et al., 1995; Sheliga et al., 1994; Sheliga, Riggio, & Rizzolatti, 1995).

In the present experiment, participants were presented with a series of non-words and direction words, and made a lexical decision by saccading to a target at the left or right of the display. Saccades following critical words (the 'up', 'neutral' and 'down' words) did not seem to be affected by perceptual simulations of spatial location. Analyses of initial saccade direction, maximum curvature, area curvature and saccade landing position showed that none of the measures were significantly affected by word direction. It seems that either perceptual simulations of spatial location do not affect saccade trajectories or Experiment 6 did not find an effect that exists.

The results of the previously presented experiments (see Chapter 4) show that the linguistic stimuli used in Experiment 6 evoke a perceptual simulation of spatial location that is prominent enough to direct attention and affect eye movements. However, there are numerous aspects of the methodology that may have masked the effect from our analyses. Recall that Sheliga et al. (1994) covertly directed attention to a spatial location by displaying a visual cue, albeit in the vertical periphery, thus it is possible that saccade trajectory deviations occur only after covertly attended visual cues. It is also notable that an explicit part of the task stated by Sheliga et al. (1994, p. 508) required participants "to direct attention to the cued box, without breaking fixation". The instructions order the participant to avoid looking at the cued location – it is perhaps no surprise that saccade trajectories showed a curved deviation in an opposing direction to the cued location. Experiment 6 did not use a visual cue to direct attention, but rather implicitly directed spatial attention to an upward or downward location via the use of linguistic stimuli. The compatibility effects shown by Kamide (2007) and Kamide et al. (in press) also used an implicit cue to direct spatial attention – the motion trajectory implied by the verb (e.g. 'throw' vs. 'roll'). Whilst the results of these studies show that dynamic perceptual simulations evoked by language can affect visually simulated motion trajectories, this was in the presence of a visual scene depicting the objects referred to in the linguistic stimuli. Experiment 6 did not include any visual depictions and required the participant to saccade to one of two predetermined response locations ('Y' or 'N' targets), rather than detect an on-screen signal.

The results of Experiment 6 suggest that perceptual simulations of spatial location do not affect saccade trajectories. However given that this is one of the first experiments to investigate linguistic effects on saccade trajectories, there is a wide scope for future research. The results of the experiments in Chapter 4 show that perceptual simulations of spatial location directly affect eye movements in the form of saccade planning. However, one major difference between the experiments in Chapter 4 and Experiment 6 presented here is the difference between saccade planning and saccade execution. Experiment 6 investigates different trajectory measures of the executed saccade, whereas those experiments in Chapter 4 measure the time it takes from word offset to launching the saccade – essentially, saccade planning.

It is interesting to note that a similar pattern of results is shown in research using hand movements as a response. Recall that the aforementioned ‘action sentence compatibility effect’ (ACE), outlined by Glenberg and Kaschak (2002), showed faster responses to a compatible (vs. incompatible) location when a sentence implied a toward or away movement. Indeed, the RTs were measured from the onset of the visual stimulus (the sentence), which appeared by pressing a centrally located button on a button box, to the moment the finger released the central button and moved in the direction of the ‘toward’ or ‘away’ button. Hence Glenberg and Kaschak (2002) were measuring the time it took to read and understand the sentence, and plan the corresponding motor movement.

However, this is not only the case with sentence comprehension. Lachmair et al. (2011) showed that participants were quicker to respond to a direction word when the response key was located in a compatible location (i.e. making a lexical decision on the word “roof” was quicker when the response required an upward rather than downward hand movement). As with Glenberg and Kaschak (2002), participants were pressing a centrally located button during the task and thus RTs were calculated from stimulus onset until the release of the centrally located button, rather than the subsequent pressing of the compatible response button which followed. These results show how perceptual simulations of vertical space seem to affect motor planning mechanisms. However, given that Lachmair et al. (2011) did not report the RTs from the release of the central button to the pressing of the ‘upward’ or ‘downward’ located response button, it is unclear whether motor execution processes were also impacted.

More recently, Dudschig and Kaup (in press) used a modified version of the procedure used by Lachmair et al. (2011). In a Stroop-like experiment, participants responded to the colour of ‘up’ and ‘down’ words by pressing the corresponding colour button. Similar to Lachmair et al. (2011), the procedure for Experiments 1 and 2 implemented by Dudschig and Kaup (in press) required participants to press a centrally located button prior to each trial beginning. Importantly, RTs were measured from word onset until the corresponding response colour response button was pressed; hence RTs included reading and understanding the word[10], and planning and executing the response. The results showed a compatibility effect whereby upward responses were faster after reading an ‘up’ (vs. ‘down’) associated word, and downward responses were faster after reading a ‘down’ (vs. ‘up’) associated word. Whilst these results do not dissociate between

[10] Reading and understanding the word was not a necessary part of the task given that the Stroop-like procedure required a response to the colour of the presented word. However, given the compatibility effect between word direction and response direction shown in the results, we can assume that participants processed word meaning at some level.

the motor response planning and execution stages outlined above, it is somewhat interesting to note that RTs which included both stages still showed an interaction between word direction and response direction. In summary, it may be the case that whilst perceptual simulations affect the planning of motor responses, execution may not be affected. Indeed, this goes some way to explaining the results show in Experiment 6.

Part II. Body Position Effects

Chapter 6 Embodied Perceptions of Space

6.1 Perceptions of space

Communicating a spatial location by saying “the box is behind the car” is ambiguous as to whether the box is on the other side of the car from where I am standing, or at the rear of the car. Upon describing the location of an object, people generally use one of three available ‘frames of reference’. Levelt (1984) proposes use of an *intrinsic* frame of reference whereby the object itself is the reference point (i.e. the car; hence the box is at the rear of the car). Alternatively, a *deictic* (also called *relative*) frame of reference is used, whereby the viewer or referrer is the reference point (i.e. the box is on the occluded side of the car from my location). Levinson (1996) introduces the *absolute* as a third frame of reference. Here, salient environmental features are used as reference, for instance the fixed “cardinal directions” (such as north/south/east/west). In this case, regardless of where I stand, the box is always East of the car.

It is suggested that we describe locations egocentrically, or in other words our perception uses a relative frame of reference as we describe locations relative to our own position (Clark, 1973; Miller & Johnson-Laird, 1976). Indeed, recent work in the field of perspective taking and audience design suggests that initial interpretations of given instructions were egocentric (e.g. Epley, Morewedge, & Keysar, 2004). However, the present work is more focused on how we internally represent the location of an object.

If mental representations of concepts are grounded in modal or multimodal systems (i.e. Barsalou, 2008; see Section 2.3 above) then our perceptual representations of spatial locations should be grounded in our sensory-motor interactions with the physical world. Myachykov, Scheepers, Fischer, and Kessler (2013) argue that our grounded representations of space are hierarchically ordered in terms of their stability and relationship with the environment. The most stable aspect of the hierarchy is termed *tropism.* Grounded by the physical world, *tropic* representations are made up of our conceptual knowledge about static components in the world; most relevant to the current work is our representation of vertical space. The concept of verticality is determined by gravity; a stable and ubiquitous presence that, Myachykov et al. (2013) argue, is unaffected by simulation constraints (e.g. bodily differences, language, social context). Hence our mental simulation of a concept like "up" is represented as 'away from gravity' regardless of any constraining factors.

The less stable aspects of this hierarchy are *embodied* and *situated* conceptual representations (Myachykov et al., 2013). Embodied representations largely encompass conceptual representations in which the bodily state is part of the conceptual knowledge. For instance, a conceptual representation of "writing" is grounded in the brain's modal systems and simulated as a movement that is more malleable and less stable than tropic representations. So, whilst the concept of writing is similar for different individuals, the sensory-motor experience and mental simulation would most likely differ for a right-handed compared with a left-handed individual. Finally, situated conceptual representations are affected by the context in which they are retrieved and are the least stable of the proposed hierarchy. For example, our conceptual representation of "hour" differs depending on the context in which it is retrieved. Hence, whilst 2 hours may seem like a long commute to work each morning, it is a relatively short time in which to read an entire novel.

Myachykov et al. (2013) argue that the tropic component of mental representations takes priority over other components and should be free from constraining factors such a body position. However, there is debate as to whether an egocentric or allocentric (i.e. relative to another's position) perspective is most prominent when mentally representing spatial locations (Clark, 1973; Kappers, 2004; McNamara, Rump, & Werner, 2003; Miller & Johnson-Laird, 1976; Tversky & Hard, 2009). Moreover, when body position is no longer aligned with the vertical axis (e.g. when lying down), egocentrism and tropism no longer overlap. Consider the thought experiment posed by Levelt (1984) in which he argues that, in Figure 6.1a and Figure 6.1b, it is acceptable for the depicted agent to say "the yellow balloon is *left* of the red balloon". On the other hand, in Figure 6.1c, the perceived verticality becomes contextualised (i.e. by the horizon line) and it now seems odd to say that "the yellow balloon is left of the red balloon" rather, "the yellow balloon is *above* the red balloon".

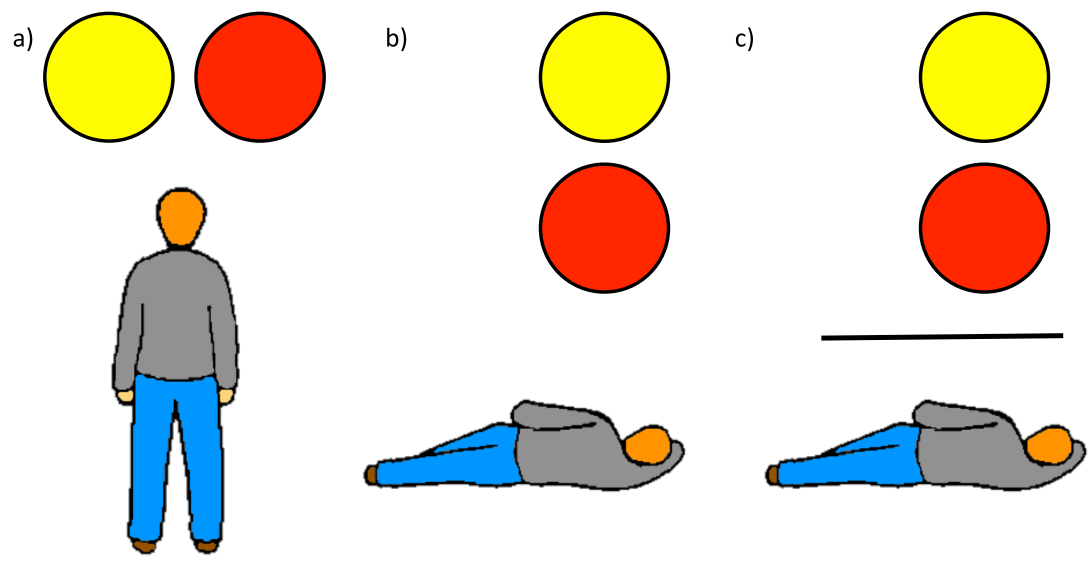


Figure 6.1. Modified and reproduced from Levelt (1984, p. 332)

There is also empirical evidence suggesting that this is the case; for instance, Friederici and Levelt (1990) showed that participants were significantly more likely to describe the position of a ball in an image relative to their own body position rather than other visual cues also in the image (i.e. pictures of [upside down] trees) – interestingly, this was true both on Earth and in a zero gravity environment (i.e. in space). More evidence comes from Carlson-Radvansky and Irwin (1993); the researchers asked participants to rate the acceptability of a sentence describing an image. In critical trials, the image displayed a fly either intrinsically or extrinsically above an object (see Figure 6.2), with the participants rating the acceptability of the sentence "the fly is above [the object]." Their results showed that the intrinsic frame of reference (relative to the donkey, hence fly #2 in Figure 6.2) was deemed significantly less acceptable than the extrinsic or deictic frame of reference (relative to our own position, hence fly #1). With regards to Myachykov et al. (2013), these results do not disambiguate between the tropic and embodied perspective and therefore do not provide evidence either way for their theory. However, it could be argued that there was no tropic perspective when Friederici and Levelt (1990) asked participants for spatial judgements in zero gravity, rather there was only an embodied or situated perspective. This would suggest that spatial judgements were made from an embodied perspective, albeit in an environment where 'tropic' is undefined.

Figure 6.2. Example stimuli from Carlson-Radvansky and Irwin (1993, p. 229) which depicts a fly in either a deictic/extrinsic location (fly #1) or an intrinsic location relative to the position of the donkey (fly #2). Is fly #1 or fly #2 'above' the donkey?

The question now is whether our conceptual representations are embodied and can thus be determined by our body position. Or rather, is it possible that we internally represent 'up' as above us, rather than just away from gravity? In this sense, whilst Myachykov et al. (2013) determine conceptual representations of vertical space as tropic, one could argue that they are embodied as our body position (or body location) contributes to our conceptual representation of space. One could go so far as to argue that perceived verticality is defined by contextual influences such as the horizon and thus are 'situated' representations (similar to Levelt, 1984; see Figure 6.1c), however it would be difficult to find a realistic situation whereby all three (tropic, embodied and situated) potential vertical representations are not aligned with one another (e.g. floating in an 'upside down' room in zero gravity).

In summary, when contemplating descriptions of spatial location using frames of reference and mental simulations of spatial location, it becomes apparent that conceptual knowledge about 'up' and 'down' may be less stable than originally proposed by Myachykov et al. (2013). Broadly, Experiments 7, 8 and 9 ask if (and how) body position affects our processing of vertical space.

6.2 Experiment 7

Experiment 7 investigates whether our conceptual representation of ‘up’ and ‘down’ is affected by the position of our body. As outlined in Section 6.1, Myachykov et al. (2013) argue that our conceptual representation of “up-ness” is a stable, tropic representation defined by gravity such that ‘up’ is away from gravity and ‘down’ is towards gravity. However, recall that Barsalou (2008) states that conceptual knowledge is grounded in our sensory-motor systems, and concept retrieval activates a perceptual simulation based on our experiential knowledge. Experiment 7 investigates whether the concept of “up-ness” is indeed a stable, tropic representation or if our perceptual simulation is affected by the position of the body, thus making it an embodied (or even situated) representation.

In order to dissociate between embodied spatial representations (those defined by our body position) and tropic spatial representations (those defined by gravity), participants were required to either lie on one side or sit upright whilst undertaking the experiment. Each participant was presented with a series of visual scenes, with critical trials depicting arrows pointing in various directions. Each trial had a concurrently presented auditory question asking, “How many arrows point [up/down/left/right]?” Participants were required to answer each question with a number – crucially, those participants who were lying down could answer from a tropic or an embodied perspective as the number of arrows pointing in each direction differed. For example, a participant lying on their left hand side may see a display showing three arrows and one distractor shape. Two arrows may point upwards and the remaining arrow may point leftwards (in a tropic sense). The participants would hear the auditory question – “how many arrows point up?” – to which they would respond dependent on their chosen reference frame. In this example, if the participant answered “2”, we could determine that they have used the tropic perspective to define ‘up’; whereas if they answered “1”, we could determine that they have used the embodied perspective as ‘up’ is now aligned with their body position. To this end, the answer to each question gave insight into whether the participant was defining the spatial location from the perspective of their body, or gravity (see Methods, Section 6.2.1, for further details).

If our mental representation of spatial location is predominantly tropic, participants lying on one side should answer the critical questions no differently to those who are upright. Indeed, those participants who are seated upright act as a control condition as they should always answer from the tropic perspective due to it being aligned with their body position. However, if our mental representation of space is embodied, such that “up” is defined as ‘above our head’, then arrows pointing ‘away from gravity’ should be defined by those who are lying down as left or right depending on which side they are lying (e.g. Figure 6.1b).

Response times were also recorded as previous research has shown that choosing a frame of reference takes longer when participants are presented with multiple options (e.g. Sun & Wang, 2010; Tamborello II, Sun, & Wang, 2011). To this end, we expect response times to be longer for those participants encountering multiple correct possibilities – i.e. those who are lying down can choose from either ‘embodied’ or ‘tropic’ reference frames.

6.2.1 Method

Participants

Sixty participants (22 Male; M=22.4 years) from the University of Glasgow participated in the study, each receiving £2 or course credits; none of these participated in any of the previous experiments. All participants had either normal or corrected-to-normal vision, were native monolingual English speakers and were right-hand dominant.

Materials

Visual stimuli

105 image slides were created using Microsoft PowerPoint software. Of the 105, 20 slides were critical trials, 80 were filler trials and 5 used as practice trials. The critical trials each consisted of three arrows and one distractor shape – either a square or circle. The position of the arrows was counterbalanced across the 20 slides. The direction the arrows were pointing was ordered so the answer to the associated question (“how many arrows point up/down/left/right?”) would allow insight into the preferred reference frame (tropic vs. embodied). The 80 filler trials were split into four groups of 20, with 20 trials containing shapes, 20 containing letters, 20 containing numbers and 20 containing currency symbols (Figure 6.3). The 20 filler slides containing ‘shapes’ were similar to the critical trials except they did not all contain 3 arrows but rather, contained a random selection of arrows, squares and circles. 5 practice slides were also created; these contained 1 ‘letter’ slide, 1 ‘number’ slide, 1 ‘currency symbols’ slide and 2 ‘shapes’ slides. An array of different stimuli was chosen in order to deter participants from guessing the aim of the experiment. Each 1024 x 768 pixel slide consisted of 4 images that appeared in a 400 x 400 pixel centrally located area. In this area, each image was centrally located in one of the four 200

x 200 pixel quarters. Figure 6.3 shows example stimuli in detail. All of the 100 experimental slides are presented in the Appendix.

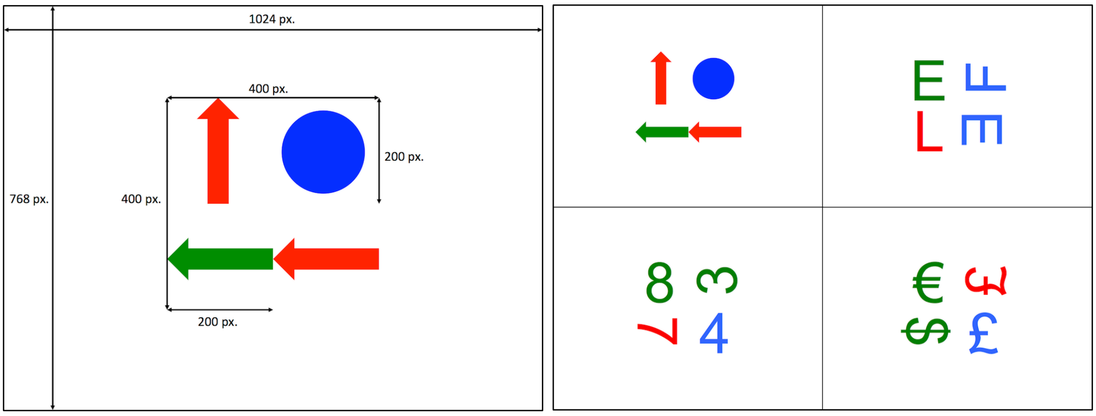


Figure 6.3. Example Stimuli. The left image shows the layout and sizes of each slide. The right image shows 4 example slides. The top left shows a critical trial whereas the other three show filler trials.

Auditory Stimuli

Auditory stimuli were created using IVONA Reader voice synthesis software. A male British-English voice[11] was used to read 30 questions, each was saved as a separate waveform audio file. Word stress was controlled so that each question had a steady tone with no obvious rising or falling intonation. The critical trials used one of four questions relating to the direction of the on-screen arrows – "How many arrows point up/down/left/right?" The filler trials consisted of two types of questions, relating to either the amount of certain objects or colour of the stimuli e.g. "How many numbers are blue?/How many L's in total?" (See Appendix for a full list of filler questions). Participants answered "zero", "one" or "two" to all 30 questions. Each audio file had the volume normalised to -6dB (peak level) using Sound Studio (Felt Tip Software). The average question length was 1703ms (1903ms for critical trials).

Apparatus

Stimuli were presented on a 13-inch Apple MacBook computer with a display refresh rate of 60Hz. A floor-mounted microphone and Audacity v2.0.5 software was used to record verbal responses. The experimental script ran using Experiment Builder software (SR Research).

[11] The voice of 'Brian' from http://www.ivona.com/en/reader

Procedure

Participants were assigned to one of three conditions – lying left, lying right or seated upright, with 20 participants in each condition. Those in the lying left condition were asked to lie on their left hand side on a mattress placed at floor level. Similarly, those in the lying right condition were asked to lie on their right hand side (see Figure 6.4 below). Those participants in the seated upright condition were seated at a desk. In all conditions, the monitor was oriented in an upright position and was placed 70cm in front of the participant. The centre of the screen was aligned with the bridge of the participant's nose. Tested individually, participants saw 105 visual stimuli (5 practice, 100 experimental) each accompanied by a concurrently presented auditory question. Each of the 100 experimental visual stimuli was presented once to each participant in a randomly determined order, hence each participant saw the 20 critical slides and 80 filler slides once. Participants had to answer the critical question verbally, and as quickly as possible. If no response was recorded, the trial terminated 5000 ms after onset. The experiment lasted no longer than 10 minutes.

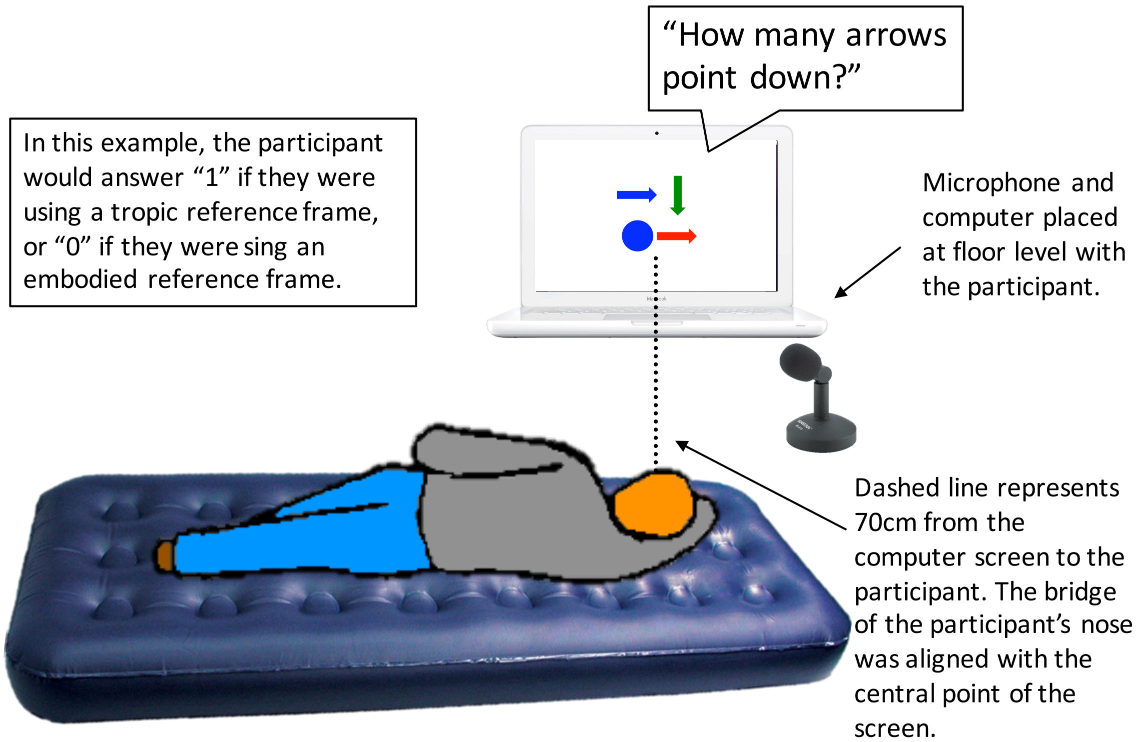


Figure 6.4. Illustration of the experimental setup for the lying down conditions in Experiment 7. In this illustration, the participant would be required to answer the question "how many arrows point down?". Their response would provide insight into whether they were using an embodied or tropic reference frame.

Figure 6.5 shows an example trial: In this example, if the participant was lying on their left hand side, we would expect their answer to the critical question to be "1" or "2", which would be coded as 'tropic' or 'embodied' respectively (any other response was coded as incorrect). Alternatively, if the participant was lying on their right hand side, we would expect their answer to be "1" or "0", which again would be coded as 'tropic' or 'embodied' respectively (again, any other response was coded as incorrect). Finally, if our participant was seated upright we would only expect their answer to be "1", with any other responses being coded as incorrect.

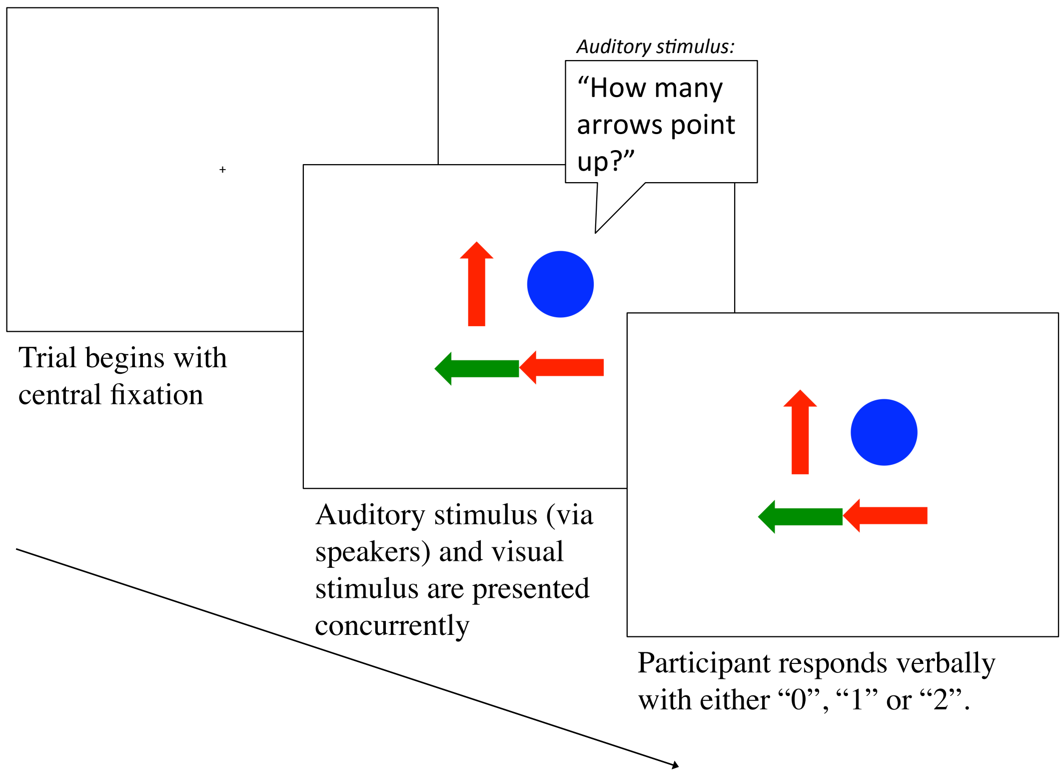


Figure 6.5. Schematic of a single critical trial. Participants' answers to the critical question provide insight into their spatial reference frame.

Data Analysis

Data were analysed offline to see if there was a difference in the proportion of responses (embodied vs. tropic) to the critical questions between the position conditions (lying left vs. lying right vs. upright). As the upright condition acted as a control, responses were coded as 'tropic' or 'incorrect' due to their being no dissociation between tropic and embodied. In the lying down conditions, responses were coded as 'tropic', 'embodied' or 'incorrect'. No inferential analyses were carried out on this response data (see below).

Response times were calculated from the onset of the critical direction word until the participant gave a verbal response. For example, from the onset of the word "up" in the question "How many arrows point up?" until the onset of the verbal response. Inferential analyses were carried out on response time data, and were analysed based on GEE (see Experiment 1 for description). As with previous experiments presented here, we implemented a Gamma regression approach by using a *Gamma* distribution and *Log* link function in the GEE model specifications. Body position ('left', 'right', 'upright') was entered as a between-subjects factor in the *by-subjects* analyses, and a within-items factor in the *by-items* analyses. All analyses assumed an exchangeable covariance matrix for repeated measures.

6.2.2 Results

Accuracy

Responses to the critical questions were highly accurate for both critical and filler trials across both conditions, with less than 3% of responses being coded as incorrect. Given the nature of the lying down conditions (i.e. participants chose one of two reference frames), both tropic and embodied responses were classified as correct. For critical trials, participants in the upright condition showed a higher proportion of correct response (M = 99.5%) than those in the lying down conditions (M = 98%), although this difference was not significant ($p > .1$). Similarly, the proportion of correct responses between the 'left side' (M = 98.5%) and 'right side' (M = 97.5%) factors in the lying down condition did not differ significantly ($p > .1$). In total, there were 18 incorrect responses across 1200 critical trials (1 from upright and 17 from lying down conditions).

Embodied vs. Tropic Responses

The proportion of tropic responses across the lying down conditions (90.63%) was lower than in the upright condition (99.5%) suggesting that participants were less likely to use tropic representation to determine their reference frame in the lying down conditions. However, across 800 critical trials in the lying down conditions there were 58 'embodied' responses, 17 'incorrect' responses and 725 'tropic' responses hence further statistical analyses were deemed inappropriate.

Response Time Analysis

Due to the small number of embodied responses in the experiment, RT analyses focused solely on tropic responses. Generalized Score Chi-Square statistics from the Gamma regression analyses in GEE show a significant main effect of body position (left vs. right vs. upright) for both participants, GS $\chi^2_{(2)} = 11.36$, $p < .01$; and items GS $\chi^2_{(2)} = 18.02$, $p < .001$. 95% confidence intervals show that tropic RTs were significantly shorter in the upright condition compared to both the left condition (by-subject contrast: -307 ± 131 ms; by-items: -307 ± 48 ms); and the right condition (by-subject contrast: -261 ± 101 ms; by-items: -261 ± 47 ms). Tropic RTs to the left and right condition did not differ significantly by-subjects, and was marginally different (p = .043) by-items (by-subject contrast: 46 ± 164 ms; by-items: 46 ± 45 ms).

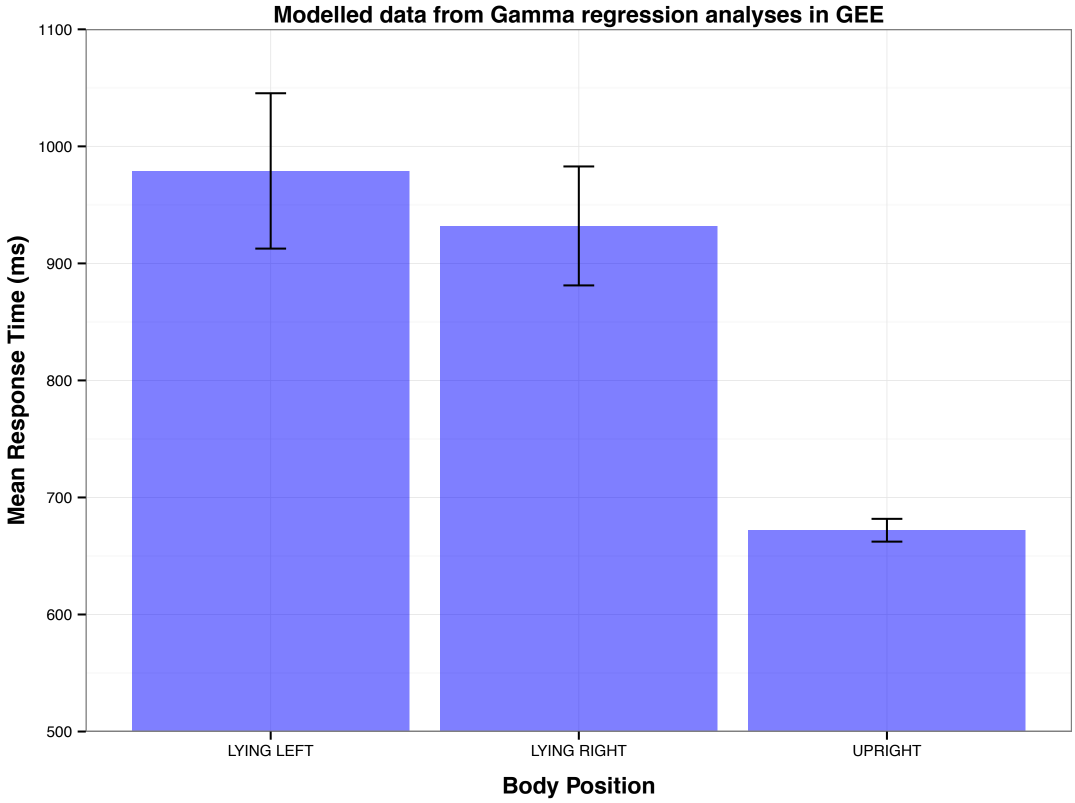


Figure 6.6. Modelled mean response times (ms) and standard errors per condition from the Gamma regression analyses in GEE.

6.2.3 Discussion

Experiment 7 sought to examine whether our mental representations of space are affected by a change in body position. It was hypothesised that a change in body position causing dissociation between vertical body plane and gravity (i.e. lying down) would allow differences in categorisation of verticality to be observed. In essence, we asked whether "up" was still defined by gravity when an individual is lying on one side. The results showed that the vast majority of spatial classifications remained consistent with the plane of gravity, with only 4 (out of 40) 'lying down' participants defining spatial location from an embodied perspective. One factor that may have influenced which conceptual representation people ultimately chose (tropic or embodied) is the context. In the experimental environment, individuals were surrounded by contextual cues – computer displays, chairs, desks, and the experimenter – all of which were in their canonical upright position. Participants may have used these contextual cues to influence their choice of response and ultimately answer from a tropic perspective. However, this assumes that choosing a spatial perspective was a somewhat conscious decision. Indeed, the response time analysis showed significantly increased RTs for participants in the lying down conditions, which could be an effect of response competition. That is, answers in both the tropic and embodied perspectives are correct, hence the participant must decide which perspective to take thereby causing a delay in responding compared to the upright condition, where there is no competition from tropic and embodied perspectives. Whilst the effect of response competition seems plausible, one might assume that it would be highly inefficient to perform such a cognitive process for every trial. Perhaps participants may have experienced response competition on the first few trials, but once the decision was made they continued to answer from that spatial perspective (note that participants also saw 5 practice trials prior to the experimental procedure beginning and hence a decision may have been made here). Indeed, the majority of participants answered consistently from one spatial perspective, suggesting that a separate perspective decision was not made at the onset of each trial. So why is there a significant increase in response times when we lie down?

The increased RTs for participants in the two lying down conditions may be a result of mental rotation. A large body of literature shows that the more an image needs to be mentally rotated, the longer it takes for participants to respond. In their seminal study, Shepard and Metzler (1971) showed that participants took longer to mentally rotate an image of a shape in order to match it to a control shape. Interestingly, the results showed a linear relationship between the angle of rotation and response times such that the greater the amount of mental rotation required, the longer the participant took to respond (a similar pattern of results is shown by Roberts Jr & Aman, 1993). In Experiment 7, regardless of whether a participant was lying on their left or right side, the image was dissociated with the ‘tropic up’ (the vertical plane of gravity) by 90°. The similar RTs shown for participants lying left and participants lying right could be explained by both sets of participants having to perform a 90° mental rotation.

In order to determine whether response competition, mental rotation or another factor is causing the difference in response times, further research is needed. In summary, Experiment 7 shows that making overt judgements on vertical spatial location is affected by body position.

6.3 Introduction to Experiments 8 and 9

The previous experiments presented in this thesis have investigated how perceptual representations of space affect processing, with particular focus on saccadic eye movements. Experiment 7 considered the effect of body position on perceptions of space and showed that when our vertical body position is dissociated with gravity, perceptual representations of spatial location or mental rotation processes seem to affect processing. Experiments 8 and 9 aim to bring the two strands of this thesis together by employing an eye-movement activated lexical decision task similar to those used in Experiments 1 and 2, but combining this with a horizontal body position design as used in Experiment 7.

By using a set of stimuli and methodology that has previously been shown to successfully investigate perceptual representations of spatial location but adding a further experimental manipulation of body position, the two experiments aim to show whether our perceptual simulations of spatial location facilitate processing to an embodied-compatible or tropic-compatible location. Experiment 7 showed that the vast majority of participants defined spatial location from a tropic perspective, however lying on one side and having lexical decision response targets in tropic-up and tropic-down locations (the top and bottom of the display) would also constitute the embodied-left and embodied-right for participants. Similarly, the reverse would also be true – a set of response targets appearing at the left and right of the display would constitute the participant's embodied-up and embodied-down. In summary, Experiment 8 essentially replicates Experiment 1 (vertical lexical decision, see Section 4.2) by showing targets at the top and bottom of the display and Experiment 9 replicates Experiment 2 (horizontal lexical decision, see Section 4.3) by showing targets at the left and right of the display, however as with Experiment 7, participants will be lying on one side.

If perceptual simulations of spatial location are defined from a tropic perspective as suggested by Experiment 7, then Experiment 8 should show faster responses to the lexical decision 'word' targets appearing in a location compatible with the cued concept (relative to the neutral baseline condition). However, if perceptual simulations of spatial location are defined from an embodied perspective, then Experiment 9 should show faster responses (vs. neutral baseline) to the 'word' targets appearing in embodied-compatible locations (i.e. the left and right of the screen).

Experiment 1 used a similar design to Experiment 8 and showed that perceptual simulations of spatial association facilitate responses to compatible locations. Experiment 7 showed that participants in a horizontal, lying down position were more likely to adopt a tropic (vs. embodied) frame of reference. Taken together, we would perhaps expect Experiment 8 to show faster saccadic responses (vs. neutral baseline) to targets appearing in a location deemed tropic-compatible (e.g. hearing an 'up' word and looking to the green target at the top of the screen should be quicker than hearing a 'down' or 'neutral' word and looking to the top of the screen).

6.4 Experiment 8

Experiment 8 uses an eye-movement activated lexical decision task similar to Experiment 1, with the 'word' and 'non-word' targets (small green and red squares) appearing at the top and bottom of the display screen. Whilst undertaking the task, participants were lying on one side and hence the response targets are deemed to be in tropic-up and tropic-down locations (or, embodied-left and embodied-right).

6.4.1 Method

Participants

Thirty-two individuals (20 Female; M=22.6 years) from the University of Glasgow participated in the study, each receiving £4 or participation credits; none of these had participated in the previous experiments. All participants had either normal or corrected-to-normal vision and were native monolingual English speakers.

Materials

The 120 direction words and 120 matched non-words outlined in the *Linguistic Stimuli* section in Chapter 3 were used in this experiment.

Apparatus

The stimuli were presented on an Apple iMac (9,1; mid-2009 model) with 20-inch viewable LCD display and a video refresh rate of 60 Hz. The screen was positioned at a distance of 70 cm from the participant, with the centre of the screen in line with the bridge of their nose. Participants' eye movements were continuously monitored using an SR Research EyeLink 1000 eye-tracker, sampling at 1000 Hz. Although viewing was binocular, only the dominant eye was tracked, as established by a variation of the Miles test (Miles, 1930; Roth et al., 2002). Stimulus presentation and data collection were controlled using Experiment Builder software (SR Research).

Procedure

Participants were lying on either their left or right side (counterbalanced between participants) on a mattress placed at floor level (see Figure 6.7 below). The procedure was almost identical to that of Experiment 1 (see Section 4.2 above). To recap, each participant was presented with 240 auditory stimuli (120 words and 120 non-words, outlined in Section 3.1) in an individually determined random order. Prior to the experimental trials, participants undertook 5 practice trials to ensure they understood the task. Each trial began with the presentation of a central fixation cross for drift correction. While the participant kept looking at the cross and 150 ms after drift correction, the sound file was played via speakers; at the onset of the sound file, a green and a red square appeared on the screen. Each square measured 10 × 10 screen pixels and appeared 200 pixels (4.70º of visual angle) above and below the central fixation cross, respectively. The participant's task was to decide, as quickly and accurately as possible, whether what they had just heard was an actual English word or not by looking at either the green square (if they thought they heard a word) or the red square (if they thought they heard a non-word). The location of the red and green square was counterbalanced across participants; 16 participants had the red square at the top and the green square at the bottom (and vice versa for the remaining 16 participants) for all 240 trials. This between-subject manipulation lowered the chances of participants figuring out the purpose of the study. Each trial terminated when a fixation was detected in one of the target areas (as outlined earlier by the dashed rectangles in Figure 4.2), or after a timeout of 3000 ms, respectively. The target areas for the trial-terminating gaze trigger were defined as the inside edge of the coloured square to the top or bottom edge of the screen (190 pixels, 4.45º) and were 800 pixels (25.59º) wide. Participants were instructed to keep their head as still as possible throughout the experiment, however the experimenter chose to recalibrate and validate (9-point fixation procedure) the eye-tracker if the inter-trial drift correct was greater than 2º of visual angle from the central ellipse. These were supplementary to the mandatory calibration and validations that occurred after every 40 trials. Finally, unlike Experiment 1 the sound files were played through the computer speakers, as headphones were deemed inappropriate for participants lying on their side.

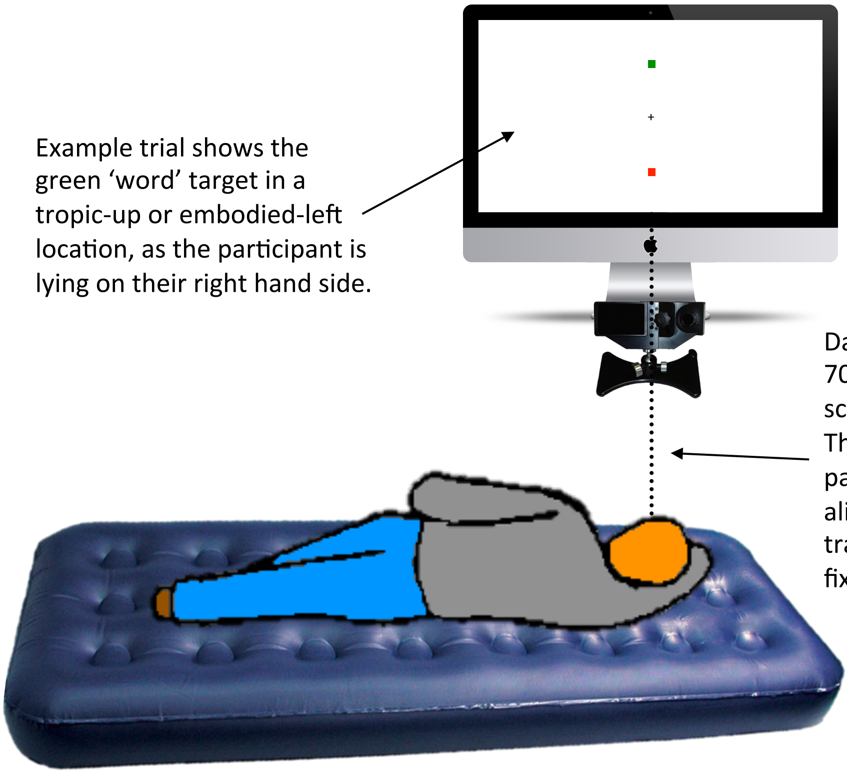


Figure 6.7. Illustration of the experimental setup for Experiment 8. In this illustration, the participant would be required to saccade to the top of the screen (green square) upon hearing a word or bottom of the screen (red square) upon hearing a non-word.

Data Analysis

Lexical decision accuracy was greater than 95% in each condition. Only the critical word trials were considered for analysis. The dependent variable of interest was the *saccade launch latency for correct 'word' decisions* (saccade towards green square), measured from the *offset* of the auditory word presentation until the eye started moving away from the central fixation cross (as determined by saccadic acceleration and velocity thresholds). Trials that contained eye-blinks, multiple saccades, off-target saccades (not landing within 100 pixels of the green square), or terminating after the 3000 ms timeout, were excluded from analysis (affecting ca. 9% of the critical trials). Saccade launch latency outliers of more than 2.5 *SD*s away from the mean of a given subject × condition combination were also removed (affecting less than 4% of the data).

Inferential analyses were based on Generalised Estimating Equations (GEE). We implemented a Gamma regression approach by using a *Gamma* distribution and *Log* link function in the GEE model specifications. The *Log* link required all data points to be positive, which was not always the case because perceivers occasionally started to move their eyes before the end of the word (recall that the latencies were determined relative to word offset). To deal with this, we added a 300 ms constant to each data point before performing the inferential analyses. In the descriptive means reported below, this constant has been removed.

Two types of analyses were performed. In the *by-subject* analysis, word direction ('up', 'neutral', 'down') was entered as within-subjects factor and target location ('tropic-up', 'tropic-down') as between-subjects factor. In the *by-item* analysis, word direction was between- and target location within-items. To account for imbalances in item-specific control variables, the three principal components identified in the principal component analysis (Section 3.1, above) were entered as additional within-subjects / between-items covariates. Analyses without these covariates yielded nearly identical results. The latter will not be reported in detail but see they grey lines in Figure 6.8 for comparison. All analyses assumed an exchangeable covariance matrix for repeated measurements.

6.4.2 Results

Table 6.1 shows Generalized Score Chi-Square statistics from the Gamma regression analyses in GEE. A first point to note is that all three of the control covariates (principal components) had a significant overall influence on saccade launch latency: Lexical difficulty was associated with a positive coefficient (+.078 ± .006 *SE* by subjects; +.073 ± .017 *SE* by items), valence-dominance with a negative coefficient (−.017 ± .005 *SE* by subjects; −.021 ± .017 *SE* by items) and arousal with a negative coefficient (−.041 ± .007 *SE* by subjects; −.040 ± .018 *SE* by-items), meaning that greater lexical difficulty led to an increase, and greater arousal, valence and dominance to a decrease in saccade launch latency times.

Effect		**By Subjects**		**By Items**	
	df	**GS χ^2**	***P***	**GS χ^2**	***P***
Word Direction (W)	2	5.23	0.73	1.42	0.50
Target Location (T)	1	0.071	0.79	1.76	0.19
W × T Interaction	2	0.82	0.67	1.45	0.49
Lexical Access*	1	27.44	< .001	13.23	< .001
Valence-Dominance*	1	7.32	< .01	1.51	0.22
Arousal*	1	18.38	< .001	4.88	< .05

Table 6.1. Inferential results (Generalized Score Chi-Squares, degrees of freedom, p-values) from the Gamma regression analyses in GEE. Control predictors (covariates) are marked with an asterisk.

With regards to our experimental manipulations, there was no significant main effect of word direction or target location and no significant word direction x target location interaction either by-subjects or by-items. Figure 6.8 shows the covariate-adjusted by-subject means and *SE*s per condition (for comparison, non-adjusted means and *SE*s are shown in grey).

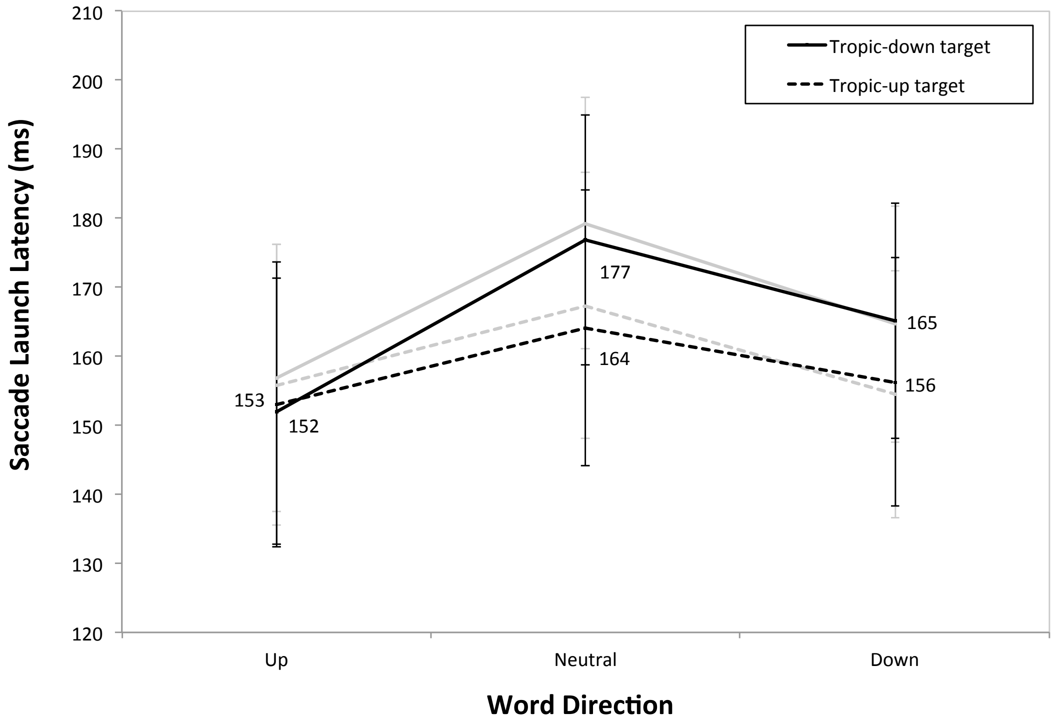


Figure 6.8. By-subject means and SEs per condition. The black lines show covariate-adjusted data and the grey lines show data from an analysis that did not include the covariates.

6.4.3 Discussion

Experiment 8 used an eye-movement activated lexical decision task in order to investigate whether perceptual simulations of spatial location are defined from a tropic perspective. Based on the results of the previously presented experiments (in particular, the facilitatory effect shown in Experiment 1 and the overwhelming amount of tropic responses in Experiment 7), it was predicted that perceptual simulations of spatial location defined from a tropic perspective would lead to shorter saccade launch latency times when the 'word' target was in a cue-compatible location. That is, when hearing an 'up' word (e.g. 'moon'), participants would be quicker to launch a saccade to the 'word' target at the top of the display compared to if they had heard a 'neutral' word (e.g. 'letter'). Similarly, saccades should be launched quicker towards the bottom of the screen (relative to the 'neutral' baseline condition) following a 'down' word (e.g. 'sewer'). The results suggest otherwise.

Firstly, all previous experiments showed a significant main effect of saccade direction, with upward saccades being consistently quicker than downward saccades (Experiments 1 & 3-5) and rightward saccades being quicker than leftward saccades (Experiment 2). The present experiment shows no significant main effect of target location. Whilst one might expect a similar effect to Experiment 2 (as these are horizontal saccades), recall that the side on which a participant was lying was counterbalanced between subjects throughout the present experiment hence both the 'tropic-up target' and 'tropic-down target' conditions consist of leftward and rightward saccades.

Crucially, the results show no significant interaction between the word direction and target location suggesting that perceptual simulations of spatial location neither facilitated nor inhibited processing of a target in a tropic-compatible location. This is discussed below in context with the results of Experiment 9 (see Section 6.6).

6.5 Experiment 9

Experiment 9 uses an eye-movement activated lexical decision task similar to Experiment 2, with the 'word' and 'non-word' targets (small green and red squares) appearing at the left and right of the display screen. Whilst undertaking the task, participants were lying on one side and hence the response targets are deemed to be in embodied-up and embodied-down locations (or, tropic-left and tropic-right).

6.5.1 Method

Participants

Thirty-two individuals (27 Female; M=23.2 years) from the University of Glasgow participated in the study, each receiving £4 or participation credits; none of these had participated in the previous experiments. All participants had either normal or corrected-to-normal vision and were native monolingual English speakers.

Materials

The 120 direction words and 120 matched non-words outlined in the *Linguistic Stimuli* section in Chapter 3 were used in this experiment.

Apparatus and Procedure

The procedure and apparatus were almost identical to Experiment 8. The crucial difference between the two experimental procedures was the location of the target squares. The squares were now located at the left and right sides of the screen. Each square appeared 200 pixels (6.50º of visual angle) left and right of the central fixation cross, respectively. The target areas for the trial-terminating gaze trigger were defined as the inside edge of the coloured square to the left or right edge of the screen (184 pixels, 5.98º) and spanned the entire height of the screen (768 pixels, 17.86º).

Data Analysis

Lexical decision accuracy was greater than 96% in each condition. Only the critical word trials were considered for analysis. The dependent variable of interest was the *saccade launch latency for correct 'word' decisions* (saccade towards green square), measured from the *offset* of the auditory word presentation until the eye started moving away from the central fixation cross (as determined by saccadic acceleration and velocity thresholds). Trials that contained eye-blinks, multiple saccades, off-target saccades (not landing within 100 pixels of the green square), or terminating after the 3000 ms timeout, were excluded from analysis (affecting ca. 10% of the critical trials). Saccade launch latency outliers of more than 2.5 *SD*s away from the mean of a given subject × condition combination were also removed (affecting less than 3% of the data).

Inferential analyses were based on Generalised Estimating Equations (GEE). We implemented a Gamma regression approach by using a *Gamma* distribution and *Log* link function in the GEE model specifications. The *Log* link required all data points to be positive, which was not always the case because perceivers occasionally started to move their eyes before the end of the word (recall that the latencies were determined relative to word offset). To deal with this, we added a 700 ms constant to each data point before performing the inferential analyses. In the descriptive means reported below, this constant has been removed.

Two types of analyses were performed. In the *by-subject* analysis, word direction ('up', 'neutral', 'down') was entered as within-subjects factor and target location ('embodied-up', 'embodied-down') as between-subjects factor. In the *by-item* analysis, word direction was between- and target location within-items. To account for imbalances in item-specific control variables, the three principal components identified in the principal component analysis (Section 3.1, above) were entered as additional within-subjects / between-items covariates. Analyses without these covariates yielded nearly identical results. The latter will not be reported in detail but see they grey lines in Figure 6.9 for comparison. All analyses assumed an exchangeable covariance matrix for repeated measurements.

6.5.2 Results

Table 6.2 shows Generalized Score Chi-Square statistics from the Gamma regression analyses in GEE. A first point to note is that all three of the control covariates (principal components) had a significant overall influence on saccade launch latency: Lexical difficulty was associated with a positive coefficient (+.040 ± .002 *SE* by subjects; +.040 ± .009 *SE* by items), valence-dominance with a negative coefficient (−.009 ± .003 *SE* by subjects; −.009 ± .009 *SE* by items) and arousal with a negative coefficient (−.019 ± .004 *SE* by subjects; −.018 ± .010 *SE* by-items), meaning that greater lexical difficulty led to an increase, and greater arousal, valence and dominance to a decrease in saccade launch latency.

Effect		**By Subjects**		**By Items**	
	df	**GS χ^2**	***P***	**GS χ^2**	***P***
Word Direction (W)	2	6.88	< .05	2.50	0.29
Target Location (T)	1	5.35	< .05	75.99	< .001
W × T Interaction	2	9.13	< .05	7.08	< .05
Lexical Access*	1	27.04	< .001	13.58	< .001
Valence-Dominance*	1	7.97	< .01	0.89	0.35
Arousal*	1	14.18	< .001	3.40	0.07

Table 6.2. Inferential results (Generalized Score Chi-Squares, degrees of freedom, p-values) from the Gamma regression analyses in GEE. Control predictors (covariates) are marked with an asterisk.

With regards to our experimental manipulations, the main effect of word direction (ca. 28 ms higher saccade launch latencies in the ‘neutral’ condition compared to the other word direction conditions) was significant within-subjects but not between-items (presumably due to reduced power in the latter case). The main effect of target location (ca. 90 ms higher saccade launch latencies to embodied-down targets than to embodied-up targets overall) was significant both within-items and between-subjects. Note that previous research (e.g. Dudschig et al., 2013; Goldring & Fischer, 1997; Miles, 1936) and experiments presented earlier in this thesis have shown a similar general disadvantage for downward saccades.

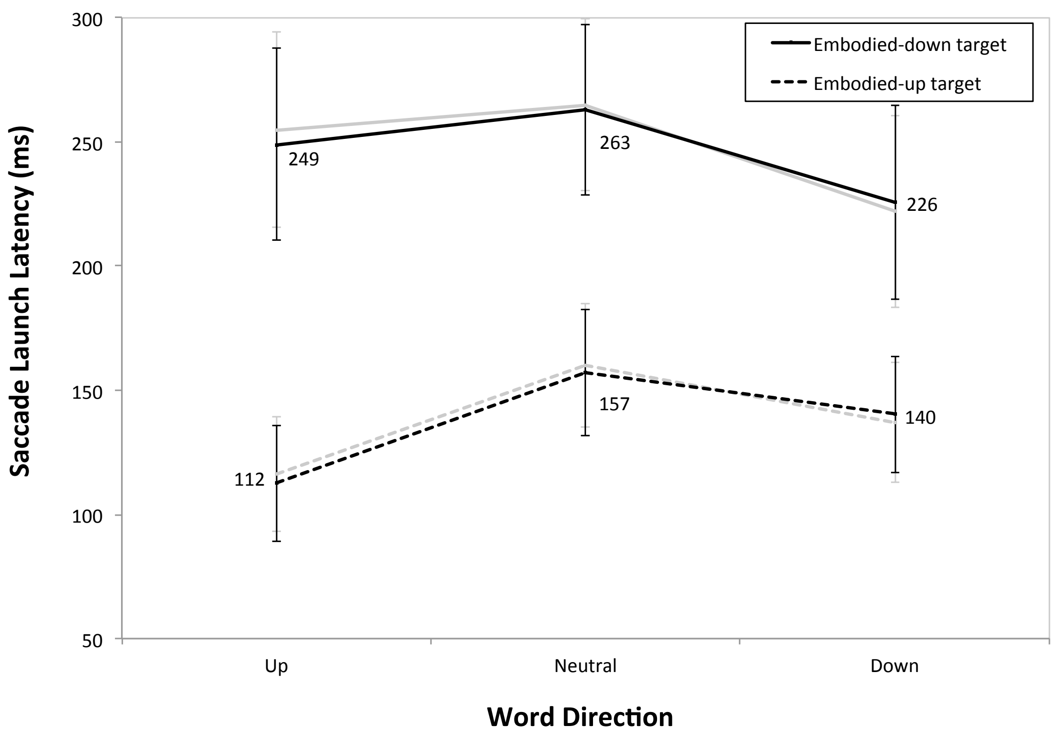


Figure 6.9. By-subject means and SEs per condition. The black lines show covariate-adjusted data and the grey lines show data from an analysis that did not include the covariates.

Most crucially, there was a significant word direction × target location interaction in both the by-subject and the by-item analysis. Figure 6.9 shows the covariate-adjusted by-subject means and *SE*s per condition (for comparison, non-adjusted means and *SE*s are shown in grey). 95% CIs for simple effects showed that saccades to embodied-up targets were launched more quickly upon hearing 'up' words like *moon* than upon hearing 'neutral' words like *letter* (by-subject contrast: 44 ± 24 ms; by-items: 48 ± 44 ms); the comparison between 'down' words like *sewer* and 'neutral' words like *letter* was not significant (by-subjects: 16 ± 18 ms; by-items: 21 ± 50 ms). Conversely, saccades to embodied-down targets were launched quicker after 'down' words like *sewer* compared to 'neutral' words like *letter* (by-subjects: 37 ± 34 ms; by-items: 41 ± 55 ms), whereas the contrast between 'up' words like *moon* and 'neutral' words like *letter* was not significant (by-subjects: 14 ± 35 ms; by-items: 17 ± 53 ms).

6.5.3 Discussion

Experiment 9 used an eye-movement activated lexical decision task in order to investigate whether perceptual simulations of spatial location are defined from an embodied perspective. Based on the results of Experiment 8, it was predicted that perceptual simulations of spatial location defined from an embodied perspective would lead to shorter saccade launch latency times when the 'word' target was in a cue-compatible location.

That is, when hearing an ‘up’ word (e.g. ‘moon’), participants would be quicker to launch an upward saccade to the ‘word’ target compared to if they had heard a ‘neutral’ word (e.g. ‘letter’). Similarly, downward saccades should be launched quicker towards the ‘word’ target (relative to the ‘neutral’ baseline condition) following a ‘down’ word (e.g. ‘sewer’).

Firstly, the significant main effect of target location replicates an effect found throughout this thesis (Experiments 1 & 3-5) and beyond (e.g. Dudschig et al., 2013; Goldring & Fischer, 1997; Miles, 1936). More specifically, saccades to ‘embodied-up target’ locations are by their definition, upward saccades and were consistently quicker than downward saccades (to ‘embodied-down target’ locations).

The significant interaction between word direction and target location shows that saccades were launched more quickly to embodied target locations that were compatible with the spatial location cued by a word. That is, upon hearing an ‘up’ word (e.g. ‘moon’), saccades towards an embodied-up target were launched significantly more quickly compared with having heard a ‘neutral’ word (e.g. ‘letter’). Similarly, saccades to embodied-down target locations were launched more quickly (relative to the neutral baseline condition) following ‘down’ words (e.g. ‘sewer’). These results suggest that perceptual simulations of spatial location evoked by the cue words facilitated processing, and therefore responses, to a compatible embodied target location. These results are discussed below in context with the results from Experiment 8.

6.6 General Discussion for Experiments 8 and 9

Experiments 8 and 9 investigated how body position affected perceptual simulations of spatial location. More specifically, if perceptual simulations of spatial location are defined from a tropic perspective, it was expected that spatial word cues would facilitate processing of a target in a tropic-compatible location. Alternatively, if perceptual simulations of spatial location are defined from an embodied perspective, we expected a similar facilitatory pattern to embodied-compatible target locations. The results are considered in context with those from Experiments 1 and 2, before focusing on the theoretical implications.

Experiment 1 in this thesis (see Section 4.2) showed that in an eye-movement activated lexical decision task, participants were quicker to launch a vertical saccade to the ‘word’ target when it was in a location compatible with the cue word. Experiment 2 (Section 4.3) used a similar design, but with horizontal targets, and showed no significant interaction between the spatial location implied by the cue word, and the target location. Experiment 8 was introduced as “essentially replicating Experiment 1”, but with participants lying on one side (similarly with Experiment 9 replicating Experiment 2). However, when considering the procedure from the perspective of the participant, or indeed from a so-called ‘embodied’ perspective, we can see that Experiment 8 could also be framed as a replication of Experiment 2 (and similarly, Experiment 9 replicating Experiment 1). That is, from the perspective of the participant the vertical tropic targets (top and bottom of the display) are actually more like embodied horizontal targets. As arbitrary as this may seem, the results to Experiments 1, 2, 8 and 9 suggest that this could be a more representative way of explaining the theoretical implications. Figure 6.10 shows the results of the horizontal lexical decision task (Experiment 2) plotted alongside the results from Experiment 8. Given the striking similarities between the two patterns of results it is perhaps more insightful to consider Experiment 8 as having embodied-horizontal targets.

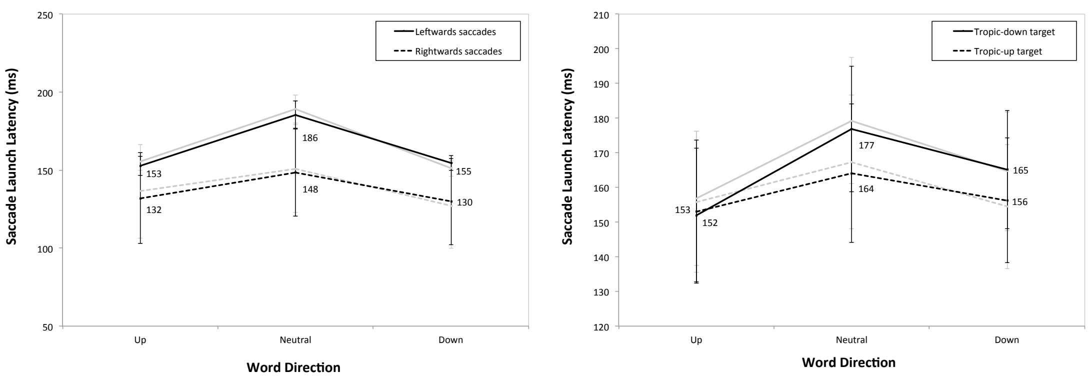


Figure 6.10. Results from Experiment 2 (horizontal lexical decision, left graph) presented alongside those from Experiment 8 (right graph).

Furthermore, Figure 6.11 shows the results of the vertical lexical decision task. (Experiment 1) plotted alongside the results of Experiment 9. Again, considering the similar patterns of results, we should continue to consider the results of Experiment 9 from an embodied perspective. Overall, the results from Experiments 8 and 9 show that perceptual simulations of space facilitate the processing of a cue in an embodied-compatible location.

Myachykov et al. (2013) argue that tropic mental representations should be free from constraining factors such as body position. The results presented here show that manipulating body position seems to affect so-called ‘tropic’ mental representations of spatial location. However, it is possible that the facilitatory pattern of results shown in Experiment 9 is due to a heavily ingrained visuo-motor association between experiential knowledge and eye movements. That is, part of our experiential knowledge associated with a concept is its spatial location, and more often than not our previous experiences involve looking towards an object. Hence part of our knowledge about the moon, is that it resides in the sky and therefore attending to the moon generally requires an upward saccade. With this in mind, the results shown in Experiment 9 may not necessarily reflect a 'rotated' perceptual simulation such that 'up' is represented as above our heads, but could instead be the result of visuo-motor association between our knowledge of the moon's location and saccading upward. In this case, the results do not disagree with the stability of a tropic representation (Myachykov et al., 2013).

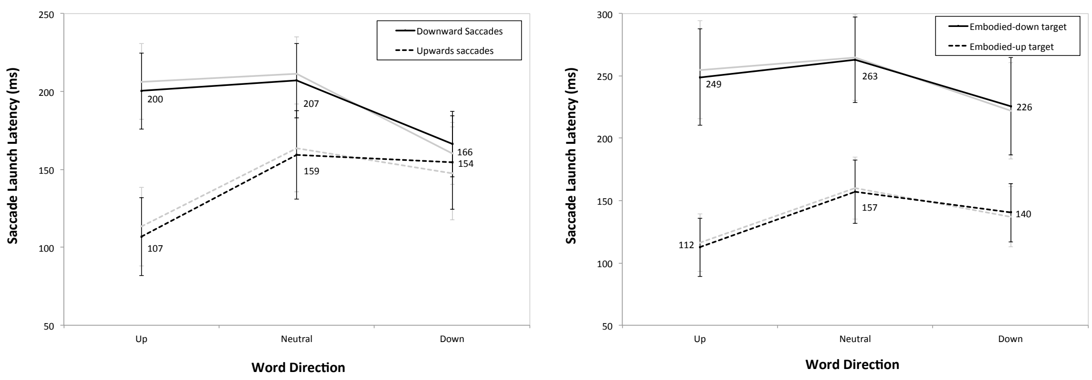


Figure 6.11. Results from Experiment 1 (vertical lexical decision, left graph) presented alongside those from Experiment 9 (right graph).

Conversely, one could argue that the results to Experiments 8 and 9 support the idea that perceptual simulations of spatial location seem to be (what Myachykov et al., 2013 would call) 'embodied' or 'situated' representations. That is, the bodily state becomes part of the conceptual representation (embodied) or the representation is affected by the context in which it is recalled (situated). It is worth pointing out, that further research is needed before making any conclusions about the results. However, it is possible to speculate – and if perceptual simulations of spatial location were embodied (or situated), then one would expect them to be affected by the position of the 'simulator'. By their definition, mental representations are internal thus a perceptual simulation of spatial location relative to our own position in space seems logical. Given that a number of previously referenced theoretical and empirical works have shown a preference for egocentric descriptions of space (Clark, 1973; Epley et al., 2004; Miller & Johnson-Laird, 1976), it is certainly plausible that our mental representations of space are also determined from an egocentric (or in this case, embodied) perspective.

If our mental representations of space are defined by our body position, why were there only a handful of participants answering in an 'embodied' perspective in Experiment

7? Recall that participants were presented with a series of arrows pointing in varying directions and answered questions such as “how many arrows point up?” The majority of these responses (over 90%) were given from a tropic perspective, with the RTs for those lying down being significantly longer than for those sitting upright. We suggested that this might be a result of response competition. In fitting with the results from Experiment 8 and 9, participants may initially represent spatial locations from an embodied perspective (resulting in the facilitatory pattern of results shown in Experiment 9), but respond from a tropic perspective. If this is the case, then the tropic responses in Experiment 7 are affected by the retrieval context and are what Myachykov et al. (2013) would call ‘situated’. Indeed, it could be the case that having to overtly communicate the response (to the experimenter) affected which response was chosen. There is a plethora of research showing that communicating spatial locations to others leads to individuals taking the perspective of the other person, rather than their own (Schober, 1993, 1995; Tversky & Hard, 2009; Tversky, Lee, & Mainwaring, 1999). This would explain why different perspectives seem to be taken between Experiment 7 and Experiment 9; however further work would be necessary to confirm this prediction. A similar design to Experiment 7 could be used with participants using a button box, or with the experimenter removed from the room therefore removing the communicative aspect of the design. If the above explanation were accurate, one would expect a lower proportion of tropic responses (and thus a higher proportion of embodied responses).

In terms of grounded cognition, a combination of the above explanations is most likely. When considering how heavily ingrained visuo-motor associations affect the present results, we argue that these associations stem from our experiential knowledge and are therefore part of a mental simulation, which becomes active upon concept retrieval. Although it *seems* as if our mental representations of vertical space are affected by body position, we argue that the results support the ‘tropic’ definition outlined by Myachykov et al. (2013). That is, the actual mental representation is unaffected by our position in space and the results shown here do not reflect a change in the way vertical space is actually represented. Instead, the results show that our visuo-motor associations of vertical space are so heavily ingrained, that responding with an upward or downward saccade is not affected by our body position. In hindsight, perhaps eye-movement methodologies do not allow for an accurate measure of perceptual simulations, but rather they measure the visuo-motor associations between conceptual knowledge and space.

In this case, one might question the rationale for using an eye-movement methodology. We chose to use eye movements as a measure of attentional shifts that may not be reflected in other methods, such as hand or arm movements (i.e. button-pressing). Visual world experiments have shown the efficacy of eye movements as a real-time measure of attention (e.g. Tanenhaus, Spivey-Knowlton, Eberhard, & Sedivy, 1995). Indeed with regards to language processing, eye movements are directed to aspects of a visual scene described by the concurrently presented auditory stimuli – and in some cases, anticipatory looks occur prior to the object's reference (Altmann & Kamide, 1999). In our experiments, the visual referent (the target square) does not share features with the auditory stimuli, other than the spatial association; however, given the close link between language processing and eye movements as a measure of attention, we felt that an eye-movement methodology was most suited to investigating our research question. It would potentially be feasible to conduct Experiments 8 and 9 using a hand movement measure (similar to Lachmair et al., 2011), but it is unclear whether we would find a similar pattern of results.

In summary, the results to Experiments 8 and 9 show that perceptual simulations of spatial location facilitate processing of a target in an embodied-compatible, but not tropic-compatible location. However, we conclude that this is due to a visuo-motor association between our experiential knowledge of a concept and saccading towards it, rather than evidence for a body-position specific 'embodied' perceptual simulation such that our representations of vertical space are defined by our position in space (i.e. 'up' is defined as above the head).

Chapter 7 Implications and Conclusions

This thesis presents a series of experiments aiming to investigate how mental simulations of spatial location impact on processing. Theories of grounded cognition propose that the meaning of a concept, which stems from acquired experiences, is encoded in sensory-motor systems (e.g. Barsalou, 1999, 2008). Of specific relevance, if a concept has a prominent vertical spatial location, then later knowledge retrieval will activate a mental simulation of vertical space. The experiments here explored the conflicting patterns of results in the literature whereby the spatial location of a concept has been shown to both facilitate (e.g. Dudschig et al., 2013) and inhibit (e.g. Estes et al., 2008) processing and responses to a compatible location. Moreover, Part 2 explored whether our mental simulations of vertical space are affected by our body position.

7.1 Summary of Experimental Findings

Experiments 1 to 5 (Chapter 4) showed that perceptual simulations of spatial location can both facilitate and inhibit processing depending on the task demands. By including a neutral baseline measure in our experiments, we can reliably conclude facilitatory or inhibitory effects. As baseline conditions were not included in previously published studies alluded to throughout this thesis, the conflicting pattern of results may have been due to misinterpretations. Experiment 1 demonstrated that perceptual simulations of spatial location facilitate the launching of a saccade towards a compatible target location. The results showed no sign of an inhibitory effect for responses to incompatible vertical locations. Similarly, Experiment 2 showed no sign of inhibitory effects for responses to incompatible horizontal target locations. As such, we conclude that perceptual simulations of spatial location activated by an ‘up’ or ‘down’ concept overlapped with the saccade planning mechanisms and therefore facilitated the launching of a saccade towards a compatible spatial target location. This explanation seems to support the theory of grounded cognition as the vertical ‘spatial location’ of a concept is part of our experiential knowledge, and hence is activated upon concept retrieval (i.e. hearing the stimulus word).

Experiment 3 attempted to investigate the previous results in terms of the perceptual-featural overlap hypothesis proposed by Estes et al. (2008). This theory assumes that the spatial location activated by a concept would only facilitate responses to compatible locations if there were ‘featural overlap’ between the cue and target. With regards to our earlier experiments, there was no featural overlap between the cue words and the green target square. However, we hypothesised that the spatial location itself may indeed be an overlapping feature and hence we should expect a facilitatory pattern of responses to a target appearing in a compatible location. Instead, the results showed a small but significant inhibitory effect, whereby saccades launched to incompatible vertical target locations were slower than the neutral baseline. Furthermore, there was no sign of a facilitatory effect. A crucial finding of Experiment 3 was that without a neutral baseline, we would have concluded a compatibility effect, as responses to compatible target locations were significantly quicker than those to incompatible target locations. This observation alone highlights the fundamental importance of including a baseline condition in similar future experiments. With regards to the differing pattern of results from Experiment 1, we began to manipulate different aspects of the task.

Experiments 4 and 5 manipulated two aspects of the task that were suspected of influencing the pattern of response behaviour; namely, the stimulus onset asynchrony (SOA) between the cue word and the response target, and the number of possible target locations. The results showed no conclusive pattern of responses at different SOAs, with no significant interaction between stimulus onset time, word direction and target location. However, Experiment 5 showed that simply replicating the design of Experiment 3 with less target locations changed the pattern of responses from an inhibitory to a facilitatory effect, similar to that shown in Experiment 1. From this, we concluded that the cued spatial location must be compatible and useful for the response behaviour to be facilitated. Nonetheless, this still does not explain the pattern of results shown by Estes et al. (2008), and the replication of these results shown by Gozli et al. (2013). Realistically, further research is needed in order to determine the intricacies that lead to differing patterns of results. We propose that one of those intricacies is the use of eye-movement responses. Given the close link between eye movements and attention, especially with regards to mental representations of linguistic stimuli (e.g. Altmann & Kamide, 1999), it may be the case that our research paradigm is measuring something different to those reporting different response behaviour (such as ‘button box’ responses or keyboard presses). It is unclear whether a methodology similar to our own using hand movement responses would show similar patterns of results. Indeed, whilst there are instances of two-target designs showing an inhibitory effect (Estes et al., 2008; Gozli et al., 2013), there are similar designs showing a facilitatory effect (Dudschig, de la Vega, De Filippis, & Kaup, 2014; Dudschig et al., 2012; Lachmair et al., 2011). Furthermore, in a recent study Dudschig and Kaup (in press) used a hand movement methodology and a 4-target design (similar to our Experiment 3). Yet, in contrast to Experiment 3, Dudschig and Kaup (in press) report a facilitatory pattern[12]. It is clear that eye-movement results are not directly generalizable to hand movement paradigms, but at this stage we are unable to decipher whether these two methods are measuring something different from one another, or if there is a methodological issue causing the discrepancy in response patterns. Indeed, to address this issue one may consider an experiment using both hand and eye-movement methods in order to allow a direct comparison between the two. Furthermore, as stated above it should be paramount that future experiments use a baseline condition in order to reliably conclude the direction of any (in)compatibility effects.

In summary, we draw two major conclusions from Experiments 1 to 5. Firstly, in order to truly show a facilitatory or inhibitory pattern of response behaviour, the use of a neutral baseline condition is an unquestionable necessity. And secondly, perceptual simulations of spatial location activated by ‘up’ and ‘down’ concepts impact on processing and response behaviour – however, the mechanisms of this impact seem task dependent.

Experiment 6 (Chapter 5) investigated whether perceptual simulations of spatial location activated by ‘up’ and ‘down’ concepts can influence the execution of a saccade. A number of saccade trajectory measures showed no significant differences in the responses to our direction words. We concluded that perceptual simulations seem only to impact on saccade planning mechanisms and not saccade execution.

Part 2 of this doctoral thesis explored how manipulating body position affected mental representations of vertical space. Experiment 7 showed that making judgements about vertical or horizontal space took significantly longer when our body is not aligned with the perceived stimulus direction. We conclude that this is most likely an effect of mental rotation, however it is possible that competing responses (in this case, tropic vs. embodied) may also affect response times. Importantly, as the vast majority of responses were in the ‘tropic’ dimension, the results seem to support the idea proposed by Myachykov et al. (2013) that tropic mental representations of space are stable and unaffected by body position or retrieval context.

Experiments 8 and 9 used a similar methodology to Experiments 1 and 2 with the added manipulation of body position. Taken together, the results show that perceptual simulations of spatial location facilitate the launching of a saccade towards an embodied-compatible location. Crucially, saccades launched towards tropic-compatible locations

[12] Note that none of these studies included a baseline condition, hence we cannot confirm whether an effect is facilitatory or inhibitory.

showed no facilitatory or inhibitory effects. From these results, our preliminary conclusion is that the facilitatory effect represents heavily ingrained visuo-motor associations between vertical space, eye movements and experiential knowledge of a concept. However, in order to confirm this further research is required, as these results do not rule out the possible effect of a 'rotated' mental simulation whereby 'up' is determined from an egocentric reference frame (i.e. up is above our head).

A final interesting point to note regarding the results presented within this thesis is the role of the principal components. In the target detection experiments (3, 4 and 5) the three principal components had no significant effect on the dependent variable (saccade launch latency). However in the lexical decision experiments using the same dependent variable (1, 2, 8 and 9), the principal component Lexical Access had a consistent positive effect, and Arousal had a consistent negative effect, on saccade launch latency. Furthermore, in Experiments 8 and 9 the principal component Valence-Dominance showed a negative effect on saccade launch latency.

Warriner et al. (2013) state that low arousal ratings indicate the stimuli word induces feelings such as 'excited' or 'stimulated', whereas high arousal scores indicate 'relaxed' or 'calm' feelings. Similarly, low valence ratings related to feelings of unhappiness (high ratings reflected happiness) and low dominance ratings relate to feeling in control (high ratings reflected feeling controlled). The four experiments presented here showing a negative effect of arousal (Exps. 1, 2, 8 and 9) suggest that the more relaxed or calm the participant feels upon processing a word, the faster their lexical decision response was. Furthermore, the two experiments (8 and 9) showing a negative valence-dominance effect suggest that words inducing feelings of unhappiness and/or feeling in control led to shorter lexical decision response times. Indeed, the results presented here could reflect an experiential component of word processing such that experiencing feelings of calm and being in control lead to better (or faster) processing capabilities, which are reflected in the motor response times. Similarly, one could argue that the feelings of being controlled and lacking autonomy (i.e. feelings induced by high dominance words) could lead to slower processing, and ultimately slower response times. This explanation seems to fit with a grounded cognition approach whereby a previously encountered feeling (or feelings) is mentally simulated upon later recall of, in this case, a word's meaning.

However, given that we do not see these covariate effects in the target detection tasks it could be argued that participants did not mentally simulate feelings and hence they did not significantly affect response times. Experiment 3 and Experiment 5 both show an effect of word direction – so why would participants mentally simulate the verticality aspect of a concept, but not the affective aspect? One alternative explanation suggests that there is a general processing advantage for emotional words over neutral words (Kousta, Vinson, & Vigliocco, 2009). In our experiments, it may be that this processing advantage is only apparent in tasks requiring 'deeper' lexical processing such as lexical decision (and not target detection). Moreover, this account may explain the significant effect Valence-Dominance had in Experiments 8 and 9 – perhaps an emotional bias existed between our word direction categories (i.e. happy could be 'up') thus the effect of valence becomes significant (as part of the Valence-Dominance covariate) when verticality is potentially conflicted. We should be aware though, that covariates were analysed as main effects only as we were only interested in controlling for any potential effects rather than examining them further hence we are unable to conclude an interaction between a principal component and an experimental factor.

7.2 Theoretical Implications

Theories of grounded cognition postulate that language processing is grounded in our sensory-motor experiences rather than solely in amodal systems (e.g. Barsalou, 1999, 2008). Furthermore, they hypothesise that language comprehension involves a mental simulation, or re-enactment stemming from our experiential knowledge. The majority of experiments in this thesis worked with the premise that comprehending a concept with a prominent vertical spatial location would lead to a mental simulation, part of which involved reactivating experiential knowledge regarding the concept's perceived location in space. For example, hearing 'moon' would activate a mental simulation of an 'upward' spatial location. The results throughout this thesis seem to support the grounded theory of language processing with evidence for perceptual simulations being activated during language comprehension. In other words, our results show that linguistic processing involves people mentally simulating at least one aspect of their experiential knowledge of a concept – in this case, the perceived spatial location. In turn, these simulations seem to directly impact response behaviour. Furthermore, conceptual processing of vertical space is unaffected by retrieval context – either in terms of the body position of the comprehender (embodied), environmental cues (situated), or a combination of both.

Indeed, the visuo-motor associations between vertical eye-movements and experiential knowledge are potentially the epitome of grounded language comprehension. Moving our eyes is part of our experiential knowledge acquired during actual interaction with a concept – such as looking up to see the moon. Hence during later concept retrieval, this motor action is reactivated as part of our mental representation. Our results show that the reactivated motor action (in this case, simulated eye-movements) affects the actual motor action required for responses in the task (saccadic eye-movements). Although we cannot be sure that a mental simulation occurs, it is highly unlikely that motor responses would interact with language processing mechanisms without a mental simulation. Furthermore, whilst the results presented throughout this thesis show that simulated eye-movements affect actual eye-movements in different ways (facilitation vs. inhibition), this too is in fitting with grounded theories. That is, theories of grounded cognition state that retrieval of a perceptual symbol (and the resulting mental simulation) is dynamic and affected by, for example, contextual influences such as task demands (Barsalou, 1999). To this end, it becomes clear how differing patterns of results seem to arise from perceptual simulations of repeated concepts such as in the experiments outlined here.

The question now becomes, can grounded theories predict rather than just explain the differing patterns of results? Since undertaking the experiments in this thesis, subsequently published research suggests that a perceptual simulation account is unable to predict, or even fully explain, the conflicting facilitatory and inhibitory patterns of results (Estes et al., 2015). That is not to say that perceptual simulation does not occur, but rather that something other than the simulation itself affects the results. Estes et al. (2015) proposes the perceptual matching hypothesis whereby 'object' and 'location' are coded separately – thus, when we are cued with the word "moon" and an image of the moon ('object' match) appears in the upper visual field ('location' match) our response is facilitated. Conversely, an object or location mismatch (e.g. an unrelated target object in the upper visual field, or an image of the moon in the lower visual field) would lead to an inhibitory pattern of results. Whilst Estes et al. (2015) provide a number of experiments supporting their theory, it fails to explain results shown throughout this thesis and beyond (e.g. Dudschig et al., 2013) whereby a facilitatory effect is shown despite an 'object' mismatch (i.e. using a small square or circle as a target).

With relevance to the present work, Estes et al. (2015) imply that eye-tracking paradigms specifically measure attentional cueing rather than the effects of perceptual simulation; however, this again fails to explain the patterns of results in this thesis. If our results were to be explained in terms of attentional cueing, one would expect that vertical saccades launched after 'up' and 'down' words would always differ to those launched after 'neutral' words as 'neutral' words should not cue attention to a vertical response location. Furthermore, how would attentional cueing explain the inhibitory pattern of results shown here in Experiment 3?

Another interesting aspect of the work presented by Estes et al. (2015) is how the visual stimulus affects responses. As alluded to above, an ‘object’ and ‘location’ match facilitates responses (e.g. an image of the moon appearing in the upper visual field), however it is interesting to consider how such a setup would affect responses in, for example, Experiments 8 and 9 presented here. Indeed, we considered using a similar methodology whereby participants would hear a word and look towards a depicted object in either a tropic-(in)compatible or embodied-(in)compatible location (i.e. hearing ‘bird’ and looking to an image of a bird of the top of the screen would be tropic-compatible, or embodied-incompatible). Ultimately, we decided against this due to the confounding factors involved with whether to rotate the image on the vertical plane – we felt that participants may take their reference frame from the (un)rotated image[13]. Still, it is interesting to consider how using a visual world design could affect responses. Estes et al. (2015) show that responses to compatible images in a compatible location are facilitated. Visual world experiments have shown that participants are quicker to fixate an object related to the auditory stimuli (i.e. a word or sentence) compared to unrelated distractors. This effect has been shown in a number of different ways. For example, with semantically related objects such as hearing ‘trumpet’ and fixating a piano, or hearing ‘Africa’ and fixating an image of a lion (Cooper, 1974; Huettig & Altmann, 2005); and, with functionally similar sensory-motor objects, i.e. hearing ‘piano’ and fixating a typewriter. Moreover, participants have been shown to map the perceptual aspects of a concept onto a compatible visual scene. For example, participants are more likely to look to objects that share a similar shape or colour to the auditory referent compared to unrelated objects (Huettig & Altmann, 2004, 2011). If it was possible to control for such potentially confounding effects, we could investigate the spatial associations between concepts, whilst also manipulating body position. Hence, a ‘visual world’ style experiment with participants lying on one side may show faster saccadic responses to objects associated with certain spatial locations, such that hearing the word ‘sky’ may lead to a faster saccadic response to an image of a bird. Crucially, the experiment would investigate whether the response is facilitated when a bird appears in a tropic- or embodied-compatible location. Such an experiment would allow insight into how internal representations of space are mapped onto a more representative and lifelike (compared to red and green squares) visual scene.

[13] For example, rotating an image of a plane (✈) may not prime a reference frame due to our experience of planes taking off and landing, however rotating an image of a bicycle (🚲) probably would prime a reference frame as bicycles are very rarely seen propped up on the front/back wheel.

7.3 Conclusion and Closing Remarks

The research presented in this doctoral thesis shows that individuals mentally simulate the spatial location of a concept during understanding, supporting the theory of Grounded Cognition (Barsalou, 1999, 2008). It is unclear whether such simulations are a necessary part of the comprehension process or whether they are a peripheral or epiphenomenal occurrence. However, regardless of whether they are required as an explicit part of the task or response, it seems that they do occur.

Furthermore, we show that the way in which mental simulations of spatial location impact on processing and response behaviour differs dependent on task demands. Whilst recent attempts have been made to explain the differences in response behaviour (e.g. Estes et al., 2015), we conclude that further work is necessary. In particular, the 'facilitation vs. inhibition' conflict is unlikely to be fully explained without the use of baseline conditions to clarify the direction of effects. Similarly, in order to understand how mental simulations of spatial location impact on language processing, future research should consider the full extent of language rather than just words.

Language is not just words; language involves numerous other aspects, for example sentences, grammar, interaction, and communication. One avenue for future research should build on the results presented in this thesis and expand to include the other fundamental aspects of language. Our research shows that the word "jump" leads to an 'up' mental simulation of spatial location; but would that still be true in an incongruent sentential context. If "the man jumps quickly off the diving board", then our 'upward' simulation, relevant at the verb "jump", becomes incongruent with the final location of the agent – he has gone downward from his original location. The automaticity of mental simulations has been outlined in this thesis and one would perhaps expect an 'upward' simulation to occur post-verb in the sentence above (i.e. shortly after "jumps"). However this may not be the case, perhaps the sentential context overrides the initial mental simulation and a congruent 'downward' simulation would occur in fitting with the described event.

Finally, future research should aim to build on the effects shown here (i.e. the 'grounded' interaction between language and eye movements) using sophisticated neuroimaging techniques such as functional magnetic resonance imaging (fMRI), magnetoencephalography (MEG), and repetitive transcranial magnetic stimulation (rTMS) if we wish to understand the dynamics of mental simulations and ultimately create a testable computational model of language comprehension.

Appendices

Appendix I – Materials for Experiments 1-6, 8 and 9.

Mean Rating refers to the mean spatial association rating from the norming procedure outlined in Section 3.1.

Word	Direction	Mean Rating
Basement	Down	-3.143
Bottom	Down	-3.591
Burying	Down	-4.040
Capsizing	Down	-3.500
Cellar	Down	-3.136
Collapsing	Down	-4.074
Crushed	Down	-3.000
Declining	Down	-3.704
Decreasing	Down	-3.519
Deep	Down	-3.727
Deflating	Down	-3.808
Depreciating	Down	-3.346
Depressed	Down	-3.227
Descending	Down	-4.222
Digging	Down	-3.481
Diver	Down	-3.182
Diving	Down	-3.769
Drain	Down	-2.818
Dripping	Down	-3.391
Drooping	Down	-3.280
Droopy	Down	-2.818
Dropping	Down	-4.074
Ducking	Down	-3.400
Fainting	Down	-3.667
Falling	Down	-4.481
Hell	Down	-4.273
Low	Down	-2.864
Mineshaft	Down	-3.727
Negative	Down	-3.636
Plunging	Down	-3.667
Quicksand	Down	-2.864
Root	Down	-2.955
Sewer	Down	-3.045
Sinking	Down	-4.111
Slumped	Down	-2.818
Slumping	Down	-3.520
Submerging	Down	-3.333
Subway	Down	-2.818

Sunken	Down	-3.409
Underground	Down	-3.714
Belt	Neutral	-0.045
Butting	Neutral	-0.185
Car	Neutral	0.273
Catching	Neutral	0.083
Direct	Neutral	0.364
Drifting	Neutral	0.107
Eastern	Neutral	-0.045
Equator	Neutral	-0.045
Forced	Neutral	-0.273
Grass	Neutral	-0.045
Hastily	Neutral	0.136
Imposing	Neutral	0.091
Letter	Neutral	-0.091
Loping	Neutral	0.043
Middle	Neutral	-0.091
Nodding	Neutral	0.077
Passing	Neutral	0.107
Payphone	Neutral	-0.136
Postbox	Neutral	0.095
Pulling	Neutral	-0.120
Ramming	Neutral	-0.160
Returning	Neutral	0.000
Rolling	Neutral	0.083
Sauntering	Neutral	0.083
Scattering	Neutral	-0.037
Serving	Neutral	0.200
Skimming	Neutral	0.160
Spreading	Neutral	0.120
Still	Neutral	0.000
Subsequent	Neutral	-0.182
Surface	Neutral	0.182
Swapping	Neutral	0.167
Swimming	Neutral	0.000
Swiping	Neutral	-0.038
Table	Neutral	0.227
Traffic	Neutral	0.045
Turning	Neutral	-0.185
Waistline	Neutral	-0.143
Wide	Neutral	0.429
Wiping	Neutral	0.000
Adding	Up	3.400
Aeroplane	Up	3.333
Ascending	Up	3.556
Building	Up	3.185

Ceiling	Up	3.591
Climbing	Up	3.786
Cloud	Up	3.409
Elevated	Up	4.000
Elevating	Up	4.240
Erecting	Up	3.852
Erupting	Up	3.708
Flying	Up	4.192
Growing	Up	3.643
Heaven	Up	4.364
Heavenly	Up	3.818
High	Up	4.000
Hoisting	Up	3.423
Inflating	Up	3.481
Jumping	Up	3.080
Launching	Up	3.625
Leaping	Up	3.259
Levitating	Up	3.360
Lifting	Up	3.346
Moon	Up	4.000
Mountain	Up	4.190
Mounting	Up	3.310
Raising	Up	3.542
Reaching	Up	3.070
Rising	Up	4.231
Rocket	Up	3.545
Satellite	Up	3.381
Sky	Up	4.182
Soaring	Up	3.769
Springing	Up	3.000
Stars	Up	4.000
Sun	Up	4.136
Tall	Up	3.318
Top	Up	3.636
Tower	Up	3.381
Towering	Up	3.591

Appendix II – Filler materials for Experiments 3, 4 and 5.

Filler words		
Absolutely	Environmental	Prisoner
Acquisition	Exploited	Promote
Addition	Facility	Protest
Address	Fight	Pursue
Agreement	Finger	Radiation
Airport	First	Relish
Ambassador	Fix	Repeat
Application	Flat	Reserve
Approach	Flavour	Result
Arrest	Focus	Return
Authority	Gene	Rival
Back	Graphic	Rose
Bath	Greece	Rule
Bend	Grip	Rules
Birth	Guest	Science
Block	Handle	Seriously
Branch	Heat	Serve
Careers	Hostage	Shape
Centre	Hotel	Sight
Cleared	Impossible	Silk
Clinic	Institute	Slipped
Colleague	Island	Smell
Collect	Joint	Smile
Commander	Lake	Soap
Communication	Loyalty	Somebody
Compare	Monday	Sport
Convince	Myth	Spot
Cook	Negotiation	Spread
Copy	Obvious	Straight
Correspondent	Occasion	Tail
Cross	Offer	Tendency
Cute	Organization	Territory
Demonstration	Otherwise	Tongue
District	Package	Toward
Dress	Past	Tradition
Dynamic	Pointing	Uncertainty
Economic	Popularity	Unlikely
Emerged	Prepared	Wanted
Engineer	Presence	Wonderful
Entire	Press	Worrying

Appendix III – Critical instructions given to participants for each experiment.

Below is a list of instructions provided to participants prior to each experiment. Lebois, Wilson-Mendenhall, and Barsalou (2015) have shown that making participants aware of the verticality manipulation can affect subsequent responses. We consistently aimed for participants to remain unaware of the 'word direction' manipulation such that the instructions never explicitly mentioned verticality as part of the task.

Two sets of instructions were given to participants for each experiment: one written, and one verbal. These are presented below.

Written Instructions

Lexical Decision Experiments (1 & 2)

This study investigates your language processing performance for words and nonwords. You will be asked to listen to a mixture of 240 words and nonwords and decide whether you think what you hear is a real word or not by looking at a target square.

Target Detection (4-targets) and SOA Experiments (3 & 4)

This study investigates your recognition memory performance after a simple target detection task. You will be asked to listen to 240 words and look as quickly as you can to a small green square that will appear above, below, or either side of the central fixation. After every 60 words, you will fill in a word recognition questionnaire.

2-Target Detection Experiment (5)

This study investigates your recognition memory performance after a simple target detection task. You will be asked to listen to 240 words and look as quickly as you can to a small green square that will appear above or below the central fixation. After every 60 words, you will fill in a word recognition questionnaire.

Saccade Trajectory Experiment (6)

This study investigates your language processing performance for words and nonwords. You will be asked to listen to a mixture of 240 words and nonwords and decide whether you think what you hear is a real word or not by looking at a 'Y' or 'N' appearing either side of the screen.

Tropic vs. Embodied Experiment (7)

This experiment investigates colour and shape perceptions during different levels of relaxation. You will be asked a series of simple questions about the colours, numbers, letters, and shapes that appear onscreen whilst in a variety of relaxing positions.

Lexical Decision Experiments (8 & 9)

This study investigates your language processing performance whilst in a relaxing position. You will be asked to listen to a mixture of 240 words and nonwords and decide whether you think what you hear is a real word or not by looking at a target square.

<u>Verbal Instructions</u>

Lexical Decision Experiments (1 & 2)

You are first required to look at the central fixation cross. You will hear a word or nonword play through the headphones, upon which two target squares will appear. If you think you hear a real English word, look to the green square appearing at the [top / bottom / left / right] of the screen. If you think what you hear is not a real word, you should look to the red square at the [top / bottom / left / right] of the screen. The squares will always be in the same place, and you should answer as quickly and accurately as possible.

Target Detection (4-targets) and SOA Experiments (3 & 4)

You are first required to look at the central fixation cross. You will hear a word play through the headphones, upon which a target square will appear. The square will either appear in the top, bottom, left, or right of the screen. You task is to look to the square as soon as it appears. Try not to predict the location of the square, as it is completely random. After every 60 words, I will give you a word recognition questionnaire and you can mark any words that you believe you heard in those 60, therefore you should try to remember as many of the words as you can.

2-Target Detection Experiment (5)

You are first required to look at the central fixation cross. You will hear a word play through the headphones, upon which a target square will appear. The square will either appear in the top or bottom of the screen. You task is to look to the square as soon as it appears. Try not to predict the location of the square, as it is completely random. After every 60 words, I will give you a word recognition questionnaire and you can mark any words that you believe you heard in those 60, therefore you should try to remember as many of the words as you can.

Saccade Trajectory Experiment (6)

You are first required to look at the central fixation cross. You will hear a word or nonword play through the headphones, upon which two target squares will appear. If you think you hear a real English word, look to the 'Y' appearing at the right of the screen. If you think what you hear is not a real word, you should look to the 'N' at the left of the screen. The letters will always be in the same place, and you should answer as quickly and accurately as possible.

Tropic vs. Embodied Experiment (7)

You should begin by looking at the fixation cross. An image showing letters, numbers, currency symbols or shapes will appear onscreen. At the same time, you will hear a question play through the speakers, such as "How many square in total?" Your task is to verbally answer each question with one of three possible answers: "zero", "one" or "two". You should answer as quickly and accurately as possible.

Lexical Decision Experiments (8 & 9)

You are first required to look at the central fixation cross. You will hear a word or nonword play through the speakers, upon which two target squares will appear. If you think you hear a real English word, look to the green square appearing at the [top / bottom / left / right] of the screen. If you think what you hear is not a real word, you should look to the red square at the [top / bottom / left / right] of the screen. The squares will always be in the same place, and you should answer as quickly and accurately as possible.

Appendix IV – Auditory and Visual Stimuli for Experiment 7

Critical Items

Auditory Stimuli (each used 5 times during 1 experimental session):

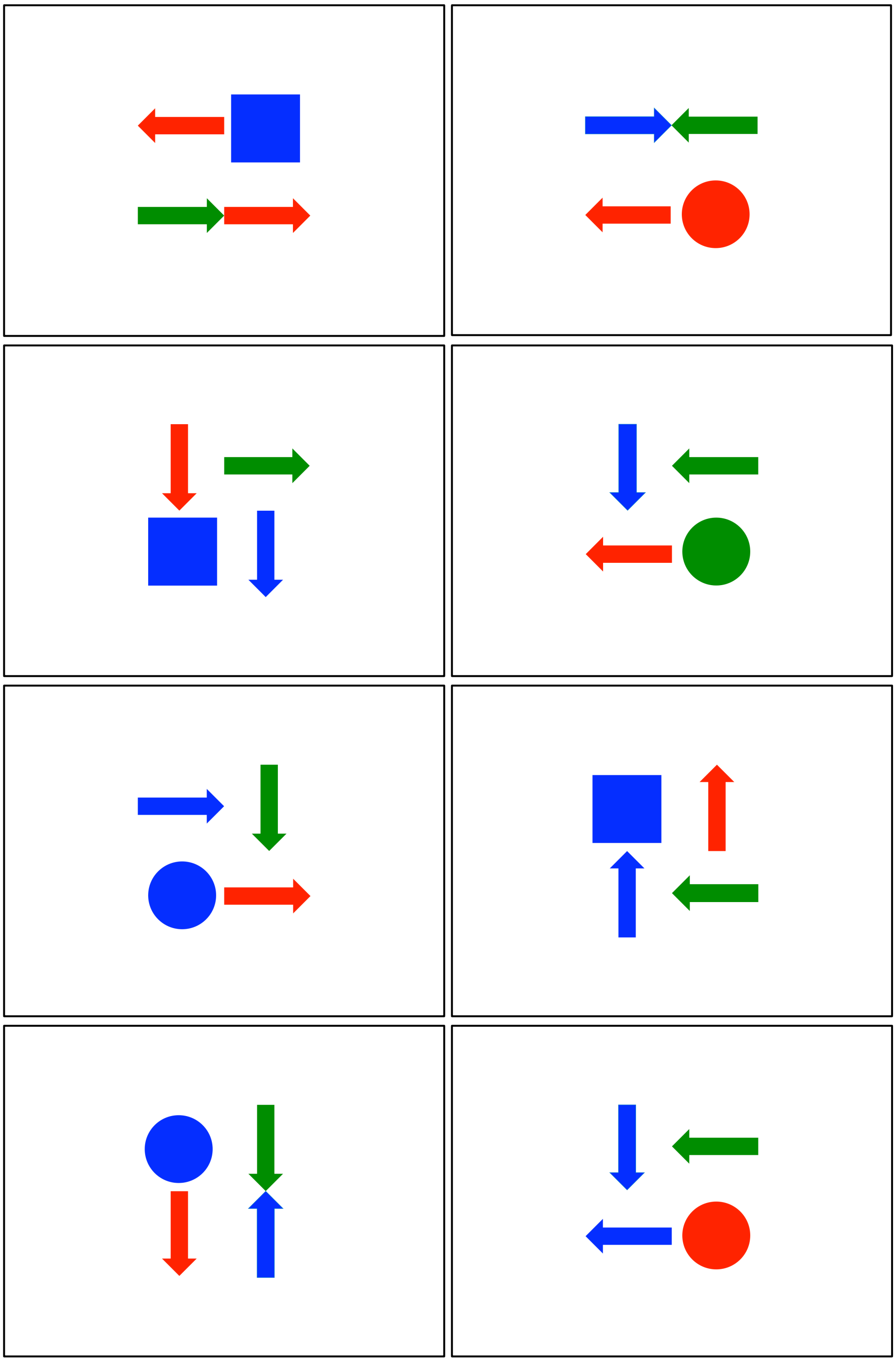
How many arrows point down?

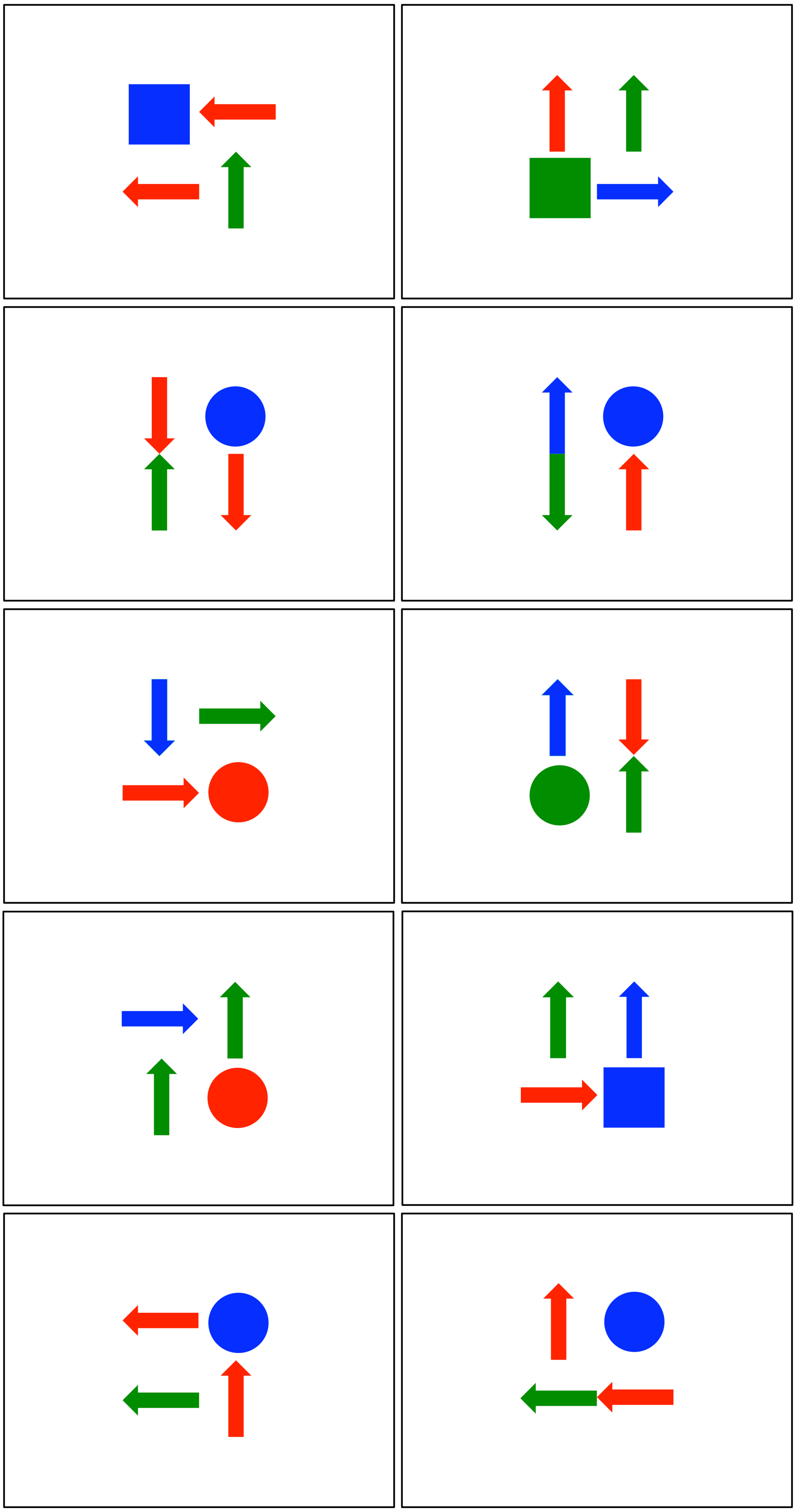
How many arrows point left?

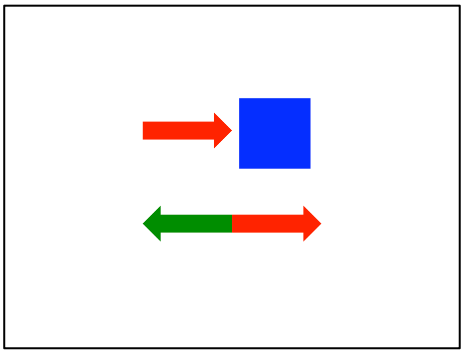
How many arrows point right?

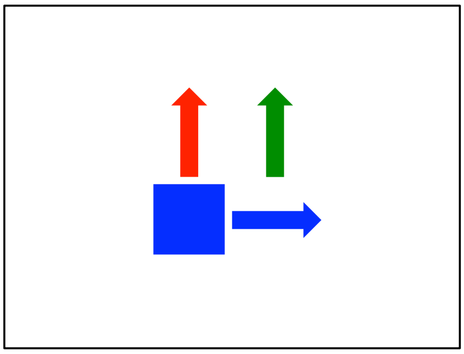
How many arrows point up?

Visual Stimuli:









<u>Filler Items</u>

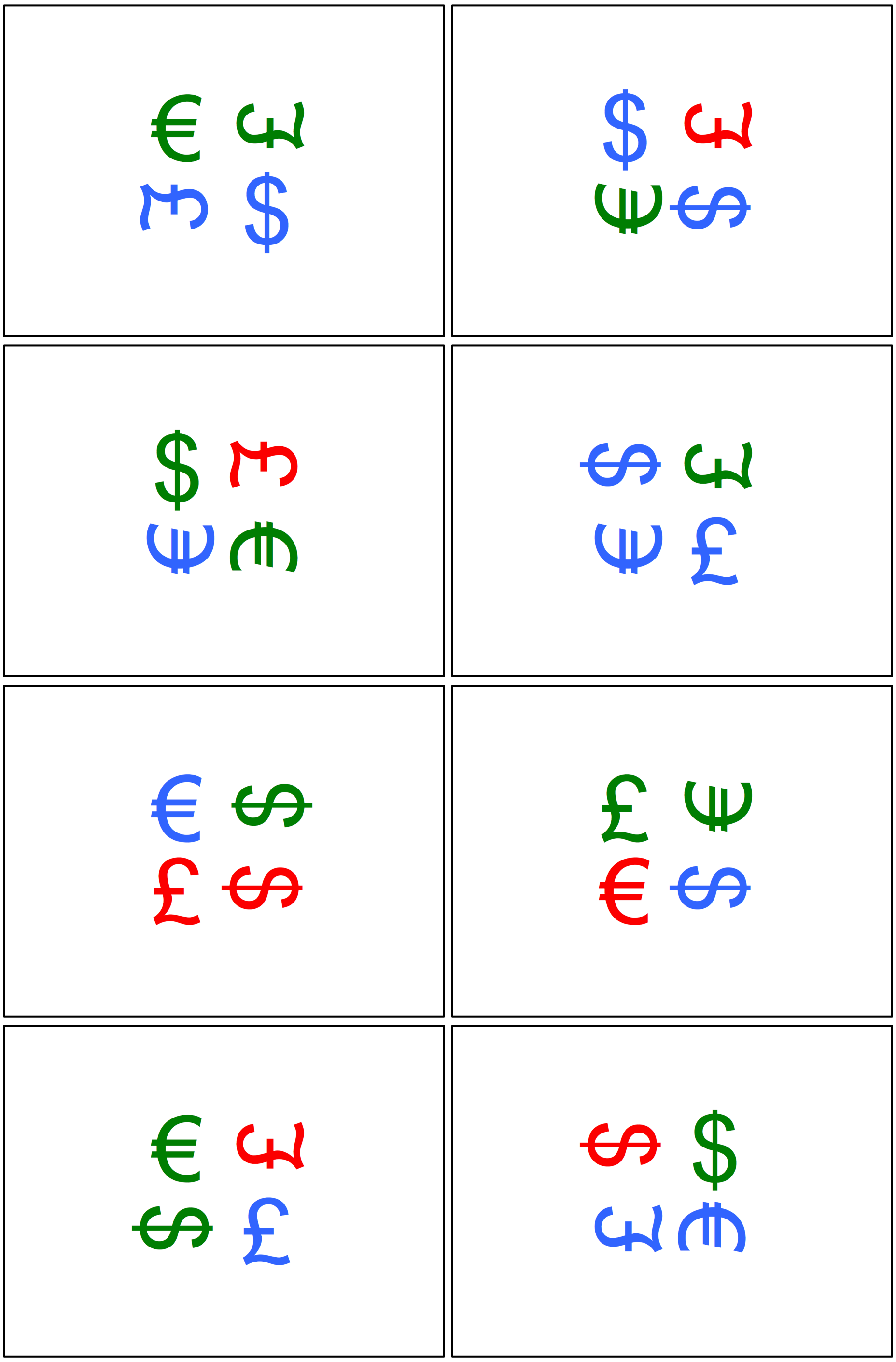
Auditory Stimuli (each used 5 times during 1 experimental session):

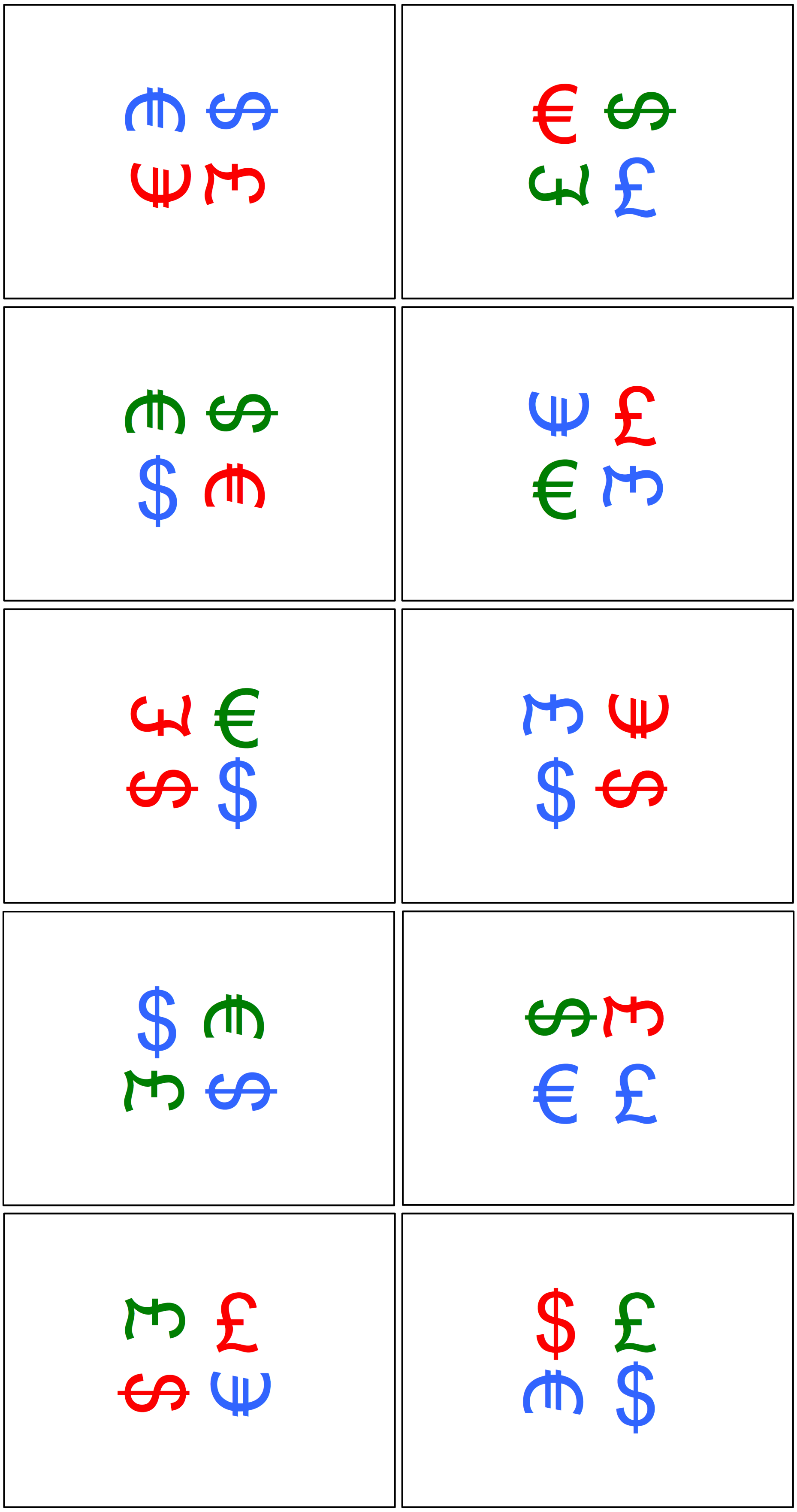
Question	**Relevant Visual Stimuli**
How many dollar signs in total?	Currency symbols
How many euro signs in total?	Currency symbols
How many pound signs in total?	Currency symbols
How many signs are blue?	Currency symbols
How many signs are green?	Currency symbols
How many signs are red?	Currency symbols
How many E's in total?	Letters
How many F's in total?	Letters
How many J's in total?	Letters
How many L's in total?	Letters
How many letters are blue?	Letters
How many letters are green?	Letters
How many letters are red?	Letters
How many 3's in total?	Numbers
How many 4's in total?	Numbers
How many 5's in total?	Numbers
How many 8's in total?	Numbers
How many numbers are blue?	Numbers
How many numbers are green?	Numbers
How many numbers are red?	Numbers
How many arrows in total?	Shapes
How many circles in total?	Shapes
How many squares in total?	Shapes

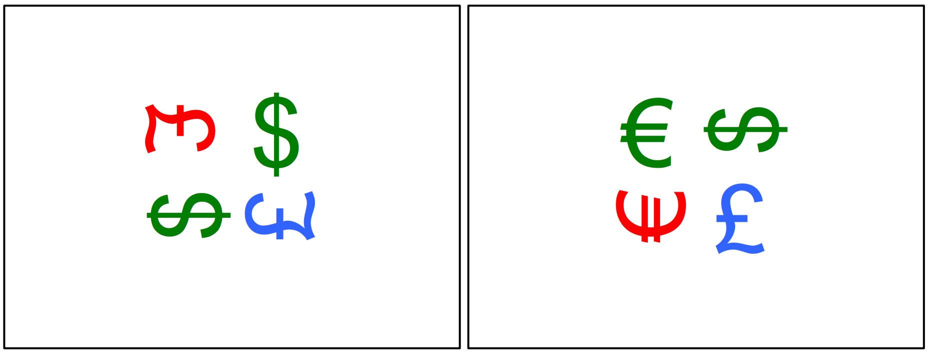
How many shapes are blue?	Shapes
How many shapes are green?	Shapes
How many shapes are red?	Shapes

Visual Stimuli:

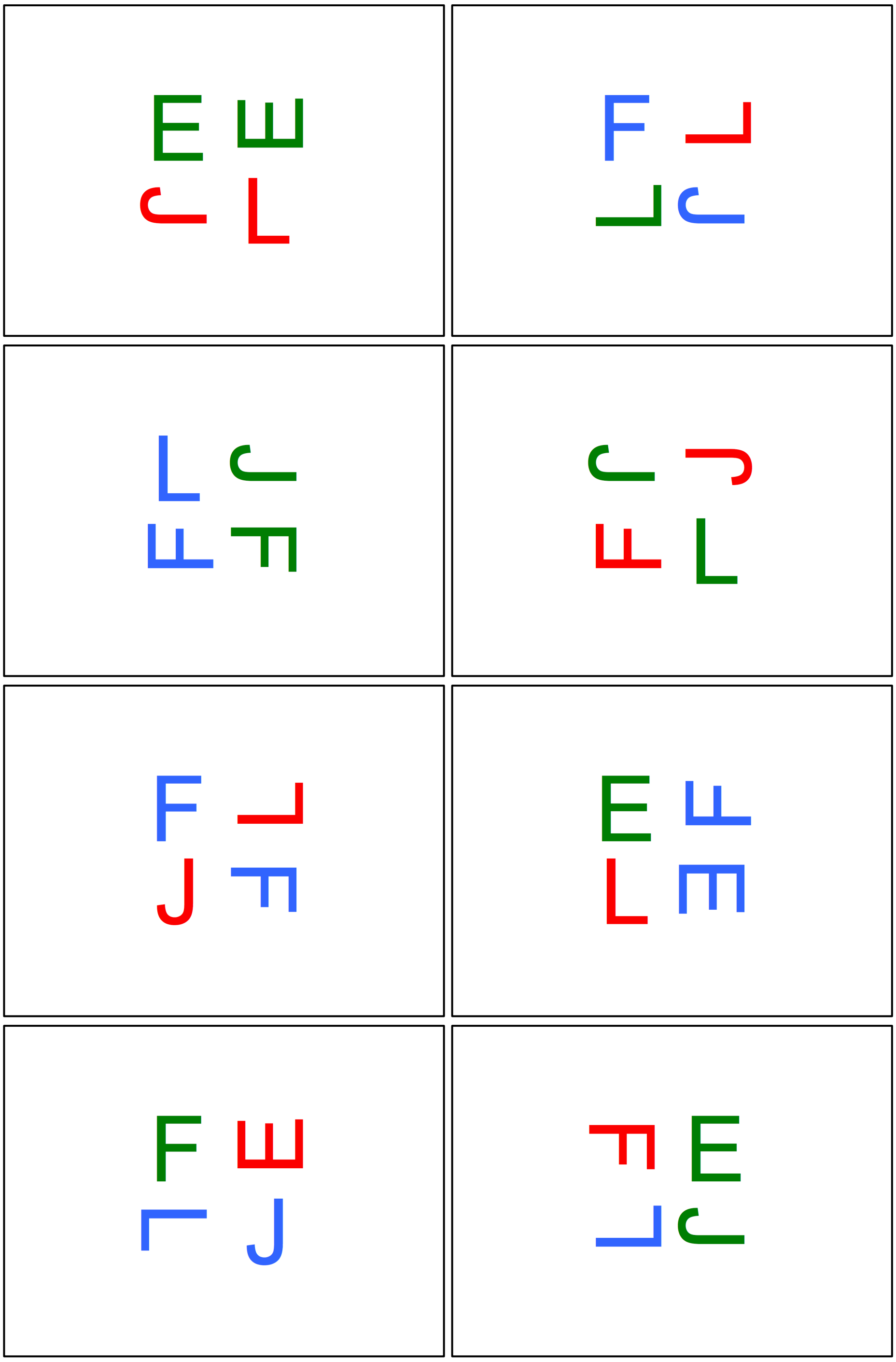
Currency Symbols

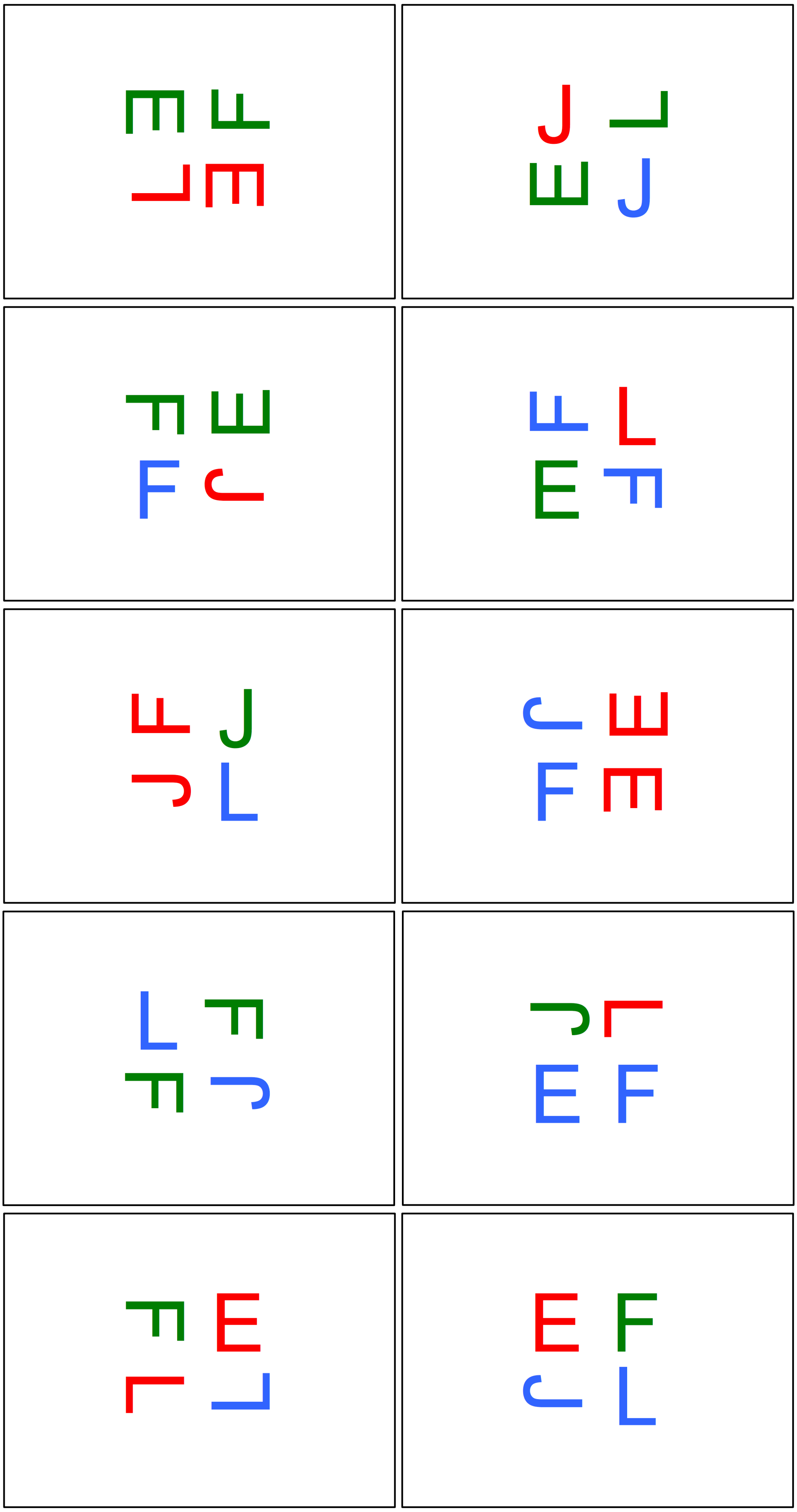


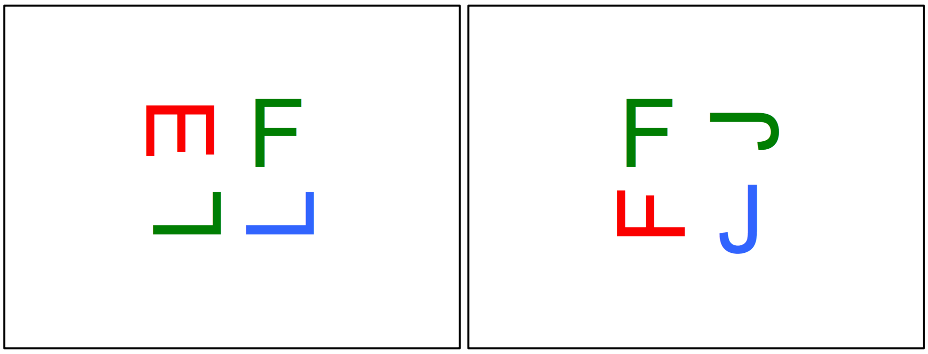




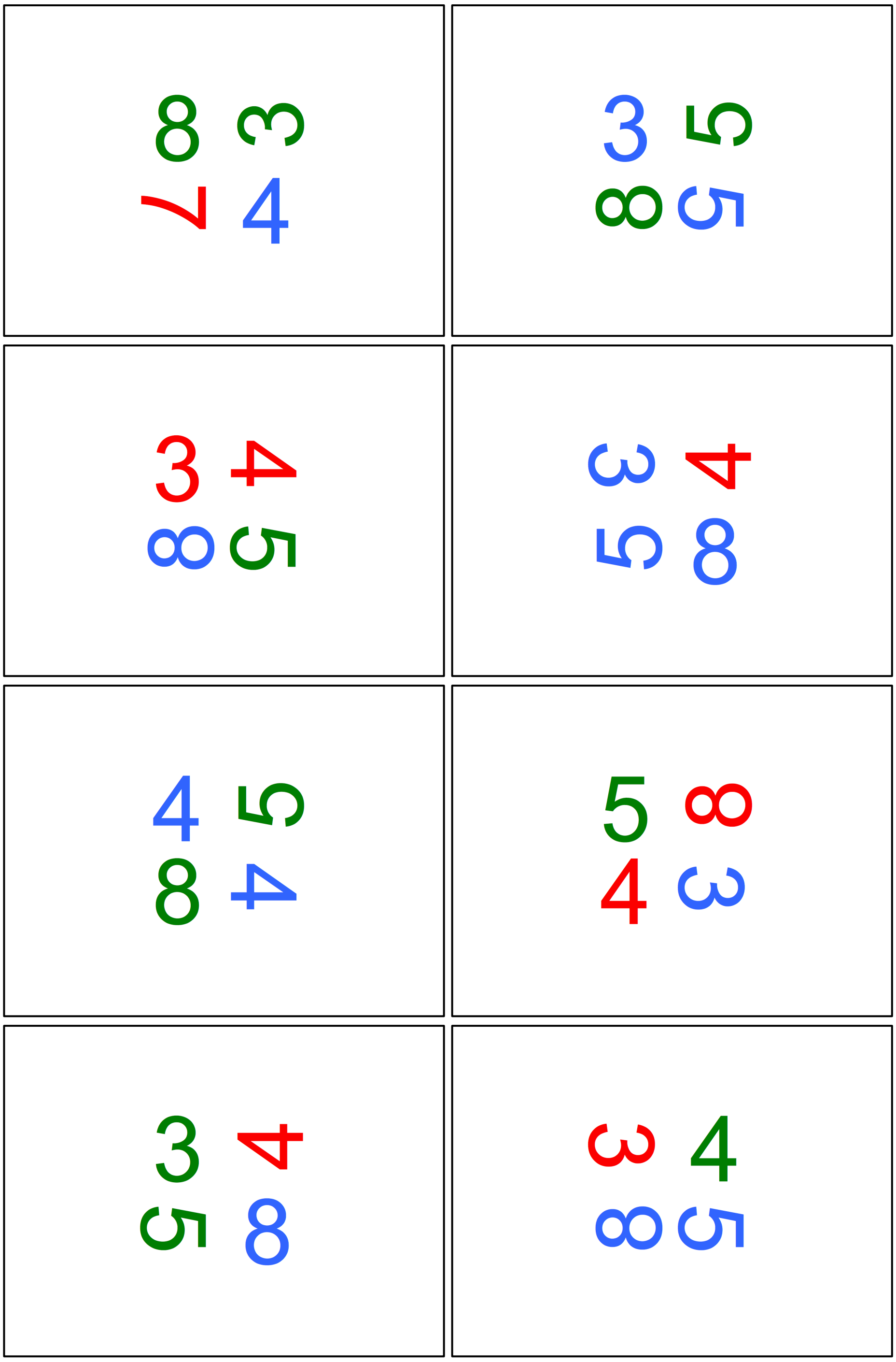
Letters

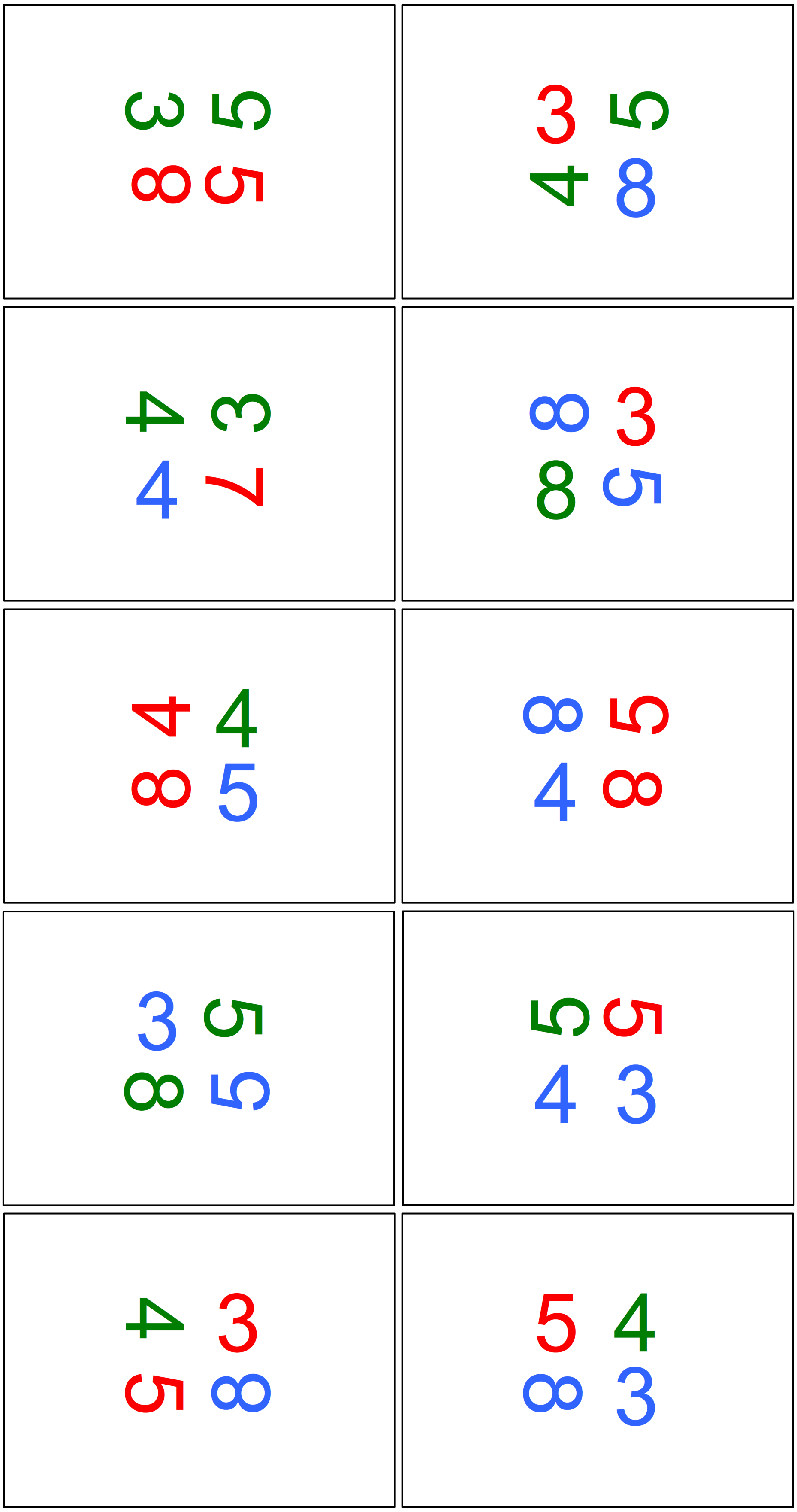


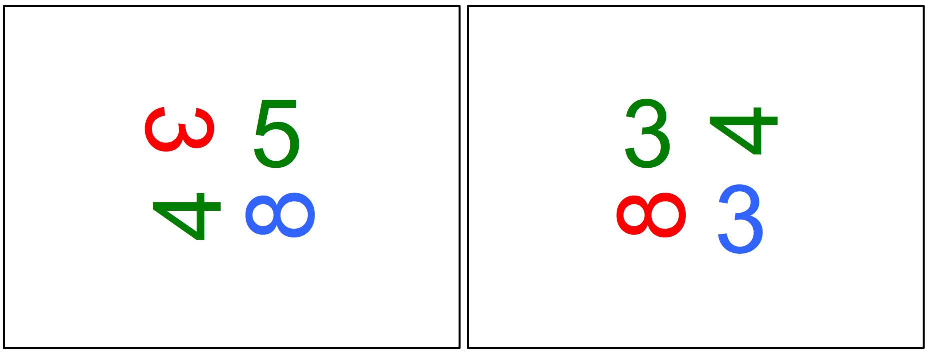




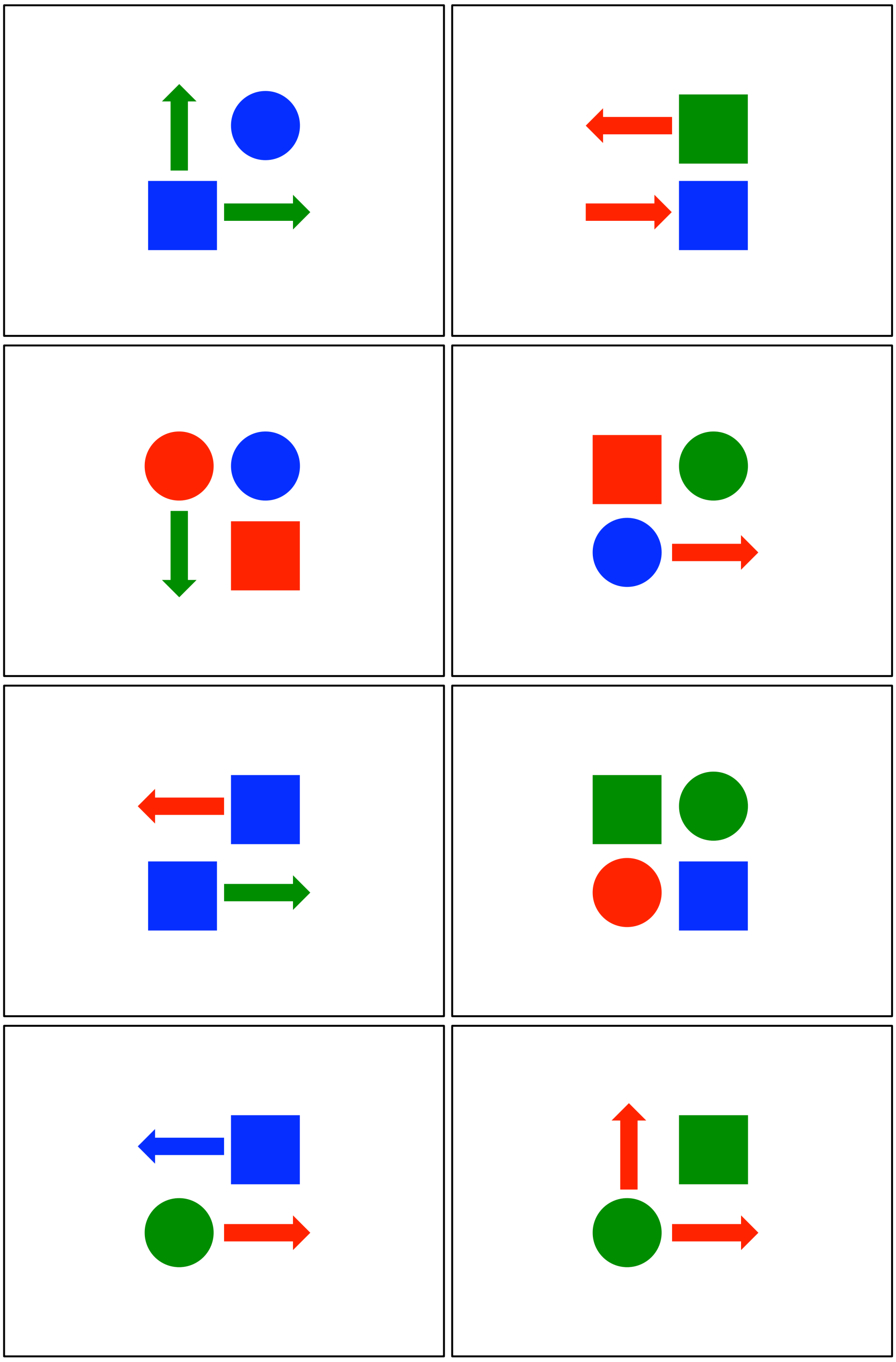
Numbers

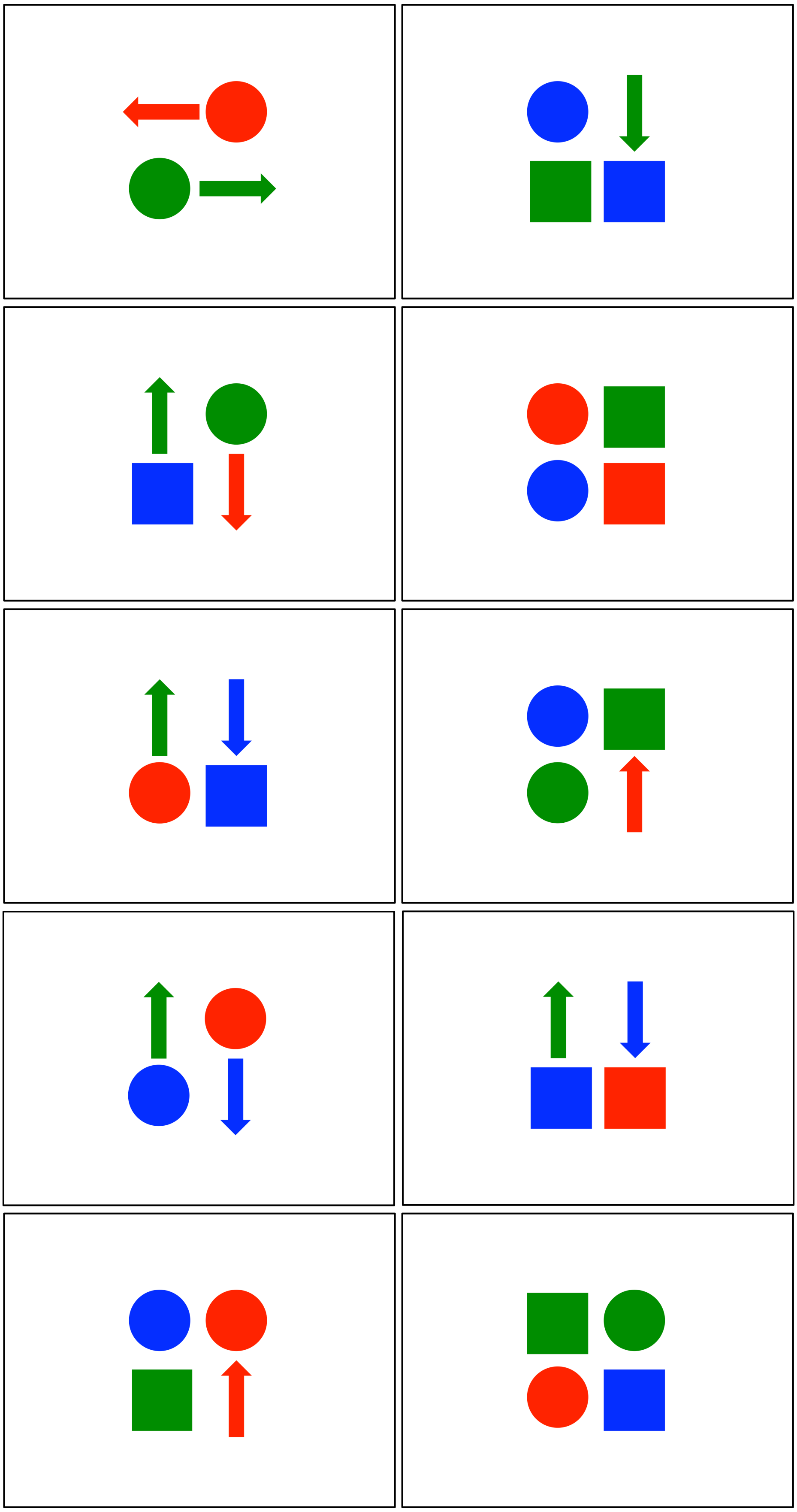


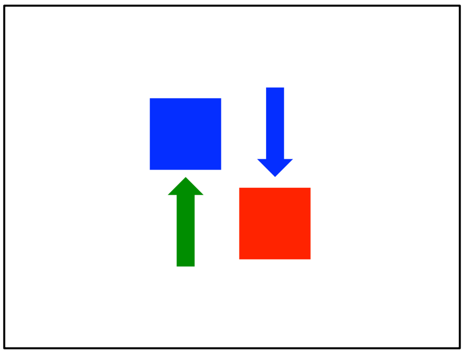


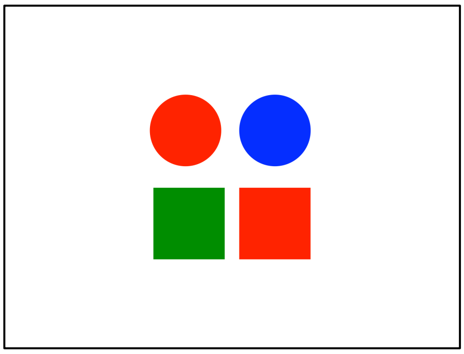


Shapes









References

Altmann, G. T. M., & Kamide, Y. (1999). Incremental interpretation at verbs: Restricting the domain of subsequent reference. *Cognition, 73*(3), 247-264.

Baayen, R., Piepenbrock, R., & Van Rijn, H. (1993). The CELEX lexical database, 1993. *Linguistic Data Consortium, University of Pennsylvania, Philadelphia*.

Balota, D. A., Yap, M. J., Hutchison, K. A., Cortese, M. J., Kessler, B., Loftis, B., . . . Treiman, R. (2007). The English lexicon project. *Behavior Research Methods, 39*(3), 445-459.

Barr, D. J., Levy, R., Scheepers, C., & Tily, H. J. (2013). Random effects structure for confirmatory hypothesis testing: Keep it maximal. *Journal of Memory and language, 68*(3), 255-278.

Barsalou, L. W. (1999). Perceptual symbol systems. *Behavioral and Brain Sciences, 22*(4), 577-660.

Barsalou, L. W. (2008). Grounded cognition. *Annu Rev Psychol, 59*, 617-645. doi:10.1146/annurev.psych.59.103006.093639

Becker, W. (1989). Metrics. In R. H. Wurtz & M. E. Goldberg (Eds.), *The Neurobiology of Saccadic Eye Movements.* (pp. 13-67). Amsterdam: Elsevier.

Bergen, B., Lindsay, S., Matlock, T., & Narayanan, S. (2007). Spatial and linguistic aspects of visual imagery in sentence comprehension. *Cognitive Science, 31*(5), 733-764.

Borghi, A. M., Glenberg, A. M., & Kaschak, M. P. (2004). Putting words in perspective. *Memory & Cognition, 32*(6), 863-873.

Carlson-Radvansky, L. A., & Irwin, D. E. (1993). Frames of reference in vision and language: Where is above? *Cognition, 46*(3), 223-244.

Casasanto, D. (2009). Embodiment of abstract concepts: good and bad in right-and left-handers. *Journal of Experimental Psychology: General, 138*(3), 351.

Casasanto, D., & Chrysikou, E. G. (2011). When left is "right" motor fluency shapes abstract concepts. *Psychological Science, 22*(4), 419-422.

Chasteen, A. L., Burdzy, D. C., & Pratt, J. (2010). Thinking of God moves attention. *Neuropsychologia, 48*(2), 627-630.

Clark, H. H. (1973). Space, time, semantics and the child. In T. E. Moore (Ed.), *Cognitive Development and the Acquisition of Language* (pp. 28-64). New York: Academic.

Cooper, R. M. (1974). The control of eye fixation by the meaning of spoken language: A new methodology for the real-time investigation of speech perception, memory, and language processing. *Cognitive Psychology*.

Decety, J., & Grèzes, J. (2006). The power of simulation: imagining one's own and other's behavior. *Brain Res, 1079*(1), 4-14.

Dehaene, S., Bossini, S., & Giraux, P. (1993). The mental representation of parity and number magnitude. *Journal of Experimental Psychology: General, 122*(3), 371.

Dodd, M. D., Van der Stigchel, S., Leghari, M. A., Fung, G., & Kingstone, A. (2008). Attentional SNARC: There's something special about numbers (let us count the ways). *Cognition, 108*(3), 810-818.

Doyle, M., & Walker, R. (2001). Curved saccade trajectories: Voluntary and reflexive saccades curve away from irrelevant distractors. *Exp Brain Res, 139*(3), 333-344.

Doyle, M., & Walker, R. (2002). Multisensory interactions in saccade target selection: curved saccade trajectories. *Exp Brain Res, 142*(1), 116-130.

Dudschig, C., de la Vega, I., De Filippis, M., & Kaup, B. (2014). Language and vertical space: On the automaticity of language action interconnections. *Cortex, 58*, 151-160.

Dudschig, C., de la Vega, I., & Kaup, B. (2015). What's up? Emotion-specific activation of vertical space during language processing. *Acta Psychologica, 156*, 143-155.
Dudschig, C., & Kaup, B. (in press). Is it all task-specific? The role of binary responses, verbal mediation and saliency for eliciting language-space associations. *Journal of Experimental Psychology: Learning, Memory, and Cognition*.
Dudschig, C., Lachmair, M., de la Vega, I., De Filippis, M., & Kaup, B. (2012). From top to bottom: spatial shifts of attention caused by linguistic stimuli. *Cognitive processing, 13*(1), 151-154.
Dudschig, C., Souman, J., Lachmair, M., de la Vega, I., & Kaup, B. (2013). Reading "Sun" and Looking Up: The Influence of Language on Saccadic Eye Movements in the Vertical Dimension. *PLOS ONE, 8*(2), e56872.
Epley, N., Morewedge, C. K., & Keysar, B. (2004). Perspective taking in children and adults: Equivalent egocentrism but differential correction. *Journal of Experimental Social Psychology, 40*(6), 760-768.
Estes, Z., Verges, M., & Adelman, J. S. (2015). Words, objects, and locations: Perceptual matching explains spatial interference and facilitation. *Journal of Memory and language, 84*, 167-189.
Estes, Z., Verges, M., & Barsalou, L. W. (2008). Head Up, Foot Down Object Words Orient Attention to the Objects' Typical Location. *Psychological Science, 19*(2), 93-97.
Evans, J. S. B., & Stanovich, K. E. (2013). Dual-process theories of higher cognition advancing the debate. *Perspectives on psychological science, 8*(3), 223-241.
Findlay, J. M., & Harris, L. R. (1984). Small saccades to double-stepped targets moving in two dimensions. *Advances in Psychology, 22*, 71-78.
Fischer, M. H., Castel, A. D., Dodd, M. D., & Pratt, J. (2003). Perceiving numbers causes spatial shifts of attention. *Nature neuroscience, 6*(6), 555-556.
Freyd, J. J. (1983). The mental representation of movement when static stimuli are viewed. *Attention, Perception, & Psychophysics, 33*(6), 575-581.
Freyd, J. J. (1987). Dynamic mental representations. *Psychological review, 94*(4), 427.
Freyd, J. J., & Finke, R. A. (1984). Representational momentum. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 10*(1), 126-132.
Friederici, A. D., & Levelt, W. J. (1990). Spatial reference in weightlessness: perceptual factors and mental representations. *Perception & Psychophysics, 47*(3), 253-266.
Garnham, A. (1987). *Mental models as representations of discourse and text*. Chichester: Ellis Horwood.
Garnham, A. (1996). The other side of mental models: Theories of language comprehension. In J. Oakhill & A. Garnham (Eds.), *Mental models in cognitive science: Essays in honour of Phil Johnson-Laird* (pp. 35-52). Hove, UK: Psychology Press.
Giessner, S. R., & Schubert, T. W. (2007). High in the hierarchy: How vertical location and judgments of leaders' power are interrelated. *Organizational Behavior and Human Decision Processes, 104*(1), 30-44.
Girardi, G., Antonucci, G., & Nico, D. (2013). Cueing spatial attention through timing and probability. *Cortex, 49*(1), 211-221.
Glenberg, A. M., & Kaschak, M. P. (2002). Grounding language in action. *Psychonomic Bulletin & Review, 9*(3), 558-565.
Glenberg, A. M., & Langston, W. E. (1992). Comprehension of illustrated text: Pictures help to build mental models. *Journal of Memory and language, 31*(2), 129-151.
Glenberg, A. M., Meyer, M., & Lindem, K. (1987). Mental models contribute to foregrounding during text comprehension. *Journal of Memory and language, 26*(1), 69-83.

Goldman, A. I. (2006). *Simulating minds: The philosophy, psychology, and neuroscience of mindreading*. Oxford: Oxford University Press.

Goldring, J., & Fischer, B. (1997). Reaction times of vertical prosaccades and antisaccades in gap and overlap tasks. *Exp Brain Res, 113*(1), 88-103.

Gozli, D. G., Chasteen, A. L., & Pratt, J. (2013). The cost and benefit of implicit spatial cues for visual attention. *Journal of Experimental Psychology: General, 142*(4), 1028.

Greene, H. H., Brown, J. M., & Dauphin, B. (2014). When do you look where you look? A visual field asymmetry. *Vision research, 102*, 33-40.

Hanley, J. A., Negassa, A., & Forrester, J. E. (2003). Statistical analysis of correlated data using generalized estimating equations: an orientation. *American journal of epidemiology, 157*(4), 364-375.

Hardin, J. W., & Hilbe, J. M. (2003). *Generalized estimating equations*: Wiley Online Library.

Harnad, S. (1990). The symbol grounding problem. *Physica D: Nonlinear Phenomena, 42*(1), 335-346.

Huettig, F., & Altmann, G. T. (2004). The on-line processing of ambiguous and unambiguous words in context: Evidence from head-mounted eyetracking. *The online study of sentence comprehension: Eyetracking, ERPs and beyond*, 187-207.

Huettig, F., & Altmann, G. T. (2011). Looking at anything that is green when hearing "frog": How object surface colour and stored object colour knowledge influence language-mediated overt attention. *The Quarterly Journal of Experimental Psychology, 64*(1), 122-145.

Huettig, F., & Altmann, G. T. M. (2005). Word meaning and the control of eye fixation: Semantic competitor effects and the visual world paradigm. *Cognition, 96*(1), B23-B32.

Hung, Y.-h., Hung, D. L., Tzeng, O. J.-L., & Wu, D. H. (2008). Flexible spatial mapping of different notations of numbers in Chinese readers. *Cognition, 106*(3), 1441-1450.

Intraub, H., & Hoffman, J. E. (1992). Reading and visual memory: Remembering scenes that were never seen. *The American journal of psychology*, 101-114.

Ito, Y., & Hatta, T. (2004). Spatial structure of quantitative representation of numbers: Evidence from the SNARC effect. *Memory & Cognition, 32*(4), 662-673.

Johnson-Laird, P. N. (1983). *Mental models: towards a cognitive science of language, inference, and consciousness* (Vol. 6): Cambridge University Press.

Kamide, Y. (2007). *Prediciting dynamic motion trajectories.* Paper presented at the CUNY Conference on Human Sentence Processing, California, USA.

Kamide, Y., Lindsay, S., Scheepers, C., & Kukona, A. (in press). Event Processing in the Visual World: Projected Motion Paths During Spoken Sentence Comprehension. *Journal of Experimental Psychology: Learning, Memory, and Cognition*. doi:10.1037/xlm0000199

Kappers, A. M. (2004). The contributions of egocentric and allocentric reference frames in haptic spatial tasks. *Acta Psychologica, 117*(3), 333-340.

Kaschak, M. P., Madden, C. J., Therriault, D. J., Yaxley, R. H., Aveyard, M., Blanchard, A. A., & Zwaan, R. A. (2005). Perception of motion affects language processing. *Cognition, 94*(3), B79-B89.

Keuleers, E., Lacey, P., Rastle, K., & Brysbaert, M. (2012). The British Lexicon Project: Lexical decision data for 28,730 monosyllabic and disyllabic English words. *Behavior Research Methods, 44*(1), 287-304.

Kingstone, A. (1992). Combining expectancies. *The Quarterly Journal of Experimental Psychology, 44*(1), 69-104.

Kousta, S.-T., Vinson, D. P., & Vigliocco, G. (2009). Emotion words, regardless of polarity, have a processing advantage over neutral words. *Cognition, 112*(3), 473-481.
Lachmair, M., Dudschig, C., De Filippis, M., de la Vega, I., & Kaup, B. (2011). Root versus roof: automatic activation of location information during word processing. *Psychonomic Bulletin & Review, 18*(6), 1180-1188.
Lakens, D., Semin, G. R., & Foroni, F. (2015). Why your highness needs the people. *Social Psychology*.
Lakoff, G., & Johnson, M. (1999). *Philosophy in the flesh: The embodied mind and its challenge to western thought*: Basic books.
Lebois, L. A., Wilson-Mendenhall, C. D., & Barsalou, L. W. (2015). Are Automatic Conceptual Cores the Gold Standard of Semantic Processing? The Context‐Dependence of Spatial Meaning in Grounded Congruency Effects. *Cognitive Science, 39*(8), 1764-1801.
Levelt, W. J. (1984). Some perceptual limitations on talking about space *Limits in perception* (pp. 323-358). Utrecht: VNU Science Press.
Levinson, S. C. (1996). Language and space. *Annual review of Anthropology, 25*, 353-382.
Lindsay, S., Scheepers, C., & Kamide, Y. (2013). To Dash or to Dawdle: Verb-Associated Speed of Motion Influences Eye Movements during Spoken Sentence Comprehension. *PLOS ONE, 8*(6), e67187.
Lu, A., Zhang, H., He, G., Zheng, D., & Hodges, B. H. (2013). Looking Up to Others: Social Status, Chinese Honorifics, and Spatial Attention. *Can J Exp Psychol*. doi:10.1037/cep0000008
Ludwig, C. J., & Gilchrist, I. D. (2002). Measuring saccade curvature: A curve-fitting approach. *Behavior Research Methods, Instruments, & Computers, 34*(4), 618-624.
Maass, A., & Russo, A. (2003). Directional bias in the mental representation of spatial events nature or culture? *Psychological Science, 14*(4), 296-301.
McNamara, T. P., Rump, B., & Werner, S. (2003). Egocentric and geocentric frames of reference in memory of large-scale space. *Psychonomic Bulletin & Review, 10*(3), 589-595.
Meier, B. P., Hauser, D. J., Robinson, M. D., Friesen, C. K., & Schjeldahl, K. (2007). What's" up" with God? Vertical space as a representation of the divine. *Journal of Personality and Social Psychology, 93*(5), 699.
Meier, B. P., & Robinson, M. D. (2004). Why the sunny side is up associations between affect and vertical position. *Psychological Science, 15*(4), 243-247.
Meier, B. P., & Robinson, M. D. (2006). Does "feeling down" mean seeing down? Depressive symptoms and vertical selective attention. *Journal of Research in Personality, 40*(4), 451-461.
Meteyard, L., Bahrami, B., & Vigliocco, G. (2007). Motion detection and motion verbs language affects low-level visual perception. *Psychological Science, 18*(11), 1007-1013.
Miles, W. R. (1930). Ocular dominance in human adults. *The Journal of General Psychology, 3*(3), 412-430.
Miles, W. R. (1936). The reaction time of the eye. *Psychological Monographs, 47*(2), 268.
Miller, G. A., & Johnson-Laird, P. N. (1976). *Language and perception*. Cambridge, MA: Harvard University Press.
Myachykov, A., Scheepers, C., Fischer, M. H., & Kessler, K. (2013). TEST: a tropic, embodied, and situated theory of cognition. *Top Cogn Sci, 6*(3), 442-460. doi:10.1111/tops.12024
Neely, J. H. (1977). Semantic priming and retrieval from lexical memory: Roles of inhibitionless spreading activation and limited-capacity attention. *Journal of Experimental Psychology: General, 106*(3), 226.

Piaget, J., & Inhelder, B. (1969). *The psychology of the child*: Basic Books.
Posner, M. I. (1980). Orienting of attention. *Quarterly Journal of Experimental Psychology, 32*(1), 3-25.
Posner, M. I., & Cohen, Y. (1984). Components of visual orienting. *Attention and performance X: Control of language processes, 32*, 531-556.
Posner, M. I., Nissen, M. J., & Ogden, W. C. (1978). Attended and unattended processing modes: The role of set for spatial location. In H. L. Pick & I. J. Saltzman (Eds.), *Modes of perceiving and processing information* (pp. 137-157). Hillsdale, N.J.: Erlbaum.
Posner, M. I., Snyder, C. R., & Davidson, B. J. (1980). Attention and the detection of signals. *Journal of Experimental Psychology: General, 109*(2), 160.
Potter, M. C., Kroll, J. F., Yachzel, B., Carpenter, E., & Sherman, J. (1986). Pictures in sentences: Understanding without words. *Journal of Experimental Psychology: General, 115*(3), 281-294.
Richardson, D. C., Spivey, M. J., Barsalou, L. W., & McRae, K. (2003). Spatial representations activated during real-time comprehension of verbs. *Cognitive Science, 27*(5), 767-780.
Roberts Jr, R. J., & Aman, C. J. (1993). Developmental differences in giving directions: spatial frames of reference and mental rotation. *Child Development*, 1258-1270.
Robinson, D. (1975). Oculomotor control signals. In G. Lennerstrand & P. Bachy-Rita (Eds.), *Basic mechanisms of ocular motility and their clinical implications* (pp. 337-374). Oxford, UK: Pergamon.
Roth, H. L., Lora, A. N., & Heilman, K. M. (2002). Effects of monocular viewing and eye dominance on spatial attention. *Brain, 125*(9), 2023-2035.
Scheepers, C. (2014). *Between-group matching of confounding variables: Why covariates remain important for analysis.* Paper presented at the Architectures and Mechanisms for Language Processing (AMLaP), Edinburgh, UK.
Schober, M. F. (1993). Spatial perspective-taking in conversation. *Cognition, 47*(1), 1-24.
Schober, M. F. (1995). Speakers, addressees, and frames of reference: Whose effort is minimized in conversations about locations? *Discourse Processes, 20*(2), 219-247.
Schubert, T. W. (2005). Your highness: vertical positions as perceptual symbols of power. *Journal of Personality and Social Psychology, 89*(1), 1-21.
Searle, J. R. (1980). Minds, brains, and programs. *Behavioral and Brain Sciences, 3*, 417-424.
Šetić, M., & Domijan, D. (2007). The influence of vertical spatial orientation on property verification. *Language and Cognitive Processes, 22*(2), 297-312. doi:10.1080/01690960600732430
Sheliga, B. M., Riggio, L., Craighero, L., & Rizzolatti, G. (1995). Spatial attention-determined modifications in saccade trajectories. *Neuroreport, 6*(3), 585.
Sheliga, B. M., Riggio, L., & Rizzolatti, G. (1994). Orienting of attention and eye movements. *Exp Brain Res, 98*(3), 507-522.
Sheliga, B. M., Riggio, L., & Rizzolatti, G. (1995). Spatial attention and eye movements. *Exp Brain Res, 105*(2), 261-275.
Shepard, R., & Metzler, J. (1971). Mental rotation of three-dimensional objects. *Science, 171*, 701-703.
Sisson, C. (1993). Dante Alighieri: The Divine Comedy: Oxford, UK: Oxford University Press.
Smit, A., & Van Gisbergen, J. (1990). An analysis of curvature in fast and slow human saccades. *Exp Brain Res, 81*(2), 335-345.
Speed, L. J., & Vigliocco, G. (2013). Eye Movements Reveal the Dynamic Simulation of Speed in Language. *Cognitive Science*, n/a-n/a. doi:10.1111/cogs.12096

Stanfield, R. A., & Zwaan, R. A. (2001). The effect of implied orientation derived from verbal context on picture recognition. *Psychological Science, 12*(2), 153-156.
Sun, Y., & Wang, H. (2010). Perception of space by multiple intrinsic frames of reference. *PLOS ONE, 5*(5), e10442.
Tamborello II, F. P., Sun, Y., & Wang, H. (2011). Spatial reasoning with multiple intrinsic frames of reference. *Experimental psychology, 59*(1), 3-10.
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.
Taylor, J. E. T., Lam, T. K., Chasteen, A. L., & Pratt, J. (2015). Bow Your Head in Shame, or, Hold Your Head Up with Pride: Semantic Processing of Self-Esteem Concepts Orients Attention Vertically. *PLOS ONE, 10*(9), e0137704.
Tucker, M., & Ellis, R. (2004). Action priming by briefly presented objects. *Acta Psychologica, 116*(2), 185-203.
Tversky, B., & Hard, B. M. (2009). Embodied and disembodied cognition: Spatial perspective-taking. *Cognition, 110*(1), 124-129.
Tversky, B., Kugelmass, S., & Winter, A. (1991). Cross-cultural and developmental trends in graphic productions. *Cognitive Psychology, 23*(4), 515-557.
Tversky, B., Lee, P., & Mainwaring, S. (1999). Why do speakers mix perspectives? *Spatial cognition and computation, 1*(4), 399-412.
Van der Stigchel, S. (2010). Recent advances in the study of saccade trajectory deviations. *Vision research, 50*(17), 1619-1627.
Van der Stigchel, S., Meeter, M., & Theeuwes, J. (2006). Eye movement trajectories and what they tell us. *Neuroscience & Biobehavioral Reviews, 30*(5), 666-679.
Van der Stigchel, S., & Theeuwes, J. (2007). The relationship between covert and overt attention in endogenous cuing. *Perception & Psychophysics, 69*(5), 719-731.
Van Dijk, T. A., & Kintsch, W. (1983). *Strategies of discourse comprehension*. New York: Academic Press.
Van Gisbergen, J., Van Opstal, A., Roebroek, J., & Noord, G. G. (1987). Stimulus-induced midflight modification of saccade trajectories. In J. K. O'Regan & A. Levy-Schoen (Eds.), *Eye movements: From physiology to cognition* (pp. 27-36). Amsterdam: Elsevier.
Warriner, A. B., Kuperman, V., & Brysbaert, M. (2013). Norms of valence, arousal, and dominance for 13,915 English lemmas. *Behavior Research Methods, 45*(4), 1191-1207.
Yarbus, A. L. (1967). *Eye Movements and Vision* (B. Haigh, Trans.). New York: Plenium Press.
Zanolie, K., van Dantzig, S., Boot, I., Wijnen, J., Schubert, T. W., Giessner, S. R., & Pecher, D. (2012). Mighty metaphors: Behavioral and ERP evidence that power shifts attention on a vertical dimension. *Brain and cognition, 78*(1), 50-58.
Zebian, S. (2005). Linkages between number concepts, spatial thinking, and directionality of writing: The SNARC effect and the reverse SNARC effect in English and Arabic monoliterates, biliterates, and illiterate Arabic speakers. *Journal of Cognition and Culture, 5*(1), 165-190.
Zwaan, R. A. (1999). Situation Models. *Current Directions in Psychological Science, 8*(1), 15.
Zwaan, R. A., Madden, C. J., Yaxley, R. H., & Aveyard, M. E. (2004). Moving words: Dynamic representations in language comprehension. *Cognitive Science, 28*(4), 611-619.
Zwaan, R. A., & Radvansky, G. A. (1998). Situation models in language comprehension and memory. *Psychological bulletin, 123*(2), 162-185.

Zwaan, R. A., Stanfield, R. A., & Yaxley, R. H. (2002). Language comprehenders mentally represent the shapes of objects. *Psychological Science, 13*(2), 168-171.
Zwaan, R. A., & Taylor, L. J. (2006). Seeing, acting, understanding: motor resonance in language comprehension. *Journal of Experimental Psychology: General, 135*(1), 1.
Zwaan, R. A., & Yaxley, R. H. (2003). Spatial iconicity affects semantic relatedness judgments. *Psychonomic Bulletin & Review, 10*(4), 954-958.