System-level attention links cognition, perception and action: Evidence from language comprehension and eye movements

Xierong Liu

Submitted for the degree of PhD

University of York
Department of Psychology

November 2009

Abstract

The research reported in this thesis attempted to establish the underlying representational substrate within which cognition, perception and action interact. A theoretical framework was adopted in which attention functions at a system level as the mediating mechanism between cognitive functions and sensorimotor responses. This was achieved by addressing two issues: a) whether representations activated by language comprehension can compete with representations involved in eye movement control; and b) whether this competition creates attentional conflict within the system and thus modulates the oculomotor response. The effects of two types of words, directional verbs and locational nouns, on two types of eye movements, pursuit and saccades, were explored in nine eye-tracking experiments. Empirical findings suggested that a) eye velocity during pursuit was systematically modulated by verb semantics; depending on whether there was agreement or conflict between representations activated by the directional verbs and the oculomotor task, eye velocity was respectively increased or decreased; b) saccadic launch latency was consistently modified by verb semantics; saccades were initiated with reduced or increased latencies when representations involved in language comprehension and eye movement control were in accordance or in conflict with each other. This collection of evidence points to a unified, attention-governed system that encompasses cognition, perception and action.

List of Tables and Figures

Acknowledgements

I would like to thank my supervisor – Gerry Altmann, and the members of my research committee – Tony Morland and Silvia Gennari, for the continuous help and support. Special thanks to Jelena Mirković.

Declaration

This thesis contains original work completed solely by the author under the supervision of Prof. Gerry Altmann.

The data reported in Chapter 2, 3, 4 and 6 were presented at the following conferences:

Conference Talks

Altmann, G. T. M., Kamide, Y., & Liu, X. (2008). Language-mediated eye movements: From the ear to the eye in less time than it takes to blink. Talk presented at the Cognitive Section, 25th Anniversary Annual British Psychological Society Conference, Southampton, UK.

Liu, X., & Altmann, G. T. M. (2008). What can a moving dot tell us about anything? Talk presented at the Experimental Psychology Society, Liverpool, UK.

Conference Posters

Liu, X., & Altmann, G. T. M. (2008). What can a moving dot tell us about anything? Poster presented at 49th Annual Meeting of the Psychonomic Society, Chicago, US.

Liu, X., & Altmann, G. T. M. (2007). What can a moving dot tell us about language? Poster presented at 14th European Conference on Eye movements, Potsdam, Germany.

Liu, X., & Altmann, G. T. M. (2007). What can a moving dot tell us about language? Poster presented at the Workshop on Embodied Sentence Processing, Saarbrucken, Germany.

Chapter 1. Literature review

Grounded cognition

The architecture underlying human cognition remains in question. The dominant view (e.g. Newell & Simon, 1972; Fodor, 1975; Pylyshyn, 1986) assumes that cognitive processing is accomplished by a modular system that is separate from other modality-specific systems in the brain, such as perception and action. Representations constructed from perceptual and motoric experiences in the modal systems are converted into abstract amodal symbols. Experiential knowledge in semantic memory is represented by these amodal symbols and cognitive functions rely on these amodal symbols and operate in isolation according to different principles from perception and action.

The assumption that cognition depends on amodal symbols held in semantic memory is rejected by the *grounded cognition* approach (e.g. Barsalou, 1999, 2007; Fincher-Kiefer, 2001; Glenberg, 1997; Lakoff & Johnson, 1980; McNeill, 1992), an alternative view that has been receiving increasing attention in recent years. The grounded cognition perspective questions the necessity of an additional amodal system and suggests that cognition is grounded in already existing perceptual and motoric mechanisms: Perception and action are tightly interlinked throughout the processing stream and cognitive functions operate upon and within the same substrates as perception and action based on identical principles.

Many accounts of grounded cognition focus on the role of *simulation* (e.g. Barsalou, 1999; Goldman, 2006). Simulation refers to the reenactment of perceptual or motoric experiences acquired through interactions with the external world. This is achieved through re-activations of the neural response patterns that have been formed during the actual perceptual or motoric experiences. As an event is experienced (e.g. walking across a grassy field), information captured by different modalities (e.g. the action of walking, the colour and smell of the grass, the feeling of relaxation) is represented by distinct activation patterns of the brain and integrated into a multimodal representation stored in memory. When the information is needed in future interactions, the relevant representations stored during the actual experiences are summoned and re-activated, in order to simulate how perceptual and motoric states associated with those specific experiences were represented by the brain. The account of grounded cognition holds simulation as the fundamental form of computation in the brain: All cognitive functions are supported by a collection of simulation mechanisms that share the same representational substrate. Therefore, instead of being kept separately, high-level cognitive functions are closely linked to low-level perceptual and motoric systems.

The account of grounded cognition has received support from evidence gathered using both behavioural and neuroscientific methods (e.g. Damasio & Damasio, 1994; Müsseler & Hommel, 1997; Tucker & Ellis, 1998; Rizzolatti & Craighero, 2004), however, the most consistent and compelling evidence comes from research into language comprehension. The grounded cognition framework has attracted a substantial amount of attention from researchers interested in the fundamental question of how language conveys meaning. This is hardly surprising given that language processing probably represents the highest level and the most complex form of human cognition. In order to process linguistic stimuli, all types of cognitive functions, including memory, abstract thinking, reasoning and prediction, have to be recruited and utilized efficiently. The traditional approach to language comprehension holds the view that language conveys meaning by using abstract symbols (i.e. words) bound by syntactic rules (e.g. Chomsky, 1980; Kintsch, 1988; Pinker, 1994). Thus, words are arbitrarily associated with their referents and bear no direct relationship with the physical or functional properties of the entities they refer to. However, according to the predictions made under the grounded cognition framework, linguistic meaning is grounded in perceptual and motoric experiences, and language comprehension is achieved through perception and action simulation by re-activating relevant sensorimotor representations stored in memory. As a result, the processing of linguistic stimuli (e.g. words or sentences) should entail the activation of related features of the corresponding referents, and equally, but more directly related to the topic of this thesis, language processing should share (at least some) neural or representational substrates with the processing of perceptual and motoric information. The research to be described below is predicated on this relationship between language comprehension and perceptual and motoric processes. Specially, it will address one specific kind of sensorimotor response that represents the collaboration between perception and action, namely the process of eye movement control.

Evidence from language research

In the language domain, evidence supporting the grounded cognition account comes mainly from neuroscientific and behavioural research. Studies using neuroscientific methods have been focusing on the question of whether the same neural mechanisms are activated during language comprehension and related perceptual or motoric processing. Meanwhile, the primary attempts of the behavioural approach have been at relating language comprehension to the mechanisms responsible for perceptual processing and generating motoric actions.

Evidence from neuroscience

Brain-imaging studies have revealed that as well as the classic language areas (e.g. Broca's), language comprehension involves the activation of cortical areas that are associated with sensorimotor processing of the object or action denoted by the linguistic stimuli (Martin, 2007; Pulvermüller, 1999). It has been shown that exposure to verbs that refer to actions, or nouns that refer to tools can activate related motor areas in the brain (e.g. Preissl, Pulvermüller, Lutzenberger, Birbaumer 1995; Pulvermüller, Lutzenberger, & Preissl, 1999). Martin and colleagues (1996) reported that naming animals selectively activated brain regions involved in visual processing while naming tools discriminatively activated premotor areas, as animals were more likely to be made distinctive based on their visual features whereas tools were frequently characterized by their manipulative or functional properties. More importantly, these language-related neural activations can be fine-grained and induced by information *implied* by the linguistic stimuli. For example, verbs referring to movements of the face, arm or leg (e.g. "*lick", " pick" or "kick"*) can activate the corresponding motor and premotor areas involved in performing these actions (Hauk, Johnsrude, & Pulvermüller, 2004). Furthermore, these activations can occurr within 170 ms after the words were presented (Pulvermüller, Shtyrov, & Ilmoniemi, 2005), indicating that the access of verb meanings through activating relevant motoric areas is likely to be an automatic process.

If language comprehension indeed recruits related brain areas responsible for perceptual or motoric processing, it can be naturally hypothesized that perceptual and motoric processes will also be under the influence of language comprehension. Using TMS (Transcranial Magnetic Stimulation), Glenberg and colleagues (2008) found that comprehending sentences describing transferring actions had a greater modulatory effect on activities in hand muscles when compared to sentences that did not describe transfer. This modulatory effect can be effector-specific: Sentences denoting foot-related actions can influence activities in foot muscles while activities in hand muscles are modulated by sentences describing hand actions (Buccino, Riggio, Melli, Binkofski, Gallese, & Rizzolatti, 2005).

Evidence from behavioural research

One of the main predictions of the grounded cognition approach is that when a sentence or word is being processed, features of objects being referred to, or relevant information in events being described are activated and represented. It has been found that the same object noun can activate different features depending on its sentential context (Zwaan, Stanfield, & Yaxley, 2002). Response latencies to pictures depicting objects of a particular shape (e.g. an eagle with open wings) were shorter after hearing sentences implying the same object shape (e.g. *The ranger saw the eagle in the sky*) compared to the ones that implied a different shape (e.g. *The ranger saw the eagle in its nest*). These results have been taken as evidence that implied perceptual characteristics in sentences can modulate perceptual simulations. The same view holds for motoric simulations: Directionality implied by sentences can affect the execution of hand movements; a phenomenon termed the *Action-Sentence Compatibility Effect* (Glenberg & Kaschak, 2002). These researchers reported that it took participants longer to make hand movements in one direction (e.g. toward the body) to respond to a sentence implying the opposite direction (e.g. the sentence "*Close the drawer*" implied motion away from the body). It is worth noting, however, that no control condition was employed in this study. Therefore, it remains unclear whether the sentences facilitated the hand movements when the directions implied were the same as the actions, or impaired the hand movements when the directions implied by the sentences were the opposite. Nonetheless, this study demonstrated that directionality implied by sentences describing actions was represented during sentence comprehension.

Objects and events described by single words can be simulated, too. Estes and colleagues (2008) presented subjects with nouns denoting objects that tended to appear at high or low spatial locations (e.g. bird, puddle) and measured response time to perceptual discrimination targets displayed either in the upper or lower visual field. It was found that when the spatial location implied by the noun was congruent with the location of the target, discrimination performance was impaired. These authors interpreted the results as the consequence of perceptual simulations: First, attention was reflexively allocated to the location implied by the word and then a perceptual simulation was activated by the word. Discrimination was hindered because the perceptual system was engaged during object simulation at the attended location and therefore less available for the discrimination task at the same location. While nouns can produce simulations of object-related features, verbs can activate action-related or motion-related features. Some researchers used the *random-dot stimulus* to measure the potential influence of verbs on motion perception (Meteyard, Bahrami, & Vigliocco, 2007). The random-dot stimulus is typically used to measure motion detection thresholds, which contains a group of moving dots with a certain proportion of them moving coherently in the same direction while the others moved randomly in different directions. It has been reported that when verbs describing motion in a specific direction (e.g. "*rise*", "*sink*") were presented, motion detection threshold was affected, depending on the congruency between the directions of the implied motion and the coherent motion in the random-dot stimulus. These authors argued that it was the simulated motion induced by verb comprehension that biased the motion detection threshold and the results could be taken as evidence supporting the view that language comprehension shares the same mechanism as perception.

Taken together, language comprehension shares, at lest some, neural and representational substrates with perceptual processing or motoric response generation. Language comprehension, and processes behind perception and action have a finely tuned relationship in which even information implied by language can provide the linkage. It is worth emphasizing that by adopting the grounded cognition view on language comprehension, not only has progress been made for psycholinguistic research, but also an immediate link between language and sensorimotor processing has been made significantly evident.

Issues of a grounded view on language comprehension

Despite being supported by abundant empirical evidence, there are several unsolved issues within the grounded cognition approach to language comprehension, especially when the main theoretical concept, simulation, is concerned. A significant number of experimental studies have demonstrated the close relationship between sensorimotor activations and language comprehension; however, few theories have been developed to explain the underpinning mechanisms. Consequently, it is not surprising that there is a mixed pattern of facilitatory and inhibitory experimental outcomes in the literature (See Bergen, 2007 for a review). For example, as described in the previous section, the ACE study by Glenberg and Kaschak (2002) found facilitatory effects for language-mediated simulations whereas Estes and colleagues (2008) reported inhibitory effects. Attempts have been made to account for the varied results in the literature (Borreggine & Kaschak, 2006) and the *Theory of Event Coding* (e.g. Hommel, Müsseler, Aschersleben, & Prinz, 2001) has been brought into the picture to provide a solution to these issues.

The Theory of Event Coding (TEC) aims to explain interactions between perception and action. This theoretical framework proposes a common, feature-based representational medium that underlies the perception of objects and action planning. Various features of the percept or action plan, such as shape or direction, are coded and temporarily integrated into "event codes" in the interaction between perception and action planning. The integration processes that eventually create event codes serve to bind features that are relevant to the object being perceived or the action being planned. Thus there are two processing stages involved in action planning. Initially, all the perceived relevant features to the action plan (e.g. the direction of the action needed) are activated. Based on these activated features, an action is then selected for execution. It is during the second stage that the relevant features are integrated into the representation of the selected action to form an *event code* thus become less available to the planning of subsequent actions.[1]

Based on a series of experiments comparable to the original ACE study, Borreggine and Kaschak (2006) suggested that the time interval between language processing and the planning of the required motoric responses was the crucial factor in determining whether language-mediated simulations would produce facilitatory or inhibitory effects. If the required action was to be planned after the relevant features were activated during simulation, priming would occur, and the planning of the action would be facilitated. However, if too much time had elapsed and the simulating processes had been completed, these activated features would have been bound into an event code for the simulated action, hence becoming less available to the to-be-planned motoric response. In short, if a common feature is shared by both the simulated and required action (e.g. the direction of the hand actions in the ACE), facilitatory effects for the to-be-planned motoric response are expected before the simulation is completed; whereas the planning of the required motoric response will be disrupted after the shared feature is bound to the finalized simulation.

The TEC provides a possible account for the mixed facilitatory/inhibitory findings in the grounded cognition literature. Furthermore, the common feature/event coding mechanism presents itself as a potential interface between language comprehension and sensorimotor responses. However, there are several limitations on establishing the link between language processing and perception and action via the TEC: First, the TEC is a theoretical framework that deals specifically with action planning (Hommel et al., 2001), thus it excludes the action execution stage and limits the interaction between language processing, perception and action to the phase *before* or *during* action planning. As a result, the TEC is not sufficient to complete the grounded view on language comprehension, as some evidence indicates that the period during which language processing can make contact with sensorimotor responses extends beyond the action planning stage. For example, Zwaan and Taylor (2006) reported that reading sentences implying clockwise/anti-clockwise rotating actions could influence *concurrent continuous* manual rotation online. Second,

---

[1] An *event code* can be considered as a unified representation of the planned action as well as the features based on which the planned action is selected.

by attempting to create the language-sensorimotor connection via the TEC, a qualitatively different mechanism along with additional assumptions have to be introduced and incorporated into the grounded cognition framework. Therefore the TEC may not necessarily provide the most parsimonious and straightforward account for the observed interactions between language comprehension and perception and action, especially given that such concepts and processes as "feature codes" and the "binding of feature codes" have yet been explicitly specified in the first place. Consequently, applying the TEC to explain how language affects sensorimotor processing essentially adds further assumptions to the concept of simulation.

In sum, the biggest challenge for the simulation view on language comprehension remains to be the question of how language processing affects sensorimotor responses, or more specifically, how language-mediated simulations introduce biases into perceptual or motoric processing.[2] In order to resolve this issue, the research presented in the current thesis proposes a common attentional mechanism to be the critical link between language comprehension and perception and action. Furthermore, this thesis specifically defines the concept of attention using the guided activation theory of cognitive control (Miller & Cohen, 2001), which has been developed through building computational models (e.g. Cohen, Dunbar, & McClellan, 1990) that are firmly based on neurophysiological findings (e.g. Desimone & Duncan, 1995). The following section draws an outline of this common attentional mechanism with respect to the guided activation theory. Additionally, given the strong tie between attention and eye movements, some relevant research on the relationship between language comprehension and eye movements is subsequently reviewed and discussed.

Attention – The missing link

Many models have been proposed to define the concept of attention (e.g. Posner, Rafal, Choate, & Vaughan, 1985; Allport, 1993; Klein, 1980; Henderson, 1992). Based on evidence from neuroscientific research (e.g. Braver & Cohen,

---

[2] Many challenges have been raised for the simulation view in the field of psycholinguistics research (Zwaan, 2009). The present thesis will, however, refrain from discussing these issues due to the focus of the present research.

2000; Desimone & Duncan, 1995; Rougier & O'Reilly, 2002) and principles of the parallel distributed processing (PDP) framework (e.g. McClelland, 1993; O'Reilly & Munakata, 2000; Rumelhart, McClelland, & PDP Research Group, 1986), Cohen and colleagues (1990) developed a model of attention using the Stroop task (Stroop, 1935), which illustrated the guided activation theory of cognitive control (Miller & Cohen, 2001).

Attention at a system level

In the Stroop task (Stroop, 1935), one featural dimension of a stimulus must be attended to while another competing but more salient feature must be ignored, for example, when subjects have to name the colour in which a word is printed when the word meaning is incongruent, such as the word *green* printed in red. In the model of the Stroop task (Cohen et al., 1990), attention is defined as the task-specific activation of the neural circuits involved in perception and action, which biases the ultimate behavioural outcome. Connections are naturally stronger in the word-naming pathway in the Stroop model, according to the assumption that written words are more frequently and consistently associated with their pronunciations than is the percept of colours with the articulation of their names. Thus, the word is naturally more likely to be read out than the colour to be named when the word is printed in an incongruent colour (e.g. the word "*green*" printed in red). In other words, behaviour is biased by the more salient pathway or representation. However, the difference between the word-naming and colour-naming pathway can be biased in the opposite direction when a task (i.e. name the colour) is implemented. In this case, the weaker dimension of the stimulus, the colour, is responded to behaviourally (i.e. named). Therefore in this model, attention is defined as the influence that the task representation has on the word and colour processing pathways.

This model is in line with the *biased competition theory* (Desimone & Duncan, 1995) derived from neurophysiological evidence. This theory proposes that at any given point, there are multiple representations active in the brain and these different representations compete for behavioural expression. Attention is viewed as the regulatory force that biases this competition in favour of some of the representations over the others by modulating their relative activation strengths. The source of bias comes from various contextual factors ranging from high-level cognitive (e.g. task demands) or low-level perceptual (e.g. contrast).

One important feature of the model is that attention does not arise from a qualitatively distinct mechanism. Instead, it is the consequence of the functional principles of the system, which encompass perception, action and cognition. Attention is the influence that one representation has on the selection process of which, or to what extent, other representations should be processed. In other words, attention functions at a *system level*.

The relevance of the guided activation theory (Miller & Cohen, 2001) and the biased competition theory (Desimone & Duncan, 1995) for the grounded cognition framework (e.g. Barsalou, 1999, 2007) lies in the functional principles of attention suggested by these two theories: The observed interactions between language comprehension and perception or action can be viewed as competition between language-mediated simulations (or representations) and representations activated during perceptual or motoric processing. The consequent biases exhibited in the sensorimotor responses reflect the ultimate outcome of this competition regulated by an attentional mechanism that governs the whole information-processing system. Although the idea of applying attention to connect language comprehension to perception and action is novel, the guided activation theory and the biased competition theory can nevertheless account for past empirical findings in the literature. The next section starts with an overview of the research on language-mediated eye movements in the grounded cognition literature, followed by an account of how attention can act as the interface between language and saccades, illustrated in the context of one particular empirical study (Altmann & Kamide, 2009). Finally, reasons why further research using a different type of eye movements is necessary are suggested.

Eye movements, attention and simulations

Saccadic eye movements have been used as a major behavioural measure to demonstrate the interaction between language comprehension and sensorimotor responses. According to the premotor account of attention (Rizzolatti, Riggio, Dascola, & Umiltà, 1987; Rizzolatti, Riggio, & Sheliga, 1994; Sheliga, Riggio, Craighero, & Rizzolatti, 1995), shifting covert attention relies on the same neural mechanism as generating saccadic eye movements. Given the close relationship between attention and saccades, it is natural to assume that attention may act as the mediating factor between language and eye movements.

The visual-world paradigm

Studies illustrating language influences on saccadic eye movements have typically adopted a method named the *visual-world paradigm*. The visual-world paradigm refers to the experimental methodology in which subjects are typically presented with pictures of objects or visual scenes and auditory linguistic stimuli, while their eye movements are monitored and recorded. Cooper (1974) was the first to adopt this paradigm and reported that the eyes seemingly moved reflexively to the objects in the concurrent visual scene when these objects were being referred to by spoken language. The visual-world paradigm has since become one of the main methodologies used in psycholinguistic research and the relationship between saccadic eye movements and language has been reliably established (see Henderson & Ferreira, 2004 for a review).

The visual-world paradigm is of special interest for the grounded view on language comprehension. This is largely due to the nature of eye movements as a type of motoric response. Eye movement is a typical example of the interaction between processes behind perception and action. As sensory organs, the eyes are distinct from the others, as the eyes move to actively perceive information from the environment but at the same time the perceived information can in turn drive the eyes to move, determine where they move to next, and how they move to that location in extrapersonal space. There is not another type of human organ that depends on the interaction and cooperation between perception and action as much as the eye. By bringing in linguistic stimuli, the visual-world paradigm has created a perfect platform for studying the interaction between language comprehension and perception, by controlling the sensory input, and action, by monitoring the eye response.

Using this paradigm, the simulation view has received even more supporting evidence. For example, Huettig and Altmann (2007) reported that upon hearing the word "*snake*", saccades were immediately launched towards a depiction of an electric cable among a visual display of four objects, indicating that object shapes were activated during the simulation of object names. More interestingly, some observations have indicated that simulation manifested not only in looks to a target object referred to, or implied by the language, but in patterns of looks across objects and the space these objects occupy. For example, in a study by Spivey and Geng (2001), subjects listened to descriptions of spatiotemporally dynamic scenes while facing a blank screen. These descriptions referred to spatial locations progressively further along a particular dimension, for example, a description of a scene containing a tall building might start from the bottom floor and then moved gradually upwards. It was found that while facing the blank screen, subjects tended to make saccades in the same direction as the spatiotemporal dynamics of the scene description. In a later study, Richardson and Matlock (2007) confirmed these findings but with sentences describing fictive motion (e.g. *The road runs through the valley*). Subjects looked at visual scenes that contained a path region while listening to sentences. More path-scanning eye movements were made if the fictive motion was implied to be difficult by a prime sentence (e.g. *The valley is covered with ruts*). However, this was not the case for sentences without fictive motion (e.g. *The road is in the valley*). Based on these results, the authors argued that sentences describing fictive motion activated mental simulations of motion, and certain features of these simulations (i.e. the difficulty level of the simulated motion) were reflected in the saccades executed during scene viewing.

Despite providing support for the simulation view, on the other hand, language research using the visual-world paradigm has also raised the fundamental question suggested earlier: Through what mechanism and process do language-induced simulations bias perception or action? More specifically, how are sensorimotor representations translated into corresponding eye movement patterns? This question is considered in the context of the following example study: Altmann & Kamide (2009) presented subjects with a visual scene containing a woman, a table and, on the floor, an empty wine glass and a bottle. The scene was shown for a few seconds and while facing a *blank screen*, subjects listened to sentences such as: "*The woman will move the glass onto the table. Then, she will pick up the bottle and pour the wine carefully into the glass*" (i.e. the "moved" condition) or "*The woman is too lazy to move the glass onto the table. Instead, she will pick up the bottle and pour the wine carefully into the glass*" (i.e. the "unmoved" condition). In the "moved" condition, more saccades were launched in the blank screen to where the tabletop had been upon hearing "the wine carefully into" compared to where the glass had been. However, the converse was true in the “unmoved” condition: More looks were directed to where the glass had been than where the tabletop had been. Under the simulation view, although the glass remained on the floor in the memory of the visual scene, the event described in the unfolding sentence (i.e. glass being moved to the table) would be simulated during the blank screen period. The simulation of the described event could be considered as the cause of the differences in the saccadic patterns between the “moved” and the “unmoved” condition. However, what simulation did was to ‘enact’ the movement of the glass to the table, which should not wipe out the memory of the glass having been on the floor. Thus there must have been *competition* between the simulated event (i.e. glass being moved to the table) and the memory of the visual scene (i.e. glass being on the floor), given that the eyes can only fixate at one location at a time. At this point, it is apparent that there is a missing link between simulated representations and the motoric responses of eye movements. The simulation view alone cannot explain how language comprehension shapes saccadic eye movements and some additional mechanism certainly needs to be incorporated into the grounded cognition framework so that whether and how a simulation can explicitly bias behavioural responses can be predicted and accounted for.

These issues can be resolved if attention is considered to be the interface between language and perception and action.[3] Considering the “move the glass” study (Altmann & Kamide, 2009), both the simulated event (i.e. glass being moved to the table) during language comprehension and the memory of the visual scene (i.e. glass being on the floor), could potentially initiate saccades to the location being referred to. According to the biased competition theory (Desimone & Duncan, 1995), the representation of the glass being on the table and the representation of the same glass being on the floor would be simultaneously active and thus compete with each other for behavioural expression. More specifically, each of these two concurrently activated representations would act to bias saccades to one location over the other. Under the guided activation theory of cognitive control (Miller & Cohen, 2001), the

---

[3] Some studies have hinted at the role of attention (e.g. Estes et al., 2008; Spivey et al., 2000). However, in these studies, attention has been treated as a peripheral factor outside the theoretical construct, rather than the mechanism that renders simulations into biases in perceptual experiences or motoric responses.

competition is regulated by attention that governs the whole system, which favours one representation over the others. During the blank screen period, the dominant task was to comprehend the auditory sentence, since the task of looking at the visual scene expired with the removal of the scene. Consequently, the representation of the glass being on the table activated by the sentence was more task-relevant and would become comparatively stronger in activation strength. Biased by the more active representation, increased numbers of saccades were launched to where the tabletop had been.

Taken together, in order to complete the picture of the relationship and interaction between language comprehension and sensorimotor processes, an intervening mechanism is needed to "translate" the product of linguistic processing into motoric responses. Attention that functions at a system level so far provides the most ideal mediating mechanism to fulfil this function, given that it is not modal-specific, it arises from the general functional principles of the system, and testable predictions can be generated from it.

Problems with the visual-world paradigm

There are several problems with using saccadic eye movements and visual scenes (or pictures of objects) to study the relationship between attention, eye movements and language comprehension under the grounded cognition framework. Some of the issues are methodological, while the others are theoretical. The first empirical problem comes from the transient nature of saccades. Typically, three or four saccades are executed each second, with durations of no more than 400 ms each. In order to uncover the time course of the attentional effect on motoric responses caused by language comprehension, a more long-lived and continuous behavioural measure would be desirable, so that the dynamic nature of the underlying representations could be better probed. Furthermore, it has been suggested that attention is only involved in saccadic programming but not execution (Rizzolatti et al., 1987). Therefore, by measuring saccadic behaviour, only the "after effect" of attentional shifts is revealed, rather than the "real-time" modulation, as attention is not sustained during saccadic execution.

More importantly, there is a theoretical issue in adopting the visual-world paradigm to investigate the interaction between language comprehension and sensorimotor responses: In typical visual-world studies, the visual stimuli used are often scenes comprising people and objects that could potentially interact. These visual scenes thus contain rich semantic and contextual information that cause them to be readily related to the linguistic stimuli presented concurrently. Thus, in visual-world studies, language-mediated eye movements reflect the interaction between the semantics of the language and the potential semantic interpretation of the scene. The observed results may not necessarily reflect the effect language processing has on the motoric system responsible for generating saccades, but rather the mapping between visual semantics and linguistic meaning. In order to directly test language influences on eye movements, the visual stimulus should ideally be devoid of semantic information, so that its interpretation is purely perceptual without mediating any semantic processing.

Pursuit eye movements as an alternative behavioural measure

Pursuit eye movement refers to the ability of the eyes to track a moving object in extra-personal space. It serves the purpose of stabilising the object image on the retina to maintain high visual acuity. Compared to saccades, pursuit eye movements are more advantageous as a behavioural measure for both methodological and theoretical reasons.

Methodological advantages

Unlike saccades, which involve fast but transient attentional shifts, pursuit is continuous, with attention sustained during its initiation (Ferrera & Lisberger, 1995), maintenance (Chen, Holzman & Nakayama, 2002), and termination (van Donkelaar & Drew, 2002).[4] Thus, pursuit offers an opportunity for the observation of the exact time course of the potential language influence on the oculomotor system and the attentional effect can in turn be studied online.

Theoretical advantages

In contrast to the pictures and scenes customarily used in the visual-world paradigm, the visual stimuli usually employed to induce pursuit eye movements are simple moving geometric shapes or patterns (e.g. a dot). These shapes and patterns convey no semantic or contextual information and require no semantic

---

[4] Generally, the initiation stage refers to the period from the onset of the moving target until the eyes have accelerated to a velocity that is close to the target velocity; while the termination stage denotes the time span between when the eyes start to decelerate and the point the eye velocity decays to zero. Pursuit maintenance refers to the period between the initiation and the termination stage.

information for their interpretation. In principle, pursuit eye movements could simply reflect an interaction between the position/velocity of the pursuit target and the position/velocity of the eye. In the pursuit case, in order to perform the oculomotor task, it is not necessary to relate linguistic meaning to either the perceived visual motion or the eye movements through semantics. Thus linguistic stimuli presented in parallel to pursuit should be incidental to the eye movement itself. If language can influence pursuit eye movements, it can only do so through the interaction between semantic representations activated by the linguistic stimuli and the visual perception of the motion or the oculomotor response *per se*. Consequently, any language effect on pursuit eye movements reflects an original relationship between language comprehension and sensorimotor responses, and can be taken as direct evidence that language processing shares the same representational substrates as perception and action.

At first glance, it seems unlikely that language processing will have any influence on pursuit eye movements, except for some potential distractor effect. However, assuming that attention functions at the level of the whole system, any competition between two representations or processing pathways can result in biases in behaviour (Cohen et al., 2004). This means that if there is any conflict between representations (or one of the representations) activated during language comprehension and representations (or one of the representations) required by the pursuit task, the outcome of these competitions should reflect in the eye movement pattern. For example, simulating directional verbs involves the activation of representations of motion in specific directions (Meteyard et al., 2007). On the other hand, in order to pursue a moving target, its motion direction must be determined. Furthermore, the ability to predict any future changes in the direction of the motion is crucial to whether the target can be successfully tracked. As a result, the representation of directionality could be one of the possible representational substrates within which the processes involved in language comprehension and pursuit eye movements could interact and compete with each other. For the benefit of the guided activation theory (Miller & Cohen, 2001), the demonstration of competition between language processing and pursuit eye movements in the absence of any semantics will reinforce the assumption that attention operates at the level of the system, indiscriminately towards all types of functioning.

Finally, investigating language effect on pursuit eye movements is also theoretically interesting for the psychophysical research of this type of eye movement itself. While there is abundant evidence demonstrating the cognitive influence on saccadic eye movements (Hutton, 2008), significantly less research has focused on the question of which cognitive factors can modulate pursuit. This is because traditionally, pursuit is thought to be a pure sensorimotor response and not susceptible to deliberate and voluntary control (Kowler, 1990). It is not until Yasui and Young (1975) demonstrated that stable pursuit relied on the internally represented target motion instead of perceived motion alone that cognitive influences on pursuit eye movements started to attract research interest. Various cognitive factors, such as anticipation, memory and attention, have been revealed to be important for pursuit in different contexts (e.g. Becker & Fuchs, 1985; Kowler & Steinman, 1979; Krauzlis & Miles, 1996; Pola & Wyatt, 1997). For example, it has been reported that the internal representation of target motion can sometimes serve as a short-term storage. Velocity and direction information can be held in this store and released later to drive pursuit response with a velocity scaled to the stored target velocity (Barnes & Asselman, 1991). More importantly, Jarrett and Barnes (2002) have found that subjects were able to store multiple levels of target velocity/direction simultaneously. If different patterns of target motion were preceded by certain symbolic cues (e.g. a square), subjects were able to associate these target velocities/directions with their corresponding symbolic cues through learning, and later to use the cues to predict the upcoming target velocity/direction and generate appropriate anticipatory pursuit responses. These results are taken as evidence that anticipation and memory are both involved in the control of pursuit eye movements. To demonstrate attentional effects during pursuit, Ferrera and Lisberger (1995) presented subjects with a single distractor at the same time as the pursuit target. When the distractor moved in a different direction from the target, the initial eye velocity was the vector average of the responses that would be made to the target and the distractor separately. Based on these results, it was proposed that before pursuit initiation, the target had to be identified by a selection process modulated by attention. The involvement of attention in pursuit is also illustrated by the finding that in the complete absence of retinal target motion, pursuit eye movements can be generated by having the subjects deliberately shifting their attention back and forth (Barnes, Goodbody, & Collins, 1995).

The general conclusion to be reached is that similar to saccades, pursuit eye movements are under the modulation of cognitive influences coming from memory, anticipation and attention. However, in order to study the cognitive effect on pursuit, the psychophysical approach tends to manipulate the stimuli at a sensory level (e.g. by including visual distractors or visual/auditory cues that can be used to predict target motion). Consequently, unlike saccadic eye movements, high-level processes involved in pursuit have not been thoroughly studied through the deployment of stimuli that are incidental to the pursuit task or associated with richer semantic contexts. More specifically, it remains unknown whether semantics carried by language can affect pursuit eye movements in the same way as saccades. The demonstration of any potential semantic modulation on pursuit eye movements would be important for both research areas on language comprehension and eye movements: Despite being a methodologically more advantageous behaviour measure, pursuit eye movements reflect the intimate interaction between perception and action that can only be related to language processing through perceived motion features or the eye movement *per se*.

Current research

The research described in the present thesis attempts to establish the missing link by introducing a pre-existing model of attention (Cohen, Aston-Jones, & Gilzenrat, 2004; Desimone & Duncan, 1995) into the grounded cognition framework in order to connect language-activated sensorimotor representations to the subsequent eye movement patterns. Language effects on two types of eye movements, pursuit and saccade, were explored in light of this model of attention. Conclusions will be drawn that point toward a cognitive system that shares representational substrates with perception and action, which is regulated by an attentional mechanism that functions at a system level.

We take the view that high-level cognition shares the same representational substrates as perception and action: According to the grounded cognition literature, language comprehension relies on the same mechanisms for perception and action. It is achieved through mentally simulating past perceptual or motoric experiences. The guided activation theory of attention (Miller & Cohen, 2001) provides an account for how language-activated mental simulations/representations can bias behavioural outcomes. However, the direct interaction between language comprehension and sensorimotor responses via a system-level attentional mechanism has yet to be demonstrated, given the limitations of the visual-world paradigm and the nature of saccadic eye movements. Another type of eye movements, pursuit, seems to be a more efficient behavioural measure to address the question of whether language can impact on the oculomotor system per se, without going via the route of overlapping semantics between linguistic and visual processing. On the other hand, despite the recently growing interest in the cognitive influence on pursuit eye movements, there is yet any demonstration of the effect from an external stimulus on pursuit other than sensory stimuli. Any potential semantic influence reflected in pursuit eye movements will provide a significant reconciliation between the psychophysical theories of oculomotor control and the cognitive accounts of the relationship between perception and action.

In order to address these issues, we investigated the effects of single word meaning on the oculomotor responses of both pursuit and saccadic eye movements. More specifically, we tested whether directional verbs (e.g. *climb, dive*) and nouns with a spatial component (e.g. *attic, basement*) could affect pursuit or saccadic eye movements. These directional verbs and spatial nouns were chosen as a result of the precedents in the literature (Estes et al., 2008; Glenberg & Kaschak, 2002; Meteyard et al., 2007). We opted for single words instead of sentences because as an initial attempt, we felt the need for language stimuli that could be the most efficiently controlled for their linguistic variables such as frequency. The logic behind the experiments came directly from the guided activation theory of attention (Miller & Cohen, 2001): The patterns of oculomotor responses should vary as a function of the competition outcome between representations simultaneously activated by the language comprehension and the oculomotor task. The linguistic and the sensorimotor stimuli always shared a single featural dimension (motion direction or spatial location) and depending on whether there was competition between the featural representations activated by the linguistic stimuli and the oculomotor task, different eye responses would be observed. For example, pursuit performance might be affected if representations of downward and upward motion were activated concurrently by the word (e.g. *dive*) and the pursuit task (e.g. pursuing a downward moving target).

Although the basic assumption was shared by all experiments reported in the present thesis, there were distinct manipulations and research questions involved in each of the experiments. The question of whether verbs implying upward or downward motion could influence pursuit response to a vertically moving stimulus was dealt with first (Chapter 2). Following the establishment of the verb semantic effect on pursuit, the focus was shifted to the role of the visual stimulus in the interaction between language and eye movements (Chapter 3). The involvement of the visual motion used to induce pursuit eye movements was either attenuated or extinguished completely to test whether the original semantic effect persisted. The subsequent chapter (Chapter 4) continued to investigate to what extent the semantic effect could be generalised by changing the dimension of the visual motion or altering the types of words used. A new paradigm was then introduced and evaluated to answer the question of whether biases in attention could be artificially created during pursuit (Chapter 5). Finally, another type of eye movements (i.e. saccades) other than pursuit was examined under the influence of language comprehension (Chapter 6).

The fundamental goal of the present research is to enable a theoretical vision of high-level cognition sharing the same representational substrates as perception and action with attention, which arises as a result of the system-level functional principles, as the mediating and coordinating force. Above all, as living organisms, our primary task is to perceive and act on the external world and any other types of functioning, high-level or not, are to make the interactions between us and the world possible and successful.

Chapter 2. The effect of verb semantics on pursuit eye movements

Introduction

This first empirical chapter describes three experiments that represent the initial attempt to address the question of whether language comprehension can affect oculomotor control in the absence of visual semantics. More specifically, we explored how the semantics of single verbs affect the psychophysics of pursuit eye movements via a common attentional mechanism. All three experiments reported below employed the same stimuli and procedures, but each with a different design and a unique group of subjects. Nonetheless, all experiments demonstrated a complex, but systematic, interaction between language and attention during the maintenance of pursuit.

The grounded cognition approach (e.g. Barsalou, 1999, 2007; Glenberg, 1997) proposes that language comprehension is achieved through simulating relevant past perceptual or motoric experiences. The visual-world paradigm has been extensively used to test the prediction that language comprehension is grounded in the mechanisms underlying perception and action (e.g. Knoeferle & Cocker, 2007; Richardson & Matlock, 2007; Spivey & Geng, 2001), given that this paradigm involves the processing of linguistic and visual stimuli, and the generation of motoric actions (i.e. eye movements). As discussed in the literature review, evidence collected using the visual-world paradigm may not necessarily reflect the direct impact of language processing on the mechanisms responsible for perception or action, but rather, the mapping between linguistic meaning and visual semantics.[5]

To our knowledge, the only attempt to directly address this issue is a study investigating the effect of verb semantics on motion detection thresholds (Meteyard, Bahrami, & Vigliocco, 2007). Subjects performed a coherent motion detection task with threshold coherent motion stimuli displayed using random dot kinematograms.[6] The motion was either coherent motion upward or downward, or random incoherent motion. While performing this task, subjects were auditorily presented with blocks of single verbs implying motion with a

---

[5] See the section titled "Problems with the visual-world paradigm" in Chapter 1 for a more detailed discussion.

[6] See Chapter 1 for further details.

direction that was either congruent or incongruent with the coherent motion. *Signal detection theory* (e.g. Wickens, 2002) was applied to the data analyses in which the value of *d'* represented the level of discriminability and a decrease in d' indicated poorer perceptual sensitivity. It was predicted that if language comprehension could impact on the earliest stage of sensory processing, there would be a change in sensitivity in motion detection. Based on the results, these authors concluded that motion detection sensitivity was modulated by the directional verbs and low-level visual perception could be affected by language comprehension.

There were several issues with this study: First, the blocked presentation of the directional verbs hindered the examination of the time course of the linguistic effect on motion perception. Furthermore, it is unclear what the consequences were for having a blocked design in a behavioural study when the presentations of the two types of stimuli (i.e. words and visual motion) were unsynchronised. According to the authors, semantic effects were expected to build up over all the verbs within a block. However, there is no precedent to demonstrate such "built-up" semantic effects, and it has yet to be determined how the accumulation of semantic effects is possible and the temporal course of it.

Second, the data collected in this study indicated that although a control condition was employed (i.e. a verb block that did not imply directionality), for the sensitivity measure (d'), the only statistically reliable difference found was between the congruent and the incongruent condition, with a lower sensitivity for the incongruent condition. Similarly, a difference in *criterion* (β) was only revealed in the comparison between the congruent and the incongruent condition, with a lower criterion for the congruent condition[7]. Due to a lack of difference between the control and experimental conditions, the direction of these effects remained unclear. Furthermore, the results related to the dependent measure, criterion, caused it to be more difficult to argue for a pure low-level perceptual account for the data: The criterion was found to be lower in the congruent condition compared to the incongruent condition while there was no difference

---

[7] Criterion reflects the subject's criterion for acting on the information provided by the senses. A lower criterion means that the subject is more inclined to make a confirmatory response when the signal is considered ambiguous.

between the incongruent and the control condition.[8] This means that compared to the incongruent and the control condition, when the verbs implied the same direction as the coherent motion, subjects were more likely to report perception of motion even if the strength of the signal remained constant. Thus the higher sensitivity in the congruent condition could simply reflect the consequence of a decrease in criterion. As a result, there was a possibility that instead of the low-level perceptual processes (i.e. motion detection sensitivity), what had been modulated by language comprehension was the high-order decision processes behind criterion setting, which in turn influenced sensitivity. Thus the question of whether semantic representations activated during language comprehension could affect low-level perceptual or motoric processing remains to be answered.

The following three experiments were designed to explore whether language comprehension could influence low-level sensorimotor processing when: a) there were no semantics present in the visual stimulus so that a semantic association between the linguistic and visual stimuli was not possible; b) no decision had to be made in order to complete the experimental task; and c) the type of eye movements measured (i.e. pursuit) is rarely under deliberate control once initiated. In all three experiments, verbs denoting upward or downward motion (e.g. *climb* and *dive*) were presented auditorily during the pursuit of a dot moving upwards or downwards. Based on the principles of the guided activation theory of cognitive control (Miller & Cohen, 2001), we predicted that pursuit eye movements would be modulated by the congruency between the directionality activated by the verbs and the directionality of the oculomotor task.[9]

The norming experiment

All experimental items were chosen based on the results gathered in a norming experiment. The experiment was conducted online with a web-based questionnaire, in which subjects had to answer the question “To what degree

---

[8] The authors did not report the statistics from the comparisons between the control and the congruent condition. This has made their results even harder to interpret for the readers.

[9] For a more detailed discussion of the guided activation theory of cognitive control (Miller & Cohen, 2001), see the section titled “Attention at a system level”, Chapter 1.

does this verb imply an action in a particular direction?” by rating on a scale of -3 to 3, with 0 indicating no directionality at all.[10] Negative numbers on the scale represented downward motion while positive numbers stood for upward motion. The questionnaire was randomly distributed to universities across the country.128 responses were received in total and four of them were incomplete and excluded from the analyses. 24 verbs, 12 indicating upward motion and 12 implying downward motion, were selected based on their mean ratings and standard deviations. There was no significant difference between the ratings for the downward verbs and the upward verbs (2.11 vs. 1.82, t (23) = 2.01, $p > .05$), suggesting that the both types of verbs implied motion in a specific direction to the same extent.[11]

Experiment 1

Method

Participants

Forty students from the University of York took part in this experiment. They participated in exchange for either course credit or £4. All participants were native speakers of English and had normal or corrected-to-normal vision.

Materials

a. Visual stimuli

The primary visual target to induce pursuit eye movements was a black dot presented in a uniform light grey (RGB: 180, 180, 180) background. The target dot subtended 0.86° on the display screen, with a ‘hollow’ centre of 0.22°, at a fixed viewing distance of 600 mm. The dot moved with a sinusoidally varying velocity peaking at 12 degree/s. This particular target velocity was generally considered to be the velocity level at which the target could be pursued with ease. For horizontal movements, target motion started from either the left or the right edge of the screen and terminated once the dot moved across the screen and reached the opposite edge. As a result, every trial contained half of a sinusoidal cycle, with a frequency of 0.15Hz. For vertical movements, the dot moved across the screen downward or upward from the top edge or the bottom edge. In order to keep the peak velocity constant, the frequency for the vertical

---

[10] See Appendix 1 for instructions given to subjects.

[11] For the purpose of this comparison, the responses to the downward verbs were converted positive by taking the absolute values of the original negative ratings.

motion was adjusted to 0.21Hz, as the distance to travel was shorter compared to horizontal target motion, due to the size of the display monitor.[12] Horizontally, target motion lasted for 3226 ms while vertically the dot took 2419 ms to travel to the opposite edge. Every trial was preceded by a fixation cross subtending 0.37°, which was presented for 1000 ms at the commencing location of target motion. Both the target dot and the fixation cross were created with Experiment Builder (SR Research).

b. Linguistic stimuli

24 motion verbs were employed as experimental items.[13] Half of them implied motion upward, for example, "*climb*". The other half implied motion downward, such as "*dive*". These verbs had previously been normed with an online questionnaire to ensure that they all implied directions in the intended dimension. No spatial adverbs (e.g. up, down, above, below) or metaphorical motion verbs (e.g. increase, decline) were used. 48 nouns and 24 verbs that did not imply directionality were also included as fillers and served as the control condition.[14]

c. The eye-tracker

Eye movements were recorded using an Eyelink II (SR Research) head-mounted eye-tracker sampling at 250Hz. Eyelink II is a video-based eye tracking system with a spatial resolution of < 0.005°. Eye position is recoded using pupil tracking in combination with corneal reflections. This is to reduce errors caused by headband slippage, muscle tremor and environmental vibration. Pupil-only tracking was occasionally applied when corneal reflection tracking was not possible.

Eye movement events, such as fixations, saccades and smooth movement, were identified by eye position changes using the online parser of the Eyelink II tracker. The default threshold and algorithm settings of the parser were used.

---

[12] This frequency change for the vertical movement should not affect the results of the experiments, as for a single trial, the target only completes half of a sinusoidal cycle.

[13] See Appendix 2 for the full list of experimental items.

[14] See Appendix 3 for the full list of control items. Unlike subsequent experiments, the non-directional control words used in this particular experiment were *nouns*, and they had *not* been matched for duration. The implications of this will be considered in the discussion section.

Saccades were detected based on three thresholds: Motion, velocity and acceleration. Dictated by the default settings of the parser, if an eye movement event had a velocity exceeding 30°/s, an acceleration above $8000°/s^2$ and an amplitude larger than 0.15°, it would be defined as a saccade. The parser would also mark any period of smooth eye movement under the velocity threshold of 70°/s as pursuit. However, the threshold for defining saccades during pursuit was an elevated value of 60°/s compared to other non-pursuit intervals.

d. The display screen

All visual stimuli were displayed on a 22-inch viewing monitor (Iiyama 514) with a resolution of 1024 x 768 pixels and a refresh rate of 120 Hz. The dimensions of the monitor screen essentially provided a viewing area of 36.82° x 28.03°.

e. The audio equipment

All linguistic stimuli were recorded by a male native speaker of British English and sampled at 44.1kHz. The sound files were presented through two loudspeakers placed on either side of the display monitor.

Design

This experiment used a two-way within-subjects design. There were two within-subjects factors of target motion direction with four levels (leftward, rightward, upward and downward) and directionality implied by the verbs with two levels (upward and downward).

All 24 directional verbs were paired with dot motion in all four directions yielding eight conditions. The non-directional control words were coupled with target motion in all four directions so that there were equal numbers of trials with dot motion in all four directions, and there were identical numbers of nouns and verbs paired with each direction.

The order of target motion directions was randomized, so that it was not possible to anticipate the direction of the target motion prior to the onset of each trial. The materials were arranged in a fixed-random sequence so that no consecutive trials belonged to the same condition.

Procedure

After written consent was collected, subjects were seated in front of the display monitor with their eyes roughly 600 mm away from the screen. Eye movements were recorded from the right eye although viewing was binocular.

No explicit task relating to the words was given and subjects listened to them passively. Subjects were told that the purpose of the study was to monitor the effects of words on pupil diameter during eye movements, and their task was to track the target dot as accurately as possible.

There were 12 practice trials before the experimental block. Before each trial, subjects were shown a centrally located dot and instructed to fixate it to correct any drift in the eye-track calibration. The dot was then replaced by a cross for 1000 ms. The cross was positioned at the starting point of the motion. As soon as the cross was replaced by the target dot, the motion would start. The words were presented 750 ms after the initiation of pursuit. A blank screen was presented for 500 ms after the target dot reached the opposite edge to its starting point and the trial was automatically terminated when the blank screen expired. All subjects were debriefed at the end of the experiment.

The eye-tracker was recalibrated using a 9-point fixation stimulus after every six trials, which resulted in 16 calibrations in total for a given experiment. Every calibration took approximately 20 s, during which the errors between the eye position and the calibration target were recorded. These errors were then verified in a validation procedure. If the variations in the errors between calibration and validation exceeded default threshold level, calibration was repeated until satisfactory performance was achieved.

Results

Eye position data (x and y values in screen coordinates) were sampled every 4 ms. All data sampled during saccades were separated from the data sampled during smooth eye movements and subjected to different analyses.

Two main dependent variables were measured for pursuit eye movement: Positional errors between the gaze position and the target dot, and eye velocity. Positional errors were calculated for each sample by subtracting the y coordinate of the dot position from the corresponding coordinate of the gaze position. During pursuit, the gaze position could sometimes be ahead of the target and result in positive errors and other times lag behind and produce negative errors. The absolute values of these positional errors were taken to avoid errors with different signs cancelling each other out. Thus, positional error was a measure of how far the eye was from the target, regardless of whether it was ahead or behind the target (but see below). Any value that was two standard deviations from the mean difference was excluded.[15] The remaining absolute positional errors were converted into visual angles and subjected to statistical tests to compare the means across conditions.

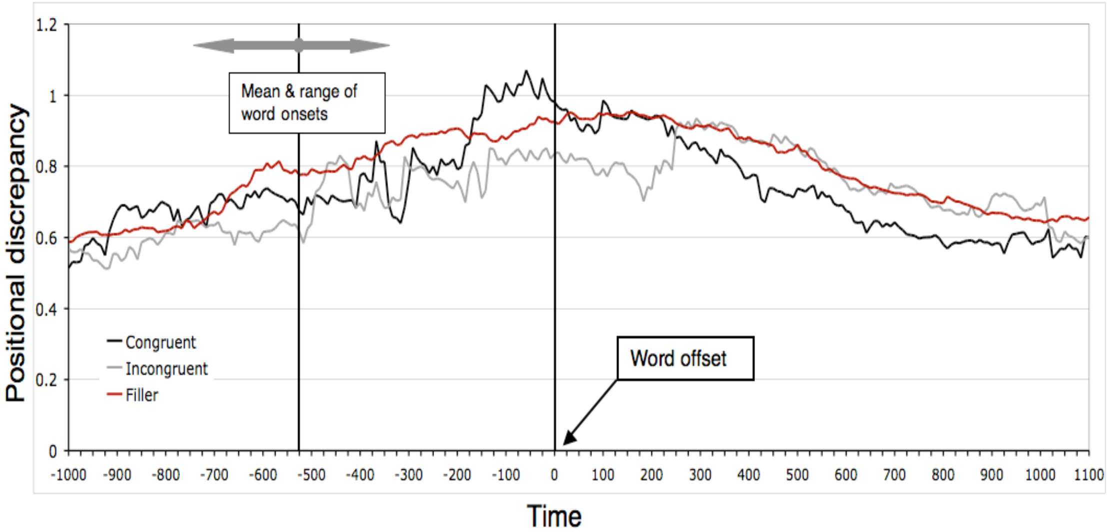


Fig.2.1. *Absolute positional errors between the gaze position and the target dot across three conditions for upward pursuit: Congruent, incongruent and control. The solid grey circle indicates the approximate point of word onset based on the mean word duration, with the leftmost arrow representing the onset of the longest word and the rightmost arrow signalling the onset of the shortest word. Time zero signals word offset.*[16]

There was no significant difference across the conditions of upward pursuit. It should be noted that the standard deviation of positional errors in the filler trials for upward pursuit was significantly larger compared to downward pursuit (F (3, 37) = 4.55, $p < .01$), indicating that upward pursuit was harder as an oculomotor task and upward pursuit performance was less consistent than downward pursuit. The lack of an effect in upward pursuit could be the result of this greater noise in the upward data. The graph (Fig. 2.1.) suggested that between 150 ms before the word offset to 250 ms after the word offset, there

---

[15] This method of exclusion was applied to all analyses reported in this thesis. The calculations for the standard deviation and the mean were conducted on all the data entered into the analysis, which included data collected from the congruent, incongruent, and the control condition but not from the filler items.

[16] Although all words were presented 750 ms after the target motion onset, the positional errors during the course of a single trial were synchronized to the word offset. This is to ensure that any effects on pursuit could be synchronized to the point at which it was certain that sufficient acoustic material had been heard to enable the recognition of the words.

could be differences among the three conditions. However, the observed difference was not statistically reliable (F (1, 39) = 1.35, $p > .05$).

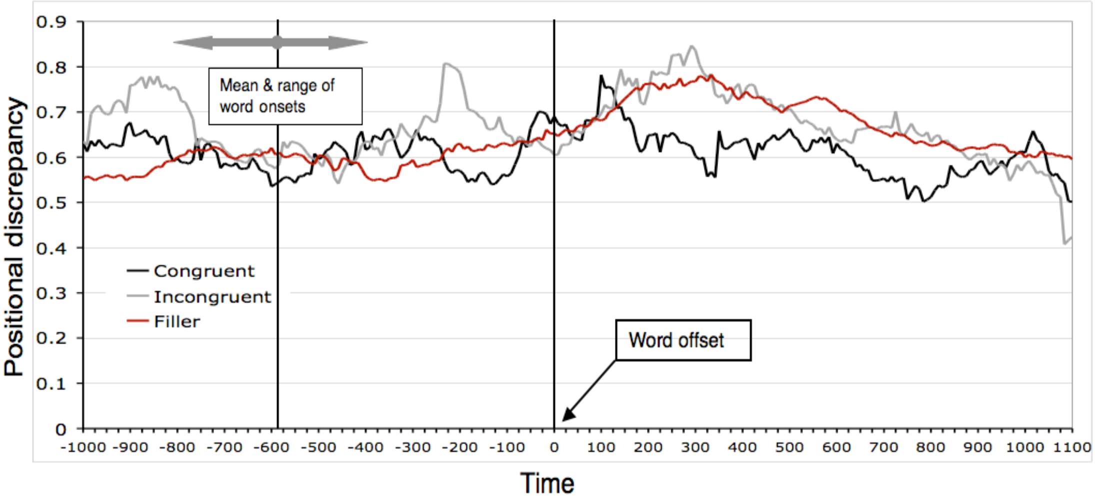


Fig. 2.2. *Absolute positional errors between the gaze position and the target dot across three conditions for downward pursuit: Congruent, incongruent and control. Positional errors were synchronized to word offset. The solid grey circle indicates the approximate point of word onset based on the mean word duration, with the leftmost arrow representing the onset of the longest word and the rightmost arrow signalling the onset of the shortest word. Time zero signals word offset.*

Inspection of Fig. 2.2. suggested that the positional errors in the congruent condition was smaller compared to the incongruent and the control condition. In addition, this decrease in positional error in the congruent condition was observed only temporarily *after* the word offset. In other words, the effect was *post-verb*.

The positional errors in a 300 ms time window between 150 ms to 450 ms after the word offset from all three conditions (congruent, incongruent and control) were subjected to a one-way within-subject ANOVA.[17] Mauchly's test indicated that the assumption of sphericity had been violated ($\chi^2(2) = 16.66$, $p < .001$); therefore multivariate test statistics were used instead. There was a significant main effect of condition (F (1, 39) = 4.12, $p < .05$) indicating that within that 300 ms time window, the positional errors from the congruent, incongruent and control condition were different. Wilcoxon's signed-rank test

---

[17] Although the difference between the congruent and the control condition seems to persist until around 1000 ms after the word offset, the difference between these two conditions in the time window from 150 ms to 1000 ms after word offset is not statistically reliable (0.71 vs. 0.78, t (39) = -1.85, $p > .05$).

revealed that the positional errors in the same time window from the congruent condition were significantly smaller than the control condition (0.63 vs. 0.75, $T = 153.00$, $p < .01$). However, no significant difference was found between the incongruent and the control condition (0.77 vs. 0.75, $T = 363.50$, $p > .05$) or between the congruent and the incongruent condition (0.63 vs. 0.77, $T = 316.00$, $p > .05$).[18]

The same results pattern also emerged from individual time point analyses: At 300 ms *after* the word offset, there was a main effect of condition (F (2, 78) = 4.25, $p < .05$). Congruent verbs significantly decreased the positional errors compared to the control condition (0.60 vs. 0.77, F (1, 39) = 4. 33, $p < .05$) while the incongruent condition did not differ from the controls (0.84 vs. 0.77, F (1, 39) < 1, $p > .05$).[19]

Finally, analyses on samples containing saccades indicated that in the same time window (between 150 ms to 450 ms after the word offset), there was no difference across these three conditions in the number of (F (2, 78) < 1, $p > .05$) or the mean amplitude of (F (2, 78) = 2.86, $p > .05$) saccades launched.

Discussion

The data from Experiment 1 demonstrated that the activation of semantic representations could interfere with the control of smooth pursuit eye movements, even though the visual environment for the pursuit task contained no semantic information, and even though the word meaning was completely incidental to the oculomotor task. However, there were several issues within this experiment: First, having 24 experimental items allocated into eight conditions resulted in only three items in each condition. Having a small number of items in each condition had in turn given rise to a considerable amount of variance in the data (See Figure 2.1 and 2.2). Second, the control and the experimental items were not matched for duration. As a result, comparisons in eye velocity between the control and the experimental conditions were not possible. This is because the velocity of the pursuit target varied constantly in a sinusoidal fashion, thus words of different durations would end at different points in the trajectory where the target/eye would be moving with different velocities. Experiment 2 was

---

[18] The unit of the mean positional discrepancies reported here is degree (°).

[19] Results regarding the effects of these vertical directional verbs on horizontal pursuit are reported in Chapter 4.

designed as a replication study with these two issues taken into consideration. In order to reduce the variance, Experiment 2 had a more powerful design by having all 24 directional verbs coupled with vertical pursuit only. Furthermore, the control items were replaced with a new set of nouns with their mean duration matched to the directional verbs.

Experiment 2

Method

Participants

A different sample of 40 students from the University of York took part in this experiment. They participated in exchange for either course credit or £4. All participants were native speakers of English and had normal or corrected-to-normal vision.

Materials

A new set of 24 nouns with matched frequency and duration to the directional verbs were selected as the control items. Another 24 nouns and 24 verbs were also included as the fillers.[20]

Design

The variables manipulated within a subject were target motion direction (upward and downward) and directionality implied by the verbs (upward and downward). All 24 directional verbs were paired with dot motion in the vertical dimension. Thus, there were only four conditions in this experiment.

---

[20] See Appendix 3 for the full list of control items.

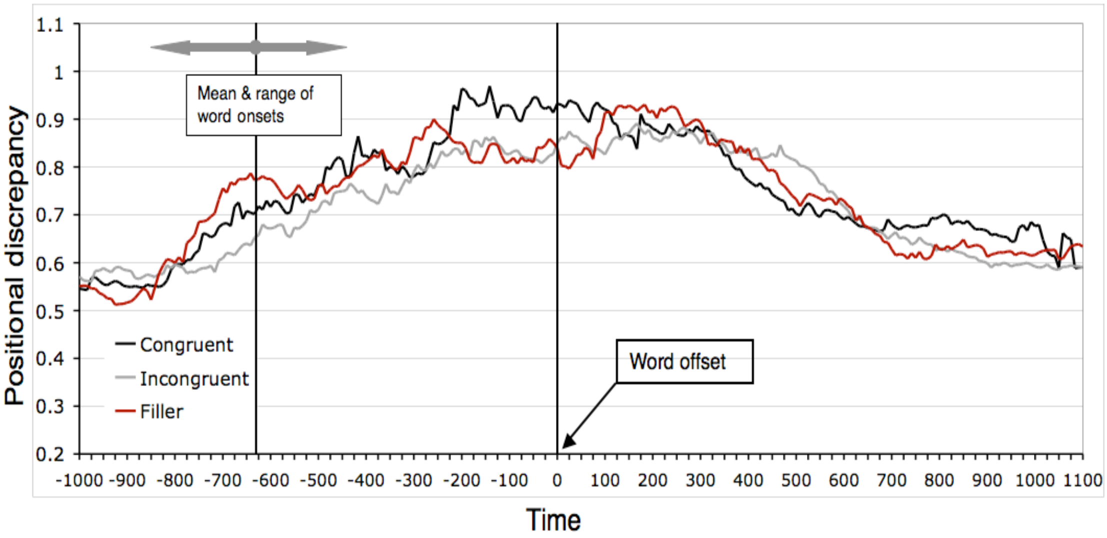


Fig.2.3. *Absolute positional errors between the gaze position and the target dot across three conditions for upward pursuit: Congruent, incongruent and control. The solid grey circle indicates the approximate point of word onset based on the mean word duration, with the leftmost arrow representing the onset of the longest word and the rightmost arrow signalling the onset of the shortest word. Time zero signals word offset.*

All effects reported here were again observed during downward pursuit only. There was no significant difference in the data from upward pursuit. These data were again considerably noisier compared to downward pursuit, with larger standard deviations of positional errors in the filler trials (F (1, 39) = 9.25, $p <$ .01).

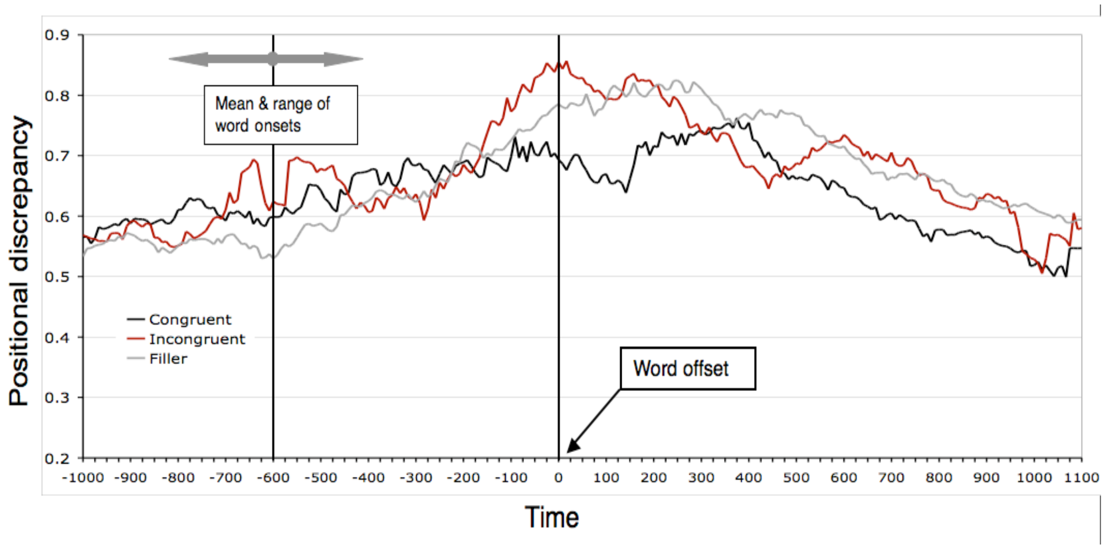


Fig. 2.4. *Absolute positional errors between the gaze position and the target dot across three conditions for downward pursuit: Congruent, incongruent and control. Positional errors were synchronized to word offset. The solid grey circle indicates the approximate point of word onset based on the mean word duration, with the leftmost arrow representing the onset of the longest word and the rightmost arrow signalling the onset of the shortest word. Time zero signals word offset.*

It is evident from Figure 2.4 that congruent words again decreased the positional discrepancies compared to the incongruent and the control condition. In a time window from 150 ms *before* the word offset to 300 ms *after* the word offset, there was a main effect of condition (F (2, 78) = 4.32, $p < .05$). Positional errors occurring in the congruent condition were significantly smaller compared to the control condition (0.70 vs. 0.78, F (1, 39) = 5.56, $p < .05$). However, there was no difference between the incongruent and the control condition (0.80 vs. 0.78, F (1, 39) <1, $p > .05$). Thus the reduction in positional errors observed in Experiment 1 was replicated, however, in a slightly different time window, which was from 150 ms *before* the word offset to 300 ms *after* the word offset. In other words, the semantic effect occurred *during* the verb.

Individual time point analyses confirmed this time window shift: At 150 ms after the word offset, there was a main effect of condition (F (2, 78) = 4.38, $p < .05$). At this moment in time, positional errors in the congruent condition were reduced compared to the control condition (0.66 vs. 0.80, F (1, 39) = 6.02, $p < .05$) while there was no difference between the incongruent and the controls (0.83 vs. 0.80, F (1, 39) < 1, $p > .05$).

Saccadic analyses once again revealed that in the same time window from 150 ms before the word offset to 300 ms after the word offset, there was no difference across conditions in the number of (F (2, 78) = 1.12, $p > .05$) or the amplitude of (F (2, 78) < 1, $p > .05$) saccades launched.

Since the experimental and the control items have been matched for duration, velocity analyses were conduced for this experiment. Eye velocity was calculated by dividing the distance travelled by the time elapsed between adjacent samples. Despite excluding all data sampled during saccadic eye movements, a period of three samples (24 ms) before and after saccade onset and offset were also excluded. This was to remove all pre-saccadic acceleration and post-saccadic velocity residue. It should be noted that unlike the positional errors, eye velocity was almost always in the direction of the motion hence the values are always positive. Therefore, there was no need to convert eye velocities into absolute values. However, for analyses using eye velocity as the dependent measure, the gaze position could impact on eye velocity differently depending on whether it was ahead or behind the target. Thus the data were split into two groups based on the viewing position of the eyes relative to the pursuit target: The "leading" cases (48.6%), where the gaze position was ahead of the target, and the "lagging" cases (51.4%), where gaze position was behind the target.[21] The mean responses across conditions in these two cases were then compared using statistical tests.

In the leading cases, compared to the control condition, congruent verbs decreased eye velocity (7.47 vs. 8.52, F (1, 37) = 8.03, $p < .05$) in the same time window (from 150 ms before the word offset to 300 ms after the word offset) while incongruent verbs increased eye velocity (8.74 vs. 8.52, F (1, 37) = 4.10, $p = .05$), although only marginally.[22] On the other hand, in the lagging cases, the congruent verbs seemed to have caused the eyes velocity to increase compared to the control condition; however, this increase was not statistically reliable (5.07 vs. 4.84, F (1, 38) = 3.97, $p > .05$). Meanwhile the incongruent words caused the eyes to decelerate significantly in contrast to the control condition (4.65 vs. 4.84, F (1, 38) = 4.15, $p < .05$).

---

[21] The proportion of leading and lagging samples were calculated over the time period from 150 ms before the word offset to 300 ms after the word offset.

[22] The unit of the velocity reported here is degree/s (°/s).

Individual time point analyses revealed the same result pattern: At 150 ms after the word offset, when the gaze position was ahead of the target (50.9 %), congruent verbs caused a deceleration in eye velocity (6.45 vs. 8.69, F (1, 35) = 6.88, $p < .05$) while incongruent verbs triggered an acceleration (8.83 vs. 8.69, F (1, 35) = 4.32, $p < .05$). However, the incongruent words decreased eye velocity in the lagging cases (4.73 vs. 5.49, F (1, 25) = 5.56, $p < .05$), which took up 49.1% of all the samples. No difference in eye velocity was found at this moment in time between the congruent and the control condition when the gaze position was behind the target (5.12 vs. 5.49, F (1, 25) = 3.78, $p > .05$).

Discussion

The results from this experiment replicated the findings of positional error differences in Experiment 1. The time window difference, that is, the effect in Experiment 2 being in an earlier time window (i.e. during the word) compared to Experiment 1 (i.e. post-verb), could be due to the more powerful design and better-controlled items of Experiment 2. The additional velocity analyses revealed complex yet systematic interactions between the eye response and language comprehension as a function of the verb semantics and the gaze position relative to the target. In the leading cases, downward verbs caused eye deceleration while upward verbs caused acceleration. However, it was almost the reversed pattern for the lagging cases. The failure to find a reliable acceleration in the congruent condition when the gaze position was lagging behind could be due to one of the following issues with Experiment 2: First, although the design was more powerful relative to Experiment 1, there were still only six items in each condition. Second, the control words were *nouns* and these nouns were not pair-wise matched to the directional verbs. Experiment 3 was designed as a replication for Experiment 2 with an even more powerful design and pair-wise matched *verbs* as the control items. An account of the velocity results of Experiment 2 is given later in conjunction with the results from Experiment 3.

Experiment 3

Method

Participants

20 participants were tested in this experiment. All were native speakers of English and had normal or corrected-to-normal vision.

Materials

The directional verbs used were identical as the previous two experiments. 24 non-directional *verbs* were selected as the control condition. They had been pair-wise matched for duration and frequency to the directional verbs. In addition, another 48 verbs were also included as fillers.[23]

Design

The directional verbs and their pair-wise matched controls were paired only with downward target motion. Upward pursuit was only performed on the filler trials. As a result, there were only two conditions in this experiment: Downward – congruent and downward – incongruent.

Results

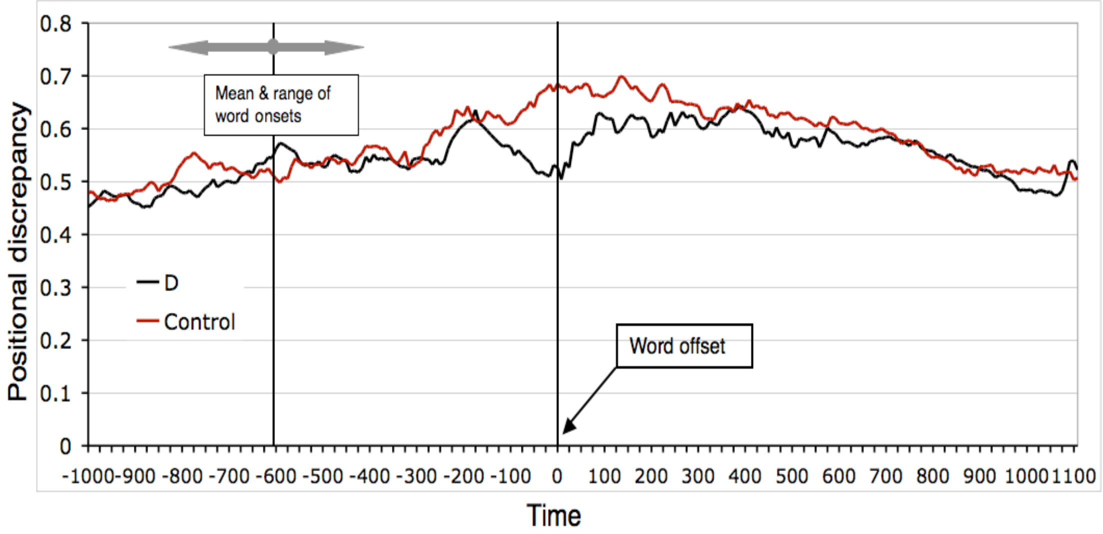


Fig. 2.5. *Absolute positional discrepancies between the gaze position and the target dot across two conditions: Downward (or congruent) verbs and their pair-wise matched controls.*

---

[23] See Appendix 3 for the full list of items.

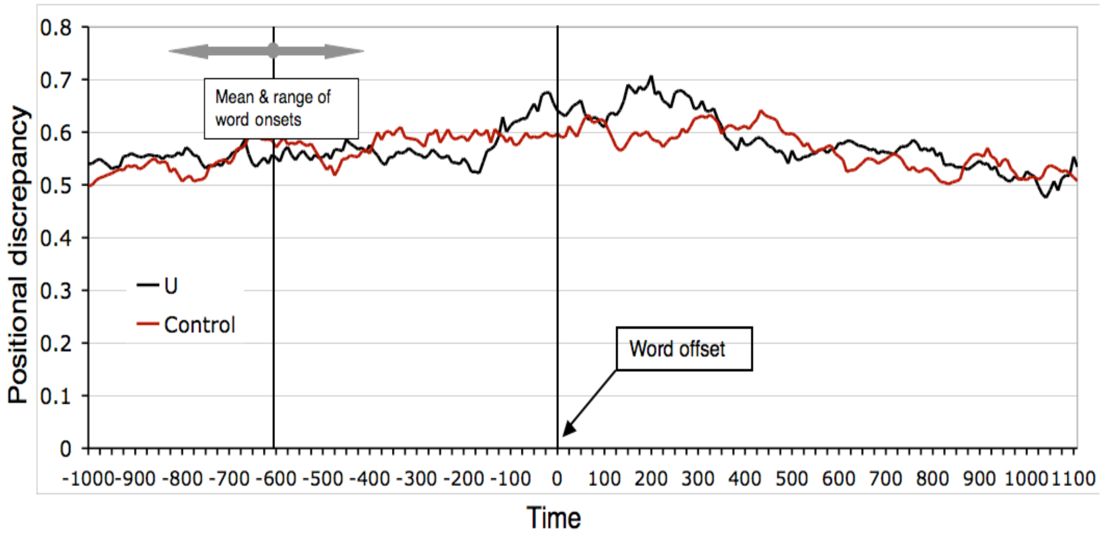


Fig. 2.6. *Absolute positional discrepancies between the gaze position and the target dot across two conditions: Upward (or incongruent) verbs and their pair-wise matched controls.*

With positional discrepancies as the dependent measure, there was a significant main effect of Condition (F (1, 19) = 4.56, $p< .05$) in a time window from 200 ms *before* the word offset to 300 ms *after* the word offset. Congruent verbs significantly decreased the positional errors compared to the control condition (0.59 vs. 0.66, F (1, 19) = 4.89, $p< .05$) while the incongruent words did not have an effect on pursuit eye movements (0.63 vs. 0.60, F (1, 19) < 1, $p >$ .05). Thus the semantic effect on pursuit eye movements seen in Experiment 1 and 2 was again replicated: Congruent verbs reduced the positional errors while incongruent verbs did not differ from the control condition. The time window in which the effect was observed in this experiment was comparable to the one in Experiment 2. However, significant differences across conditions revealed in individual time point analyses were shifted from 150 ms after word offset, as in Experiment 2, to the actual word offset. At time 0, there was a main effect of condition (F (1, 19) = 5.02, $p < .05$). Compared to their pair-wise matched controls, congruent words decreased the positional errors (0.53 vs. 0.68, F (1, 19) = 4.46, $p < .05$) but not the incongruent words (0.64 vs. 0.60, F (1, 19) <1, $p >$ .05). This shift time compared to Experiment 2 could be due to the fact that the durations of the experimental items and their controls were pair-wise matched in this experiment but not in Experiment 2. This means that in the current experiment, the target velocity in the experimental conditions and the target velocity in the control conditions were more comparable at each moment in time, given the target motion was sinusoidal. By having pair-wised matched controls, the possible confounding factor of target velocity was eliminated, and this greater sensitivity might have led to the shift in the time course of the effect that was observed in this study relative to the previous one. Finally, in the same time window where differences in positional errors were observed, no difference was found in the number of or the amplitude of saccades launched between the congruent and its control condition (F (1, 19) < 1, $p > .05$; F (1, 19) < 1, $p > .05$), or between the incongruent and its control condition (F (1, 19) < 1, $p > .05$; F (1, 19) = 2.09, $p > .05$).

Additional velocity analyses revealed the two-way interaction between eye response and verb semantics hinted by the results of Experiment 2. In the same time window as the positional analyses (from 200 before the word offset to 300 ms after the word offset), congruent words caused the eyes to decelerate (4.94 vs. 5.06, F (1, 19) = 4.45, $p < .05$) when the gaze position was leading the target (51.2%) while incongruent words accelerated the eyes (5.23 vs. 5.04, F (1, 19) = 4.67, $p < .05$). However, the reversed pattern was observed in the lagging cases (48.8%): The congruent words increased eye velocity (5.49 vs. 5.18, F (1, 18) = 4.72, $p < .05$) while the incongruent words decreased eye velocity (4.86 vs. 5.25, F (1, 18) = 4.66, $p < .05$).[24]

The same pattern was found at word offset: In the leading cases (51.6%), congruent words decreased eye velocity (5.77 vs. 6.31, F (1, 19) = 4.41, $p < .05$) while incongruent words increased eye velocity (6.54 vs. 6.40, F (1, 19) = 4.89, $p < .05$). On the other hand, in the lagging cases (48.4%), congruent words accelerated the eyes (6.85 vs. 6.72, F (1, 18) = 4.55, $p < .05$) while incongruent words caused a deceleration in eye velocity (6.71 vs. 6.98, F (1, 18) = 4.42, $p < .05$).

Discussion

These results demonstrated again that cognitive processes implicated in the activation of semantic representations could affect pursuit eye movements, even though the pursuit task was confined to a visual environment that required no semantic information for its interpretation. This was not simply a distractor

---

[24] The proportion of leading and lagging samples were calculated over the time period from 200 ms before the word offset to 300 ms after the word offset.

effect, as the concurrent language influenced pursuit eye movements as a function of its semantic content.

In order to perform the pursuit task, relevant representations were activated to guide and regulate the eye response. However, semantic representations were at the same time active during language comprehension upon hearing the words. According to the biased competition theory (Desimone & Duncan, 1995), when multiple representations are simultaneously active, they compete for behavioural expression. The guided activation theory of cognitive control (Cohen & Miller, 2001) adopted this idea and suggested that attention could be considered as the modulatory influence that biases the outcome of the competition between concurrently active representations. Based on the seemingly complex results from the velocity analyses in Experiment 2 and Experiment 3, we propose that language exerts its effect on pursuit via a common attentional mechanism as outlined below.

In order to pursue a target, its velocity and position must be sampled. Therefore the target must be attended to, at least some of the time (Chen et al., 2002). Using a dual-task paradigm in which subjects had to pursue a moving stimulus while trying to detect a target appearing in the periphery, it was found that the sensitivity to the detection target was the highest when it appeared *directly* ahead of or behind the pursuit target (van Donkelaar & Drew, 2002). These results indicate that not only is the target represented and attended to during pursuit, but so is the *immediate space* in which the target motion occurs. Thus in the present study, when the gaze position was leading the target during downward pursuit (i.e. when the gaze position was spatially *below* the target), the pursuit task demanded a representation of the moving target, as well as the space *above* the gaze position, in which the target motion occurred. When verbs implying downward motion were heard, a representation of downward motion, as well as the space *below* the gaze position was activated by the directional semantics carried by the verbs. The representation dictated by the oculomotor task and the representation activated by language comprehension would compete with each other and the semantic representation would act to bias attention "away" from the pursuit target. The deceleration observed in the pursuit response rose as the consequence of this competition, as reduced eye velocity was frequently observed when attention to the pursuit target was compromised (e.g.

Kerzel, Souto, & Ziegler, 2008). On the other hand, processing verbs implying upward motion formed a representation of the part in space *above* the gaze position, which coincided with and boosted the representation activated by the pursuit task, thus causing the eye to accelerate. This finding is in line with the observation that eye velocity was increased if pursuing while performing a task associated with the target, which aimed at increasing attention to the target (Shagass, Roemer, & Amadeo, 1976).

The results obtained from when the eyes were behind the target dot in the "lagging" cases can be interpreted in the same way: When the gaze position was lagging behind the target during downward pursuit, the area around the target dot *below* the gaze position was critical to the oculomotor task. In this case, downward verbs caused the eyes to accelerate rather than to decelerate, as the semantic representation activated by the verbs "highlighted" the part in space that was in accordance with the pursuit task. Therefore there was no competition with the representation of the pursuit target, which was instead given a boost in activation strength. This "boosted" target representation in turn caused the eyes to accelerate. By the same logic, deceleration occurred for upward verbs, because the representation activated by the upward semantics would create competition with the pursuit task thus causing the eyes to slow down.

Given that pursuit is essentially a sensorimotor response that involves both perceptual and motoric processing; there is the question of whether the semantic effects seen here reflect the interaction between language comprehension and the perceptual component of pursuit, or the mechanism responsible for generating the motoric response. Although this question has not been directly tested in the current study, we believe that verb semantics most likely influence pursuit eye movements through the perceptual mechanism involved. If verb semantics had a direct impact on the motoric response, a much simpler data pattern might be expected: Verbs implying motion in a direction that is congruent with pursuit will always cause the eyes to accelerate while incongruent verbs will always lead to eye deceleration, regardless of whether the gaze position is leading or lagging behind the target. However, this pattern was not found in current data; the interaction with gaze position (leading vs. lagging) suggested that the influence of language comprehension on pursuit interacted with attentional factors and was not simply motoric.

The velocity analyses for Experiment 3 (and to a certain extent, Experiment 2 also) have revealed a set of complex yet systematic interactions between verb semantics and pursuit eye movements: These interactions are complex because they cannot be considered in simple facilitation/inhibition terms. Meanwhile they are also dynamic and constantly changing on a moment-by-moment basis depending on the gaze position relative to the target. However, these interactions are also systematic because they can be accounted for and predicted using simple competition/bias principles. Furthermore, and perhaps most remarkably, the relationship between such distinct systems for language comprehension and oculomotor control can be accounted for without introducing a third mechanism that is qualitatively different from the ones responsible for language processing or eye movement control.

General discussion

At first glance, these three studies reported here support the grounded view of language comprehension (e.g. Barsalou, 1999; 2007; Glenberg, 1997). Our results confirm the predictions from the grounded view in two ways: First, language comprehension is not isolated from perceptual and motoric processes and is not based on completely different principles. Instead, it is tightly linked to the mechanisms responsible for the control of sensorimotor responses and these sensorimotor mechanisms are, at the same time, sensitive to the modulatory effects of language processing. Second, upon hearing a word, representations are automatically activated and these representations directly reflect features and properties of the denoted object or event.

Although it seems that our results agree with the grounded cognition literature, our studies differ fundamentally from its precedents. Unlike the previous visual-world studies (e.g. Spivey & Geng, 2001), we have demonstrated language effects on a sensorimotor task that is confined in a visual environment devoid of semantics and contextual information, and when the task itself can be achieved without being interpreted semantically. Therefore our results reflect the direct impact of verb semantics on the sensorimotor systems without being confounded by any possible semantic mappings between the linguistic and visual stimuli. Furthermore, the eye response pattern observed in the present study cannot be accounted for by any underlying decision processes (cf. Meteyard et al., 2007), as the mechanisms responsible for accelerations/decelerations in pursuit eye movements are most likely insensitive to deliberate control. Thus these results reported here are the first demonstrations of language semantics directly influencing the sensorimotor mechanism and biasing the behavioural outcome.

Our studies also contribute to the grounded cognition framework by explicitly illustrating the possible mechanism behind the interaction between language comprehension and perceptual or motoric processing. There are several advantages of applying the biased competition theory (Desimone & Duncan, 1995) and the guided activation theory (Cohen & Miller, 2001) to linking cognitive functions to perception and action. Each of these advantages is discussed below:

Both the biased competition theory and the guided activation theory share the view that the source of bias (i.e. attention) can be either bottom-up (e.g. driven by a stimulus) or top-down (e.g. regulated by task demands). However, neither theory emphasizes the difference between "high-level" and "low-level" processes, since the general functional principles proposed by both frameworks are at a system level, and there is, therefore, no need to make distinctions between high-level and low-level processing. This view is in line with the current direction of cognitive research using eye movements as a measure (Hutton, 2008; Kowler, 1990). As reviewed by Kowler (1990), it is becoming increasingly difficult to distinguish between "high-level cognitive functions" and "low-level perceptual or motoric processes". Many concepts that used to be representative of high-level cognitive functions, such as attention and working memory, have been found to share substrates with perceptual or motoric processing. On the other hand, some supposedly strictly sensorimotor responses, such as pursuit eye movements, have been found to be under the control of high-level cognitive factors, for example, as demonstrated here, language comprehension. It is evident that the traditional boundary between high-level vs. low-level processing has been substantially blurred by recent research findings[25]. Thus it is beneficial and convenient to consider human information processing as

---

[25] Although there seems to be no need to distinguish between "high-level" and "low-level" processes, such terms are nonetheless used in the subsequent text for consistency with the existing literature.

a system that is based on the same general operational principles throughout. The view that attention is the modulatory force that rises from the general functional principles of the system provides the grounded cognition framework with the missing link that readily connects “high-level cognitive functions” to “low-level perceptual or motoric processes” and promotes the perspective that cognition is grounded in perception and action.

Both the biased competition theory and the guided activation theory emphasize the “bias” in the system that eventually leads to different behavioural outcomes. The term “bias” indicates that the modulatory effect of attention is *graded* in nature rather than “winner-take-all”. It has been explicitly pointed out that there is always some information flowing along a particular processing pathway even in the absence of any modulatory representation (e.g. task demands). Although some limited degree of activation may not be strong enough to generate an overt response, it may be enough to influence processing in other pathways. More specifically, in the Stroop task, colour naming is still delayed by the processing in the word pathway even when there is no explicit task associated with the word. Apart from the Stroop effect, this graded bias assumption is also in line with other established effects. For example, it has been reported that distractors presented in the periphery can affect saccadic curvature even when the distractor does not play any role in the actual task (Van der Stigchel & Theeuwes, 2008; McSorley, Haggard, & Walker, 2006). The graded bias assumption has provided intriguing insights for one of the unresolved major issues related to the core concept of simulation: If simulation involves the reactivation of the exact neural substrates activated during the actual experiences, and if language comprehension is accomplished through simulation, there must be some additional processes involved that inhibit any actual behavioural response during language comprehension. In simpler terms, if simulating an event is the same as actually experiencing the event, when the word “kick” is heard, a kicking action must be inhibited, at least for some of the time. It seems laborious and unnecessary to introduce an additional mechanism of inhibition into the framework given that the graded bias principle, which functions at a system level, can provide the most parsimonious account: Language comprehension indeed relies on and activates the same mechanisms responsible for perception and action. However, the activation of the appropriate pathways by language may not be strong enough to elicit an overt behavioural response. Nonetheless, the activations caused by language processing may still be enough to influence information processing in other pathways and therefore *biases* any behavioural responses to be generated. This graded bias account not only can address questions such as why there is not always a kicking action when the word "kick" is heard, but also issues such as why there are always looks towards the distractor objects/regions in visual-world studies (e.g. Altmann & Kamide, 2009).

Finally, by incorporating the concept of attention into the grounded cognition framework, a broad range of established phenomena in the grounded cognition literature can be accounted for under the same theoretical framework. One of the widely applied frameworks in the literature is the *theory of event coding* (Hommel et al, 2001).[26] However, the *theory of event coding* (TEC) is limited to the extent that it focuses exclusively on the relationship between the late stage or the "end result" of perception and the planning stage of actions. Thus perception at a sensory level and action at an execution level, which are both essential components in the grounded cognition framework, are left out completely by TEC. In contrast, as pointed out by Cohen and colleagues (2004), it is unclear that any mental processing can occur entirely independent of attention, and thus attention encompasses the whole range of mental processing. This is advantageous for the grounded cognition framework, as attention pervades the whole processing stream of the same three components (i.e. perception, action and cognition) while its specific functional principles can motivate a range of clear and testable predictions.

The current study not only has significant implications for the grounded framework for language comprehension, but also for psychophysics research on pursuit eye movements. From a methodological point of view, experiments reported here provide the first instances in which potential influence of gaze position relative to the target during pursuit has been considered in eye velocity analyses. To our knowledge, relative gaze positions have been largely overlooked in the pursuit literature. Our method of splitting eye velocity data

---

[26] See Chapter 1 for details of this theory.

based on whether the gaze position is ahead or behind the target has brought a new perspective on how pursuit eye movement data can be analysed.

From a theoretical point of view, while there is abundant evidence demonstrating language effect on saccadic eye movements (e.g. Spivey-Knowlton et al., 1995; Tanenhaus et al., 1995), the present study is the first demonstration of language-mediated semantic effects on pursuit eye movements. Past research on the cognitive influence of pursuit eye movements typically focused on the distractor effect during pursuit initiation and maintenance, or the effect of learning with repetitively presented stimuli, which had since demonstrated the involvement of factors such as attention, memory and anticipation (cf. Barnes, 2008). Models of the control mechanism underlying pursuit eye movements have been built with specific components to account for the cognitive involvement (e.g. Barnes & Asselman, 1991; Robinson, Gordon, & Gordon, 1986). For example, in the model developed and revised by Barnes and colleagues (Barnes & Collins, 2008; Bennett & Barnes, 2003), the internal loop is designated exclusively for explaining how cognitive functions affect pursuit eye movements, with the node β being under the influence of cognitive factors such as attention (Figure 2.5). Nonetheless, this model cannot generate outcomes comparable to the semantic effects revealed by the present research, given that it is impossible to produce such complex interaction patterns by simply altering the value of one parameter. In addition, instead of a single node/parameter, it is perhaps more plausible to model the involvement of attention in pursuit eye movements at a system level, as attention can modulate the pursuit system in more than one way and is constantly engaged in the control of this type of eye movements (Chen et al., 2002). The present research not only demonstrates a semantic influence on pursuit eye movements, but also raises new questions for the psychophysics research on the pursuit system.

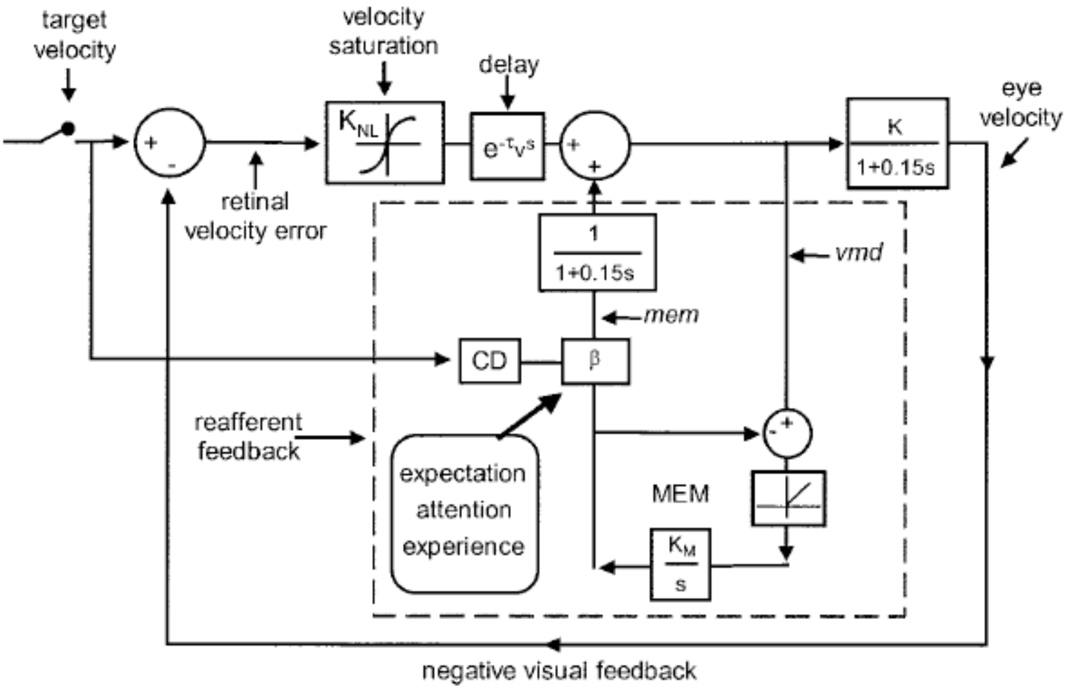


Fig. 2.7. *A model of pursuit eye movements reproduced from Bennett & Barnes (2003). The outer loop, which is illustrated by the solid line, encapsulates how motion perceived by the retina is "translated" into eye velocity. The internal loop, which is depicted by the broken line, represents the component that accounts for cognitive effects on pursuit eye movements. The node ß is thought to be under the influence of expectation, attention and experience. ß gates the output of the MEM loop, where an internal representation of target motion is stored. This stored representation of target motion is crucial for driving and regulating pursuit during the maintenance stage. Thus, according to this model, "high-level" cognitive functions can only influence pursuit eye movements through altering the value of ß to control the release of the stored target motion information.*

One of the issues left unresolved is whether language affects pursuit eye movements at a sensory or mental level. Yasui and Young proposed (1975) that stable pursuit relied on an internally constructed signal representing target velocity, instead of perceived motion alone. Subsequent research has shown that this internal representation of target motion can be stored and then released with appropriate timing to drive pursuit, so that the visual signal does not have to be sampled all the time (Becker & Fuchs, 1985; Barnes & Asselman, 1991). Based on current results, it is difficult to determine whether the semantic effect is due to an interaction between language comprehension and the direct visual perception of target motion, or the stored mental representation abstracted from perceived motion. Since there is evidence indicating that location information stored in memory can drive saccadic eye movements along with concurrent spoken language (Altmann, 2004), it is not implausible to hypothesise that language can affect pursuit eye movements through interacting with the stored mental presentation of target motion. This proposal has been tested in two studies described in the following chapter, in which the perceptual sampling of target motion was either reduced or extinguished temporarily during pursuit so that the system relied more or entirely on the internally represented motion. Potential semantic effects from the same group of directional verbs were examined under these conditions.

Conclusion

These results so far have demonstrated that the semantics of single verbs can affect pursuit eye movements. Featural representations are automatically activated during the comprehension of these verbs. Depending on whether there is competition between the featural representations activated by language and task-relevant representations activated by pursuit demands, eye velocity is systematically modulated via a shared attentional mechanism, which has risen inasmuch language comprehension and sensorimotor responses are likely to operate based on common functional principles.

In sum, the three experiments reported in this chapter have demonstrated acceleration/deceleration in eye velocity during smooth pursuit as a function of the competition between language processing and eye movement control under the regulation of a shared attentional mechanism. These results have provided the missing link between simulation and motoric output for the grounded view on language comprehension and posed challenges for the research on eye movements.

Chapter 3. The effect of verb semantics on pursuit eye movements during extinction and linear target motion

Empirical findings reported in Chapter 2 demonstrated the influence of directional verb semantics on pursuit eye movements as a function of the relative gaze positions to the target (i.e. the leading vs. lagging cases). The present chapter describes two studies in which pursuit relies, predominantly or entirely, on the stored mental representation of, rather than the direct visual perception of target motion. Thus the interactions between language comprehension and pursuit eye movements reported in the previous chapter can be explored when the sampling of information related to gaze position is diminished or removed completely. In Experiment 4 and 5, the original semantic effects were examined during *linear target motion* and *pursuit extinction*. During linear pursuit, target velocity stays constant and there is less demand to perceptually sample target motion frequently. Extinction refers to the period during which the pursuit target is temporarily extinguished and there is no visual motion available. Therefore the potential relevance of gaze position was attenuated and eliminated, respectively, in Experiment 4 and 5. The same directional verbs used in Experiment 1-3 were presented auditorily during linear target motion or extinction. Potential semantic effects on pursuit response were measured a) when target motion was not sampled all the time and b) in the absence of target motion. The results revealed that interactions between verb semantics and pursuit eye movements under these two conditions could be accounted for under the same theoretical framework proposed in the previous chapter.

Experiment 4

Introduction

Much research has been carried out on the *open-loop* stage of pursuit eye movements, which refers to the period before the system is influenced by any visual feedback (Barnes, 2008). Comparatively, only limited attention has been paid to pursuit *maintenance* (or steady-state pursuit). This is due to the fact that, for the interest of psychophysical research, any potential modulatory influence on pursuit from perceptual or cognitive factors can be more conveniently observed and studied in the absence of visual feedback. However, this imbalance in pursuit literature has left many unresolved issues regarding how pursuit is sustained and controlled once it has entered the steady state.

One of these issues concerns whether the frequency of the sampling rate is constant across the initiation and maintenance of pursuit. Evidence has indicated that target motion information can be abstracted very rapidly (e.g. within 80-100 ms after target motion onset) during open-loop pursuit (Carl & Gellman, 1987). This is also the case with eye velocity information, since a recent study has suggested that eye velocity may be sampled within the first 100 – 150 ms of target motion (Barnes & Collins, 2008). This information related to target and eye velocity sampled during the open-loop stage is then stored, and applied along with visual feedback to drive and regulate pursuit response during the maintenance stage (Barnes & Asselman, 1991; Young & Stark, 1963), which is essentially a *closed-loop* response because of the involvement of visual feedback. However, whether target or eye velocity information is re-sampled and compared against the stored target or eye velocity information sampled initially remains unclear. Indirect evidence has demonstrated the possible involvement of perceptual processing by showing that visual attention is involved in both the initiation and maintenance stages of pursuit: During a primary pursuit task, Chen and colleagues (2002) presented subjects with a secondary visual discrimination task in the periphery of the pursuit target motion. It was revealed that both eye acceleration during the initial stage, and sustained eye velocity during the maintenance stage could be impaired by the presence of the secondary task, but with the sustained eye velocity being less susceptible to the influence of the distractor task. These authors inferred from these results that visual attention was required by both pursuit initiation and maintenance; however, less attention was demanded in the maintenance stage compared to the initial stage. Based on these results, it can be speculated that at least some perceptual re-sampling of target motion information is carried out during pursuit maintenance, but perhaps at a lower frequency.

The predictive nature of pursuit responses also points to the same assumption: The stored target velocity information may drive pursuit eye movements, but it may also be used to predict future target motion in order to diminish the delay within the visuomotor processing system when responding to an external stimulus (Barnes & Asselman, 1991). With repeatedly presented predictable stimuli, the pursuit system relies more on the internally stored representation of target motion than visual feedback. This is simply because the internally constructed signal is of a predictive nature and not influenced by the processing delay in the visual system, thus more efficient at reducing the system delay in contrast to visual feedback (Bennett & Barnes, 2003). Thus it would not be surprising if target motion were sampled at a lower frequency during the maintenance stage of pursuit response, especially when the target motion was predictable.

Two types of predictable target motion have been used in the pursuit literature: Linear and periodic. Linear motion refers to when the target velocity stays constant while periodic motion refers to when the target velocity oscillates in a periodic fashion, such as sinusoidally (i.e. with a velocity profile in the shape of a sine wave). Abundant evidence has already demonstrated that pursuit of targets moving with a constant velocity utilizes stored target motion information (e.g. von Noorden & Mackensen, 1962; Lisberger & Fuchs, 1978; Krauzlis & Miles, 1996b). In the absence of visual feedback, the stored target motion information can sustain smooth eye movements for up to 4 s (Becker & Fuchs, 1985). The ease of maintaining linear motion pursuit is hardly unexpected, given that the motion pattern is simple and only one level of velocity information has to be kept. However, for sinusoidal velocity, evidence indicates that when a tracked target temporarily disappears, smooth eye movement may continue for up to only 1.5 s (Whittaker & Eaholtz, 1982). This means that at least some information related to target motion is sampled and stored during sinusoidal pursuit; but given that the velocity profile for sinusoidal motion is much more complex than for linear motion, it may be more difficult for such information to be sampled and stored. Nonetheless, it has been shown that motion imitating sinusoidal waveforms can be internally generated and stored, and later applied to maintain smooth eye movements (Barnes, Barnes & Chakraborti, 2000). The processes behind the sampling and storage of sinusoidal motion information have been compared to sequence learning: Barnes and Schmid (2002) found that up to four motion sequences could be learnt after one or two presentations, and anticipatory eye movements with appropriate velocities and directions could be generated prior to each component within the motion sequence. This claim was supported by the demonstration that smooth eye velocity during extinction did not depend on pre- or post- occlusion target velocity but a dynamic internal representation of target motion that evolved with time (Orban de Xivry, Missal, & Lefèvre, 2008).

If the construction of an internal target motion representation for sinusoids is achieved through processes similar to sequence learning, the target must be perceptually attended to more frequently during sinusoidal pursuit more than linear pursuit. This assumption is also hinted at by the difference between the durations in which smooth eye movements can be sustained in the absence of perceptual feedback during linear and sinusoidal pursuit (4 s vs. 1.5 s). In other words, during linear pursuit maintenance, the system relies much less on perceptual feedback, as the target velocity does not change so that there is no need to re-sample target motion information. On the other hand, sinusoidal pursuit maintenance depends more heavily on visual feedback, since target velocity varies constantly. This is reflected in our results reported in Chapter 2 in which gaze positions relative to the target played a significant role in regulating eye velocity.

Based on this, we predicted that during linear pursuit maintenance, smooth eye movements would mainly rely on the internal representation of target motion instead of the perceptual sampling of the actual target. In order to test this prediction, an eye-tracking experiment was carried out in which subjects performed a pursuit task with linear motion while listening to the same set of directional verbs used in previous studies reported earlier. If linear pursuit maintenance depended predominantly on the internal representation of target motion as predicted, we expected to find representations activated by language to interact solely with target motion. More specifically, congruent verbs should cause eye acceleration while incongruent verbs should produce eye deceleration, regardless of the relative gaze position. However, if linear pursuit maintenance still relied on perceptual feedback to a considerable extent, relative gaze position to the target would play a role in modulating eye velocity thus any potential semantic effects would resemble more closely to the ones revealed by Experiment 2 and 3.

Method

Participants

28 students from the psychology department participated in this experiment. They were offered either course credit or £4 for their effort. All participants were native speakers of English and had normal or corrected-to-normal vision.

Materials and Procedures

The materials and procedures involved in this study were similar to the ones in Experiment 2, except that the pursuit target moved at 12 °/s instead of sinusoidally.

Design

All 24 experimental items (i.e. directional verbs) and their 24 controls were presented during upward or downward motion. All 48 filler words were paired with leftward or rightward target motion.

Results

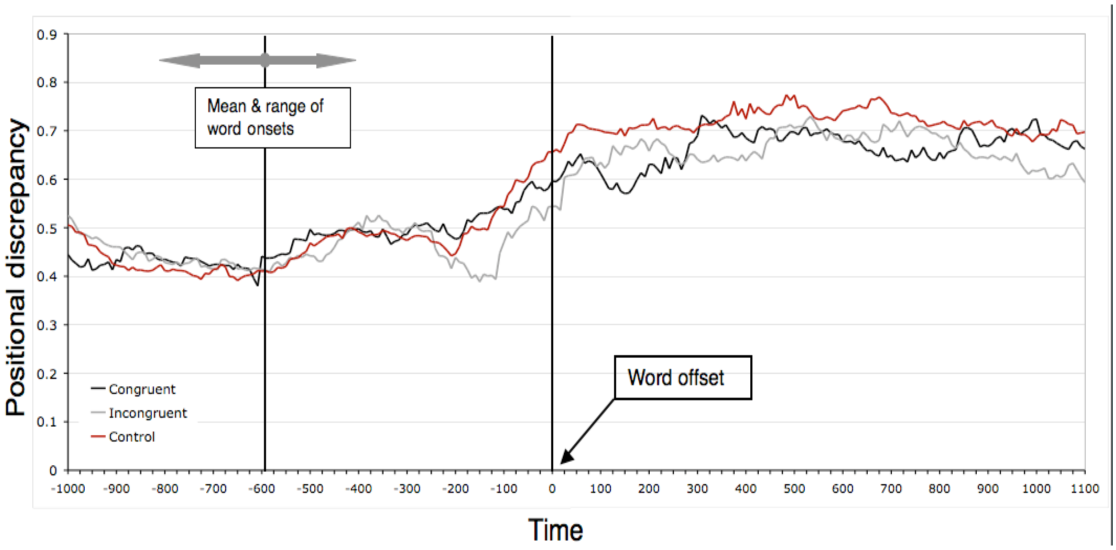


Fig. 3. 1. *Absolute positional errors between the gaze position and the target dot across three conditions for upward pursuit: Congruent, incongruent and control. Positional errors were synchronized to word offset for reasons described elsewhere.*[27] *The solid grey circle indicates the approximate point of word onset based on the mean word duration, with the leftmost arrow representing the onset of the longest word and the rightmost arrow signalling the onset of the shortest word. Time zero signals word offset.*

---

[27] See Experiment 1, Chapter 2 for details.

Eye positional error and velocity were measured as dependent variables. These variables were calculated using methods described in Experiment 1. Once again, all effects reported here came from downward pursuit. Figure 3.1 suggested no clear divergence among the three lines representing the congruent, incongruent and control condition. The standard deviation of positional errors in the filler trials for upward pursuit was once again significantly larger than for downward pursuit (*F* (1, 55) = 8.77, *p* < .01).

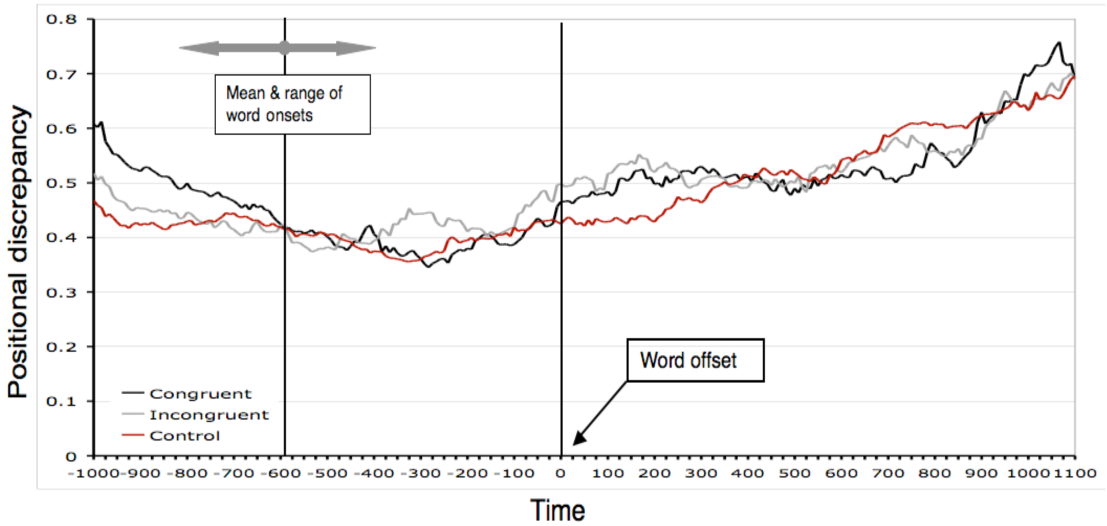


Fig. 3. 2. *Absolute positional errors between the gaze position and the target dot across three conditions for downward pursuit: Congruent, incongruent and control. Positional errors were synchronized to word offset for reasons described elsewhere. The solid grey circle indicates the approximate point of word onset based on the mean word duration, with the leftmost arrow representing the onset of the longest word and the rightmost arrow signalling the onset of the shortest word. Time zero signals word offset.*

Initial inspection of the graph above suggested that the positional discrepancies between the gaze position and the pursuit target were modulated by both congruent and incongruent words. However, as revealed by a one-way within-subject ANOVA, the within-subject factor of Condition did not modulate positional errors significantly (*F* (2, 54) = 2.29, *p* >.05) in a time window from 100 ms before the word offset to 350 ms after the word offset. However, within this period of pursuit, incongruent words significantly increased the positional errors compared to the controls (1.39 vs. 1.12, *F* (1, 27) = 4.89, *p* < .05) while congruent words had no effect (1.31 vs. 1.12, *F* (1, 27) < 1, *p* > .05). At 150 ms after the word offset, there was no main effect of Condition (*F* (2, 54) = 2.39, *p* >

.05). Pair-wise comparisons suggested that at this moment in time, positional errors in the incongruent condition were larger than in the control condition (1.42 vs. 1.13, F (1, 27) = 5.16, $p < .05$), but there was again no difference between the congruent and the control condition (1.34 vs. 1.13, F (1, 27) < 1, $p > .05$). [28]

With eye velocity as the dependent measure, no effect of Condition or difference between conditions was revealed in either the time window (i.e. from 100 ms before word offset to 350 ms after word offset) analyses (F (2, 54) = 1.76, $p > .05$) or the individual time point (i.e. at 150 ms after word offset) analyses (F (2, 54) = 2.01, $p > .05$).

In order to test whether there was any interaction between verb semantics and relative gaze position, the data from this time window (i.e. from 100 ms before word offset to 350 ms after word offset) were separated into leading (53.2%) and lagging (46.8%) cases. No difference in velocity was found across conditions for the leading cases (F (2, 54) = 2.14, $p > .05$) or the lagging cases (F (2, 54) < 1, $p > .05$).

Discussion

Although it seemed that the incongruent verbs had an effect on pursuit eye response to linear motion, this effect was hard to interpret due to the absence of a main effect from the variable Condition. Furthermore, there was no significant semantic modulatory effect on eye velocity during linear pursuit. Taken together, no definitive conclusions could be made based on these data regarding whether language comprehension could affect pursuit eye response to targets moving at a constant velocity.

A possible reason for this null effect is that the distinction between the two predictions proposed earlier may be more ambiguous than suggested: During linear pursuit maintenance, although the system relies more heavily on the internally stored motion information, the target must be perceptually attended to occasionally. Thus gaze position relative to the pursuit target should have at least some transient effects on eye velocity. However, it is conceivable that the representation activated by the target sometimes deviates from the stored representation of target motion, for example, when the gaze position is ahead of

---

[28] Additional analyses revealed no difference across the conditions in the number of (F (2, 54) < 1, $p > .05$) or the mean amplitude of saccades launched (F (2, 54) = 3.04, $p > .05$).

the target dot during downward pursuit. Furthermore, the representations activated by the occasional re-sampling of the target and the internally stored representation of target motion may not be always simultaneously active, but instead, alternatively activated in an intermittent fashion. In other words, under these circumstances, it is difficult to determine in what kind of interactions that verb semantics are involved in thus no clear pattern of results has been revealed in the present study.

Despite the null effect, there were several interesting observations made in this experiment: First, the mean positional error in the current study was smaller compared to Experiment 2 (0.51 vs. 0.73), indicating that subjects were more efficient at pursuing a moving target with a constant velocity and linear pursuit was indeed a less demanding task than sinusoidal pursuit. Second, positional errors decreased as the target approached the mid point of its motion trajectory in the present experiment, however in Experiment 2, the errors increased while the target got closer to the mid point of its trajectory (See Figure 2.2 and 3.2). This might be due to the fact that the oculomotor system was confronted with the switch from accelerating to decelerating around the mid point of target trajectory during sinusoidal pursuit while the halfway point during linear pursuit was at its steadiest. Thus the errors between the eyes and the target at the mid point were the biggest during sinusoidal pursuit but the smallest during linear pursuit. Finally, the time window for the semantic effect observed from Figure 4.1 in this experiment was almost identical to the one in Experiment 2, which suggested that despite having been selected in a somewhat arbitrary way, the time windows chosen for statistical analyses were likely to be the period in which the observed semantic effect truly located.

In short, no definitive conclusion can be made from these data due to the null results. Thus it remains unclear whether language comprehension can affect linear pursuit. Nonetheless, two interesting notions can be inferred from these data: a. Linear pursuit is a comparatively effortless task, especially after it enters the steady state; b. the selections of time windows for the analyses from previous experiments are relatively reliable.

As suggested earlier, the null effect revealed in the present experiment could be the consequence of gaze position having some transient effect on eye velocity. In order to eliminate this possibility, the pursuit target was extinguished temporarily in the following experiment (i.e. Experiment 5). During this target extinction period, pursuit had to rely entirely on the mental representation of target motion and no information regarding relative gaze position could be sampled.

Experiment 5

Introduction

One of the goals of pursuit eye movements is to match eye velocity to retinal target velocity in order to eliminate any retinal velocity error. This is achieved through a two-stage process: The first stage generally refers to the first 100 ms of the smooth response after pursuit is initiated (Carl & Gellman, 1987). During this stage, the eye response is not under the modulation of the error feedback mechanism, as the delay in visual processing dictates that within this time period the movement of the eye does not change the retinal velocity error. This stage is thus also referred to as the *open-loop* phase. Following the initial 100 ms, the retinal velocity error is detected and corrected, and eye velocity eventually settles to an average that is close to target velocity. This stage is termed the *steady-state* pursuit or pursuit *maintenance*.

Yasui and Young (1975) proposed that steady-state pursuit relied on an internally constructed signal representing target velocity instead of perceived motion alone. This proposal was supported by the finding that when the target suddenly disappeared during pursuit maintenance (i.e. extinction), smooth eye movements continued at a reduced velocity instead of terminating abruptly or completely breaking down into a string of saccades (Becker & Fuchs, 1985).[29] Subsequent research has shown that an internal representation of target motion can be stored and then applied with appropriate timing to drive pursuit (e.g. Barnes & Asselman, 1991). Models have been developed to reflect the role of the internal representation of target motion and how it is stored and released (e.g. Krauzlis & Lisberger, 1994; Krauzlis & Miles, 1996b; Robinson et al., 1986).

In the model developed and revised by Barnes and colleagues (Figure 3.3), the pursuit system receives both retinal and extraretinal input (Barnes & Asselman, 1991; Barnes & Collins, 2008; Bennett & Barnes, 2003). Retinal input

---

[29] This is only the case when the target is expected to reappear. Smooth eye velocity decays to zero if subjects do not anticipate target reappearance.

refers to visual motion signals such as velocity and position, which is represented by the outer loop in the model. Extraretinal input meanwhile is illustrated by the inner loop in which the internal presentation of target motion is stored. The activity of the inner loop is gated by the node ß, which is in turn regulated by cognitive factors such as attention and expectation. In other words, the inner loop is the component in this model that can function during extinction to sustain smooth pursuit and account for any potential cognitive influence on pursuit eye movements.

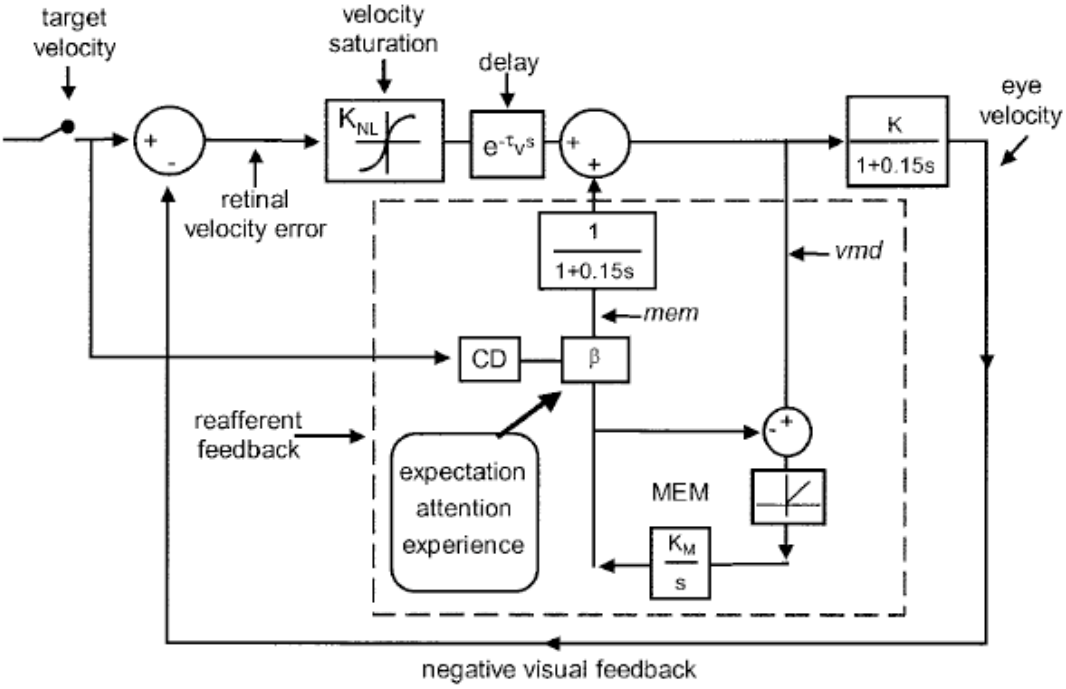


Fig. 3.3. *A model of pursuit eye movements reproduced from Bennett & Barnes (2003). The outer loop, which is illustrated by the solid line, represents the component that is regulated by retinal (or perceptual) input. The inner loop highlighted by the broken line is where the internal representation of target motion (MEM) is stored. The release of the stored target motion information is gated by ß, which is under the influence of cognitive factors.*

As discussed in the previous chapter, the inner loop and the node ß are not sufficient for explaining the complex interactions between verb semantics and pursuit eye movements observed in Experiment 2 and 3. Nonetheless, it is important to establish whether language comprehension can affect pursuit eye movements through the inner loop at the level of mental representations. This is for two reasons: First, it has been demonstrated that verb semantics can impact on direct motion perception, despite the equivocal results (Meteyard et al., 2007).[30] It is worth noting that the visual stimuli used in the Meteyard study were random dot kinematograms, which displayed coherent motion that was close to the threshold of human motion detection. Thus whether verb semantics can affect the perception of unambiguous motion (e.g. motion stimuli used in Experiment 1-3) remains unknown. Nonetheless, it is possible that the semantic effect on pursuit eye movements reported in Chapter 2 arises solely as the consequence of an interaction between language comprehension and the direct perceptual sampling of target motion during pursuit maintenance. Second, some studies in the saccadic literature indicate that language interfaces with the perception of the visual world at a mental level. For example, in a "*blank screen*" study (Altmann, 2004), subjects were shown visual scenes composed of several objects for a few seconds. The scenes were then taken away and subjects listened to sentences while facing a *blank* screen. It was found that when the objects from the scene were referred to by the sentence, saccades were directed to the locations of where these objects had previously been, even though the screen was blank. This finding has been used to support the proposal that language-mediated eye movements are not dependent upon a concurrent visual stimulus, but a mental representation of that visual stimulus. Moreover, the stored mental representation of the visual environment is not static and can be updated by non-visual information. In another "blank screen" study (Altmann & Kamide, 2009), saccade patterns demonstrated that representations constructed based on the visual perception of the external world could be updated by linguistic information, and the updated version was in turn used to guide eye movements.

The aim of the present study (Experiment 5) was to address the question of whether language comprehension could interact with pursuit eye movements at the level of mental representations. The procedures and stimuli were identical to Experiment 2, except that the pursuit target went through an extinction period of 600 ms during the steady state. The onset of the extinction period was synchronized to 200 ms before the word offset so that the semantic effects on pursuit could be investigated in the absence of visual motion.[31] Based on the

---

[30] See Chapter 2 for a discussion of this study.

[31] Since semantic effects induced by directional verbs were observed from 200 ms before the word offset in Experiment 3, this time point was chosen as the onset of extinction.

findings from the previous chapter, it could be anticipated that representations activated by language would compete with representations activated during pursuit extinction. We predicted that directionality implied by motion verbs would interact directly with the mentally represented target motion during extinction, as the distinction between "leading" and "lagging" would no longer be applicable given that the target would be absent.

Method

Participants

Thirty undergraduate students took part in this experiment. They participated in exchange for either course credit or £4. All participants were native speakers of English and had normal or corrected-to-normal vision.

Materials and Procedures

The materials and procedures involved in the present experiment were identical to the ones in Experiment 2, except for the extinction procedure. The onset of the extinction period (i.e. the disappearance of the target dot) was synchronized to 200 ms before the word offset. Thus during the experiment, subjects were able to learn that the target always disappeared during the word, however, they were unable to predict the exact timing of extinction onset. The target dot reappeared 600 ms after extinction onset. Therefore the extinction period encompassed a time window of 600 ms starting at 200 ms before the word offset and terminating at 400 ms after the word offset.

Subjects were instructed to pursue the target to the best of their ability during extinction as if it was still visible. They were encouraged to imagine that the target was moving behind and temporarily occluded by an invisible wall. None of the subjects reported the task to be particularly difficult.

Design

All 24 experimental items (i.e. directional verbs) and their 24 control items were paired with target motion in the vertical dimension (i.e. upward or downward). There were also 48 filler items that were presented during leftward or rightward horizontal pursuit.

Results

Eye positional error and velocity were measured as dependent variables. The method of calculation for these variables reported in Experiment 1 was applied to the data in the current study. It is worth noting that the positional errors between the gaze position and the target during extinction were calculated by subtracting the y coordinate of where the dot would have been from the y coordinate of the gaze position.

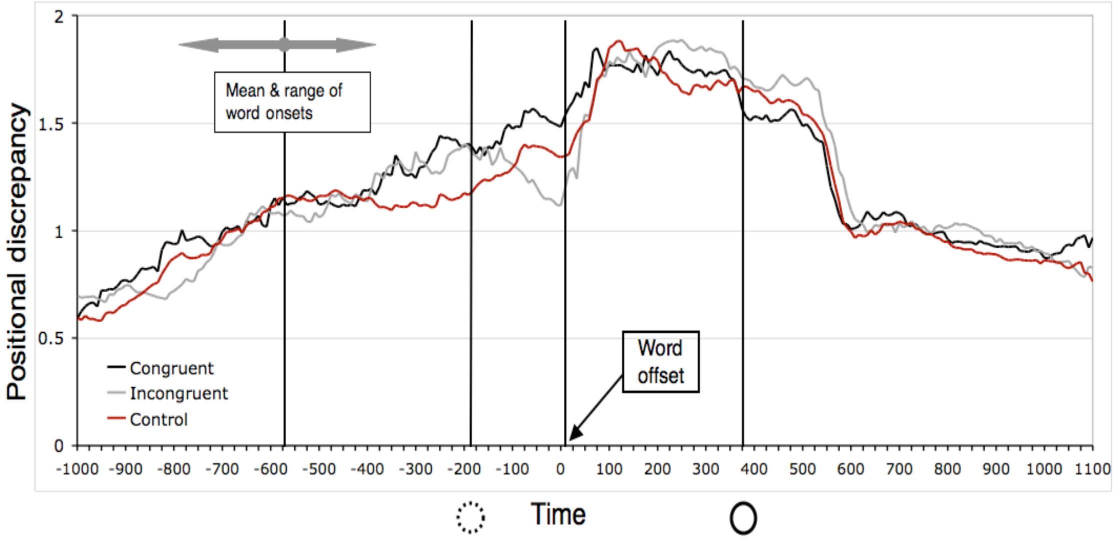


Fig. 3. 4. *Absolute positional errors between the gaze position and the target dot across three conditions for upward pursuit: Congruent, incongruent and control. Positional errors were synchronized to the word offset for the same reason mentioned previously.*[32] *The solid grey circle indicates the approximate point of word onset based on the mean word duration, with the leftmost arrow representing the onset of the longest word and the rightmost arrow signalling the onset of the shortest word. Time zero signals word offset. The circle depicted by the broken line indicates the onset of extinction (i.e. 200 ms before word offset) while the one illustrated by the solid line signals where the target reappears (i.e. 400 ms after word offset).*

All effects reported here were only observed during downward pursuit. There was no statistically significant difference across the conditions of upward pursuit. The graph hinted at a difference between the congruent and the control condition in a time window from 400 ms before the word offset to around 150 ms after the word offset. However, the divergence between these two conditions happened before the onset of the extinction period. Furthermore, this difference was not statistically reliable (1.47 vs. 1.33, t (29) = 2.03, $p > .05$). Similar to experiments reported previously, the standard deviation of positional errors in the filler trials for upward pursuit was considerably larger than for downward pursuit (F (1, 59) = 8.64, $p < .01$). This was not surprising given that the same pattern

---

[32] See Chapter 2 for the details.

was also found in Experiment 1 and 2. The lack of an effect in upward pursuit could be attributed to the large amount of noise in the upward data.

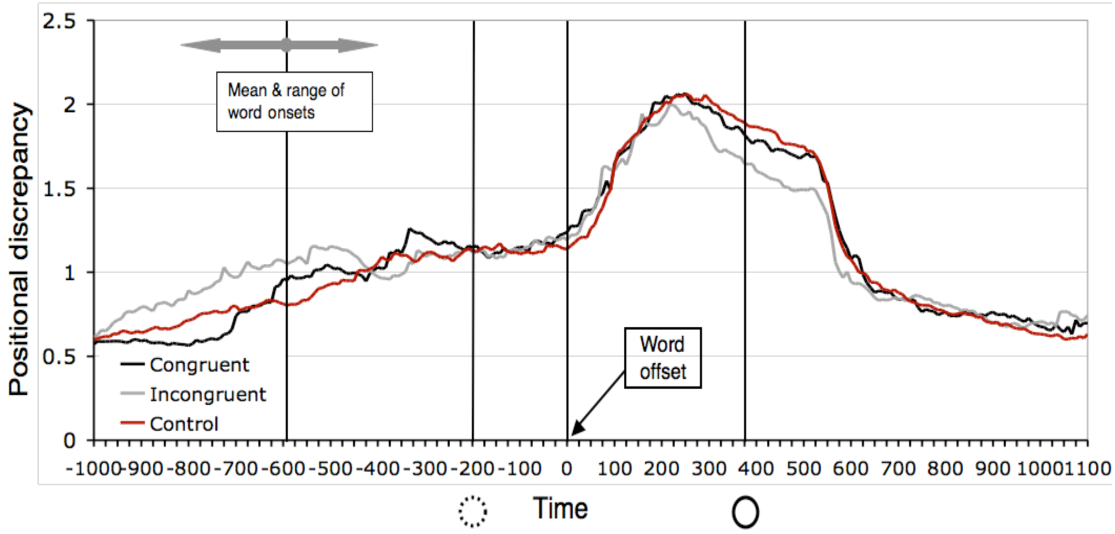


Fig. 3. 5. *Absolute positional errors between the gaze position and the target dot across three conditions for downward pursuit: Congruent, incongruent and control. Positional errors were synchronized to word offset for the same reason mentioned previously. The solid grey circle indicates the approximate point of word onset based on the mean word duration, with the leftmost arrow representing the onset of the longest word and the rightmost arrow signalling the onset of the shortest word. Time zero signals word offset. The circle depicted by the broken line indicates the onset of extinction (i.e. 200 ms before word offset) while the one illustrated by the solid line signals where the target reappears (i.e. 400 ms after word offset).*

The general eye movement pattern during extinction was observed in the present study: Due to delays in the visual processing stream (Becker & Fuchs, 1985), eye velocity only began to decelerate 190 ms after the target had disappeared, and then pursuit broke down into strings of saccades interspersed with smooth eye movements. These saccades launched during extinction tended to overshoot, as indicated by the predominant amount of leading samples during the extinction period (87.3%). These overshooting saccades led to the rapid increase in positional errors starting from 200 ms after the target was extinguished (Figure 3.5). It was not until around 100 ms after the target reappeared that smooth eye movements were restored and positional errors were reduced.[33]

In the time window from 400 ms after the extinction onset (i.e. 200 ms after word offset) to 150 ms after the target reappeared (i.e. 550 ms after word offset), incongruent words temporarily decreased the positional discrepancies compared to the congruent and the control condition. There was a main effect of condition (F (2, 58) = 4.12, $p < .05$). Pair-wise comparisons revealed that the positional errors in the incongruent condition were reliably smaller compared to the control condition (1.68 vs. 1.88, F (1, 29) = 4.94, $p < .05$).[34] However, no difference was found between the congruent and the control condition (1.82 vs. 1.88, F (1, 29) = 2.26, $p > .05$). Individual time point analyses revealed the same result pattern: At extinction offset (i.e. the point when the dot reappeared), there was a main effect of condition (F (2, 58) = 4.05, $p < .05$). Pair-wise comparisons indicated that the positional error at this time point was significantly smaller in the incongruent condition than in the control condition (1.51 vs. 1.77, F (1, 29) = 4.83, $p < .05$) while the error in the congruent condition was not different from the control condition (1.71 vs. 1.77, F (1, 29) = 2.06, $p > .05$). Finally, the number of (F (2, 58) < 1, $p > .05$) or the amplitude of (F (2, 58) < 1, $p > .05$) saccades launched did not differ across the three conditions.

As the dot was invisible during extinction, eye velocity data were not separated based on the relative gaze position to the pursuit target (cf. Chapter 2) and analyses were conducted directly across the three conditions. In the same time window used for positional error analyses (i.e. from 200 ms after extinction onset to 150 ms after target reappearance), the within-subject variable of condition had an effect on eye velocity (F (2, 58) = 11.71, $p = .001$). During this period, eye velocity was decreased significantly by the incongruent words in contrast to the control words (1.92 vs. 2.08, F (1, 29) = 16.67, $p = .001$). However, the congruent verbs did not change eye velocity relative to the controls (2.10 vs. 2.08, F (1, 29) = 3.22, $p > .05$). At extinction offset (i.e. when the target became visible again), eye velocity was significantly modulated by Condition (F

---

[33] Figure 3.4 illustrates roughly the same eye response pattern. However, due to the significant amount of variance in upward data, the general eye movement pattern during extinction demonstrated by Figure 3.4 is not as clear as Figure 3.5.

[34] The unit of the mean positional discrepancies reported here is degree (º) while the unit for the mean eye velocity is degree/s (º/s).

(2, 58) = 4.10, $p < .05$). At this time point, eye velocity in the incongruent condition was smaller compared to the control condition (1.83 vs. 2.46, F (1, 29) = 7.54, $p < .01$) while there was no difference between the congruent and the control condition (2.14 vs. 2.46, F (1, 29) = 2.02, $p > .05$).

Discussion

During downward pursuit, when there was no visual motion available and subjects were performing the pursuit task in a blank screen, incongruent directional verbs (i.e. upward verbs) decreased eye velocity in comparison to the downward verbs and the non-directional words. As the gaze positions during extinction were largely ahead of the target, eye velocity reduction caused by this group of upward verbs also led to a decrease in positional errors in the same time window.

The guided activation theory of cognitive control (Miller & Cohen, 2001) suggested that only task-relevant representations would be active in the course of any type of processing. During extinction, the oculomotor task would be to maintain smooth eye movements in a particular direction. Thus the only representation dictated by the pursuit task when the target was invisible was the representation of target motion in the right direction, which in the case of the present experiment was either downward or upward. Based on the biased competition theory (Desimone & Duncan, 1995), when a directional verb was processed, the semantic representation activated would interact with the representation of target motion, and eye velocity would be modulated depending on whether these two representations were in competition or accordance with each other. Therefore, for example, eye velocity during extinction was decreased by upward verbs when pursuit was downward, as there was competition between the representation activated by the verbs and the representation demanded by the oculomotor task. The principles of the biased competition theory (Desimone & Duncan, 1995) and the guided activation theory (Miller & Cohen, 2001) could be once again applied to account for the interactions between language comprehension and pursuit eye movements: When there were more than one representation simultaneously active within the system, they competed for expression and each biased the behavioural outcome associated with another under the modulatory force of attention.

Although the results of the present experiment could be explained using the same account as Experiment 3 described in Chapter 1, there were some differences between these studies: In the positional discrepancy analyses, congruent words decreased the errors between the gaze and the target in Experiment 3; however, it was the incongruent words that reduced the positional errors in the current experiment. This disparity in the language effects on pursuit rose as the consequence of the different tasks involved in these two experiments. When the target remained visible all the time, such as in Experiment 3, the task was to pursue the target dot. This task entailed matching eye velocity and gaze position to retinal target velocity and position as closely as possible. Consequently, representations of the target and the immediate space ahead of the target were required to be activated by the oculomotor task. Taking the relative gaze position (leading or lagging) into consideration, the representations activated by the congruent (i.e. downward) verbs competed with the representations demanded by the task. This competition caused eye deceleration in the leading cases, which in turn decreased the positional errors. However, representations activated by the downward verbs were in accordance with the ones required by the pursuit task in the lagging cases. Therefore the downward verbs increased eye velocity in the lagging cases and resulted in smaller positional errors. When the target became invisible in an extinction period, such as in the present experiment, the task had changed from having to track the target to simply maintaining smooth eye movements in a specific direction at a reasonable velocity. Accordingly, a representation of motion in the correct direction became crucial to this particular oculomotor task and the space ahead of the target was no longer of relevance.[35] Directionality carried by incongruent (i.e. upward) verbs was in conflict with represented downward motion ordered by the oculomotor task. This conflict gave rise to eye deceleration and decreased the positional errors.

Another difference in the results between Experiment 3 and the present experiment was that the time window in which the semantic effect was observed shifted from during the word (i.e. from 200 ms *before* the word offset to 300 ms *after* the word offset) in Experiment 3 to post-word (i.e. from 200 ms *after* the

---

[35] Such a variable in fact would not be present during extinction.

word offset to 550 ms *after* the word offset) in the present study. In other words, the semantic effect arrived much later in this experiment compared to the previous study. This time window shift could also be attributed to the task distinction between these two studies. Given pursuit in the absence of visual motion was more effortful and requires more attentional and volitional control (Mitrani & Dimitrov, 1978; Pola & Wyatt, 1997), the task representation during extinction was more likely to receive a boost in activation strength and biasd the oculomotor outcome to a more substantial extent compared to when the target remained constantly visible. Due to this more robust task representation, pursuit eye movements during extinction could be less susceptible to the influence from any external factors that could potentially bias the oculomotor response. As a result, the semantic effect motivated by language was delayed until target reappearance was anticipated (i.e. 200 ms after the word offset/200 ms before target reappearance), which was presumably accompanied by a decline in the activation strength of the task representation relevant during extinction.[36]

Finally, unlike in Experiment 3, no eye acceleration caused by congruent (i.e. downward) verbs was found in the present study. This was expected considering the nature of pursuit eye movements as a type of sensorimotor response. The generation of smooth eye movements in the absence of visual motion was traditionally thought to be impossible. Although volitionally generated pursuit eye movements had been observed (Kowler & Steinman, 1979a; 1979b), the eye velocities were generally very low (e.g. < 2º/s). In the case of extinction, pursuit eye movements could only be sustained at a reasonable velocity when target reappearance could be predicted and anticipated (Beck & Fuchs, 1985; Barns & Asselman, 1991). Although these studies had demonstrated relatively higher eye velocity in the absence of visual feedback, the timing of target reappearance was explicitly marked by external cues (e.g. auditory tones) thus could be accurately predicted. The failure to find eye accelerations caused by congruent verbs in the present study could be due to the fact that it was very difficult for any external factor to cause the pursuit system to produce higher velocities during extinction.

---

[36] Some evidence (e.g. Bennett & Barnes, 2003) has revealed an increase in eye velocity 200 – 300 ms before target reappeared during extinction, indicating that target reappearance was anticipated.

Taken together, the semantic effects demonstrated in this experiment were in line with our previous findings and could be accounted for under the same functional principles and assumptions proposed by the biased competition theory (Desimone & Duncan, 1995) and the guided activation theory of cognitive control (Miller & Cohen, 2001). The distinctions between the result patterns generated from having the target constantly visible and temporarily invisible further confirms the predictions regarding the role of task (Cohen et al, 1990): a. Different task demands activate different sets of representations, which in turn bias behavioural responses in one way or another; b. As well as the nature of the task, the activation strength of the task can also have distinct impacts on behaviour via the attentional mechanism that functions at a system level.

Relating to language research, the same account can be used to provide an alternative explanation for the results obtained through the "blank screen" paradigm (e.g. Altmann, 2004). For example, in the "move the glass" study (Altmann & Kamide, 2009), more saccades were launched in the blank screen to where the tabletop had been when the auditory sentence was "*The woman will move the glass onto the table. Then she will pick up the bottle and pour the wine carefully into the glass*" compared to "*The woman is too lazy to move the glass onto the table. Instead, she will pick the bottle and pour the wine carefully into the glass*."[37] Based on the assumptions of the biased competition theory (Desimone & Duncan, 1995) and the guided activation theory (Miller & Cohen, 2001), the initial task of visually scanning the picture should construct and activate a representation of the scene in which the glass is on the floor. However, the later task of sentence comprehension should activate a second representation of the same scene but with the glass being on the tabletop in the "moved" case. These two representations will most likely compete with each other to bias the behavioural outcome, that is, saccadic landing position. Since the task of sentence comprehension is more temporarily relevant and the task of visually scanning the picture will have probably already expired, the accordingly stronger representation of the glass being on the table biases saccades in the blank screen towards the location where the tabletop was. It is worth pointing out that although these findings from the blank screen study are comparable to our

---

[37] See Chapter 1 for a more detailed description of this study.

findings, the blank screen study has essentially demonstrated the mapping of spoken language, or linguistic meaning, onto the mental representation of an external world, which can only be interpreted through perceived visual semantics. On the other hand, the motion representation activated during pursuit extinction contains no semantic information but only low-level motion parameters. As a result, instead of an interaction between semantics perceived through different channels, our results reflect the interface between language comprehension and *stored* low-level motion information.

The represented motion information evidently plays an important role in driving and regulating pursuit eye movements when visual feedback is not available. The issue of what *kind* of motion information that is represented should be clarified. According to Yasui and Young (1975), pursuit in the absence of visual feedback is sustained by a continuous contribution from an efference copy of the eye movement itself and under normal circumstances (i.e. when the target is visible all the time), pursuit eye movements are produced by a summation of visual feedback and this efference copy. This proposal is supported by the findings that smooth eye movements can be maintained at a lower velocity when visual feedback is cut off unexpectedly (Becker & Fuchs, 1985). However, models built based on this principle (e.g. Robinson et al., 1986) cannot account for the observations that there is often a recovery of eye velocity before expected target reappearance (Bennett & Barnes, 2003) and eye velocity can increase to be higher than the level achieved prior to reappearance if the target velocity is expected to increase at the end of extinction (Bennett & Barnes, 2004). An alternative proposal suggests that instead of being driven by a moment-to-moment efference copy derived from the eye movement itself, pursuit in the absence of visual feedback is sustained by stored information related to target velocity, which is derived from a pre-motor drive signal and can be sampled independently from actually making an eye movement (Barnes, Grealy, & Collins, 1997). Models built with this added internal storage component (e.g. Bennett & Barnes, 2003) can successfully account for predictive behaviour exhibited during pursuit eye movements. This means that the semantic effects observed in the present study during target extinction did not arise as the consequence of language comprehension impacting *directly* on the mechanism responsible for generating motoric responses, which is in accordance with our previous findings that the direction of the actual eye movement does not play a role in the regulation of eye velocity, but instead, eye velocity varies as a function of the relative gaze position to the pursuit target.

In conclusion, this study has demonstrated that language comprehension can influence pursuit eye movements in the absence of any visual motion. We argue that this is achieved through competition between representations activated by both the linguistic and oculomotor task, which operates under the regulation of an attentional mechanism that functions at a system level. Not only do these findings provide further support for the proposal that language comprehension and sensorimotor responses share the same representational substrate and operate based on identical principles, they also further elucidate the role of task in these seemingly complex cross-modal interactions.

General discussion

Both studies reported in the current chapter illustrated the critical role played by *task*. Data collected in the linear pursuit experiment (i.e. Experiment 4) hinted at the possibility that task difficulty could determine, at least sometimes, the extent to which perceptual feedback and the internally represented target motion would be involved in the control of pursuit eye movements. When the oculomotor task was less effortful, such as during linear pursuit, the eye movement system relied predominantly on the internally constructed motion rather than on regularly sampling the actual target motion. On the other hand, if the target motion was relatively complex (e.g. sinusoidal), the oculomotor task became more demanding and the target motion was sampled more frequently. The extinction experiment (i.e. Experiment 5) highlighted the role of task through the demonstration of different patterns of interactions between language and pursuit under different oculomotor tasks: During extinction, the task switched from visually pursuing the target to maintaining a reasonable eye velocity in a particular direction due to the absence of the target. As a result, verb semantics interacted only with the representation of target motion, which was the sole task-relevant representation during extinction. Taken together, these two experiments have pointed to the conclusion that the oculomotor task determines which, and to what extent representations should be activated during pursuit. Furthermore, the competitions between representations activated by language comprehension and the oculomotor task can be regulated by task difficulty/relevance.

Given that we have established the effect of verb semantics on vertical pursuit eye movements, we are now faced with two issues: The first is whether these semantic effects can be generalized, since the visual and linguistic stimuli employed so far have remained the same (i.e. vertical motion and directional verbs). The second is that the leading vs. lagging case separation is not a controlled experimental manipulation. Instead, these leading and lagging samples emerged from the natural eye response. These two issues are addressed in the next two empirical chapters. Chapter 4 describes two experiments in which the first one involved replacing the visual stimulus with horizontal target motion, and the second one explored the effect of nouns with a locational component (e.g. attic, basement) on pursuit eye movements. Chapter 5 reports one experiment in which the leading and lagging cases were artificially created by having the subjects attending to a distractor that was placed either above or below the target during pursuit tracking.

Chapter 4. The generalisation of semantic effects on pursuit eye movements

Both experiments included in this chapter (Experiment 6 and 7) address the question of whether the semantic effects on pursuit eye movements reported in Chapter 2 and 3 can be generalised. Research described so far has demonstrated the consequent influence on smooth eye velocity originated from competition between concurrently activated representations. However, due to previous experimental manipulations, these competing language-related and pursuit-related representations always directly corroborated or contradicted with each other (i.e. representations activated by directional verbs that implied motion in either the same or the opposite directions as representations involved in pursuit eye movements). The following two experiments explored the possibility of similar semantic effects on pursuit when the concurrently activated representations were not directly in agreement or conflict with each other. Experiment 6 used the same linguistic stimuli and procedures as Experiment 2; however, the vertical target motion used to induce pursuit eye movements was replaced with horizontal motion (i.e. the target moved leftward or rightward). In this case, representations activated by the pursuit task would involve motion in a different *dimension* compared to representations activated by the vertical motion verbs. On the other hand, Experiment 7 used the same visual stimuli and procedures as Experiment 2; however, the vertical directional motion verbs were replaced with nouns with a spatial component (e.g. *attic*, *basement*). As a result, although these locational nouns might nevertheless activate representations of high or low spatial *locations*, these representations lack a motion element and thus might not interact with representations involved in the pursuing of a *moving* target. Experiment 6 generated reliable findings that could be explained using the same account described in the previous chapters. No clear pattern of results emerged from Experiment 7. Some possible reasons for these null effects, both methodological and theoretical, will be discussed. The absence of a significant effect in this experiment nonetheless motivates a set of assumptions, which will be tested in the experiment described in Chapter 6.

Experiment 6

Introduction

The majority of pursuit research has measured horizontal instead of vertical pursuit. This may be due to the fact that when the head is fixed, the human horizontal visual field is larger than the vertical one. Despite being the same type of eye movements, there are several differences between horizontal and vertical pursuit. A few studies have converged onto the finding that tracking performance is superior during horizontal than vertical pursuit (Collewijn & Tamminga, 1984; Baloh, Yee, Honrubia, & Jacobson, 1988; Grant, Leigh, Seidman, Riley, & Hanna, 1992). In contrast, it has also been reported in a later study that when pursuit initiation was concerned, subjects were better at initiating vertical smooth eye movements than horizontal ones (Rottach, Zivotofsky, Das, Averbuch-Heller, Discenna, Poonyathalang, & Leigh, 1996). Neuroscientific research supported the behavioural findings by providing evidence for the claim that horizontal and vertical pursuit are governed by different neural circuits (e.g. Keller & Heinen, 1991; Krauzlis & Lisberger, 1994). It has even been suggested that horizontal pursuit may be superior to vertical pursuit for evolutionary reasons, i.e. the pursuing of everyday motion objects tends to be in the horizontal dimension.

The distinctions between horizontal and vertical pursuit eye movement raise the question of whether the semantic effects demonstrated in the previous experiments can be generalized from vertical pursuit to horizontal pursuit. However, the difficulty in testing whether verb semantics can affect horizontal pursuit is that verbs implying horizontal motion in a specific direction are scarce.[38] Thus it is not possible to evaluate whether horizontal directional verbs can affect horizontal pursuit eye movements.

Nonetheless, the question of whether vertical directional verbs are capable of impacting on horizontal pursuit can be explored. Although there is no visual motion or eye velocity in the vertical dimension during horizontal pursuit, it is still possible for vertical directional verbs to modulate horizontal smooth eye movements. A series of experiments extensively tested the effect of vertically moving distractors on horizontal pursuit eye movements (Spering, Gegenfurtner,

---

[38] The only examples are "*read*" and "*write*", indicating rightward motion in the usage of English language.

& Kerzel, 2006). The common task for these experiments was to pursue a horizontally moving target while ignoring distractors that moved vertically into the periphery. Several interesting findings were uncovered: The horizontal smooth pursuit component decelerated considerably in response to a vertically moving distractor presented during the steady state. Eye movements also deviated vertically away from the moving distractor, even when subjects were instructed to pursue the horizontal target and ignore the distractor. However, when two vertical distractors moved simultaneously in opposite directions, the vertical deviation effect disappeared. These effects were interpreted as the consequences of inhibitory processes directed towards the distractor: The vertical distractor automatically attracted attention and a motoric response was programmed in response to it but was not executed, as the task dictated the system to pursue the target and ignore the distractor. This inhibition of the distractor response biased the eyes towards the opposite direction to the distractor. However, when there were two distractors present and moving in opposite directions, the processes to inhibit them both counteracted each other and the vertical deviation effect was diminished.

If vertically moving distractors can interfere with horizontal pursuit, vertical directional verbs, such as the experimental items used in the preceding studies reported, should also be able to have some kind of effect on horizontal smooth eye movements. However, there may be a critical difference between these directional motion verbs and vertically moving distractors: Although these verbs can activate representations of the motion denoted, as demonstrated in the previous chapters, these activations may not be strong enough to elicit an overt behavioural response but still able to bias any ongoing or future behavioural responses.[39] In other words, unlike the visual distractors, the comprehension of directional motion verbs may not be able to cause the vertical deviation effect during horizontal pursuit, especially if there is no explicit task associated with them. Nonetheless, these verbs may still be able to interfere with the horizontal pursuit task, given the motion representations activated by these verbs can regardless compete with the attentional demands of the oculomotor task.

---

[39] See the paragraph (General discussion, Chapter 2) regarding the question "why there is not always a kicking action every time when the word *kick* is heard" for more details of this assumption.

The experiment described below explored the effect of auditorily presented vertical motion verbs on horizontal pursuit eye movements. This study employed the same stimuli and procedures as Experiment 2, except that all the directional verbs implying motion in the vertical dimension were presented during leftward or rightward horizontal pursuit. Based on our previous findings, we predicted: a) that the verbs would cause pursuit deceleration, regardless of the direction they implied (i.e. upward or downward); b) language comprehension should not cause eye velocity in the vertical dimension during horizontal pursuit, thus the errors on the y axis between the gaze position and the pursuit target should not be modulated by verb semantics.[40]

Method

Participants

The subjects for this experiment were 30 undergraduate students from the University of York. They were rewarded with either course credit or £4. All subjects were native speakers of English and had normal or corrected-to-normal vision.

Materials and Procedures

See the material and procedure section for Experiment 2, Chapter 2, for details.

Design

All 24 experimental items (i.e. vertical directional verbs) and their 24 controls items were coupled with leftward or rightward target motion. There were also 48 filler words presented during vertical pursuit.

Results

There was no difference in the standard deviations of positional errors in the control trials between leftward and rightward pursuit (F (1, 29) < 1, p > .05). Therefore the data collected during leftward and rightward pursuit were collapsed together as horizontal pursuit.

---

[40] The question of whether vertical directional verbs could affect horizontal pursuit was, in fact, addressed by two of the conditions in Experiment 1 (Chapter 2). However, no semantic effect was revealed when positional errors were measured (F (1, 39) = 2.30, p < .05). This could be due to a lack of power in the experimental design as discussed previously.

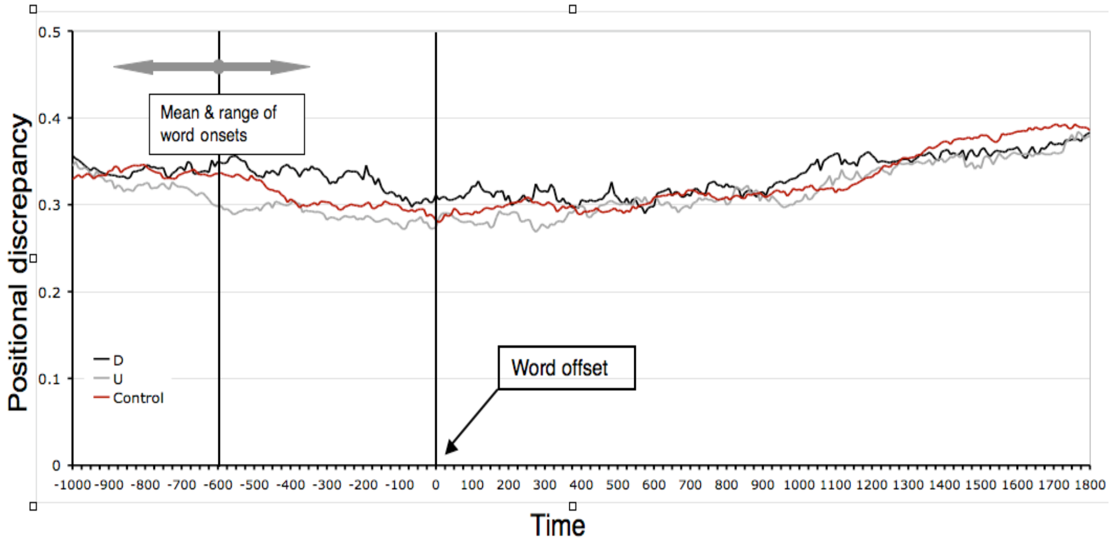


Fig. 4.1. *Absolute positional errors between the gaze position and the target dot on the y-axis across three conditions for horizontal pursuit: Downward verbs, upward verbs and control. Positional errors were synchronized to word offset for the same reason mentioned previously.*[41] *The solid grey circle indicates the approximate point of word onset based on the mean word duration, with the leftmost arrow representing the onset of the longest word and the rightmost arrow signalling the onset of the shortest word. Time zero signals word offset.*

As can be seen from the graph, vertical verbs did not influence the eye position on the y-axis. In the same time window as from previous experiments (i.e. from 150 ms before the word offset to 300 ms after the word offset), this observation was confirmed by statistical analyses (F (2, 58) < 1, p > .05). However, there was a main effect of word type when eye velocity was treated as the dependent variable (F (2, 58) = 17.35, p < .001). Pair-wise contrasts revealed that both downward and upward verbs caused eye deceleration relative to the controls (10.93 vs. 11.17, F (1, 29) = 12.85, p < .01; 10.80 vs. 11.17, F (1, 29) = 5.32, p < .05).

Discussion

Eye velocity during horizontal pursuit was decreased by vertical directional motion verbs, irrespective of whether they denoted upward or downward motion. However, these verbs did not bias eye movements away from the directions they implied in the vertical dimension.

The grounded cognition approach suggests that language comprehension is achieved through simulating relevant past experiences by reactivating the

---

[41] See Chapter 2 for details.

neural substrates involved during the actual experiences. However, a kicking action is not always generated every time the word "*kick*" is heard. We argue that this is because the modulatory effect on attention induced by language is *graded* and sometimes may only bias but not generate explicit behaviour. The results from the present study provided support for this assumption: The representations (or simulations) activated by the directional verbs did not generate eye velocity in the vertical dimension during horizontal pursuit. This could be due to the fact that the task was to pursue the horizontally moving target while no explicit task was associated with the verbs, thus the activation strength of the semantic representations was not strong enough to behaviourally bias the eye velocity direction and pursuit trajectory. Yet these motion representations activated by language (i.e. representations of vertical motion) nonetheless created competition with the representations relevant for the ongoing pursuit task (i.e. representations of horizontal target motion). This competition caused the eyes to decelerate, as the directional verbs and the pursuit task activated representations of motion in different spatial dimensions.

The results of this experiment are comparable to the study by Spering and colleagues (2006) described in the introduction. Those researchers also found horizontal eye decelerations both when the distractor moved upward and downward. However, no vertical deviation effect was observed in the present study. This could be due to the possibility that visually presented moving distractors were more likely to elicit a behavioural response (e.g. eye movements) compared to semantic representations activated by language. The outcome of the present study can also be related to several other studies in which the dependent measure was saccadic eye movements. For example, Salverda and Altmann (under review) presented subjects with visual scenes comprising several objects and the task was to saccade to a target dot as soon as it appeared at a random location in the scene. However, prior to the onset of the dot, a word was presented auditorily which might or might not refer to one of the objects in the scene. It was found that saccadic latency to the target dot was prolonged when one of the objects in the scene was named by the word than when none of the objects was named, although no eye movement was executed towards the named object. Under the account developed in our study, when an object in the scene was named, a boost in activation strength was received by the representation of that object, which in turn interfered with the oculomotor task (i.e. saccading to the target dot). However, since no explicit task was associated with the word, no saccade was actually directed towards the location of that named object.

Furthermore, our study is also comparable to the research on saccadic curvature: Saccade trajectories are almost never straight (Yarbus, 1967). Saccade curvature refers to the extent to which saccade trajectory deviates from a straight line. If a competing distractor is presented simultaneously with a saccade target, saccade trajectory tends to deviate away from the distractor (van der Stigchel & Theeuwes, 2008). In addition, visual stimuli with a strong emotion component can also bias saccade trajectory. It has been reported that when executing a saccade to bypass emotionally salient visual images in the periphery, saccade trajectories deviate away from visual scenes conveying negative emotions (Calvo, Nummenmaa, & Hyönä, 2008). These observations can all be explained by our account: Only when the distractor is salient enough or task-related does it become capable of biasing behaviour explicitly.

Finally, it is worth noting that the notion of task-relevance (Miller & Cohen, 2001; Cohen et al., 2004) is emphasized once again in the account developed for the present study. This has naturally led to the prediction that deviations in the vertical dimension should occur during horizontal pursuit when there is some kind of task associated with the verbs (e.g. lexical decision). Although no definitive conclusion can be drawn based on current data, some evidence has indicated that this is the case. When a stationary distractor was flashed during steady-state pursuit, it only evoked smooth eye movements in the direction of the flash when there was a task related to it (Blohm, Missal, & Lefèvre, 2005).

To summarise Experiment 6: Vertical directional verbs can affect horizontal pursuit eye movements, although they were not sufficient to induce any explicit behavioural response. These findings confirmed our assumption that the modulatory effect on attention mediated by language is graded and hinted again at the role of task in the interaction between different processing pathways involved in cognitive control.

Experiment 7

Introduction

The grounded cognition approach (e.g. Barsalou, 1999; 2007; Goldman, 2006) holds the view that linguistic meaning is grounded in perceptual and motoric experiences, and language comprehension is achieved through perception and action simulation by re-activating relevant sensorimotor representations stored in memory. One important prediction derived from this assumption is that the representations activated during language comprehension should directly bear an analogue relationship with their referents so that semantic representations should illustrate certain features of the denoted object or event. This prediction is supported by our previous experiments: Featural representations of motion direction implied by verbs are automatically activated when the verbs are heard. These language-mediated featural representations compete with the representations involved in ongoing pursuit and consequently bias the eye movement response to a moving stimulus.

As well as verbs, nouns can also activate certain features of their referents. Some research has been devoted to a class of nouns denoting objects that are typically found at specific locations. For example, clouds and roofs are typically found at relatively high locations while worms and basements are typically underfoot. Zwaan & Yaxley (2003) showed subjects vertically presented word pairs in which the words were either in an iconic relation with their referents (e.g. the word "*attic*" presented on top of "*basement*") or in a reversed iconic relation (e.g. the word "*basement*" presented on top of "*attic*"). It was found that the word pairs in a reversed iconic relation were responded to more slowly in a semantic-relatedness judgement task compared to the ones in a correct iconic relation. Furthermore, a study using single nouns has demonstrated similar findings in which the word "*eagle*" was responded to more quickly when it was presented at the top rather than at the bottom of a display while the reaction time to the word "*snake*" was shorter when it was at the bottom (Šetić & Domijan, 2007).

In a more recent study, Estes and colleagues (2008) presented subjects with nouns denoting objects associated with high or low locations (e.g. "*head"* or "*foot"*) prior to a perceptual discrimination task in the higher or lower visual field.[42] The results indicated that performance on the discrimination task was impaired when the location implied by the noun coincided with the location where the discrimination target appeared. The authors have attributed the observed effect to attention orienting and perceptual simulation: Attention is first automatically shifted to the location implied by the word, followed by the "running" of a perceptual simulation at that location. Thus the perceptual system was engaged at the attended location by object simulation and consequently less available for the discrimination task at that same location.

Under the biased competition theory (Desimone & Duncan, 1995) and the guided activation theory (Miller & Cohen, 2001), an alternative explanation (or at least, a more elaborated explanation) can be provided for these results. Instead of being the consequence of a perceptual system being engaged by another task, the impaired discrimination performance can be attributed to competition between the representation activated by the word and the representation activated by the discrimination task, which both encode the same spatial location. Nonetheless, regardless of whichever account, these results demonstrate that locational information implied by objects nouns (e.g. "*head*" implies high locations while "*foot*" implies low locations) is activated when the word is being processed.

Given that directionality implied by motion verbs can affect pursuit eye movements, a natural question to follow is whether locational information carried by object nouns can also influence pursuit eye movements. In the experiment reported below, we addressed this question by presenting nouns, during vertical pursuit, denoting objects that typically appear at high or low locations. The visual stimulus used to induce pursuit eye movements was a dot moving across the computer screen sinusoidally (cf. Experiment 1-3). All the experimental items (i.e. locational nouns) were selected based on the results of a norming experiment. We predicted that if the location information carried by nouns could affect pursuit eye movements, interaction patterns similar to the ones reported in Chapter 2 would be observed in which eye velocity was modulated as a function of the relative gaze position and word meaning.

---

[42] See the section titled "*Evidence from behavioural research*", Chapter 1, for details of this study.

The norming experiment

The locational nouns were chosen based on the results gathered in a norming experiment. The experiment was conducted online with a web-based questionnaire, in which subjects had to answer the question "*To what degree is this entity likely to be found in a particular spatial location*?" by rating on a scale of -3 to 3, with -3 signifying "*deep below the ground*", 3 indicating "*way above your head, high in the sky*" and 0 representing "*between your head and your feet*".[43] The questionnaire was randomly distributed to universities across the country. 28 responses were received in total. 24 nouns, 12 indicating objects at high locations (e.g. *attic*) and 12 denoting objects at low locations (e.g. *basement*) were selected based on their mean ratings and standard deviations.[44] There was no significant difference between the ratings for the high and low nouns (2.45 vs. 2.21, F (1, 23) < 1, p > .05), suggesting that high locations were implied by the high nouns to the same extent as low locations implied by the low nouns.[45]

Method

Participants

Thirty subjects took part in this experiment in exchange for course credit or £4. All were native speakers of English and had normal or corrected-to-normal vision.

Material and Procedure

The material and procedures involved in this experiment were comparable to those in Experiment 3, except that 24 locational nouns and their matched controls were added into the experimental items along with the directional verbs and their controls.

Design

All 48 experimental items (i.e. 24 directional verbs and 24 locational nouns) and their 48 pair-wise matched controls (i.e. 24 non-directional verbs and 24 non-locational nouns) were paired with downward pursuit, given null results

---

[43] See Appendix 3 for instructions given to subjects.

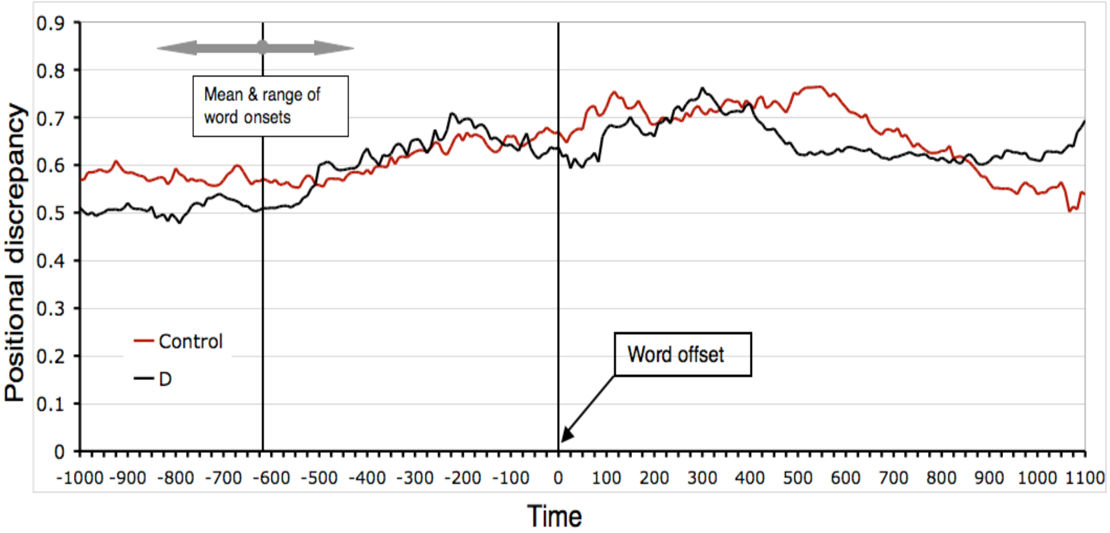
[44] See Appendix 4 for the complete list of locational nouns.

[45] In the raw data, high nouns received positive ratings while low nouns received negative ratings. However, the negative ratings were converted to positive values prior to any statistical analyses.

had been repeatedly generated by upward pursuit. In order to limit the duration of the experiment for practical reasons, the variable of Item was created as a between-subject factor, which meant that half of the subjects were presented with 12 directional verbs, 12 locational nouns and their pair-wise matched controls while the other half of the subjects listened to the other 12 directional verbs, 12 locational nouns and their pair-wise matched controls. The 24 directional verbs and 24 locational nouns were selectively assigned into two groups so that there was no difference in rating, frequency and duration between these two groups of experimental items. All 48 filler items were presented during upward pursuit. Thus an individual subject was in total presented with 12 directional verbs, 12 locational nouns and their 24 matched controls during downward pursuit and 48 filler words during upward pursuit.

Results

a. Verbs



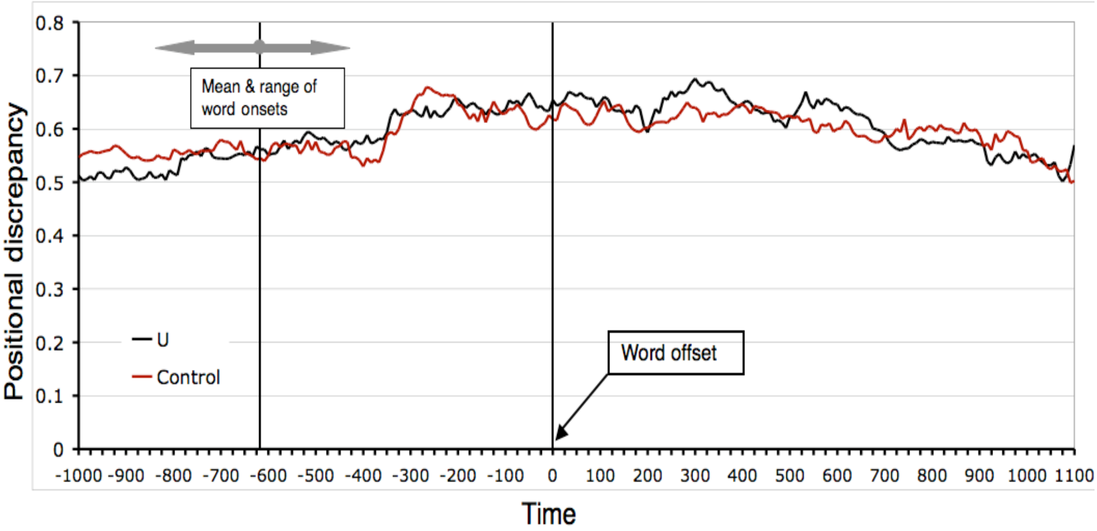


Fig. 4.2. & 4.3. *Absolute positional errors between the gaze position and the target between downward verbs and control (top), and between upward verbs and control (bottom) for downward pursuit. Positional errors were synchronized to word offset for the same reason mentioned previously.*[46] *The solid grey circle indicates the approximate point of word onset based on the mean word duration, with the leftmost arrow representing the onset of the longest word and the rightmost arrow signalling the onset of the shortest word. Time zero signals word offset.*

---

[46] See Chapter 2 for the details.

These two graphs indicated that there was no clear modulatory effect on positional discrepancies produced by the directional verbs in this experiment, except for between downward verbs and their controls in a time window from 400 ms to 750 ms after the word offset. This observation was confirmed by statistical analyses (F (1, 29) = 4.43, $p < .05$).[47] However, no reliable patterns of semantic effect emerged from velocity analyses conducted over the same time window. There was no difference in eye velocity between the downward verbs and their controls in the leading (54.3%) cases (6.18 vs. 6.35, F (1, 29) = 3.22, $p > .05$) or the lagging (45.7%) cases (5.17 vs. 5.44, F (1, 29) = 2.98, $p > .05$). Neither there was any difference between the upward verbs and their matched controls in the leading samples (6.13 vs. 6.08, F (1, 29) < 1, $p > .05$) or the lagging samples (5.06 vs. 5.20, F (1, 29) < 1, $p > .05$).

---

[47] As observed in previous experiments, neither the downward verbs (F (1, 29) < 1, $p > .05$; F (1, 29) = 1.17, $p > .05$) nor the upward verbs (F (1, 29) < 1, $p > .05$; F (1, 29) = 2.16, $p > .05$) affected the number or the amplitude of saccades launched during pursuit.

b. Nouns

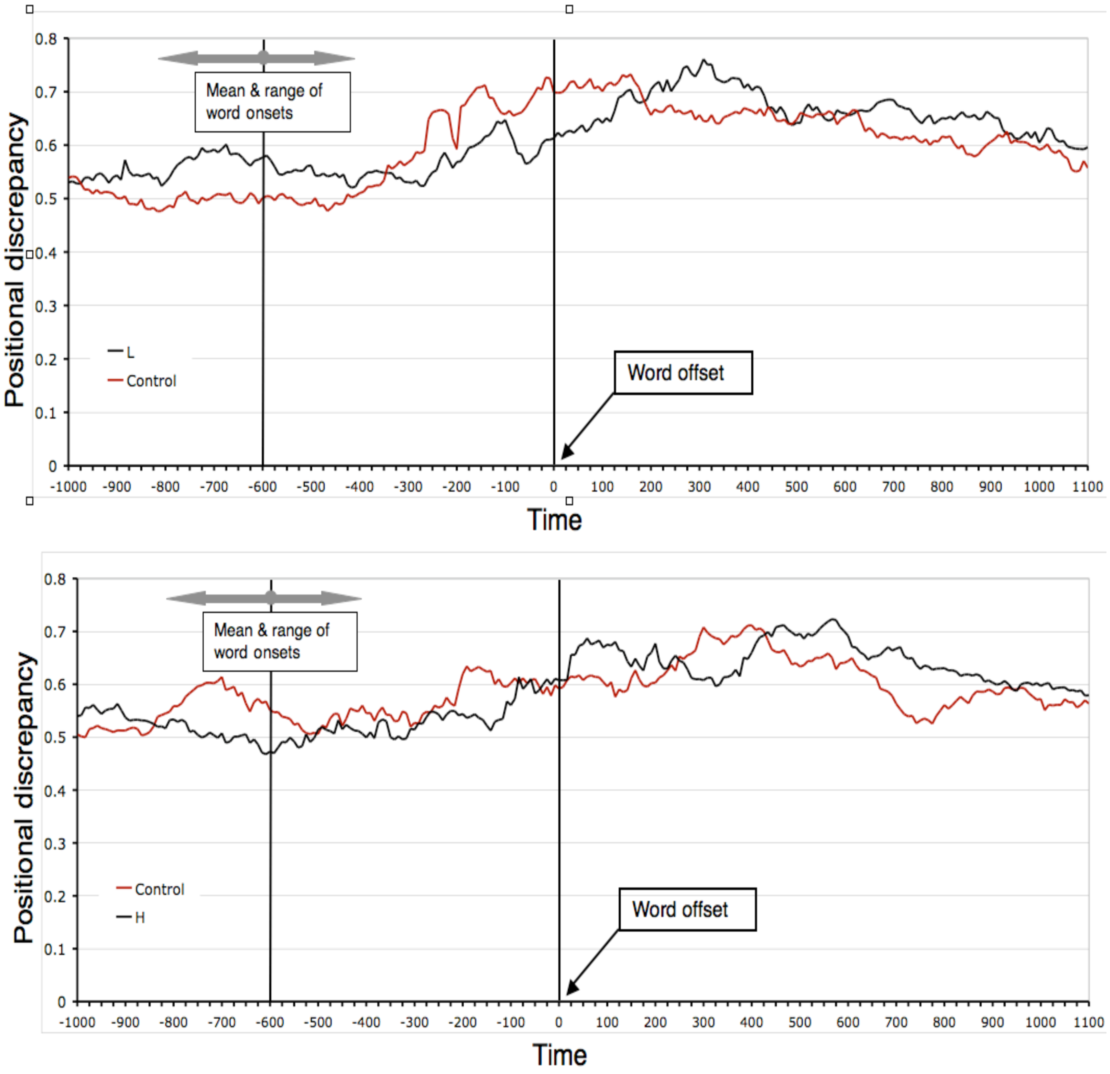


Fig. 4.4. & 4.5. *Absolute positional errors between the gaze position and the target between low nouns and control (top), and between high nouns and control (bottom) for downward pursuit. Positional errors were synchronized to word offset for the same reason mentioned previously.*[48] *The solid grey circle indicates the approximate point of word onset based on the mean word duration, with the leftmost arrow representing the onset of the longest word and the rightmost arrow signalling the onset of the shortest word. Time zero signals word offset.*

Inspections of these graphs suggested that low nouns such as "*basement*" might have modified the positional errors during pursuit, but not high nouns such as "*attic*". This observation was reflected in the statistical analyses: Low nouns decreased the positional discrepancies relative to their controls in a time window

---

[48] See Chapter 2 for the details.

from 100 ms before the word offset to 200 ms after the word offset (0.61 vs. 0.71, *F* (1, 29) = 4.38, *p* < .05). There was no difference between the high noun condition and the control condition.[49] Furthermore, after splitting eye velocity data into the leading (52.9%) and lagging (47.1%) cases, no reliable difference was found across the conditions in this time window. There was no difference in eye velocity between the high nouns and their control condition in the leading cases (6.80 vs. 6.75, *F* (1, 29) < 1, *p* > .05) or the lagging cases (5.43 vs. 5.45, *F* (1, 29) < 1, *p* > .05). Similarly, the low nouns did not differ from their control items in the leading samples (6.60 vs. 6.75, *F* (1, 29) = 3.55, *p* > .05) or the lagging samples (5.24 vs. 5.27, *F* (1, 29) < 1, *p* > .05).

Discussion

The results from the positional analyses hinted at the possibility that object nouns might be able to activate spatial representations of locations at which the denoted objects are likely to be found. However, the lack of a difference in eye velocity across conditions made the findings from positional analyses less reliable and harder to interpret. Furthermore, it seemed that the positional error reduction caused by directional verbs seen in previous experiments had been replicated in the current study. Nonetheless, the error reduction was only observed in a much later time window compared to previous experiments (i.e. from 400 ms after the word offset vs. during the word) and the semantic effects on eye velocity induced by directional verbs were absent from the present experiment.

It is worth noting that, regarding the directional verbs, although the semantic effects emerged “late” in the positional analyses and was completely absent from the velocity analyses, it was unlikely that the null effects were due to the possibility that directional verbs did not affect pursuit eye movements, as the influence of verb semantics on pursuit had been replicated in several experiments with different samples of subjects, distinct groups of control items and even across experimental manipulations. Thus, it is also premature to conclude that locational nouns cannot induce similar semantic effects that are comparable to

---

[49] Neither the low nouns (*F* (1, 29) < 1, *p* > .05; *F* (1, 29) < 1, *p* > .05) nor the high nouns (*F* (1, 29) < 1, *p* > .05; *F* (1, 29) = 1.35, *p* > .05) affected the number or the amplitude of saccades launched during pursuit.

the ones demonstrated using directional verbs in the preceding experiments. The question remains, then, why we did not replicate the earlier effects (i.e. semantic effects observed in Experiments 1-3) in this study; most likely, it is due to a lack of power in the experimental design and the additional between-subject variance introduced by splitting the items across two groups of subjects.[50]

The decrease in positional errors caused by *low* nouns (e.g. "*basement"*) indicates that some kind of spatial representations must have been activated during the comprehension of these locational nouns. Furthermore, this modulatory effect on eye position triggered by low but not high nouns is similar to the pattern observed in directional verbs, in which the downward verbs decreased the positional errors while upward verbs did not affect eye position. This calls for the questions of first, what kind of representations are activated by verbs and nouns and second, whether directional verbs and locational nouns can activate certain similar representations, since the semantics of both types of words contain a spatial component. As suggested in the discussion section of Experiment 3, when a directional motion verb is being processed, a representation of motion in the implied direction, as well as the space in which the motion occurs, should be activated as a part of the comprehension process.[51] However, although location nouns can also activate representations of particular areas in space where the denoted objects are likely to be found, there is no motion element involved in their semantics. As a result, the spatial representation associated with directional verbs is to provide the *dynamic* background in which *motion* occurs while the spatial representation activated by locational nouns reflects the *static* environment in which an *object* appears.[52] The former emphasizes the motion-related parameter of *direction* whereas the latter highlights the object-related property of *location*. We propose it is such

---

[50] Similar to the downward verbs, the low nouns decreased the positional errors compared to the control condition. Based on this finding, it is plausible to predict that these locational nouns may have the same effects on eye velocity as directional verbs, providing that the experimental design is as powerful as the one in Experiment 3, with only locational nouns as the experimental items.

[51] See Chapter 2 for details.

[52] The spatial elements associated with nouns can be modulated if the nouns are not in isolation but accompanied by modifiers. For example, representations of different parts in space should be activated by phrases such as "ant *head*" vs. "giraffe *head*".

distinctions between these semantic representations that define whether a given word is a verb that denotes motion and actions, or a noun that refers to entities. This inherently agrees with some predictions that are derivable from the grounded cognition principles: If linguistic meaning is constructed from past perceptual and motoric experiences, separate word classes such as verbs and nouns should be naturally distinguishable based on the distinct features and properties associated with the denoted objects or actions.[53]

Although directional verbs and locational nouns may activate different representations that accentuate either direction or location, it is nonetheless possible that both classes of words can influence pursuit eye movements. During the tracking of a moving target, there are two oculomotor tasks involved: First, the mechanism responsible for the smooth eye movement component needs to match eye velocity to retinal target velocity as closely as possible. Second, the positional errors between the retinal target image and the fovea need to be minimized. Thus both motion-related and object-related features, such as direction and location, are determined to be both relevant by the oculomotor tasks. These task-relevant features can in turn interact with the directional or locational information carried by the verb or noun semantics and cause the pursuit system to be susceptible to the influence of language.

Given information related to both motion and location are relevant and interlinked for the pursuit task, the assumption that directional verbs and locational nouns can activate spatial representations containing different features cannot be tested by measuring their influence on pursuit eye movements. As an alternative motoric response, saccadic eye movements may provide the solution to this problem. A model of saccade generation proposed by Findlay and Walker (1999) has illustrated the saccadic system as two parallel information-processing and command-generating pathways that extend vertically through various processing stages, ranging from high-level cognitive control to the generation of low-level motoric signals.[54] These two pathways reflect the separation of mechanisms controlling WHEN and WHERE information. The WHEN pathway decides whether and when a saccade should be launched while the WHERE pathway is responsible for calculating the metrics (e.g. amplitude) of a planned

---

[53] A similar idea has been proposed earlier by Langacker (1986).

[54] See Chapter 6 for a more detail description and evaluation of this model.

saccade given the location of the target. The decision on whether a saccade should be generated in the WHEN pathway is made based on a single individual signal generated by a competitive interaction between a *fixate* centre, which insists on maintaining a fixation, and a *move* centre, which dictates eye movements. In contrast, the WHERE pathway consists of a range of topographic mappings that result in a saliency map, in which the most salient location is eventually selected as the saccadic landing position. Given the distinction between the WHEN and the WHERE pathway, it can be predicted that the motion-related semantics carried by directional verbs may bias the competition between the fixate centre and the move centre thus modulating saccadic launch latency while it is possible for the locational information implied by the high and low nouns to impact on the saliency map within the WHERE pathway and affect saccadic landing position. These predictions were tested in an experiment described in Chapter 6, in which subjects listened to directional verbs and locational nouns prior to producing saccadic eye movements to targets presented in the periphery.

In short, the results from Experiment 7 hinted at the possibility that locational nouns could influence pursuit eye movements, however, the lack of a semantic effect on eye velocity made the experimental outcome hard to interpret. Nonetheless, the null effect observed in this experiment has led to clear and testable predictions regarding the interaction between saccadic eye movements, verbs and nouns.

General discussion

The two experiments described above were designed to address the question of whether the interaction between directional verbs and pursuit eye movements reported in preceding studies could be generalised, such as when the pursuit was horizontal rather than vertical (Experiment 6) and when locational nouns were employed as the linguistic stimuli (Experiment 7). The only statistically reliable effect was revealed by Experiment 6, which indicated that although the semantic representations activated by vertical directional verbs were not sufficient to generate any explicit behavioural response in the vertical dimension, these representations nonetheless competed with representations required by the horizontal pursuit task, and consequently influenced horizontal eye velocity. Experiment 7 produced mainly null effects; as a result, it remained inconclusive whether locational nouns could produce any semantic effects on pursuit eye movements.

Despite having generated statistically non-significant results, these two studies reported in this chapter contributed to our current line of research, as the assumptions and predictions inspired by the experimental outcomes accorded well with the theoretical frame work we had adopted (Desimone & Duncan, 1995; Miller & Cohen, 2001). At any given moment, there are multiple representations or processing pathways simultaneously active within the system. These different representations or pathways compete for behavioural expressions under the regulatory force of attention that governs the whole system. Attention adjusts the relative activation strengths of these competing representations based on certain contextual factors (e.g. task relevance) and consequently biases the competition in favour of certain representations over the others. The representations that "lose out" on the competition are not externally expressed, but may still be able to bias behavioural responses. Experiment 6 suggested that although simply hearing verbs such as "*dive*" and "*climb*" was not sufficient to induce eye movements in the vertical dimension, directional semantics carried by these motion verbs could nonetheless cause eye deceleration in the horizontal dimension. Furthermore, based on the results from Experiment 7, it could be predicted that the same set of linguistic items (i.e. directional verbs and locational nouns) would impact on saccadic eye movements in a different ways compared to pursuit, as the saccadic task places a considerably heavier burden on the end location of the eye movement (i.e. saccadic landing position) than smooth pursuit. In conclusion, these two experiments may not have provided a satisfactory and unambiguous answer to the question of whether the original semantic effect on pursuit can be generalized, the data gathered in these studies collaboratively pointed towards the reoccurring theme of the previous chapters regarding the competition between concurrent representations, the modulatory force of attention and the importance of task.

As discussed at the end of Chapter 3, there were two issues remaining after replications of the effect of verb semantics on pursuit eye movements: The first issue was whether this effect could be generalized, which the current chapter had attempted to address. The other issue was whether the same pattern of results would persist when the leading and lagging samples were artificially manipulated instead of occurring naturally during the pursuing of a moving target. In the experiment reported in the following chapter, subjects performed a perceptual detection task while pursuing a moving target. The secondary detection task dictated that either a location above or below the pursuit target had to be attended to and was used to tentatively simulate the leading vs. lagging situations. The assumption underpinning this paradigm was that when the location above the pursuit target was attended to, it would be equivalent to the leading cases, in which the gaze position was below the target; whereas the lagging cases would be similar to when the location below the pursuit target was attended to.

Chapter 5. Tracking while attending to locations above and below the target: The artificial "leading" and "lagging" cases

There were two aims for this experiment (Experiment 8) described in the current chapter: a) to introduce and evaluate a new experimental paradigm in which a perceptual detection task was incorporated into pursuit, so that attention was allocated to the pursuit target as well as to a detection target placed above or below the pursuit target; b) to use this new paradigm to artificially create situations in which attention is directed above or below the pursuit target, thereby allowing us, in principle, to explore further the "leading" and "lagging" effects observed in Experiments 1 - 3. Subjects pursued a moving dot while anticipating a dot placed either above or below the pursuit target to change colour on some of the trials. The same set of linguistic items from Experiment 2 (i.e. directional verbs and their matched controls) was only presented during *catch* trials in which the detection target did not change colour. Filler words were presented in non-catch trials and subjects pressed a button as soon as the colour change was detected. Both eye movement and reaction time data were collected and discussed. Finally, this paradigm was evaluated in relation to the results of the current study.

Introduction

As revealed by Experiment 2 and 3, it is evident that eye velocity during pursuit can be modulated by a conjunction of verb semantics and gaze position relative to the target. During downward pursuit, when the gaze position is ahead of, or spatially below the target, downward verbs (e.g. *dive*) can cause eye deceleration while upward verbs (e.g. *climb*) produce eye acceleration. However, when the gaze position is behind, or spatially above the target, the reverse occurs: Eye velocity is reduced by downward verbs but increased by upward verbs. These findings have been taken as evidence to support the proposal that language semantics can systematically modulate eye velocity during pursuit through competition between semantic representations activated by language comprehension and motion representations demanded by the oculomotor task.

Although these results have been replicated, an issue remains: The leading and lagging cases, based on which the data were separated, did not occur as a part of the experimental manipulation. In other words, the gaze position happened to be ahead or behind the target during pursuit as a function of the natural response from the oculomotor system to a moving stimulus. Although there was no difference between the proportions of the leading and lagging samples, the distribution of samples containing gaze positions being below or above the target would have inevitably varied across trials and subjects. Thus it was hard to determine the temporal relationship between the acoustic life of the word and the occurrence of leading and lagging samples.

In order to resolve this issue, we developed a new experimental paradigm (the *multiple-dot* paradigm), which was essentially a standard pursuit task combined with a perceptual detection task.[55] In this paradigm, instead of a single target, a pattern such as the one illustrated in Figure 5.1 moved either upward or downward across the screen. The dot at the centre served as the pursuit target while the detection target was either the dot on the top or at the bottom. At a random point during pursuit, the detection target would change colour and a button-press response was required when the event was detected. In order to perform the detection task, the detection target, which was either the dot above or below the pursuit target, had to be attended to during pursuit, as well as the pursuit target, until the colour change took place. By including a detection task at a location either above or below the pursuit target, the space immediately above or below the pursuit target became task-relevant. The "leading" cases, in which representations of the pursuit target and the space above the gaze position were task-relevant, and the "lagging" cases, in which representations of the pursuit target and the space below the gaze position were task-relevant, were thus artificially created and controlled for by employing the colour change detection task. However, the colour change only occurred in some of the trials. During those trials in which the colour change did not happen (i.e. the *catch* trials) and the detection target had to be attended to all the time, directional verbs such as "*dive*" and "*climb*" were presented auditorily. The aim of these catch trials was to explore directly the potential semantic effect induced by directional verbs when the space above or below the pursuit target remained task-relevant through the entire course of pursuit tracking.

---

[55] Details of this paradigm are later described in the Method section.

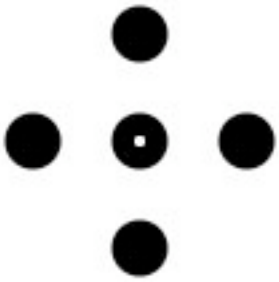


Fig. 5.1. *An illustration of the visual stimulus for the multiple-dot paradigm. The dot at the centre with a "hollow" middle is the pursuit-tracking target while the dot above or below functions as the colour-change detection target. In order to induce pursuit eye movements, the pattern moved together as a group with the same velocity profile.*

The idea of incorporating a secondary perceptual task into smooth pursuit is not novel. A number of studies have employed secondary tasks to study pursuit responses under conditions of divided attention (e.g. Hutton & Tetally, 2005; Kathmann, Hochrein, & Uwer, 1999; Kerzel et al., 2008) while some other researchers embedded perceptual tasks into the pursuit target to investigate the effect of increased attention to the target on pursuit performance (Shagass et al., 1976; Sweeney, Haas, Liu, & Weiden, 1994). Finally, secondary perceptual tasks have been integrated into pursuit tracking in order to explore the involvement of attention in pursuit eye movements (Chen et al., 2002; van Donkelaar & Drew, 2002).

The perceptual detection task in our paradigm was designed to interact with directional semantics conveyed by single verbs. Unlike previous studies that measured solely the outcome of having a secondary task, our paradigm aimed to test the interaction between the attentional consequences of performing the perceptual task and language processing. The detection target remained visible all the time and the response required was to react to a change in the state of the target. Thus at least some *continuous* attention was allocated to the detection target during pursuit. This was not the case for other studies, in which the target for the secondary task was typically presented very briefly and attentional resources were predominantly allocated to pursuit tracking. Finally, to our knowledge, there is only one other study in which the sensory detection target was presented within the motion trajectory of the pursuit target (von Donkelaar

& Drew, 2002).[56] However, only reaction time to the detection target was measured in that study but not the eye movement itself. In the current study, both the eye movement response to the pursuit target and reaction times to the detection target were measured.

Using the multiple-dot paradigm, the present study aimed at replicating the semantic effect on pursuit eye velocity observed in the past experiments under artificially created leading and lagging conditions. Subjects performed the pursuit tracking task and the perceptual detection task currently, while listening to directional motion verbs. We predicted that during downward pursuit, when the detection target was positioned *above* the pursuit target, the detection target, as well as the area *above* the pursuit target would be represented. As a result, similar to the leading cases in previous studies, downward verbs would cause eye deceleration while upward verbs would increase eye velocity. However, the reverse was expected when the dot below the pursuit target functioned as the detection target: The area below the pursuit target was dictated to be relevant by the perceptual detection task. Thus a response pattern previously seen in the lagging cases was expected in which downward verbs would cause eye acceleration while upward verbs would decrease eye velocity.

Method

Participants

Forty undergraduate students took part in this experiment in exchange for either £4 or one course credit. All subjects were native speakers of English and had normal or corrected-to-normal vision.

Materials

The linguistic materials used in this experiment were identical to the ones in Experiment 2. The visual stimulus, however, was the pattern illustrated in Figure 5.1, instead of a single dot. The dot at the centre (i.e. the pursuit target) occupied 0.75° on the display screen, with a “hollow” centre of 0.11°, at a viewing distance of 600 mm. The other four dots were of the same size and each of them was 1.49° centre-to-centre away from the pursuit target. The value 1.49° was chosen as the mean peak positional errors occurred in previous experiments never exceeded 1°. Furthermore, human subjects can react to detection targets

---

[56] See Chapter 2 for details of this study.

presented < 2° ahead or behind the pursuit target with relatively short latencies (van Donkelaar & Drew, 2002). At some point during the trial, the detection target changed its colour from black (RGB: 0, 0, 0) to dark blue (RGB: 0, 51, 112). This colour change was subtle so that the detection of the change required the target to be carefully attended to. The four dots in the periphery moved with the pursuit target as a group and the motion parameters were identical to the ones reported in Chapter 2. Reaction times to the colour change were collected using a response pad (SR Research).

Procedure

The main experimental procedures for the current study were similar to Experiment 2 and 3, except for the additional perceptual detection task. Subjects were instructed to pursue the dot as accurately as possible while attending to the dot either above or below the pursuit target in anticipation of a possible colour change. The detection target only changed its colour on non-catch trials at 450 ms, 1050 ms or 1600 ms after the motion onset. Since all words were presented 750 ms after the motion onset, the colour change occurred either before the word onset, during the word or after the word offset. Subjects were required to press a button on the response pad as soon as the colour change was detected. It was not possible to predict the arrival of the colour change, nor could subjects identify a given trial as a non-catch trial before the colour change occurred. This meant that during the catch trials, the detection target was attended to until long after the word offset towards the end of the trial.

Design

The study employed a blocked design, in which half of the subjects were instructed to always attend to the dot above the pursuit target while the other half always considered the dot underneath as the detection target. In order to limit the duration of the experiment, the experimental items (i.e. directional verbs) and their controls were split into two groups and presented to two different groups of subjects. The separation of the items was based on their rating, duration and frequency so that there was no difference between the two groups. All directional verbs and their controls were presented during downward pursuit in catch trials that did not involve the colour change. Additional filler items (i.e. non-directional nouns and verbs) were also included for counterbalancing purposes. Thus an individual subject was presented with 24 downward catch trials that were paired with 6 downward verbs, 6 upward verbs and their 12 controls, as well as 24 downward non-catch trials, 24 upward catch trials and 24 upward non-catch trials during which filler items were presented.

Results

a. Reaction time analyses

During downward pursuit, subjects performed on the detection task at ceiling level with a mean hit rate of 93.36%. Reaction times to the colour change in non-catch trials were marginally faster when the detection target was placed in front of the pursuit target compared to behind (438 ms vs. 455 ms, t (38) = 2.02, $p = .05$).[57]

b. Eye movement analyses

Discrepancies between gaze and target position and eye velocity were measured as dependent variables. Details on the method of calculation were reported in Chapter 2. Since all experimental items and their controls were presented during downward pursuit, the following results reported reflect the data collected from downward pursuit only.

For the positional analyses, the between-subject factor of whether the location above or below the pursuit target was attended to was not entered into the analyses.[58] Thus the congruency of the directional verbs was determined by whether the implied motion direction was in accordance or in conflict with the direction of pursuit. The data were analyzed in such a way in order to simulate how positional analyses were conducted in previous experiments. As indicated by a lack of clear divergence among the three curves in the graph below, no difference was found across conditions with positional error as the dependent measure.

---

[57] Only reaction time data from downward pursuit were analysed, since all experimental items were paired with pursuit in this direction.

[58] No difference was found in positional errors between the *above* or *below* condition (0.74 vs. 0.70, F (1, 38) = 2.77, $p > .05$).

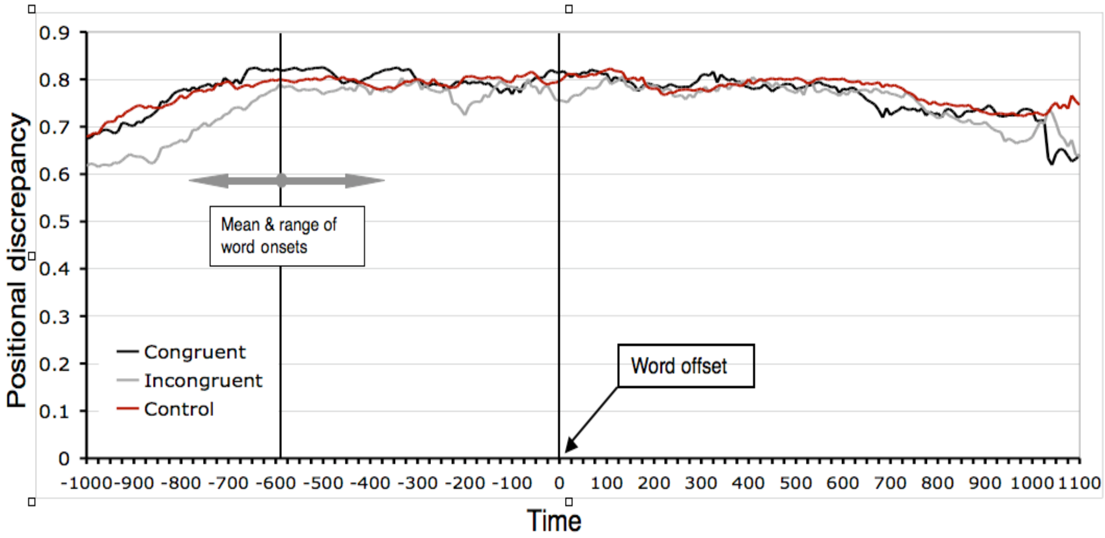


Fig. 5.2. *Absolute positional errors between the gaze position and the target dot across three conditions for downward pursuit: Congruent, incongruent and control. Positional errors were synchronized to word offset for reasons described elsewhere.*[59] *The solid grey circle indicates the approximate point of word onset based on the mean word duration, with the leftmost arrow representing the onset of the longest word and the rightmost arrow signalling the onset of the shortest word. Time zero signals word offset.*[60]

For eye velocity analyses, the eye movement data were split based on whether the colour change was allocated to the dot above or below the pursuit target. This separation was thought to reflect the procedure used in previous studies of dividing the data into leading and lagging cases. When the space above the pursuit target was attended to, which was assumed to imitate the leading cases, there was a main effect of Condition (i.e. word type) in a time window from word offset to 200 ms after the word offset (F (2, 58) = 7.42, $p < .01$). Pair-wise comparisons indicated that downward verbs increased eye velocity compared to the controls (7.13 vs. 7.05, F (1, 19) = 11.60, $p < .001$) while upward verbs did not differ from the control condition (6.97 vs. 7.05, F (1, 19) = 2.81, $p > .05$). Regarding the condition in which the space below the pursuit target had to be attended to (i.e. the lagging case equivalent), there was again a main effect of word type in a slightly earlier time window from 150 ms before the word offset to 100 ms after the word offset (F (2, 58) = 6.30, $p < .05$). In this condition, upward verbs again failed to modulate eye velocity in contrast to the

---

[59] See Experiment 1, Chapter 2 for details.

[60] The distinctively flat nature of these curves, compared to the curves seen in earlier studies, will be discussed below.

control words (6.05 vs. 6.08, F (1, 19) = 2.92, $p > .05$). However, downward verbs caused eye acceleration compared to the controls (6.22 vs. 6.08, F (1, 19) = 10.50, $p = .001$).

Analyses were also carried out on eye velocity data after the data were spilt based on whether the gaze position was ahead (i.e. the original "leading" cases) or behind (i.e. the original "lagging" cases) the pursuit target, irrespective of the location of the detection target. In the leading cases (55.3%), neither the downward verbs (7.08 vs. 7.03, F (1, 39) = 2.99, $p > .05$) nor the upward verbs (6.97 vs. 7.03, F (1, 39) = 3.21, $p > .05$) differed from the control items. The same pattern was also found for the lagging cases (44.7%): Eye velocity in the downward verb (6.14 vs. 6.10, F (1, 39) = 2.46, $p > .05$) or the upward verb condition (6.11 vs. 6.10, F (1, 39) < 1, $p > .05$) was not significantly compared to the control condition.

Discussion

These results did not confirm our predictions, in fact, they turned out to be partially contradictory to what was expected: When the detection target was located in the space above the pursuit target, instead of decreasing eye velocity as predicted, downward verbs led to eye acceleration. In the other condition where the detection task was placed below the pursuit target, downward verbs again caused eye acceleration, but in this case as predicted. Furthermore, there was no longer a two-way interaction between eye velocity and verb semantics, as the upward verbs in both the "above" and "below" conditions did not have any modulatory effect on pursuit eye movements.

The reaction time analyses revealed a slight advantage for colour-change targets located in front of the pursuit target compared to when they were behind the pursuit target. This finding is in line with the report that detection targets appearing immediately ahead of a pursuit target are more rapidly responded to than if they are presented behind (van Donkelaar & Drew, 2002). These reaction time results confirmed our assumption that, in addition to representing the target during pursuit, the space immediately ahead of the target is also critical for the oculomotor task.

Although no statistically significant difference was discovered in the positional analyses, it should nonetheless be noted that compared to Experiments 1-3, positional errors in the current study followed a different pattern across the time course of a given trial. More specifically, the magnitude of the errors stayed relatively constant throughout the trial (Figure 5.2), whereas in previous experiments, the errors increased as the target was approaching its peak velocity and decreased towards the termination stage of pursuit (Figure 2.1, 2.2, 2.3 and 2.4). However, it is worth noting that the peak error observed in the present experiment did not differ from that observed in previous experiments. The disparity in the patterns of positional errors between the current and the previous studies could be attributed to the influence of the detection task on the perceptual sampling processes involved in pursuit eye movements. As proposed earlier, pursuit relies more heavily on perceptual feedback during the initiation and termination stage. Once the eye has entered the maintenance stage, a stored representation of target motion becomes the predominant force for driving and modulating pursuit.[61] When a detection target has to be attended to, the perceptual sampling processes inevitably suffer from the secondary task that shares the same available resources.[62] Thus pursuit eye movements during the initiation and termination stage were compromised to a greater extent compared to the maintenance stage, as perceptual sampling played a more important role in pursuit control in the former two stages compared to the latter. This was precisely reflected in the positional error analyses, with an increase in the magnitude of the errors during the early and late part of pursuit due to the detection task; and the peak error during the maintenance stage remaining unchanged relative to the previous findings.

The eye velocity analyses revealed surprising results. Contrary to our predictions, downward verbs increased eye velocity irrespective of whether the detection target was placed above or below the pursuit target. Regardless of condition, upward verbs did not seem to play any role in modulating eye velocity during pursuit in the current experiment. These unexpected results could be attributed to either methodological or theoretical reasons (or both). These possibilities are discussed in turn.

---

[61] See Experiment 5, Chapter 4 for more details on this assumption.

[62] The term “resources” is used here, given that what is being shared when multiple tasks operate currently remains debatable. Some researchers consider it to be visual and cognitive attention (e.g. Chen et al., 2002) while others may conceptualise it as the privilege of certain mental processes to be expressed behaviourally.

An assumption underling the application of the multiple-dot paradigm was that the gaze position should almost always coincide with the pursuit target, or at least should oscillate approximately around the target. In other words, the detection targets would fail to simulate the attentional state during the leading and lagging samples of the previous studies if the gaze position strayed too far away from the pursuit target. For instance, when the dot above the pursuit target had to be attended to during downward pursuit, if the gaze position fell too far behind the pursuit target, it would in fact end up being *above* the detection target. In this case, both the detection and the pursuit target would be spatially below the gaze position, and the resulted attentional state of the system would no longer correspond to the leading cases in the previous studies. Although this could potentially be an issue for the multiple-dot paradigm, the data indicated that it was unlikely that the results from the current study had been affected in this way. Positional analyses (Figure 5.2) revealed that the peak positional error between the gaze position and the pursuit target was no greater then 0.9°. Since the distance separating the pursuit and the detection target was 1.49°, it was improbable that the gaze position would fall beyond the detection targets. Thus gaze position appears to have oscillated around the pursuit target within the "invisible" boundaries created by the detection targets.

The unexpected results are not likely the consequences of the paradigm. We propose instead that these seemingly puzzling results can be accounted for under the biased competition theory (Desimone & Duncan, 1995) and the guided activation theory of cognitive control (Miller & Cohen, 2001). Furthermore, the eye movement response observed in the present experiment is especially pertinent to the recurrent themes regarding the role of *task* and *graded* activations. As the consequence of including an additional detection task, perceptual sampling processes involved in pursuit, as suggested earlier, were largely compromised, since the detection target had to be attended to all the time in the catch trials. Thus it can be naturally assumed that under such conditions, pursuit eye movements should predominantly rely on the mental representation of target motion instead of perceptual feedback. Furthermore, as suggested by some psychophysics research (e.g. Hutton & Tetally, 2005; Kerzel et al., 2008), the employment of a concurrent secondary task in the periphery of pursuit tended to cause the pursuit task to become more demanding, which was typically evidenced by reductions in eye velocity. Therefore it can be speculated that under the influence of external or contextual interference, whatever drives pursuit eye response should receive an extra boost from the oculomotor task, which dictates that smooth eye movements in reaction to a moving stimulus must be maintained. This means that in the case of the current study, the internal representation of target motion was more depended upon in order to maintain smooth eye responses, and its activation strength was most likely heightened by task demands, which might be comparable to the case of target extinction. As a result, downward verbs caused eye acceleration regardless of the location of the detection target, as these verbs shared the same directional features with the sole representation that was critical for the oculomotor task. However, any potential impairment produced by the upward verbs might have been masked by the overpowering representation of downward motion and not expressed behaviourally.

Although the current study is comparable to the extinction study (Experiment 5) in terms of the role played by the internal representation of target motion, the results from these two experiments are different: In the extinction experiment, downward verbs did not affect pursuit eye velocity while upward verbs decreased eye velocity. However, in the current experiment, regardless of the detection target location, downward verbs caused eye acceleration while upwards did not influence pursuit. This raises the question of why distinct patterns of results were revealed given that the internally represented target motion was the predominant driving force for pursuit eye movements in both studies. The answer lies in the crucial difference between these two experiments: The pursuit target was absent during extinction whereas it stayed visible all the time in the present experiment. When the pursuit target was absent, it was exceedingly difficult for the system to generate any eye acceleration, even when representations activated by language were in agreement with the demands of the oculomotor task.[63] On the other hand, if the pursuit target remained visible all the time, such as in the current experiment, eye acceleration was possible, especially given that language-activated motion representations shared the same directionality as relevant representations for the pursuit task.

---

[63] See the Discussion section in Experiment 5, Chapter 3 for more details of this account.

Another issue remains for our theoretical account: The sampling rate of target motion seemed only to be compromised by the detection task but not by the auditorily presented linguistic stimuli (i.e. the words), given that both the detection task and the words could be potentially “distracting” for the perceptual sampling of the target. The answer for this issue lies in the role played by *task*. The presence of the detection target was only able to impair the perceptual sampling during pursuit and force the system to rely on the internal representation because it was associated with an explicit task (i.e. to press a button when a colour change was detected). In other words, the detection target could only compromise the perceptual sampling of the pursuit target when there was an associated explicit task. However, since the auditory words were not associated with any explicit task, the influence of these words was only reflected in *biases* in external behaviour, instead of giving rise to any changes within the control mechanism of pursuit.

In conclusion, the present experiment failed to replicate the leading vs. lagging effects and generated results that were partially contradictory to our predictions. The multiple-dot paradigm was not an effective method to artificially simulate the attentional state created as a consequence of the gaze position being above or below the pursuit target. This is presumably for the reason that the detection target was associated with an explicit task, which could compete with the oculomotor task and potentially modulate the control of the eye movement. Despite the surprising results, the speculations based on the data collected in the current experiment point once again at the importance of task and are consonant with the attentional framework used to account for our previous findings.

Chapter 6. The effect of verb and noun semantics on saccadic eye movements

Studies reported in previous empirical chapters have demonstrated the influence of verb semantics on pursuit eye movements under a range of experimental manipulations, as well as having hinted at the potential semantic effects on smooth pursuit mediated by nouns. The experiment described in this chapter asks whether verb or noun semantics can affect *saccadic* eye movements. Subjects made saccadic eye movements in response to a stationary target presented either above or below the fixation location, after listening to either verbs implying upward or downward motion, or nouns referring to objects that typically appear in high or low locations. Two saccadic parameters, launch time and landing position, were measured as dependent variables. The results will be discussed in relation to our previous findings from the pursuit experiments. Finally, a model of saccade generation will be evaluated in light of the results from the present study.

Introduction

Compared to smooth pursuit, a considerately larger amount of research has been devoted to demonstrating and accounting for cognitive influences on saccadic eye movements. This contrast between these two types of eye movements has arisen for two reasons: First, saccades are relatively more under deliberate control and can be initiated voluntarily in the absence of an external visual stimulus. Furthermore, from a functional point of view, saccades are the main type of eye movements that we use to explore the visual environment, guide our other bodily movements and even conveying our intentions in social situations. One line of evidence illustrating high-level cognitive impact on saccadic eye movements has come from research on psycholinguistics. Using the *visual-world* paradigm, it has been established that linguistic information can direct saccades to specific objects in a visual display (e.g. Tanenhaus et al., 1995) or alter patterns of saccades across objects and the space these objects occupy (e.g. Richardson et al., 2007). These demonstrations of saccadic modulation mediated by language processing have been taken as primary evidence by the grounded cognition approach to support the assumption that language comprehension is grounded in mechanisms responsible for perception and action.[64]

As discussed earlier (i.e. in Chapter 1), the semantic effects on saccadic eye movements manifested through the visual-world paradigm may essentially reflect the interaction between linguistic meaning and visual semantics, rather than the direct impact of language processing on the motoric system responsible for saccade generation. We have since demonstrated that pursuit eye velocity can be modulated by single word meaning when eye movements are confined in a visual environment completely devoid of semantics, and moreover, when the oculomotor task itself (i.e. pursuing a moving target) can be interpreted and achieved in the absence of semantic processing. This has led to the question of whether similar types of modulating force originated from language comprehension can be revealed in saccadic eye movements using a comparable paradigm, in which there exists no semantic relationship between the visual and the linguistic stimulus.

Based on the accumulated research on saccades, it is conceivable that language processing can directly modulate saccadic eye movements: First of all, a close functional relationship between saccades and attention has been proposed. For example, the *premotor theory* of attention (Rizzolatti et al., 1994) argues that saccadic planning, but not execution, is equivalent to shifting spatial attention to a specific location. According to this theory, the programming of an eye movement enhances the activity of certain neurons, which function to construct topographic “pragmatic maps” that convert spatial information to oculomotor responses. The orientation of spatial attention to specific locations is the consequence of this increase in activation strength in certain regions of these pragmatic maps. In support of this model, Rizzolatti and colleagues (1987) instructed subjects to fixate centrally and presented them with several boxes in the periphery. When one of these boxes was cued, attention was shifted to that location and a response was required as soon as a target appeared. Invalid trials were also included, in which the target was presented at a location that was not cued. It was found that reaction latencies were longer when the target appeared in the opposite hemifield to the cue. These authors suggested that this

---

[64] See the section entitled “The visual-world paradigm”, Chapter 1, for a more detail description of related studies and an evaluation of the paradigm.

phenomenon, termed the *meridian effect*, arose because a planned saccade in response to the cue in the opposite hemifield had to undergo a change in direction, which was more time consuming compared to simply adjusting a saccade programme when the cue and the target appeared in the same hemifield.

In contrast to the premotor theory, which argues that saccadic planning causes shifts in attention, the Visual Attention Model (VAM) suggested by Schneider (1995) proposes that the programming of a saccade may be the *consequence* of a shift in attention. Deubel and Schneider (1996) reported that subjects were unable to perform a discrimination task presented at a specific location unless a saccade was about to be launched to the same location. Along similar lines, Hoffman and Subramaniam (1995) found that when subjects were instructed to attend to one location while making a saccade to another, discrimination performance was superior at the saccade destination instead of the attended location. These findings suggest that saccades to a specific location have to be preceded by an attentional shift to that location.

Attention plays an essential role in the framework of the biased competition theory (Desimone & Duncan, 1995) and the guided activation theory of cognitive control (Miller & Cohen, 2001). When more than one representation or processing pathway is active within the system, attention regulates the competition between these simultaneously activated representations and thus modulates the behavioural outcome. Given the close relationship between saccades and attention, and our previous findings in pursuit eye movements, it is unlikely that language processing should not affect saccadic eye movements, especially when considering the assumption that attention functions at a system level that encompasses all aspects involved in perception, action and cognition (Cohen et al., 2004).

The study reported here aimed at testing the prediction that the meaning of single words can modulate saccadic eye movements in the absence of visual semantics. Subjects were required to saccade from a central fixation point to a target presented either above or below the fixation location. However, prior to the onset of the saccade target, directional verbs that implied either upward or downward motion, and locational nouns that conveyed either high or low locations were presented auditorily. Based on previous evidence, we predicted that both verbs and nouns would affect saccades launched to the target and the semantic effect might be manifested in either saccadic launch latency or landing position, which were both measured in the present experiment.

Method

Participants

32 students took part in this experiment in exchange for either £4 or one course credit. All subjects were native speakers of English and had normal or corrected-to-normal vision.

Materials

The linguistic (i.e. directional verbs such as "*dive*" and "*climb*, and locational nouns such as "*attic*" and "*basement*") and visual stimuli were identical to the ones used in Experiment 7 (Chapter 4).

Procedure

At the start of every trial, subjects were required to fixate on a cross displayed centrally. 200 ms later, a word was presented through loudspeakers and subjects were instructed to maintain their fixation during the entire acoustic life of the word. At word offset, a dot appeared at a location 7° above or below the fixation cross and subjects were required to launch a saccade to this dot as quickly as possible. The dot stayed on the screen for 1500 ms, followed by a blank screen for 500 ms that led to the termination of the trial.

Design

All subjects were presented with all 24 directional verbs and their 24 matched controls, and all 24 locational nouns and their 24 matched controls. Half of these experimental items and their controls were paired with targets presented above the fixation cross, while the other half with targets displayed below the fixation. In addition, there were 96 filler words, which were also paired with targets either above or below the fixation point.

Results

a. Launch latency

Only the first saccade in the direction of the target after word offset was included in the analyses when launch latency was calculated as the dependent variable. In the current experiment, launch latency referred to the period between word offset/target onset and when the eyes started to move.

As indicated by Figure 6.1 and 6.2, for downward saccades (i.e. when the target appeared below the fixation location), directional verbs had a significant effect on saccadic launch latencies (F (3, 93) = 3.26, $p < .05$) Relative to their respective controls, downward verbs did not modulate saccadic latency (182 vs. 184, F (1, 31) < 1, $p > .05$) while upward verbs increased the latency (191 vs. 178, F (1, 31) = 12.15, $p = .001$).[65] Analyses revealed that locational nouns also significantly modified launch latency (F (3, 93) = 3.80, $p < .05$). Launch latency was shorter after hearing nouns implying high locations compared to their controls (188 vs. 207, F (1, 31) = 13.90, $p = .001$) while nouns implying low locations did not have an effect (197 vs. 198, F (1, 31) <1, $p > .05$).

For upward saccades (i.e. when the target appeared above the fixation cross), there was no main effect of directional verbs (F (3, 93) = 2.31, $p > .05$). However, pair-wise comparisons revealed that while downward verbs did not affect launch latency (172 vs. 169, F (1, 31) < 1, $p > .05$), upward verbs marginally decreased the latency relative to their controls (161 vs. 169, F (1, 31) = 4.17, $p = .053$). Regarding the locational nouns, there was again no main effect of word type (F (2.28, 70.83) < 1, $p > .05$).[66] Furthermore, although the high nouns did not have any influence on the launch latency (171 vs. 174, F (1, 31) < 1, $p > .05$), the low nouns increased the latency compared to their matched controls (174 vs. 168, F (1, 31) = 4.81, $p < .05$).

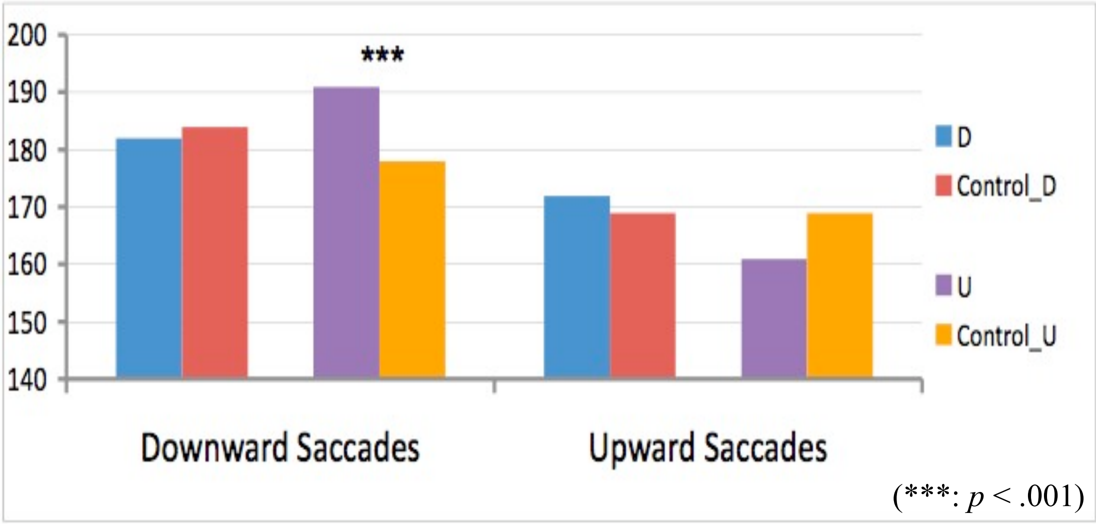
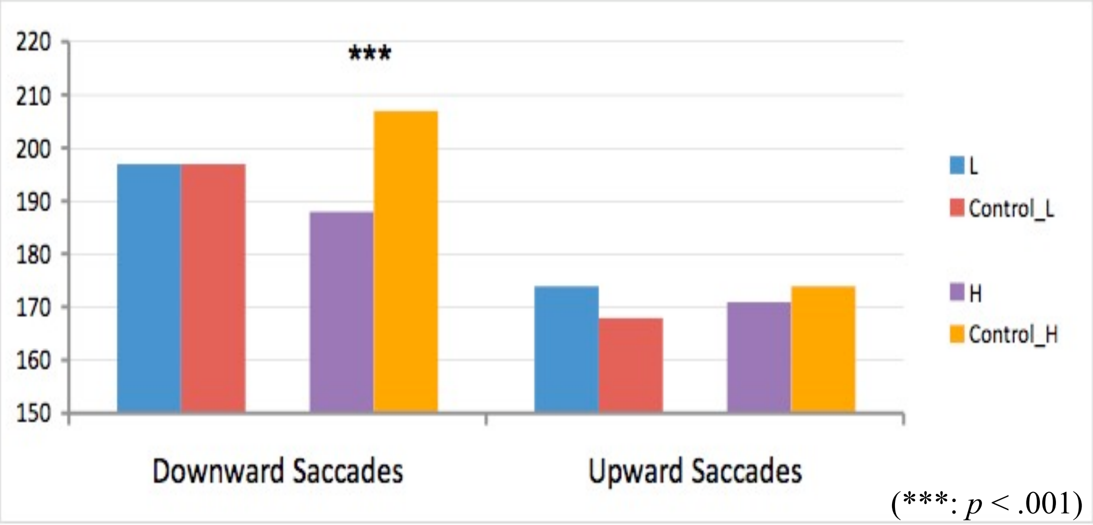


Fig. 6.1. *Launch latency for both downward and upward saccades in verb conditions.*

---

[65] Saccadic launch latency was measured in milliseconds.

[66] When the assumption of sphericity was violated, the Huynh-Feldt correction was applied.



(***: $p < .001$)

Fig. 6.2. *Launch latency for both downward and upward saccades in noun conditions.*

Additional analyses (Figure 6.3) indicated that there was no difference between the control conditions for the downward verbs and the upward verbs (184 vs. 178, t (31) = 2.37, $p > .05$), neither was there any difference between the control conditions for the high nouns and the low nouns (207 vs. 197, t (31) = 2.68, $p > .05$). However, the eyes were faster to saccade upward rather than downward (174 vs. 197, t (31) = 19.00, $p < .001$). When the data were re-analyzed according to congruency (i.e. whether the word implied the same direction/location as the saccadic direction/landing position), there was a main effect of condition for the directional verbs (F (2, 126) = 5.19, $p < .01$). Congruent verbs did not affect the launch latency (171 vs. 175, F (1, 63) = 1.33, $p > .05$) while incongruent verbs generally increased the latency in contrast to the control items (181 vs. 175, F (1, 63) = 4.11, $p < .05$). However, no main effect of condition was found for locational nouns (F (1.56, 98.36) < 1, $p > .05$). Furthermore, pair-wise contrasts revealed no difference between the congruent and the incongruent nouns and their controls (184 vs. 186, F (1, 63) < 1, $p > .05$; 181 vs. 186, F (1, 63) = 2.97, $p > .05$).

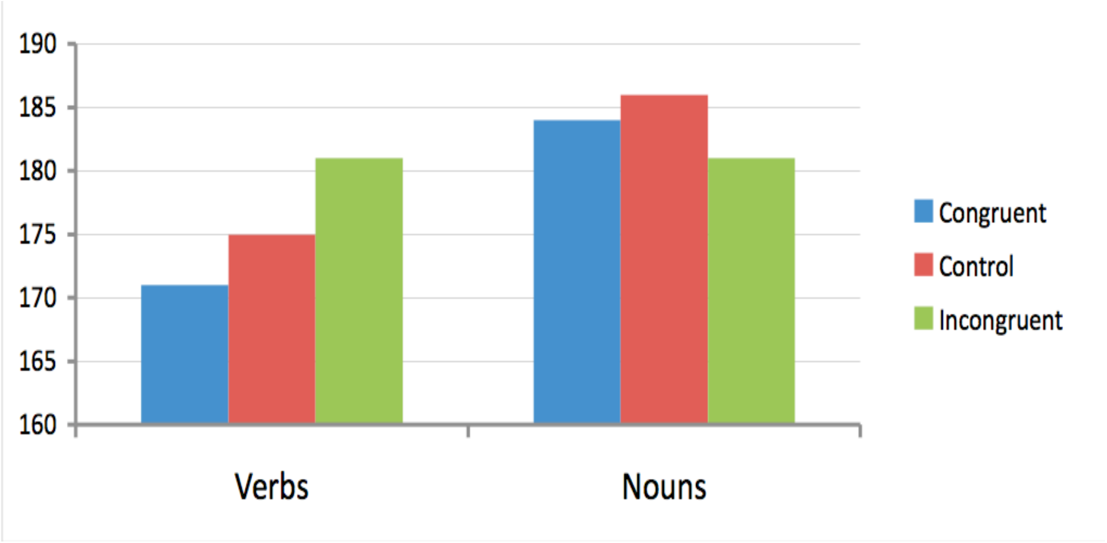


Fig. 6.3. *Launch latency for both verbs and nouns across three conditions: Congruent, control and incongruent.*

To summarise the launch latency results: Regarding downward saccades, upward verbs increased the latency while downward verbs did not have an effect. For upward saccades, upward verbs marginally decreased the latency whereas downward verbs did not influence the latency.[67] For the locational nouns, high nouns decreased the latency for downward saccades but they did not affect upward saccades. On the other hand, low nouns increased the latency for upward saccades but not for downward saccades.[68] Although distinct control conditions were employed for separate word groups, the differences found in saccadic latency were unlikely to be due to the variation between these control conditions. Finally, when only word type (noun vs. verb) and congruency were considered as the only independent variables, incongruent verbs decreased the saccadic latency while the rest of the word groups failed to have any effect.

b. Landing position

Saccadic landing position in the present experiment was defined as the distance between the fixation location and the end location of the saccade in the vertical dimension. Analyses reported below involved the landing positions of the first and the second saccades launched after the word offset in the direction

---

[67] Although there was a difference in planned pair-wise comparisons, no main effect of word type was found.

[68] No main effect of word type was found.

of the target. Landing positions of the second saccades were included in the analyses for the reason that the first saccades might not always direct the eyes to the location of the target (Becker, 1989).

For downward saccades, there was no main effect of verb semantics on the first (F (2.43, 75.44) = 1.15, p > .05) or the second saccade launched (F (1.51, 46.72) < 1, p > .05). Similarly, no main effect on downward saccades was found for noun semantics on the first (F (1.50, 46.63) <1, p > .05) or the second saccade (F (1.47, 45.40) < 1, p > .05).

	First	**Second**
D vs. Control_D	8.19 vs. 7.98 F (1, 31) = 1.92, p > .05)	8.38 vs. 7.61 F (1, 31) = 1.03, p > .05
U vs. Control_U	8.22 vs. 7.94 F (1, 31) = 2.08, p > .05	7.79 vs. 7.84 F (1, 31) < 1, p > .05
L vs. Control_L	7.84 vs. 7.60 F (1, 31) < 1, p > .05	8.33 vs. 7.26 F (1, 31) < 1, p > .05
H vs. Control_H	8.14 vs. 7.99 F (1, 31) = 1.02, p > .05	8.33 vs. 8.17 F (1, 31) < 1, p > .05

Table 6.1. *Planned pair-wise comparisons for downward saccades. In this table, "D" and "U" stand for downward and upward verbs, respectively, while "L" and "H" stand for low and high nouns.*[69]

Regarding upward saccades, no semantic effect was revealed regardless of the word type or whether it was the first or the second saccade. The directional verbs did not modulate the landing position for upward saccades (F (3, 93) < 1, p > .05; F (2.57, 79.72) = 1.51, p > .05). Furthermore, no semantic effect was found for the locational nouns (F (3, 93) = 2.51, p > .05; F (3, 93) < 1, p > .05).

---

[69] The unit of landing positions was degree (°).

	First	**Second**
D vs. Control_D	7.71 vs. 7.60 F (1, 31) < 1, p > .05	8.26 vs. 8.31 F (1, 31) < 1, p > .05
U vs. Control_U	8.22 vs. 8.15 F (1, 31) < 1, p > .05	8.22 vs. 8.15 F (1, 31) < 1, p > .05
L vs. Control_L	7.62 vs. 7.61 F (1, 31) < 1, p > .05	8.30 vs. 8.27 F (1, 31) < 1, p > .05
H vs. Control_H	7.49 vs. 7.27 F (1, 31) = 2.12, p > .05	8.31 vs. 8.15 F (1, 31) = 2.72, p > .05

Table 6.2. *Planned pair-wise comparisons for upward saccades. In this table, "D" and "U" stand for downward and upward verbs, respectively, while "L" and "H" stand for low and high nouns.*

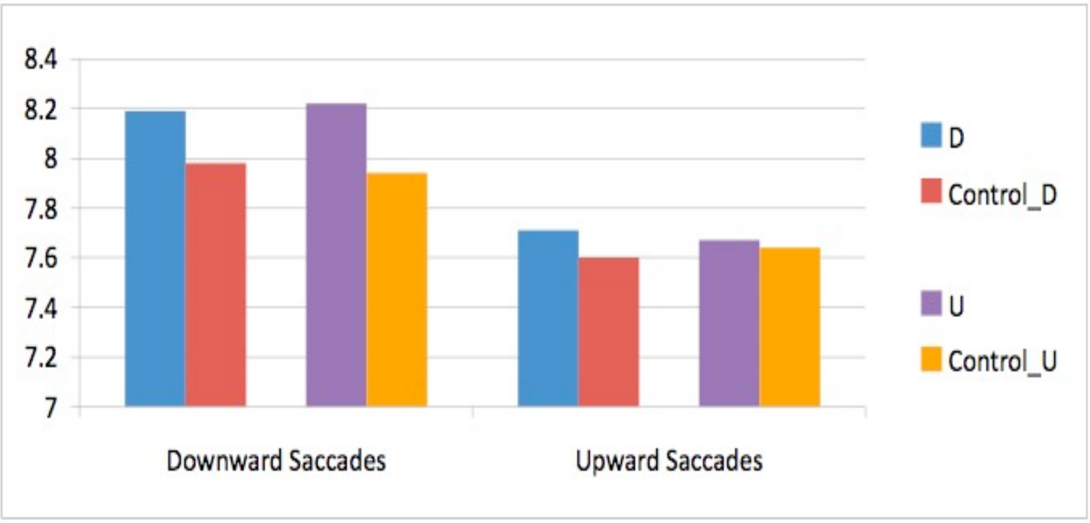


Fig.6.4. *Landing position of the first saccade launched for both downward and upward saccades in verb conditions.*

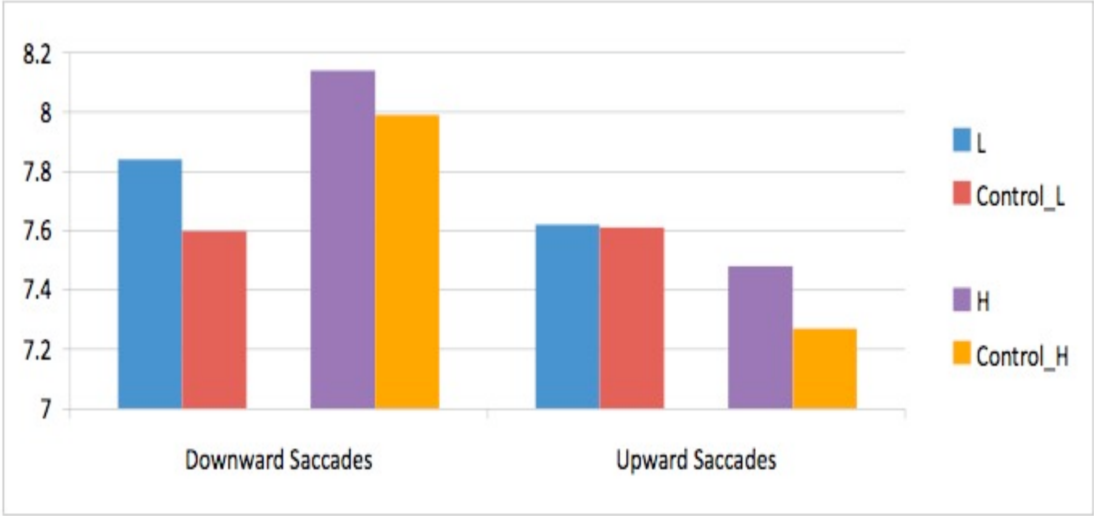


Fig.6.5. *Landing position of the first saccade launched for both downward and upward saccades in noun conditions.*

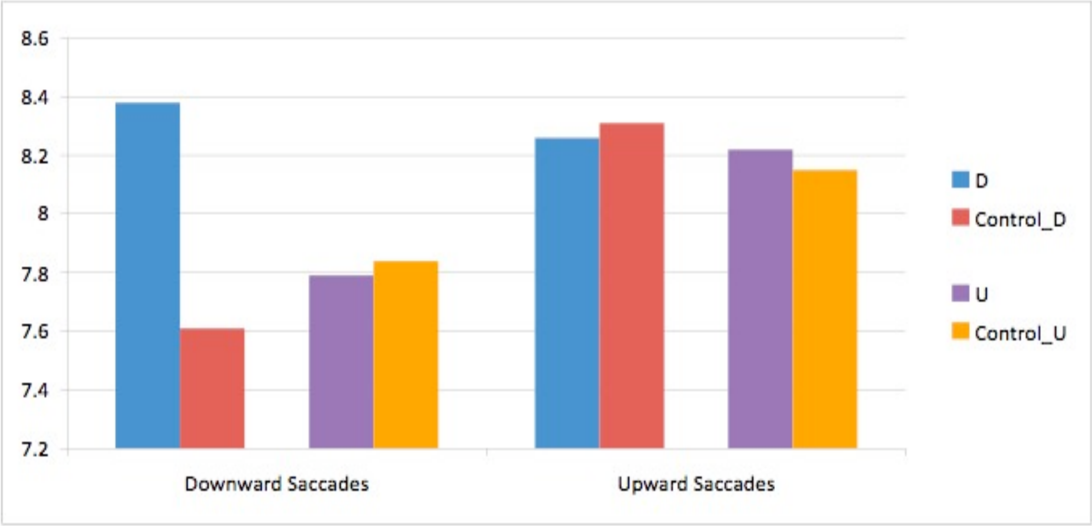


Fig.6.6. *Landing position of the second saccade launched for both downward and upward saccades in verb conditions.*

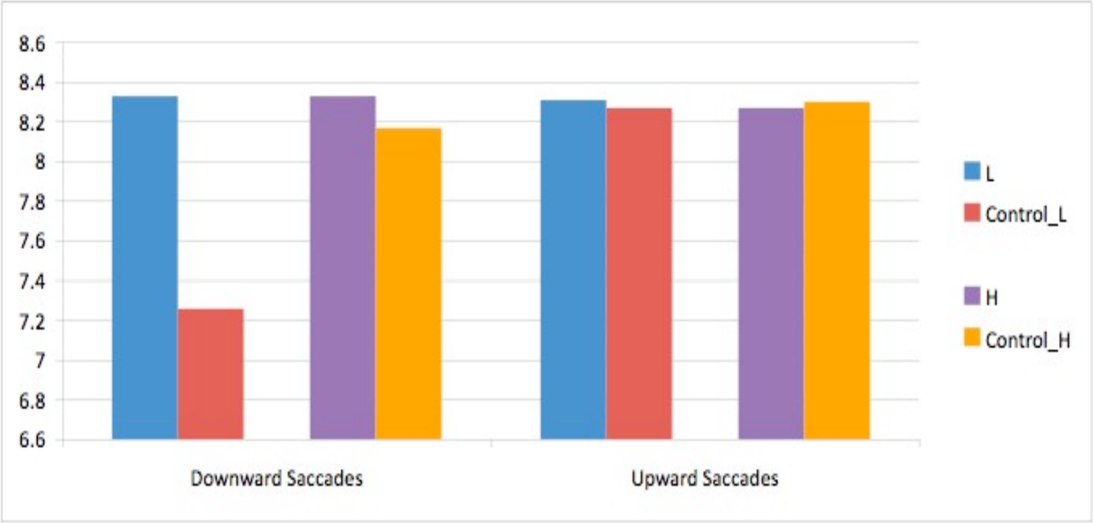


Fig.6.7. *Landing position of the second saccade launched for both downward and upward saccades in noun conditions.*

Additional analyses suggested that there was no difference between the control conditions for the downward and the upward verbs when the first (7. 98 vs. 7.94, *t* (31) < 1, *p* > .05) or the second saccade was concerned (7.61 vs. 7.84, *t* (31) < 1, *p* > .05). Similarly, the separate control conditions for the high and the low nouns did not differ from each other (7.99 vs. 7.60, *t* (31) = 1.37, *p* > .05; 8.17 vs. 7.26, *t* (31) < 1, *p* > .05). When the direction of the saccades was considered, the landing positions for downward and upward saccades were generally identical (7.61vs. 8.31, *t* (31) < 1, *p* > .05).

In sum, no semantic effect on saccadic landing position was revealed regardless of the word type. This was the case for both the first and the second saccade launched after word offset.

Discussion

These results suggested that saccadic launch latency was modulated by upward but not by downward verbs. When saccades were directed upward, upward verbs decreased the launch latency. However, upward verbs increased the latency of downward saccades. The results for the locational nouns were relatively more complicated: Nouns implying high spatial locations decreased the latency of downward saccades while nouns suggesting low locations increased the latency of upward saccades. Furthermore, for saccadic landing positions, no semantic effect from either the verbs or the nouns was found. The verb and noun related results are discussed in turn.

First of all, the facilitation/impairment observed in saccadic launch latency caused by upward verbs can be accounted for using the biased competition theory (Desimone & Duncan, 1995) and the guided activation theory (Miller & Cohen, 2001), which have also been adopted to explain the semantic effect on pursuit eye movements. When an upward verb was being processed, a representation of upward motion, as well as the space above the fixation location was activated by the directional semantics carried by the verb. After the target was presented, the target itself and its location had to be encoded and represented in order to successfully programme and launch a saccade towards it. If the same part in space was represented by both the comprehension of the verb and the oculomotor task (i.e. when the target appeared above the fixation), an increase in activation strength was received by the representation required by the saccadic task and the latency to initiate an eye movement was consequently shortened. However, competition would be created if opposite parts in space were activated by the verb and the saccadic task (i.e. when the target was presented below the fixation) and the competition between the representations activated by language comprehension and the representations dictated by the oculomotor task hence led to an increase in saccadic latency.

It was puzzling that no semantic effect on saccadic latency was found for the downward verbs. The explanation could lie in the *uniqueness point* of these downward verbs. For a spoken word (e.g. *beaker*), it sometimes shares the first few phonemes with other words (e.g. *beetle*). The uniqueness point of a word is defined as the phoneme in the word where it diverges from all other words in the language. The uniqueness point can affect the speed of spoken word recognition (Marslen-Wilson, 1990). Subjects were faster to decide whether a spoken item was a real word or a non-word if the uniqueness point was early in the word compared to when it was late. In our experimental manipulation, the saccadic target always appeared at the word offset. Thus it was plausible that the subjects would learn to use the word offset as a pre-cue for the upcoming target. If there was less variance in the position of the uniqueness point for downward verbs compared to upward verbs, the downward verbs would be unambiguously recognised faster and the offset of the downward verbs would be anticipated earlier and more accurately. Therefore the onset of the saccadic target was more likely to be precisely predicted after hearing a downward verb compared to an upward verb. This increased predictability of target onset in the downward verb condition might have interfered with the effect of verb semantics and led to the null results observed. However, post-hoc analyses on the variances in the positions of uniqueness point for downward and upward verbs did not reveal a significant difference and excluded this possibility.[70] It is worth noting at this point, nonetheless, that the null effects from the downward verbs do not necessarily indicate that these verbs cannot affect eye movements, especially given that the semantic effect on pursuit caused by the same group of verbs is robust. At the moment, however, it remains undetermined for what reasons the downward verbs have failed to modulate saccadic launch latency in the present study.

Regarding the locational nouns, the increased latency caused by the low nouns for upward saccades can be accounted for under similar principles as used for the upward verbs: When a low noun was being processed, a representation of the denoted objects was activated, as well as the space below the fixation point. As a result, the representation activated by language was in conflict with the representation demanded by the programming of an upward saccade, which contained the encoded information about the space above the fixation point. The increased latency reflected the cost of this conflict. On the other hand, it was surprising that the high nouns *decreased* the latency of downward saccades, as the reverse pattern was expected. Based on current data, an explanation was not possible and further research will required to further elucidate this observation.

No semantic effect on saccadic landing position was found for either the verbs or the nouns. A possible reason for these null results may lie in the experimental manipulation: The distance between the saccadic target and the fixation location was constant throughout the experiment (i.e. 7°). Thus the distance the eyes had to travel was already planned before the target onset and did not need to be adjusted from trial to trial. The representation activated by language, which was relatively general (i.e. the space below or above the fixation location) and varied from trial to trial, might not be able to compete with the

---

[70] An F-test for variances was carried out and no difference was found between the downward the upward verbs F (11) = 1.02, $p > .05$.

specific and well-learnt representation (i.e. 7° above or below the fixation location) required by the oculomotor task. Thus saccadic landing position was not affected by word semantics. This problem can be resolved in future research by varying the distance between the target and fixation location from trial to trial.

The possible explanation for these landing position results reflects a major distinction between saccades and pursuit: Saccadic eye movements are ballistic and cannot be modified by new information that arrives less than 70 ms before the onset of the movement (Becker & Jürgens, 1979). On the other hand, pursuit is controlled *in real time* by both perceptual and internally stored information. Eye velocity and movement trajectory can be modified by various factors during ongoing pursuit with short response latency. Given these distinct characteristics of saccades and pursuit, pursuit eye movements may be considered as a more appropriate behavioural measure compared to saccades when general issues, such as whether cognitive functions can affect sensorimotor responses, are concerned. One of the major advantages of using pursuit as the dependent measure is that the timing of the experimental stimulus is less likely to become a confounding factor. This is especially the case when the experimental manipulation is language, which is complex as a stimulus and the temporal course of language processing is difficult to control (cf. the earlier discussion of uniqueness point as a confounding variable). As a result, saccades are not an effective behavioural measure, at least not for studying the language effect on motoric responses in *real time*. However, pursuit eye movements provide a solution for this problem, as the experimental outcome will not be influenced so long as all the experimental stimuli are presented during pursuit maintenance.

Despite its disadvantages, it is nevertheless important to demonstrate the effect of cognitive functions on saccadic eye movements. First, the grounded cognition approach (e.g. Barsalou, 1999) assumes that cognitive processing shares the same representational or neural substrates as perception and action. Given that saccades are one of the major types of bodily movements used to explore and interact with the external environment, the demonstration of cognitive influence on saccadic eye movements provides fundamental support for the assumptions proposed by the grounded cognition framework. Second, the guided activation theory (Miller & Cohen, 2001) argues that attention modulates the competition between concurrent representations and pathways and functions at a system level. Based on this assumption, there is no reason why language comprehension should only affect pursuit but not saccadic eye movements. Thus evidence of semantic effects on the saccadic system would also provide essential support for the system-level hypothesis. Finally, a long-standing debate within eye movement research is about whether pursuit and saccades are controlled by disparate mechanisms (Krauzlis, 2005). The more recent view on this issue considers pursuit and saccades to have a similar functional structure and the dissimilarities between these two types of eye movements reflect different outcomes of the same system, rather than the presence of independent mechanisms of control (Krauzlis, 2003). The fact that both pursuit and saccades are susceptible to the influence of language comprehension provides support for this view.

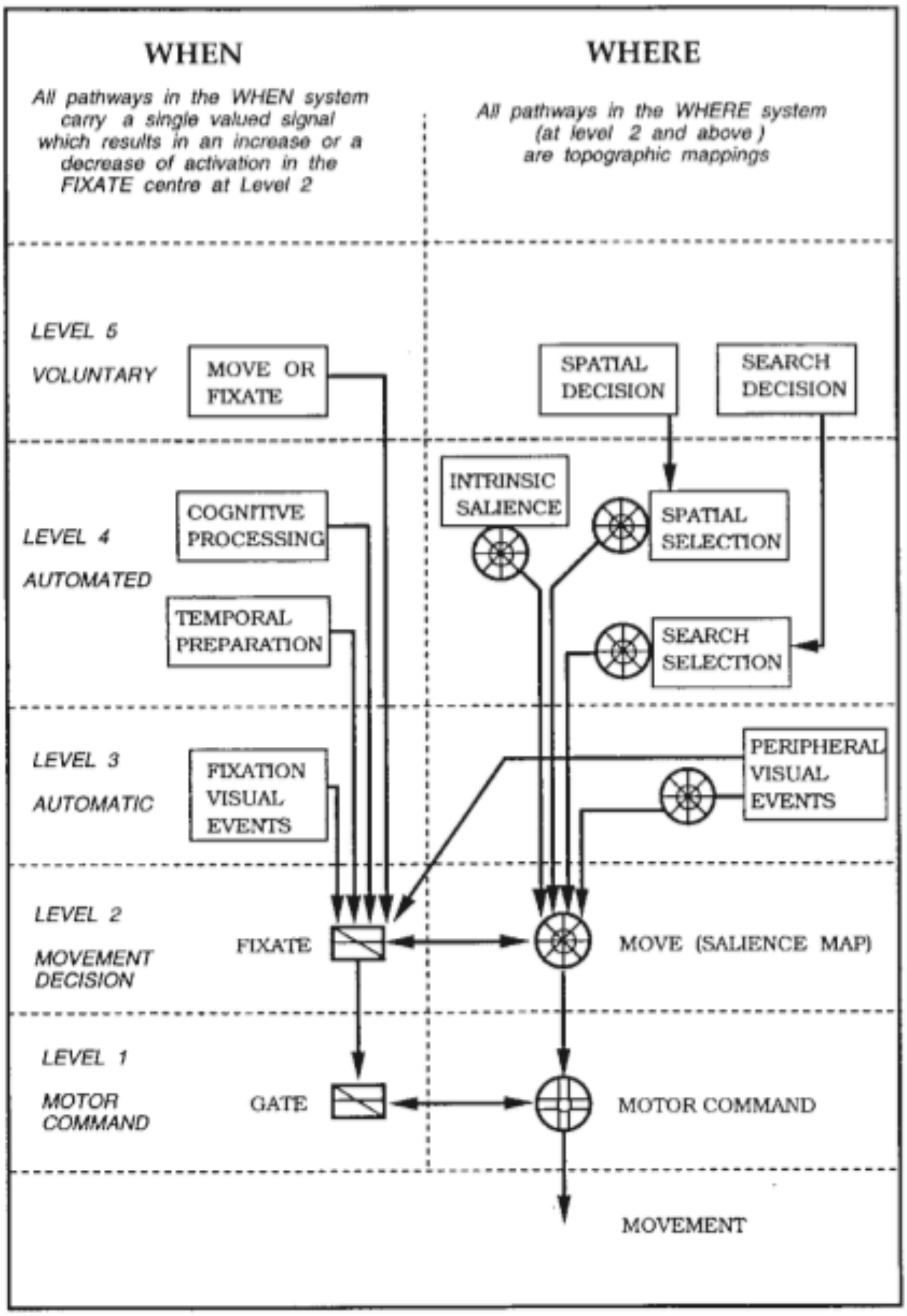


Fig.6.8. *The model of saccade generation. Reproduced from Findlay & Walker (1999).*

Unlike pursuit, cognitive influences are more explicitly incorporated into models of saccadic eye movements. For example, a model proposed by Findlay and Walker (1999) contains a component that deals with cognitive processing involved in saccade generation (Figure 6.10). This framework takes into account both the functional properties and the underlying physiological mechanisms of saccades. There are two separate pathways responsible for processing WHERE and WHEN information. The WHERE pathway is a distributed network that aims at constructing a saliency map in which the saccadic target location is selected. The product of the WHEN pathway is a single signal with varied activity strength that simply determines when the eyes should start moving. Horizontally, each band represents a processing level and interactions between the two vertical streams only occur in the lower levels, which are assumed to be less susceptible to voluntary control. Level 4 is where in the model the influence from cognitive processing comes into processes behind saccade generation. There are several problems with this arrangement: First of all, according to this model, cognitive processing can only influence saccadic eye movements through the WHEN but not the WHERE pathway (see Figure 6.10), as cognitive processing is confined in Level 4 of the WHEN pathway and there is no direct connection between the two vertical pathways at this level. This is unlikely to be the case given that evidence suggests cognitive processing can affect saccadic landing position and trajectory (e.g. Calvo, Nummenmaa, & Hyönä, 2008; Langton & Bruce, 1999). Alternatively, it is possible that some components within the WHERE pathway are under cognitive influences themselves (e.g. search decision). Thus cognitive processing should be involved in more levels than one, and more connections should be drawn between the higher levels. This is reflected in the data of the current experiment: The spatial information carried by the word semantics interacted with the spatial processing (WHERE) involved in saccade generation. These interactions then in turn modulated saccadic launch latency (WHEN).

In conclusion, we have demonstrated that the meaning of single words can affect saccadic launch latency but not landing position. These results can be partially accounted for under the principles of the guided activation theory. However, further research is needed in order to elucidate fully the interactions between language comprehension and saccadic eye movements.

Chapter 7. General discussion: The system has its own solution

Background and aims

The primary aim of this thesis was to address two fundamental issues: First, whether cognitive functions could influence sensorimotor responses; and if there was any interaction between cognition and sensorimotor responses, how the interaction was achieved. More specifically, this thesis focused on the question of whether the meaning of single words could affect pursuit or saccadic eye movements in situations where the visual environment did not contain any semantic information and no connection could be readily made based on semantics between the linguistic stimulus, the visual stimulus and the oculomotor task.

The grounded cognition approach to language comprehension (e.g. Barsalou, 1999; 2007; Glenberg, 1997) has proposed that language processing relies on the same set of mechanisms responsible for perception and action. This theoretical framework suggested simulation as the key process behind language comprehension. Simulation refers to the reenactment of past perceptual or motoric experiences, which is achieved through reactivating parts of the system that have been involved in the actual perceptual or motoric experiences. Thus processes behind the comprehension of a word or a sentence are composed fundamentally of activations of features of the objects referred to or information about the events described. Empirical studies using neuroscientific methods have shown that language comprehension can generate the same patterns of brain activations as the actual sensorimotor processing of the object or action denoted by the linguistic stimuli (e.g. Hauk et al., 2004; Martin, 2007). Meanwhile, behavioural research suggests that language comprehension can influence perceptual processing or the generation of motoric actions (e.g. Estes et al., 2008; Glenberg & Kaschak, 2002; Meteyard et al., 2007). Despite being supported by abundant neuroscientific and behavioural evidence, two major issues remain for the grounded view on language comprehension: First, although language comprehension recruits the same cortical regions as perceiving a stimulus or performing an action in reality, language processing does not necessarily equate to perception and action. For example, the word "*kick*" activates the brain areas responsible for controlling foot-related actions (Hauk et al., 2004), nonetheless, a kick action is not always generated every time when the word "*kick*" is heard. Second, it remains unclear how language-induced simulations can affect perceptual or motoric processing. There is a missing link between simulations as the product of language comprehension and their impact on behavioural outcomes. For instance, the grounded cognition approach is supported by evidence collected using the visual-world paradigm (Cooper, 1974), with which studies have demonstrated that language can influence saccadic eye movements (e.g. Chambers et al., 2004; Spivey et al., 2000), the epitome of the interaction between perception and action. However, it is still undetermined how the sensorimotor representations activated by language are translated into different eye movement patterns, and whether the interaction between language and eye movements revealed by the visual-world paradigm simply reflects an association between linguistic meaning and visual semantics.

In order to resolve these issues, we introduced the concept of attention into the grounded cognition framework. The functional principles of attention have been adopted from the biased competition theory (Desimone & Duncan, 1995) and the guided activation theory of cognitive control (Miller & Cohen, 2001): Information processing in the cortical system is competitive. At any given moment, there are always multiple representations or processing pathways that are simultaneously active and compete for behavioural expression. Attention acts to bias these competitions by modulating the relative activation strengths of the concurrently active representations. Thus the more favoured representation will have a more significant impact on the to-be-generated behavioural outcome. According to Cohen and colleagues (2004), attention is no longer defined and constrained by a qualitatively different mechanism; but instead, it arises from the functional principles of the system, which encompasses perception, action and cognition.

Experiments presented in this thesis attempted to establish the link through attention, which functions at a system level, between semantic representations activated by language comprehension and their consequent impact on sensorimotor responses. The effect of single word semantics on eye movements was adopted as the specific experimental manipulation and behavioural measure. The question of whether verbs implying upward or downward motion could influence pursuit response to a vertically moving stimulus was first dealt with (Chapter 2). Following the establishment of the verb semantic effect on pursuit, the focus was shifted to the role of the internally represented target motion in the interaction between language and eye movements (Chapter 3). The involvement of the visual motion used to induce pursuit eye movements was either attenuated or completely extinguished to test whether the original semantic effect persisted. The subsequent chapter (Chapter 4) continued to investigate to what extent the semantic effect could be generalised by changing the dimension of the visual motion or altering the type of words used. A new paradigm was then introduced and evaluated to answer the question of whether biases in attention could be artificially created during pursuit (Chapter 5). Finally, another type of eye movements other than pursuit (i.e. saccades) was examined under the influence of language comprehension (Chapter 6). The results of these experiments are summarised in the following section.

Overview of empirical findings

Both pursuit and saccadic eye movements were concerned in this thesis, although pursuit was the behavioural measure for the majority of the empirical studies (Experiment 1-8, Chapter 2-5). There are two commonalities among the findings related to pursuit eye movements: First, no semantic effect of either the verbs or the nouns was found during upward pursuit and there was consistently more variance in the set of data collected from upward pursuit compared to downward pursuit. Second, semantic effects reflected in smooth eye velocity were generally observed during the acoustic life of the word or shortly after the word offset.

Experiment 1-3 explored the question of whether directional semantics carried by verbs implying upward or downward motion (e.g. *climb*, *dive*) could affect vertical pursuit eye movements. Experiment 2 and 3 were replication studies for Experiment 1. Compared to their precedents, each of these replication studies had a more powerful design, and a set of control items that were more closely matched to the ones in the experimental conditions. The data collected in these three experiments indicated that verb semantics could influence the positional errors between gaze locations and the pursuit target through modulating eye velocity during pursuit. During downward pursuit, when the gaze position was ahead of the pursuit target, downward verbs caused eye deceleration while upward verbs led to eye acceleration. However, the reverse was true when the gaze position was behind the pursuit target: Downward verbs created eye acceleration while upward verbs caused the eyes to decelerate.

Experiment 4 and 5 focused on the potential semantic effect from the same set of vertical directional verbs on vertical pursuit when the velocity profile of the pursuit target was changed from sinusoidal to linear (Experiment 4) and when there was no visual motion available (Experiment 5). As revealed by Experiment 4, upward verbs increased the positional errors. However, neither type of verbs modulated eye velocity. In Experiment 5, upward verbs decreased the positional errors compared to downward verbs and the non-directional controls during the temporal disappearance of the pursuit target (i.e. extinction). This was caused by eye deceleration in response to the upward verbs. On the other hand, downward verbs did not modulate eye velocity during extinction.

The subsequent two experiments (i.e. Experiment 6 and 7) were mainly concerned with the generalisation of the original verb semantic effect. Experiment 6 tested whether the verb semantic effect would persist when pursuit eye movements were executed in the orthogonal dimension (i.e. horizontal) while the pursuit target moved vertically in Experiment 7 but the linguistic stimuli were changed from directional verbs to locational nouns (e.g. *attic*, *basement*). When vertical directional verbs were presented during horizontal pursuit (Experiment 6), these verbs did not influence positional errors occurred during horizontal pursuit and no eye velocity or displacement was induced in the vertical dimension. However, both downward and upward verbs decreased eye velocity in the horizontal dimension. Experiment 7 failed to replicate the verb semantic effect observed in Experiments 1 -3. Furthermore, although the low nouns (e.g. *basement*) decreased the positional errors, no influence on eye velocity was found from these nouns.

The last pursuit-related experiment, Experiment 8, introduced the multiple-dot paradigm in an attempt to artificially create samples in which covert attention was directed to a location above or below the gaze position. A pattern of five dots moved upward or downward on the computer screen. Subjects tracked the centre dot as the pursuit target, while attending to the dot located above or below the pursuit target in anticipation of a sudden colour change. Both button-pressing latency in response to the colour change and pursuit tracking performance were measured. Subjects performed on the colour-change detection task at a ceiling level. On the other hand, the eye movement data were comparatively complex for interpretation: Regardless of whether the dot above or below the pursuit target was attended to, upward verbs did not modulate pursuit response while downward verbs caused eye acceleration in both cases.

These eight pursuit experiments suggest that the semantic modulation of sinusoidal pursuit eye movements caused by vertical directional verbs is reliable and consistent (Experiment 1, 2, 3, 4 & 5). However, results yielded by the generalisation studies are more equivocal (Experiment 6 & 7). Finally, the new multiple-dot paradigm proposed in Experiment 8 did not prove to be an effective and reliable method to artificially simulate the attentional state during pursuit when the gaze position is ahead or behind the target.

The influence of directional verbs and locational nouns on saccadic eye movements was studied in Experiment 9, which was the only experiment in the current thesis that measured saccades as the primary dependent variable. When saccadic launch latency was concerned, upward verbs decreased the latency of upward saccades but increased the latency of downward saccades, while downward verbs did not affect saccade latency regardless of the saccadic direction. Both high (e.g. *attic*) and low nouns (e.g. *basement*) modulated launch latency: High nouns decreased the latency for downward saccades but did not affect upward saccades. On the other hand, low nouns decreased the latency for upward saccades but did not impact on downward saccades. Regarding saccadic landing positions, no semantic effect was found for either word type, irrespective of whether it was the first or second saccade launched.

Theoretical implications

Implications for language comprehension

The grounded cognition approach to language comprehension (e.g. Barsalou, 1999; 2007; Glenberg, 1997) suggests that language processing relies on the same representational and neural substrates as perception and action. Furthermore, language comprehension is achieved by simulating relevant perceptual or motoric experiences stored in memory. This means that during language comprehension, perceptual or motoric representations directly related to the objects or events described by language should be activated, and this prediction has been substantiated by empirical evidence (e.g. Estes et al., 2008; Glenberg & Kaschak, 2002; Meteyard et al., 2007; Zwaan et al., 2002). The current research not only provides additional evidence in support of the grounded cognition framework, it is also theoretically constraining for this framework in the following ways:

First, we have shown that semantic representations activated during language comprehension can interact with representations activated by a low-level sensorimotor task, which is largely independent from deliberate control. The uniqueness of our experimental paradigm and stimuli has led the empirical findings to be distinct in several ways: Compared to the visual-world studies (e.g. Chambers et al., 2004; Spivey et al., 2000), the visual environment in our experiments contained no semantics and the construction of connections between linguistic meaning and visual semantics is not possible. Furthermore, no explicit task related to the linguistic stimuli was given in these experiments and there was no need to interpret the oculomotor task semantically. This indicates that the interaction between language and eye movements demonstrated here truly reflects the direct impact of language comprehension on oculomotor control while excluding the possibility of taking an indirect semantic route. Compare to typical behavioural studies in the grounded cognition literature (e.g. Glenberg & Kaschak, 2002; Kaschak et al., 2005; Zwaan & Taylor, 2006), the required behavioural response was not directly or explicitly associated with the linguistic stimuli in our experiments. For example, in the study that established the *Action Compatibility Effect* (Glenberg & Kaschak, 2002), hand actions were generated in response to the sentences presented.[71] On the other hand, the pursuit or saccadic responses produced in our experiments were fully independent from the concurrent linguistic stimuli. Thus the mapping between the stimulus (i.e. the words) and the response (i.e. eye movements) can be disregarded as a potential contributing factor to the semantic effects demonstrated in the experiments described here. Finally, unlike many studies in the embodied cognition literature (e.g. Glenberg & Kaschak, 2002; Meteyard et al., 2007; Zwaan et al., 2002; Zwaan & Taylor, 2006), no decision process was needed in reaction to either the

---

[71] See the section entitled “Evidence from behavioural research”, Chapter 1 for details of this study.

visual or the linguistic stimulus in our experiments. This was especially the case for the pursuit experiments given that there was no task associated with the linguistic stimuli presented while pursuit eye movements were substantially less susceptible to deliberate control compared to saccades and hand movements. As a result, the conclusion that language comprehension *per se* affects oculomotor responses can be derived from our data with more confidence. Overall, our results truly reflect the interaction between the mechanisms responsible for language processing and the oculomotor system, which are completely devoid of any potential confounding factors such as semantic associations, the mapping between stimulus and response, and decision processes.

Despite demonstrating the "pure" language effect on eye movements for the first time, this thesis also provides a possible solution to one of the major issues of the grounded cognition approach to language comprehension: Why a kicking action is not always generated every time when the word "*kick*" is heard. Instead of a separate and possibly qualitatively different inhibitory mechanism, the current thesis proposes that this problem can be resolved by assuming that the activation strength of the sensorimotor representations initiated by language comprehension is *graded*. The results from the experiments reported in Chapter 2 (i.e. Experiment 1, 2 and 3) indicate that the directional semantics of motion verbs can activate certain representations that consequently bias the pursuit response. Furthermore, Experiment 6 (Chapter 3) revealed that vertical motion verbs were not sufficient to generate eye displacement or velocity in the vertical dimension during horizontal pursuit; nonetheless these verbs were able to interfere with the ongoing oculomotor response in the horizontal dimension. The results from these experiments in combination suggest that representations required by language comprehension are qualitatively comparable to representations activated during actual perceptual or motoric processing, such as perceiving a visual distractor moving in the vertical dimension during horizontal pursuit. However, the activation strength of language-mediated representations may not always be strong enough to produce any overt behavioural response (e.g. due to the lack of an associated task). Returning to the "kick" issue: The processing of the word indeed activates the same representations or neural substrates as when performing a kicking action, however, the strength of these activations is graded and may not be sufficient to generate a kicking action in every context.

Implications for eye movement research

Little attention has been paid to cognitive influences on pursuit eye movements until recently (see Kowler, 1990 for a review). To our knowledge, experiments described in this thesis are the first attempt to demonstrate the effect of language semantics on smooth pursuit. These findings are of special theoretical interest for smooth pursuit, which has been considered as being substantially less susceptible to deliberate control. Our results indicate that language influence on pursuit is not dependent on any additional processes that involve volition or decision-making. Instead, the control mechanism for pursuit eye movements is *directly* under the modulation of linguistic information.

As stated in the previous section, the semantic effects on eye movements (i.e. both pursuit and saccade) shown in this thesis reflect the direct impact of linguistic meaning on the oculomotor system bypassing other types of potential connections or processes. Thus our empirical findings have posed significant challenges for models built for both types of eye movements. Regarding pursuit eye movements, most models are not explicit about influences from cognitive factors. For those that do contain components under the influence of cognitive functions, these "cognitive components" typically need further development in order to account for the complex interactions between language and pursuit demonstrated here.[72] The exact component (or components) that enters into the competition with language-related representations remains to be pinpointed. Considering saccadic models, the stages in which cognitive functions are permitted to integrate into the system with other types of perceptual information are limited.[73] Taken together, our data indicate that either new components need to be introduced into current eye movement models to account for the potential cognitive influences, or modifications for these models are called for so that more complex cognitive influences can be explained.

---

[72] See the General Discussion section in Chapter 2 for a description of an example model of pursuit.

[73] See the Discussion section in Chapter 6 for a description of an example model of saccades.

Given our experimental procedures, the empirical findings reported here also revealed certain functional properties about pursuit maintenance. The maintenance period during pursuit has not been studied as extensively as pursuit initiation. This is because the maintenance stage is under the control of both perceptual feedback and an internal representation of target motion (e.g. Barnes & Asselman, 1991; Krauzlis & Lisberger, 1994; Yasui & Young, 1975). Previous research tended to control for the perceptual feedback so that the internal mechanism can be studied in isolation. As a result, the dynamics between the perceptual feedback and the internally represented target motion remained unclear. The present thesis has provided some insight into this issue. The secondary perceptual detection task included in Experiment 8 (Chapter 5) increased positional errors at the onset of pursuit compared to previous studies (i.e. Experiments 1-3) but not during the steady state. In other words, the additional perceptual detection task impaired the initiation but not the maintenance stage during pursuit. These results implied that the pursuit system relied more on perceptual feedback at the initiation stage of pursuit compared to pursuit maintenance. The possible underlying factor behind the difference of perceptual involvement between pursuit initiation and maintenance appears to be *task-relevance*: In order to initiate appropriate pursuit response to a moving stimulus, the motion properties of that stimulus must be perceptually sampled. However, once the eyes enter the maintenance stage, the oculomotor task has switched to sustaining smooth eye movements based on already sampled target-related information. Consequently, perceptual feedback is less involved in pursuit maintenance than initiation. The importance of *task-relevance* (or *task-requirements*) in determining the dynamics of perceptual feedback and the internal representation of target motion during pursuit has been confirmed by other experiments reported in this thesis. The results from Experiment 5 (Chapter 4) hinted at the possibility that when the oculomotor task was less effortful (e.g. pursuing a target moving at a constant velocity), the eye movement system depended less on perceptual feedback, compared to when the target had a more varied velocity profile and was harder to pursue (e.g. sinusoidal motion). Thus the relative gaze position to the target became less crucial and did not interact with verb semantics. Perhaps the strongest support for the hypothesis of task-relevance came from Experiment 4 (Chapter 3) that employed the extinction procedure. The transient target disappearance during extinction altered the task from pursuing the target to maintaining smooth eye movement in a particular direction. Therefore perceptual target information became task-irrelevant in the case of extinction, and the internal representation of target motion came to be the driving force for pursuit and went onto compete with currently presented linguistic stimuli. Altogether, the findings from these three experiments point to the conclusion that the dynamics between perceptual feedback and the internal representation of target motion during pursuit is not constant. Task-relevance (or task-requirements) seems to be responsible for determining to what extent each source of control is dependent upon at a particular phase during pursuit.

Implications for cognition, perception and action: Attention is the missing link

The question of how simulations come into contact with perception and action and in turn bias behavioural responses is largely overlooked in the grounded cognition literature. Furthermore, this question is not only confined to the grounded cognition framework and language comprehension, it can also be expanded into the more general issue of how human cognition shapes perception and action.

In order to address this issue, the present thesis has chosen to focus on the relationship between language comprehension, which represents the most complex form and the highest level of cognition, and eye movements, which can be viewed as the epitome of the interaction between perception and action. The theoretical contribution of the current research is mainly reflected by the attempt to set up the linkage connecting language comprehension and eye movement control through interpolating attention between these two mechanisms. The competition principle proposed in the attentional framework by Miller and Cohen (2001) provides the most parsimonious and systematic account in terms of how simulations come to modulate perception and action: Sensorimotor representations activated during language comprehension compete with concurrently active representations involved in perceptual and motoric processing. The consequence of this competition may be an increase or a decrease in the activation strength of the representations required by the perceptual or motoric task, which ultimately leads to biases in the percept or the action generated.

Not only can the guided activation attentional framework provide an explanation for the interactions observed between language and perception and action, it also has two major advantages with respect to the bond between cognition and perception and action. First, the biased competition theory assumes that the competition principle between pathways governs the entire information-processing system, which means that attention defined in the guided activation framework functions at a system-level. Thus there is no need to make the distinction between "high-level" and "low-level" processing. Furthermore, it is not necessary to consider the question of whether "high-level" cognitive functions can affect "low-level" sensorimotor processing and it should no longer be surprising that language comprehension can modulate perception and action. Second, the guided activation theory suggests that attention arises as the bias from the general functional principles of the entire system. As a result, there is no demand for a separate and qualitatively different mechanism in order for attention to be defined and its role illustrated. In sum, experiments described in the present thesis essentially demonstrated *competition* between *graded* sensorimotor representations activated by language and *task-relevan*t representations required by the oculomotor response, which were regulated by attention functioning at a *system-level*.

Based on current evidence available, cognition, perception and action seem to be intimately connected with each other, instead of being confined in isolated compartments. Although not being directly addressed by the present thesis, it can be speculated that cognition always to a certain extent determines the end product of perceptual or motoric processing. As living organisms, humans constantly perceive information from the environment and the perceived information is in turn used to guide our actions. However, during these processes, cognition is never "switched off". This means that it is never the perceived sensory information per se that determines how we interact with the external world, but rather, our interpretation of it and our interpretations are more often than not biased by our cognition. The actions we perform are likewise not simply reactions to stimuli in the external world, but instead responses generated in the context of cognitive processing at that moment in time. Cognition, perception and action coordinate as a unified system to make our interactions with the environment successful and all components of this system are organized and regulated by the system's own rules.

Future directions

The major weakness of the research reported in this thesis is that all the interactions are unidirectional: The experiments focused on demonstrating the influence that language comprehension has on eye movement control but not the reverse. Consequently, it remains unknown whether processes behind language comprehension can be modulated by a concurrent oculomotor task. In order to argue for a unified system composed of cognition, perception and action, it is crucial to provide empirical evidence that indicates the interactions between language comprehension and eye movements are bidirectional. However, half of the picture is currently missing in the present thesis. This issue can be resolved by modifying the experiments reported here so that responses to the linguistic stimuli are measured as a dependent variable. For example, given the experiments included in Chapter 2 (Experiment 1, 2, and 3), a lexical decision task for the auditory words can be employed as well as the pursuit task. This is to address research questions such as whether the response latency to the directional verbs in the lexical decision task is modulated as the consequence of any potential competition between language comprehension and the pursuit task. Following the same method, all experiments reported in this thesis can be altered and in turn provide the chance to fully explore the question whether the execution of oculomotor responses can influence language processing.

Another issue that deserves some consideration in the future is to what extent language comprehension impacts on the saccadic system. The only saccadic experiment (Experiment 9) generated ambiguous and unsystematic results, thus leading onto difficulties in making firm conclusions compared to the pursuit experiments. Further research studying the semantic effect on saccadic eye movements will provide significant theoretical contributions in two ways: First, as another major type of voluntary eye movements, saccades ultimately reflect the interaction between perception and action. Demonstrations of linguistic influence on saccadic eye movements will inevitably become another substantial building block for the unified system envisioned in the present thesis. Second, by exploring the interaction between language comprehension and saccades, the long-standing question of whether saccadic and pursuit eye movements are governed by the same mechanism (See Krauzlis, 2003 for a review) can be further elucidated. More specifically, if the same pattern of interactions can be verified between language processing and both saccades and pursuit, these two types of eye movements are more likely to be different behavioural outcomes of the same system.

In order to investigate further into the relationship between language processing and saccadic eye movements, existing studies can be modified to generate additional experiments. One of the reasons that the saccadic experiment reported here (Experiment 9) failed to generate clear results may be that too many different conditions have been included in the same experiment. The design of this experiment can be modified in the future so that only the effects of either the directional verbs or the locational nouns on saccades are tested in a single experiment. Furthermore, Experiment 9 did not reveal any semantic effect from either the verbs or nouns on saccadic landing positions.[74] This null result may have arisen as the consequence of the method used for this particular experiment: The saccadic target, which remained visible after its onset, served as a considerably strong visual cue for saccadic landing positions and any potential linguistic influence might not be able to compete with such a salient visual cue. In future research, this experiment can be repeated but with the saccadic target extinguished before any saccade is launched. Therefore subjects will essentially perform a “remembered saccade” task and direct their eyes to the remembered location of the target. This paradigm may provide more insight into whether language can affect saccadic landing positions. This and the other suggestion above are the two major ways in which the research presented in this thesis can be developed further.

---

[74] See Chapter 6 for details.

Appendices

Appendix 1: Instructions for the norming experiment (directional verbs)

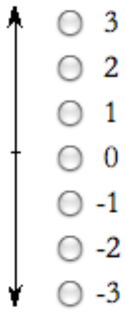
In this experiment, you will see a total of 40 words, each of which refers to an action.

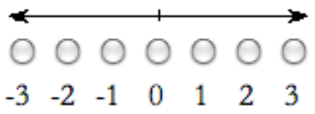
Each verb is accompanied by two rating scales, one horizontal, and the other vertical. Your task is to decide to what degree the action denoted by the word is likely to happen in a particular direction.

You can do this by selecting a rating on EITHER the horizontal scale OR the vertical scale. The ratings on the vertical scale go from -3, indicating strongly a downward action, to 3, indicating strongly an upward action. The ratings on the horizontal scale go from -3, this time indicating strongly a leftward action, to 3, indicating strongly a rightward action.

Here is an example:

Bury





You might think that the action "bury" is likely to happen in a downward moving

direction.
To answer, you need to choose the appropriate number on the scale (in this case, the vertical scale) and click on the circle next to it. In this example, you might want to click on -3 or -2, indicating strongly a downward moving direction.

Here is a summary of the rating scheme you could use: -3 - 'strongly downward (for the vertical scale) or leftward (for the horizontal scale) motion'; -2 - 'downward (V) or leftward (H) motion; -1 - 'slightly downward (V) or leftward (H) motion and conversely, 3 indicates 'strongly upward (V) or rightward motion (H)' and so on. If you think the action will not happen in any particular direction, please select '0'.

It's a hard judgment to make - so just try to be consistent. There's no right or wrong answer.

Appendix 2: List of experimental items (directional verbs)

Downward verbs	**Upward verbs**
Dive	Climb
Drown	Lift
Plunge	Rise
Sink	Raise
Drop	Grow
Fall	Rocket
Submerge	Ascend
Plummet	Heighten
Descend	Arise
Lower	Levitate
Bury	Elevate
Collapse	Escalate

Appendix 3: Lists of control items (nouns, non-directional verbs and non-locational nouns)

Experiment 1.

Nouns

Time
Town
Shop
Choice
Wall
Ball
Water
Job
Voice
Water
Child
Cup
Cat
Company
Home
Bed
Glass
Milk
House
Stone
Club
Life
Bottle
Week

Experiment 2, 4, 5, 6 & 8.

Nouns

Biscuit
Sofa
Sword
Laundry
Clock
Umbrella
Skin
Voice
Gadget
Jigsaw
Lion
Passport
Card
Food
Club
Necklace
Ocean
Advisor
Room
Dessert
Glossary
Town
Career
Curtain

Experiment 3.

Non-directional verbs

Abolish
Spawn
Locate
Worry
Separate
Spin
Reprint
Fold
Hunt
Borrow
Send
Forbid
Wait
Dazzle
Drench
Evict
Frighten
Drive
Abrogate
Formulate
Recreate
Sell
Refer
Wash

Experiment 7 & 9.

Non-directional verbs	Non-locational nouns
Abolish	Breath
Spawn	Locker
Locate	Shirt
Worry	Remnant
Separate	Yard
Spin	Symbol
Reprint	Dancer
Fold	Pot
Hunt	Filter
Borrow	Consultant
Send	Pain
Forbid	Shard
Wait	Drummer
Dazzle	Limb
Drench	Noise
Evict	Pole
Frighten	Trainee
Drive	Spider
Abrogate	Vineyard
Formulate	Platform
Recreate	Taxi
Sell	Yoghurt
Refer	Waiter
Wash	Glass

Appendix 4: Instructions for the norming experiment (locational nouns)

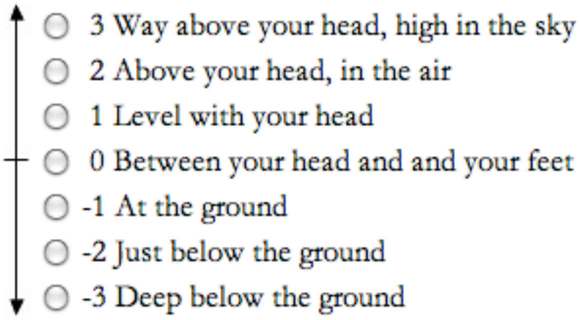
In this experiment, you will see a total of 60 words, each of which refers to an entity.

Each word is accompanied by a rating scale. Your task is to decide to what degree the entity denoted by the word is likely to be found in a particular spatial location.

You can do this by selecting a rating on the scale. The ratings on the scale go from -3, indicating some place very low; to 3, indicating some place very high.

Here is an example:

Mole



○ 3 Way above your head, high in the sky
○ 2 Above your head, in the air
○ 1 Level with your head
○ 0 Between your head and and your feet
○ -1 At the ground
○ -2 Just below the ground
○ -3 Deep below the ground

You might think that a **Mole** is likely to be found somewhere below the ground. To answer, you need to choose the appropriate number on the scale and click on the circle next to it. In this example you might want to click on -2 or -3, indicating somewhere below the ground.

It's a hard judgment to make - so just try and be consistent. There's no right or wrong answer.

Appendix 5: List of experimental items (Locational nouns)

Low nouns	**High nouns**
Sewer	Aerial
Submarine	Balloon
Basement	Ceiling
Miner	Roof
Tunnel	Satellite
Fish	Bird
Subway	Loft
Cellar	Chimney
Anchor	Cloud
Worm	Moon
Coal	Sky
Root	Planet

References

Allopena, P. D., Magnuson, J. S., & Tanenhaus, M. K. (1998). Tracking the time course of spoken word recognition using eye movements: Evidence for continuous mapping models. *Journal of Memory and Language, 38*, 419-439.

Allport, A. (1993). Attention and control: Have we been asking the wrong questions. A critical review of the last twenty five years. In D. E. Meyer & Kornblum, S. (Eds.), *Attention and Performance XIV* (pp. 183-218). Cambridge: MIT Press.

Altmann, G. T. M. (2004). Language-mediated eye movements in the absence of a visual world: The "blank screen paradigm". *Cognition, 93*, 79-87.

Altmann, G. T. M., & Kamide, Y. (2009). Discourse-mediation of the mapping between language and the visual world: Eye movements and mental representation. *Cognition, 111*, 55-71.

Baloh, R. W., Yee, R. D., Honrubia, V., & Jacobson, K. (1988). A comparison of the dynamics of horizontal and vertical smooth pursuit in normal human subjects. *Aviation Space and Environmental Medicine, 59*, 121-124.

Barnes, G. R. (2008). Cognitive processes involved in smooth pursuit eye movements. *Brain and Cognition, 68*, 309-326.

Barnes, G. R., & Asselman, P. T. (1991). The mechanism of prediction in human smooth pursuit eye movements. *Journal of Physiology, 439*, 439-461.

Barnes, G. R., Barnes, D. M., & Chakraborti, S. R. (2000). Ocular pursuit responses to repeated, single-cycle sinusoids reveal behaviour compatible with predictive pursuit. *Journal of Neurophysiology, 84*, 2340-2355.

Barnes, G. R., & Collins, C. J. S. (2008). The influence of briefly presented randomised target motion on the extra-retinal component of ocular pursuit. *Journal of Neurophysiology, 99*, 831-842.

Barnes, G. R., Goodbody, S. J., & Collins, S. (1995). Volitional control of anticipatory ocular pursuit responses under stabilized image conditions in humans. *Experimental Brain Research, 106*, 301-317.

Barnes, G. R., Grealy, M., & Collins, S. (1997). Volitional control of anticipatory ocular smooth pursuit after viewing, but not pursuing, a moving target: Evidence for a re-afferent velocity store. *Experimental Brain Research, 116*, 445-455.

Barnes, G. R., & Schimid, A. M. (2002). Sequence learning in human ocular smooth pursuit. *Experimental Brain Research, 144*, 322-335.

Barsalou, L. W. (1999). Perceptual symbol systems. *Behavioural and Brain Science, 22*, 577-660.

Barsalou, L. W. (2007). Grounding symbolic operations in the brain's modal systems. In G. R. Semin & E. R. Smith (Eds.), *Embodied Grounding: Social, Cognitive, Affective, and Neuroscientific Approaches*. New York: Cambridge University Press.

Becker, W. (1989). Metrics. In R. H. Wurtz, & M. E. Goldberg (Eds.), *The neurobiology of saccadic eye movements* (pp. 13-67). Amsterdam: Elservier.

Becker, W., & Fuchs, A. F. (1985). Prediction in the oculomotor system: Smooth pursuit during transient disappearance of a visual target. *Experimental Brain Research, 57*, 562-575.

Becker, W., & Jürgens, R. (1979). An analysis of the saccadic system by means of double step stimuli. *Vision Research, 19*, 967-983.

Bennett, S. J., & Barnes, G. R. (2003). Human ocular pursuit during the transient disappearance of a visual target. *Journal of Neurophysiology, 90,* 2504-2520.

Bennett, S. J., & Barnes, G. R. (2004). Predictive smooth ocular pursuit during the transient disappearance of a visual target. *Journal of Neurophysiology, 92,* 578-590.

Bergen, B. K. (2007). Experimental methods for simulation semantics. In M. Gonzalez-Marquez, I. Mittleberg, S. Coulson, & M. Spivey (Eds.), *Methods in cognitive linguistics.* Amsterdam: John Benjamins.

Blohm, G., Missal, M., & Lefèvre, P. (2005). Direct evidence for a position input to the smooth pursuit system. *Journal of Neurophysiology, 94*, 712-721.

Borreggine, K. L., & Kaschak, M. P. (2006). The action-sentence compatibility effect: It's all in the timing. *Cognitive Science, 30*, 1097-1112.

Braver, T. S., & Cohen, J. D. (2000). On the control of control: The role of dopamine in regulating prefrontal function and working memory. In S. Monsell & J. Driver (Eds.), *Attention and Performance XVIII: Control of cognitive processes* (pp. 713-737). Cambridge, MA: MIT Press.

Buccino, G., Riggio, L., Melli, G., Binkofski, F., Gallese, V., & Rizzolatti, G. (2005). Listening to action related sentences modulates the activity of the motor system: A combined TMS and behavioural study. *Cognitive Brain Research, 24*, 355-363.

Calvo, M. G, Nummenmaa, L., & Hyönä, J. (2008). Emotional scenes in peripheral vision: Selective orienting and gist processing, but not content identification. *Emotion, 8*, 68-80.

Carl, J. R. & Gellman, R. S. (1987). Human smooth pursuit: Stimulus-dependent responses. *Journal of Neurophysiology, 57*, 1446-1463.

Chambers, C. G., Tanenhaus, M. K., & Magnuson, J. S. (2004). Actions and affordances in syntactic ambiguity resolution. *Journal of Experimental Psychology: Learning, Memory & Cognition, 30*, 687-696.

Chen, Y., Holzman P. S., & Nakayama, K. (2002). Visual and cognitive control of attention in smooth pursuit. *Progress in Brain Research, 140*, 255-265.

Chomsky, N. (1980). *Rules and representations*. New York: Columbia University Press.

Cohen, J. D., Aston-Jones, G., & Gilzenrat, M. S. (2004). A system-level perspective on attention and cognitive control: Guided activation, adaptive gating, conflict monitoring, and exploitation vs. exploration. In M. I. Posner (Eds.), *Cognitive neuroscience of attention* (pp. 71-90). New York: Guilford Press.

Cohen, J. D., Dunbar, K., & McClelland, J. L. (1990). On the control of automatic processes: A parallel distributed processing model of the Stroop effect. *Psychological Review, 97*, 332-361.

Collewijn, H., & Tamminga, E. P. (1984). Human smooth and saccadic eye movements during voluntary pursuit of different target motions on different backgrounds. *Journal of Physiology, 351*, 217-250.

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, 6,* 84-107.

Damasio, A. R., & Damasio, H. (1994). Cortical systems for retrieval of concrete knowledge: The convergence zone framework. In C. Koch & J. L. Davis

(Eds.), *Large-Scale Neuronal Theories of the Brain*. Cambridge: MIT Press.

Desimone, R., & Duncan, J. (1995). Neural mechanisms of selective attention. *Annual Review of Neuroscience, 18*, 193-222.

Deubel, H., & Schneider, W. X. (1996). Saccade target selection and object recognition: Evidence for a common attentional mechanism. *Vision Research, 36*, 1827-1837.

Estes, Z., Verges, M., & Barsalou, L. W. (2008). Head up, foot down: Object words orient attention to the object's typical location. *Psychological Science, 19*, 93-97.

Ferrera, V. P., & Lisberger, S. G. (1995). Attention and target selection for smooth pursuit eye movements. *Journal of Neuroscience, 15*, 7472-7484.

Fincher-Kiefer, R. (2001). Perceptual components of situation models. *Memory and Cognition, 29*, 336-343.

Findlay, J. M., & Walker, R. (1999). A model of saccade generation based on parallel processing and competitive inhibition. *Behavioural and Brain Sciences, 22*, 661-721.

Fodor, J. A. (1975). *The language of thought*. Cambridge: Harvard University Press.

Ferrera, V. P., & Lisberger, S. G. (1995). Attention and target selection for smooth pursuit eye movements. *Journal of Neuroscience, 15*, 7472-7484.

Glenberg, A. M. (1997). What memory is for. *Behavioural and Brain Sciences, 20*, 1-55.

Glenberg, A. M., & Kaschak, M. P. (2002). Grounding language in action. *Psychonomic Bulletin & Review, 9*, 558-565.

Glenberg, A. M., Sato, M., Cattaneo, L., Riggio, L., Palumbo, D., & Buccino, G. (2008). Processing abstract language modulates motor system activity. *Quarterly Journal of Experimental Psychology, 61*, 905-919.

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

Grant, M. P., Leigh, R. J., Seidman, S. H., Riley, D. E., & Hanna, J. P. (1992). Comparison of predictable smooth ocular and combined eye-head tracking behaviour in patients with lesions affecting the brain stem and cerebellum. *Brain, 115*, 1323-1342.

Hauk, O., Johnsrude, I., Pulvermüller, F. (2004). Somatotopic representation of action words in human motor and premotor cortex. *Neuron, 41*, 301-307.

Henderson, J. M. (1992). Visual attention and eye movement control during reading and picture viewing. In K. Rayner. (Eds.), *Eye Movements and Visual Cognition* (pp. 260-283). Berlin: Springer-Verlag.

Henderson, J. M., & Ferreira, F. (2004). *The interface of language, vision and action: Eye movements and the visual world.* New York: Psychology Press.

Hoffman, J. E., & Subramaniam, B. (1995). The role of visual attention in saccadic eye movements. *Perception and Psychophysics, 57,* 787-789.

Hommel, B., Müsseler, J., Aschersleben, G., & Prinz, W. (2001). The theory of even coding (TEC): A framework for perception and action. *Behavioural and Brain Sciences, 24*, 849-937.

Huettig, F., & Altmann, G. T. M. (2007). Visual-shape competition during language-mediated attention is based on lexical input and not modulated by contextual appropriateness. *Visual Cognition, 15*, 985-1018.

Hutton, S. B. (2008). Cognitive control of saccadic eye movements. *Brain and Cognition, 68*, 327-340.

Hutton, S. B., & Tegally, D. (2005). The effects of dividing attention on smooth pursuit eye tracking. *Experiments in Brain Research, 163*, 306-313.

Jarrett, C. B., & Barnes, G. R. (2002). Volitional scaling of anticipatory ocular pursuit velocity using precues. *Cognitive Brain Research, 14*, 383-388.

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*, 79-89.

Kathmann, N., Hochrein, A., & Uwer, R. (1999). Effects of dual task demands on the accuracy of smooth pursuit eye movements. *Psychophysiology, 36*, 158-163.

Kerzel, D., Souto, D., & Ziegler, N. E. (2008). Effects of attention shifts to stationary objects during steady-state smooth pursuit eye movements. *Vision Research, 48*, 958-969.

Kintsch, W. (1988). The role of knowledge in discourse comprehension: A construction-integrated model. *Psychological Review, 95*, 163-182.

Klein, R. (1980). Does oculomotor readiness mediate cognitive control of visual attention? In R. S. Nickerson (Eds.), *Attention and Performance VIII* (pp. 259-276). Hillsdale: Lawrence Erlbaum.

Knoeferle, P., & Crocker, M. W. (2007). The influence of recent scene events on spoken comprehension: Evidence from eye movements. *Journal of Memory and Language, 57*, 519-543.

Kowler, E. (1990). The role of visual and cognitive processes in the control of eye movement. In E. Kowler (Eds.), *Eye movements and their role in visual and cognitive processes*. Amsterdam: Elsevier.

Kowler, E., & Steinman, R. M. (1979a). The effect of expectations on slow oculomotor control – I. Periodic target steps. *Vision Research, 19*, 619-632.

Kowler, E., & Steinman, R. M. (1979b). The effect of expectations on slow oculomotor control – II. Single target displacements. *Vision Research, 19*, 633-646.

Krauzlis, R. J. (2003). Recasting the smooth pursuit eye movement system. *Journal of Neurophysiology, 91*, 591-603.

Krauzlis, R. J. (2005). The control of voluntary eye movements: New perspectives *Neuroscientist, 11*,124-137.

Krauzlis, R. J., & Lisberger, S. G. (1994). A model of visually-guided smooth pursuit eye movements based on behavioural observations. *Journal of Computational Neuroscience, 1*, 265-283.

Krauzlis, R. J., & Miles, F. A. (1996a). Initiation of saccades during fixation or pursuit: Evidence in humans for a single mechanism. *Journal of Neurophysiology, 76*, 4175-4179.

Krauzlis, R. J., & Miles, F. A. (1996b). Transitions between pursuit eye movements and fixation in the monkey: Dependence on context. *Journal of Neurophysiology, 76*, 1622-1638.

Lakoff, G. & Johnson, M. (1980). *Metaphors we live by*. Chicago: University of Chicago Press.

Langacker, R. W. (1986). An introduction to cognitive grammar. *Cognitive Science, 10*, 1-40.

Langton, S. R. H., & Bruce, V. (1999). Reflexive visual orienting in response to the social attention of others. *Visual Cognition, 6*, 541-567.

Lisberger, S. G., & Fuchs, A. F. (1978). Role of primate flocculus during rapid behavioural modification of vestibuloocular reflex I Purkinje cell activity during visually guided horizontal smooth pursuit eye movements and passive head rotation. *Journal of Neurophysiology, 41*, 733-763.

Marslen-Wilson, W. (1990). Activation, competition, and frequency in lexical access. In G. Altmann (Eds.), *Cognitive models of speech processing* (pp. 148-172). Cambridge, MA: MIT Press.

Martin, A. (2001). Functional neuroimaging of semantic memory. In R. Cabeza & A. Kingstone (Eds.), *The handbook of functional neuroimaging of cognition* (pp. 153-186). Cambridge, MA: MIT Press.

Martin, A. (2007). The representation of object concepts in the brain. *Annual Review of Psychology, 58*, 25-45.

Martin, A., Wiggs, C. L., Ungerleider, L. G., & Haxby, J. V. (1996). Neural correlates of category-specific knowledge. *Nature, 379*, 649-652.

McClelland, J. L. (1993). Toward a theory of information processing in graded, random, and interactive network. In D. E. Meyer & S. Kornblum (Eds.), *Attention and Performance XIV: Synergies in experimental psychology, artificial intelligence, and cognitive neuroscience* (pp. 655-688). Cambridge, MA: MIT Press.

McNeill, D. (1992). *Hand and mind*. Chicago: University of Chicago Press.

Meteyard, L., Bahrami, B., & Vigliocco, G. (2007). Motion detection and motion verbs: Language affects low-level visual perception. *Psychological Science, 18*, 1007-1013.

Miller, E. K., & Cohen, J. D. (2001). An integrative theory of prefrontal cortex function. *Annual Review of Neuroscience, 24*, 167-202.

Mitrani, L., & Dimitrov, G. (1978). Pursuit eye movements of a disappearing moving target. *Vision Research, 18*, 537-539.

Müsseler, J., & Hommel, B. (1997). Blindness to response-compatible stimuli. *Journal of Experimental Psychology: Human Perception and Performance, 23*, 861-872.

Newell, A., & Simon, H. A. (1972). Human problem solving. NJ: Prentice Hall.

Orban de Xivry, J., Missal, M., & Lefèvre, P. (2008). A dynamic representation of target motion drives predictive smooth pursuit during target blanking. *Journal of Vision, 8*, 1-13.

O'Reilly, R. C., & Munakata, Y. (2000). *Computational explorations in cognitive neuroscience: Understanding the mind by simulating the brain.* Cambridge, MA: MIT Press.

Pinker, S. (1994). *The language instinct.* New York: HarperCollins.

Pola, J., & Wyatt, H. (1997). Offset dynamics of human smooth pursuit eye movements: Effects of target presence and subject attention. *Vision Research, 37*, 2579-2595.

Posner, M. I, Rafal, R. D., Choate, L. S., & Vaughan, J. (1985). Inhibition of return: Neural basis and function. *Cognitive Neuropsychology, 2*, 211-228.

Preissl, H., Pulvermüller, F., Lutzenberger, W., & Birbaumer, N. (1995). Evoked potentials distinguish between nouns and verbs. *Neuroscience Letters, 197*, 81-83.

Pulvermüller, F. (1999). Words in the brain's language. *Behavioural and Brain Sciences, 22*, 253-279.

Pulvermüller, F., Lutzenberger, W., & Preissl, H. (1999). Nouns and verbs in the intact brain: Evidence from event-related potentials and high-frequency cortial responses. *Cerebral Cortex, 9*, 498-508.

Pulvermüller, F., Shtyrov, Y., & Ilmoniemi, R. (2005). Brain signatures of meaning access in action word recognition. *Journal of Cognitive Neuroscience, 17*, 884-892.

Pylyshyn, Z. W. (1986). *Computational cognition: Toward a foundation for cognitive science*. Cambridge: MIT Press.

Richardson, D. C., & Matlock, T. (2007). The integration of figurative language and static depictions: An eye movement study of fictive motion. *Cognition, 102*, 129-138.

Rizzolatti, G., & Craighero, L. (2004). The mirror-neuron system. *Annual Review of Neuroscience, 27*, 188-194.

Rizzolatti, G, Fogassi, L., & Gallese, V. (2001). Neurophysiological mechanisms underlying the understanding and imitation of action. *Nature Reviews Neuroscience, 2*, 661-670.

Rizzolatti, G., Riggio, L., Dascola, I., & Umiltá, C. (1987). Reorienting attention across the horizontal and vertical meridians: Evidence in favour of a premotor theory of attention. *Neuropsychologia, 25*, 31-40.

Rizzolatti, G., Riggio, L., & Sheliga, B. M. (1994). Space and selective attention. In C. Umiltá & M. Moscovitch (Eds.), *Attention and Performance XV* (pp. 231-265). Cambridge: MIT Press.

Robinson, D. A., Gordon, J. L., & Gordon, S. E. (1986). A model of the smooth pursuit eye movement system. *Biological Cybernetics, 55*, 43-57.

Rottach, K. G., Zivotofsky, A. Z., Das, V. E., Averbuch-Heller, L., Discenna, A. O., Poonyathalang, A., Leigh, R. J. (1996). Comparison of horizontal, vertical and diagonal smooth pursuit eye movements in normal human subjects. *Vision Research, 36*, 2189-2195.

Rougier, N. P., & O'Reilly, R. C. (2002). Learning representations in a grated prefrontal cortex model of dynamic task switching. *Cognitive Science, 26*, 503-520.

Rumelhart, D. E., McClelland, J. L., & PDP Research Group. (1986). *Parallel distributed processing: Explorations in the microstructure of cognition.* Cambridge, MA: MIT Press.

Schneider, W. X. (1995). VAM: A neuro-cognitive model for visual attention control of segmentation, object recognition, and space-based motor action. *Visual Cognition, 2*, 331-375.

Šetić, M., & Domijan, D. (2007). The influence of vertical spatial orientation on property verification. *Language and Cognitive Processes, 22*, 297-312.

Shagass, C., Roemer, R. A., & Amadeo, M. (1976). Eye-tracking performance and engagement of attention. *Archives of General Psychiatry, 33*, 121-125.

Spering, M., Gegenfurtner, K. R., & Kerzel, D. (2006). Distractor interference during smooth pursuit eye movements. *Journal of Experimental Psychology: Human Perception and Performance, 32*, 1136-1154.

Spivey, M. J., & Geng, J. J. (2001). Oculomotor mechanisms activated by imagery and memory: Eye movements to absent objects. *Psychological Research, 65*, 235-241.

Spivey, M. J., Tyler, M., Richardson, D. C., & Young, E. (2000). Eye movements during comprehension of spoken scene descriptions. *Proceedings of the Twenty-second Annual Meeting of the Cognitive Science Societ*y, (pp. 487-492). Mawhah, NJ: Erlbaum.

Stroop, J. R. (1935). Studies of interference in serial verbal reactions. *Journal of Experimental Psychology, 18*, 643-662.

Sweeney, J. A., Haas, G. L., Liu, S., & Weiden, P. J. (1994). Selective effects of antipsychotic medications on eye-tracking performance in schizophrenia. *Psychiatry Research, 54*, 185-198.

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*, 1632-1634.

Tucker, M., & Ellis, R. (1998). On the relations between seen objects and components of potential actions. *Journal of Experimental Psychology: Human perception and Performance, 24*, 830-846.

van der Stigchel, S., Theeuwes, J. (2006). Differences in distractor-induced deviation between horizontal and vertical saccade trajectories. *NeuroReport, 19*, 251-254.

van Donkelaar, P, & Drew, A. S. (2002). The allocation of attention during smooth pursuit eye movements. *Progress in Brain Research, 140*, 267-277.

von Noorden, G. K., & Mackensen, G. (1962). Pursuit movements of normal of amblyopic eyes – an electro-ophthalmographic study 1 physiology of pursuit movements. *American Journal of Ophthalmology, 53*, 325-336.

Whittaker, S. G., & Eaholtz, G. (1982). Learning patterns of eye motion for foveal pursuit. *Investigation Ophthalmology and Visual Science, 23*, 393-397.

Wickens, T. D. (2002). *Elementary signal detection theory*. New York: Oxford University Press.

Yarbus, A. L. (1967). Eye movements and vision. New York: Planum Press.

Yasui, S., & Young, L. R. (1975). Perceived visual motion as effective stimulus to pursuit eye movement system. *Science, 190*, 906-908.

Young, L. R., & Stark, L. (1963). Variable feedback experiments testing a sampled data model for eye tracking movements. *IEEE Transactions, HFE-4*, 38-51.

Zwaan, R. A. (2009). Mental simulation in language comprehension and social cognition. *European Journal of Social Psychology, in press*.

Zwaan, R. A., Stanfield, R. A., & Yaxley, R. H. (2002). Language comprehenders mentally represent the shapes of objects. *Psychological Science, 13*, 168-171.

Zwaan, R. A., & Taylor, L. J. (2006). Seeing, acting, understanding: Motor resonance in language comprehension. *Journal of Experimental Psychology: General, 135*, 1-11.

Zwaan, R. A., Yaxley, R. H. (2003). Spatial iconicity affects semantic relatedness judgements. *Psychonomic Bulletin & Review, 10*, 954-958.