Eye movements under the control of working memory:

The challenge of a reading-span task

Dissertation

zur Erlangung des akademischen Grades
Doktor der Philosophie (Dr. phil.)
im Fach Psychologie

eingereicht bei der
Humanwissenschaftlichen Fakultät
der Universität Potsdam
im Jahr 2011

von
Anja Gendt

Danksagung

Eine wissenschaftliche Arbeit kann nur einen begrenzten Blick auf die komplexen Zusammenhänge eines großen Ganzen geben. Sie ist wie ein Puzzle, bei dem je nach Zusammensetzung der Teile neue interessante Bilder erscheinen. Auf der Suche nach meinem wissenschaftlichen Bild haben mich viele Menschen begleitet, inspiriert und unterstützt.

Jemand, der scheinbar nie daran zweifelte, dass es ein interessantes Bild werden wird, ist mein Mentor Prof. Reinhold Kliegl, dem ich an dieser Stelle für seine Unterstützung herzlich danken möchte. Ein Vortrag Reinhold Kliegls und seine Begeisterung für die kognitive Psychologie trugen einst zu meiner Entscheidung bei, dieses Fach zu studieren. Umso dankbarer bin ich, dass er mir die Möglichkeit gab im Rahmen seiner Arbeitsgruppe zu forschen. Sein Enthusiasmus für die Kognitionswissenschaften allgemein und speziell für statistische Analysen scheint unerschöpflich, ist ansteckend und fordert den eigenen wissenschaftlichen Anspruch heraus.

Ebenfalls herzlichen Dank möchte ich meinem Zweitgutachter Prof. Klaus Oberauer aussprechen. Dafür, dass er sich bereit erklärte diese Arbeit zu begutachten, aber vor allem dafür, dass er mir als Studentin das Gebiet der Arbeitsgedächtnisforschung so anschaulich und gleichzeitig mit einem sehr hohen Anspruch vermittelte. Damit waren schon einige Puzzleteile gelegt.

Bei der Suche weiterer passender Teile waren Irrwege nicht ausgeschlossen. Dank ganz lieber Mitstreiter deren Puzzlesteine ähnlich meiner waren, habe ich den rechten Pfad schnell wieder gefunden. Ein ganz herzlicher Dank für viele tolle inspirierende Gespräche geht an André Krügel, Christiane Wotschack, Daniel Schad, Elke Lange, Eike Richter, Hans Trukenbrod, Jochen Laubrock, Konstantin Mergenthaler und Michael Dambacher. Auch wenn sich bereits ein recht farbenfrohes Bild ergab, taten sich doch auch unerwartet schwierige Stellen auf. Drei, die nicht nur den wissenschaftlichen Erste-Hilfe Koffer für mich bereit hielten sind Annekatrin Hudjetz, Katrin Göthe und Sarah Risse. - *Habt vielen lieben Dank vor allem auch für Eure so hilfreichen Vorschläge zur Überarbeitung. Damit habt ihr wesentlich dazu beigetragen das die Farben des Bildes nun viel deutlicher in Erscheinung treten!*

Ein großer Dank geht an alle studentischen Hilfskräfte und vor allem an Petra Schienmann, ohne deren verlässliche und gut organisierte Datenerhebung, es erst gar keine Puzzleteile zum Sortieren gegeben hätte.

Aufgrund meiner eigenen Ressourcenlimitiertheit hat das außeruniversitäre DASEIN mitunter sehr gelitten. Meiner Familie und allen Freunden an dieser Stelle – *Danke für Euer Verständnis!*

Ganz besonders möchte ich mich bei meiner Mutter und meinen Schwiegereltern bedanken - *Ohne Eure Unterstützung und beständige Kinderbetreuung wäre diese Arbeit nie fertig geworden und die Teilnahme an den mir so wichtigen Konferenzen wäre kaum zu realisieren gewesen.* Vor allem danke ich meinem Mann, Michael, für seine liebevolle Unterstützung, und die hilfreichen Tipps beim Programmieren - *auch wenn Du schon viel zu müde dafür warst* - und meiner kleinen „Murmelmaus“ Sophie, deren Umarmung mich alles rundherum vergessen lässt.

Contents

List of Figures

List of Tables

Chapter 1

Theoretical Background

Reading is a highly complex task that involves a number of processes in the human brain. It requires continuous practice and experience before the basic cognitive processes involved become automated. Therefore, the understanding of the text depends on the reader's prior knowledge and experience. In addition, the complexity of the text plays an important role for comprehension. If we try to imagine how we read fiction in comparison to technical literature, we get the feeling that there is a difference between how we read each type of text.

Buswell (1922) already found longer reading times for words in technical literature in contrast to words in fiction. But how can these differences be explained? The words written in both texts are principally the same. However the number of unknown or infrequent words in technical literature makes the text more complex. During reading, previous statements in the text have to be remembered, while new information from currently read words has to be integrated. Highly complex texts are more difficult to remember than easy texts. As a consequence individuals read longer or reread a sentence.

The described individual differences can be explained by the construction of working memory (WM). This is a part of our memory in which we do not only keep information for a limited time, but we also operate with it. So, we could say we "work" with the given information and are able to recall it. How much we are able to remember and how many operations we are able to carry out, depends on the one hand on the material (light fiction vs. technical literature) and on the other hand on our personal abilities, resources or capacities. This is due to the fact that our cognitive system is limited. Thus, the reader's ability to remember and integrate information while reading a sentence depends on his working-memory capacity.

The thesis focuses on the question, to which extent memory processes modulate eye guidance. During reading, the eyes constantly move over the text. Where and when the eyes move during normal reading was observed in an extensive number of studies. Many of them documented that, for example, fixation durations depend on the word's frequency, length and predictability. It was often discussed that WM also has an effect on the interword eye-movement behaviour, but systematic research linking computational models of eye-movement control to theories of WM are missing. Evidence for very global relations of eye-guidance and WM come from psycholinguistic research. Here the influence of WM was explored for sentence comprehension. Thereby, the question of the influence on the interword fixation patterns remains unconsidered. The implementation of higher order cognitive processes in models of eye-movement control could help to disentangle some differential effects. Basically it is quite conceivable that different memory requirements are responsible for some heterogeneous results, which speak partly for sequential and partly for parallel models of eye-movement control.

By investigating eye-movements during a reading-related WM task, the current work takes a first step towards understanding the importance of working-memory processes for interword eye-movements. In the following section I will give a critical overview about previous results of WM influences on reading. For the theoretical background I begin with a short overview about basic principles of eye-movement control. Subsequently, some fundamental working-memory models and several theories of cognitive limitations are set in contrast to each other.

1.1 Basic eye movements

Depending on the anatomy of the eye, only the information in the fovea (or fovea centralis) can be seen clearly, and therefore the eyes move constantly to observe the environment. The fovea is the part of the eye in the retina, which allows for maximum acuity of vision. It extends out to 2° of visual angle, which corresponds to 6 - 8 characters during reading, depending on font size (Balota & Rayner, 1991). Contrary to our impression during reading, the eyes do not move continuously along a line of text. Smooth movements are accomplished only when the eyes follow an object. Contrary short rapid movements are necessary to guide the eyes from one stationary object to another one. This is the case during reading when the eyes have to be guided from word to word. These short rapid movements are called *saccades* and they are intermingled with short stops, called *fixations,* on the object. Information is received only during the fixation when the eyes remain relatively still. During a saccade we are effectively ‘blind’, as no visual input is obtained from static objects (Martin, 1974). There is considerable variability in the number and duration of fixations and saccades, which depends on the skill of the reader and on the text complexity (Rayner & Pollatsek, 1989). Individual differences become apparent if we think of reading beginners versus older adults, who have become experts through years of reading. Skilled readers' fixation duration has a mean of 200 - 250 milliseconds and the distance the eye moves during each saccade is an average of 7 - 9 characters. Nearly 20 percent of words receive more than one fixation consecutively, they are then *refixated.* An equal percentage of words receive no fixation. These words are then *skipped*. In addition, not all saccades are forward saccades. In 10 - 15 percent of the cases, a regressive saccade (*regression*) back to previously passed text passages is initialized (Rayner, 1998). A regression to a word already fixated defines the *second-pass,* whereas the first fixation and sum of all fixations until the word is left for the first time, defines the *first-pass*. The basic saccadic eye movements in reading are illustrated in Figure 1.1.

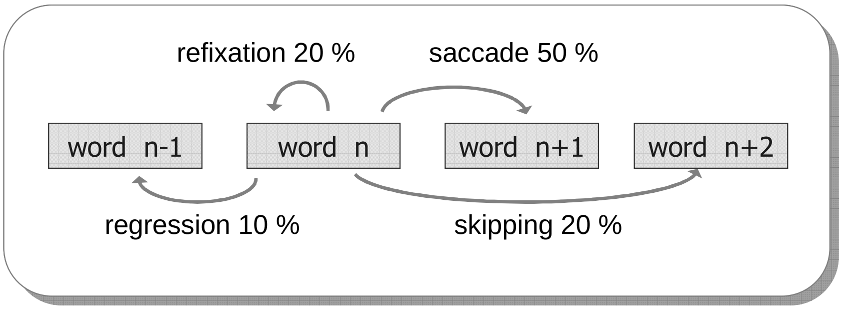

Figure 1.1: Types of saccadic eye movements and the probability of their occurrence, by starting from word n.

The main and frequently asked question is: What determines *where* and *when* we move our eyes? *Oculomotor theories* characterize low-level visuomotor factors as determining factors, like word boundaries or word length (e.g., O'Regan & Lévy-Schoen, 1987; O'Regan, 1990, 1992; McConkie, Kerr, Reddix, & Zola, 1988). Whereas *cognitive theories* suppose that lexical processing plays an important role, and thus according to this, also the words' frequency, predictability and moreover the processes of attentional control (e.g., Just & Carpenter, 1980; Morrison, 1984; O'Regan, 1979). Theoretical assumptions of both, oculomotor and cognitive theories, are considered in models of eye-movement control. To the present state of research, assumptions of working-memory theories are not integrated in models of eye guidance. Given that attentional control is the important underlying mechanism in both research fields, it stands to reason that working-memory and eye-movements are interrelated through attention processes. At this point I focus on the implementation of visual attention in models of eye guidance. Afterwards I bridge to theories of WM.

1.1.1 Different perspectives of attentional control

Models of eye-movement control differ mainly in one point, how they allocate attention. Sequential attention-shift (SAS; e.g., E-Z Reader: Reichle, Rayner, & Pollatsek, 2003) models assume that a word can only be processed if attention is focused on that word. Contrarily in guidance-by-attentional-gradient models (GAG;

e.g., SWIFT: Engbert, Nuthmann, Richter, & Kliegl, 2005; Glenmore: Reilly & Radach, 2003) attention is spatially distributed to a region beyond the fixated word and thus to a region where acuity of vision is reduced. Readers are also able to take up information from this parafoveal region which extend to 5° (or 15 - 20 characters). SAS and GAG models can be discriminated by the presence of parafoveal-on-foveal effects, which is that fixation duration on the current word is related to the difficulty of the word to the right of fixation (Drieghe, Brysbaert, & Desmet, 2005; Inhoff, Eiter, & Radach, 2005; Kennedy & Pynte, 2005; Richter Engbert, & Kliegl, 2006). It implies a preprocessing of the upcoming word before the eyes move on to that word or skip it. Corpus analytical results gave evidence (Kliegl, Nuthmann, & Engbert, 2006) for the coherence: Current single fixation durations were influenced by the frequency and predictability of the upcoming word. Only GAG models with the assumption of distributed attention can account for such effects, but parafoveal-on-foveal findings are inconsistent so far.

1.1.2 Distributed processing

Kennedy and colleges (Kennedy, 2000; Kennedy, Pynte, & Ducrot, 2002; Kennedy & Pynte, 2005; Pynte & Kennedy, 2006) demonstrated that only if the fixated word (n) is short, the word frequency of the word to the right of fixation (n + 1) influences the current gaze duration. Therefore the relation between the extent of preprocessing of word n + 1 and the properties of word n is a dynamical one. If that is the case, it could explain why some authors found evidence for parafoveal-on-foveal effects whereas others failed. The view of a dynamical relation is based on the assumption of a perceptual span. It is defined as the region of text from which useful information can be extracted (for a review, see Rayner, 1998). In consequence word n + 1 can only be preprocessed if it falls into the perceptual span, which is the case if word n is short enough. The interrelation between word length and perceptual span is illustrated in Figure 1.2. The perceptual span has been functionally approximated from moving-window studies (McConkie & Rayner, 1975; Rayner, 1975). The text outside a ‘window’ of normal text was replaced by strings of Xs. Reading times depended on window size and hence the validity of the upcoming text. Longer reading times were observed for small window sizes, where the upcoming text was invalid. For Latin alphabetical orthographies, like English or German, the perceptual span is estimated to extend from 3 characters to the left to 14 characters to the right of fixation. This asymmetry is not hardwired, but instead reflects attentional demands linked to reading direction (Pollatsek, Bolozky, Well, & Rayner, 1981). The size of the perceptual span seems to depend on the reading skill. For example several results suggest a smaller perceptual span for children in contrast to skilled readers (e.g., Taylor, 1965; Fisher & Lefton, 1976; Hochber, 1970).

Henderson and Ferreira (1990) provide evidence for a dynamical modulation of the perceptual span depending on the difficulty of the word in the fovea (see also Schroyens, Vitu, Brysbaert, & d'Idewalle, 1999). The results showed that fixation duration of word n + 1 was reduced, if word n was a highly frequent one. Word n + 1 was preprocessed only if the complexity of the word n was low. Evidence for the influences of neighbour words can not only be seen in parafoveal-on-foveal effects, but also in *skipping* and *spillover* effects. The upcoming word can only be skipped if it is preprocessed. Spillover effects emerge if the processing of a word is not completed before the eyes move on to the next word. The effects of word variables can then spill over to the subsequent fixation duration (Rayner & Duffy, 1986; Kliegl et al., 2006).

In sum, distributed processing, as defined by GAG models, means that fixation durations reflect processing demands of the fixated word as well as processing of neighbouring words. The current section focused on the influences of parafoveal-on-foveal effects and their dynamical modulation through foveal load. In chapter 3 it will be argued that the size of the perceptual span also depends on WM processes. First evidence will be presented that variations in working-memory load also lead to variations in parafoveal preprocessing.

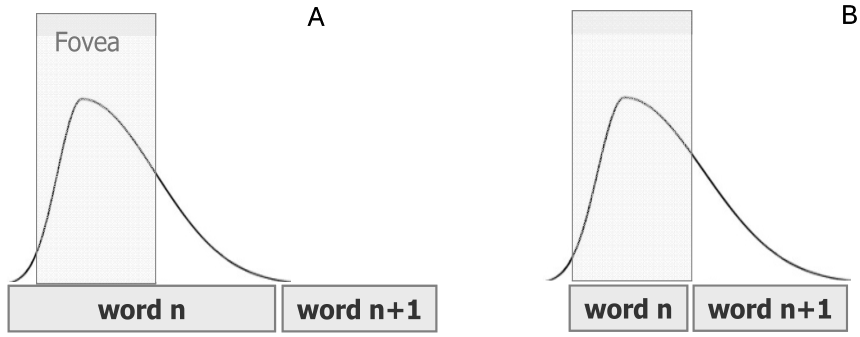

Figure 1.2: Illustration of the interrelation between word length and perceptual span: A) Word properties of word n+1 does not affect fixation durations on a long word n, because in this case word n+1 does not fall into the perceptual span and hence it can not be preprocessed. B) Word properties of word n + 1 affect fixation durations on a short word n, because in this case word n + 1 falls inside the perceptual span so that it can be preprocessed.

1.2 The working memory - an overview

Up to this point only little information about the construction of WM were given. More details and an overview of several theoretical models are supported in the current section. In the introduction the WM was described as a system which simultaneously stores and processes information. This is a common description, which is often used independently from a concrete theoretical assumption. However, the relation between storage and processing depends on the specific WM model. All models basically define WM as all processes and structures, which are essential for the current task or goal. It coordinates and evaluates strategies or processes through additional control mechanisms, like attention or executive control. As already mentioned above, attention also plays an important role in eye guidance. I suppose that attention is the underlying control mechanism, which associates models of eye-movement control with processes of WM. For the description of this coherence, one concrete WM model is not necessarily needed. Yet, it is essential for the understanding of the WM and its limitations, to give a short overview of several models and their underlying assumptions. Thereby I concentrate on model conceptions which are discussed to be important for language comprehension and reading.

1.2.1 Attentional processes as similarity of all working-memory models

Attention and activation make up the key focus of WM in concepts of cognitive psychology. The kind of attention resources is controversially discussed. Daneman and Carpenter (1980) for example act on the assumption of a uniform system, whereas Baddeley (1986) postulates relatively independent and domain specific components. These assumptions go back to the view of an independent system, which strictly separates and coordinates information from short-term and long-term memory (see Atkinson & Shiffrin, 1968). In contrast Cowan (1999) emphasizes information to be a kind of functional condition. Whether information lies in WM depends on the concrete condition, and thereby attention plays a crucial role in his model.

1.2.1.1 Limitations in attentional resources

In the sufficiently known three-component model from Baddeley and Hitch (1974) a central executive serves as an attentional control system. It controls the information flow from short-term to long-term memory. In addition it consists of two modality-specific subsystems on the level of short-term memory, the visual-spatial sketchpad for visual information, and a phonological loop. The phonological loop is storage for speech-coded information that decays in the order of 2 seconds, but the information can be refreshed by subvocal rehearsal. Baddeley (1986) considered the central executive to be a pure attentional system, without an own storage capacity. Later he revised his model (Baddeley, 2000) by adding the episodic buffer as a memory module for the central executive. This was due to research in dual-task paradigms, which points out that the central executive fulfills also own memory and process functions (Toms, Morris, & Ward, 1993). In contrast to Baddeley, further developments by Logie (1995) center on the visual-spatial subsystem. At this point I only want to mention that Logie supposed that the sensori input first reaches the long-term memory. Only if information is needed, is it transferred by the central executive to one of the subsystems. As described above, Baddeley's model was developed with a strong focus on verbal material, and for this reason it was often used to explain results of sentence-comprehension studies.

Just and Carpenter (1992) even go a step further. The assumption of their model roughly corresponds with the linguistics part of Baddeley's (1986) central executive. But the authors postulate an independent WM for language comprehension. According to their READER-model the WM combines information from long-term memory with the textual information which is essential for sentence comprehension. The capacity in that model is also restricted, and thus processes of storage and processing are limited. For example, common information is called up automatically and hardly requires capacity. In contrast specific information, like the meaning of rare foreign words, must be strategically inserted. Thereby much capacity is needed and that results in a reduced productivity of WM. In other words, limited processing resources are responsible for differences in sentence comprehension.

Recently Caplan, Waters, and DeDe (2008) reviewed the results of sentence comprehension studies and concluded that the findings actually support specialised resources for text comprehension. Thus, among other things, different resources for interpretive and post-interpretive processes are assumed (see also Waters & Caplan, 1996).

1.2.1.2 The focus of attention

In rather economical conceptions the WM is argued to be a part of long-term memory. In the models of Cowan (1999), Engle (1996) and Oberauer (2002) the activation of contents in long-term memory is of outstanding importance. Cowan (1988, 1995) postulates a close connection between WM and the attention system. Moreover, he proposed that knowledge is passively presented in long-term memory. A part of the knowledge can be activated by external stimuli, goals or other contents. Within this *activated part of long-term memory* a *focus of attention* is embedded. Information automatically obtains entrance into the focus, if it diversifies or if it receives attention. The information in the focus is conscious. Cowan (1990) emphasizes a limited capacity for the focus to 4 elements. If elements that have to be maintained are outside the focus, they are susceptible to decay. Therefore, according to Cowan, limited resources can be explained by limitations of capacity and decay.

Oberauer (2002) specified the Cowan model by adding a third component. The part, which Cowan calls focus of attention, is equated as region of direct access with a limitation to 4 elements. In contrast, the focus of attention is more restricted and holds only one element at a time. This view was supported by results of memory updating tasks, where participants showed longer reaction times for object switch than for no switch conditions (Garavan, 1998; Oberauer, 2002, 2003). Hence, it needs time to shift the attentional focus to the next object to be processed. Among other things, the degree of activation of an object determines the speed and accuracy of the object switch.

Moreover, the efficiency of WM depends on the individual ability to inhibit irrelevant stimuli (Engle, 1996; Engle, Kane, & Tuholsky, 1999). Neurocognitive models also postulate this connection and argue that the amount of inhibition is controlled by the dopaminergic neurotransmitter system (e.g., Braver, Barch, & Cohen, 1999; Dreher & Burnod, 2002; Durstewitz und Seamans, 2002). The link to additional WM concepts of other research fields is only possible by parsimonious and rather process oriented conceptions as provided by Cowan, Oberauer, or Engle. Accordingly, WM is understood as a system which flexibly guides attention, or in the words of Oberauer (2010, p. 278, l. 31): „ (…) WM is an attentional system“. This understanding of WM also leaves much room for the implementation in models of eye guidance.

1.2.2 Three theoretical assumptions of capacity limitations

As already evident from the models described above, the WM is limited in its capacity to 2 till 4 elements, depending on the material or the task (Alvarez & Cavanagh, 2004; Cowan, 2005). The term capacity limit refers to the observation that people's performance declines rapidly with an increase in memory demand. As a result, information is lost, forgotten or only incorrectly be recalled. A suggestion for the cause of forgetting was made with three approaches of explanation.

Capacity limits were traditionally explained by resouce limitations (e.g. Just & Carpenter, 1992). Theories predicted a limited pool of resources which must be shared for all memory representations and processing tasks. The more representations have to be maintained the less resource is available for each particular one. In the case of high memory demands, resources must be divided. This reduces the probability of correct recall for individual representations.

The second theory proposed that WM traces decay rapidly over time despite consolidation and storing (Baddeley, Thomson, & Buchanan, 1975; Barrouillet, Bernardin, & Camos, 2004; Page & Norris, 1998, Kieras, Meyer, Mueller, & Seymour, 1999). The model of time-based decay was supported by different findings. One of them was, for example, that not the number of processing steps in a complex task was important, but the time each step required (Barouillet et al., 2004). In consequence, memory declines when each processing step requires more time.

According to the third proposal, forgetting is caused by interference of similar representations. Thus, active representations interfere with each other, which can lead to the confusion of whole items or partial overwriting of the remaining features of item representations. The first form refers to the fact that in a serial list of items presented; sometimes the serial positions are interchanged. As a result the interchanged items are recalled for incorrect positions within the list. This is especially the case for neighbouring positions (Burgess & Hitch, 1999) and phonologically similar words (Henson, Norris, Page, & Baddeley, 1996). A second form is based on the principles of the feature model of Nairne (1990) and Neath (2000). Thus, items are represented by a set of features. For example a red triangle is represented by the features of colour, “red”, and the shape, “triangle”. In line with the idea of the model, a second object with the same colour overwrites the colour feature, which results in a loss of this feature for the triangle. This in turn degrades its representation and reduces its recall probability.

Recently the theories of time-based decay and interference were highly discussed and compared to each other. Most of the studies gave strong evidence for the interference approach (e.g., Lewandowsky, Duncan, & Brown, 2004; Lewandowsky, Geiger, Morrell, & Oberauer, 2010; Oberauer & Lewandowsky, 2008; Oberauer & Kliegl, 2001).

1.2.3 Individual differences in working-memory capacity

The WMC is not only dependent on the material or the task, as previously described, but in addition it greatly differs between individuals. This was established by a large body of studies, which used a variety of tasks. One of the first tasks to measure WMC was the "reading span" by Daneman and Carpenter (1980). Subjects read a sequence of sentences while they have to remember the last word of each sentence. At the end of a trial, all words have to be remembered in the correct serial order. The reading-span was often used in language-comprehension literature to classify the readers due to their WMC. It combines a memory span measure with a concurrent processing task. This dual-task paradigm, referred to as "complex-span", became a commonly used experimental design. In addition, also tasks that do not have this dual-task nature have been shown to be good measures for WMC (Oberauer, Süß, Schulze, Wilhelm, & Wittmann, 2000). Depending on task results, subjects were roughly divided into high or low-spans. Thus, high-spans are defined as subjects with a high WMC where as low-spans have a low WMC.

One important factor for the differences between subjects seems to be the efficiency of inhibitory processes (see Redick, Heitz, & Engle, 2007 for a review). High spans in contrast to low-span individuals are thus better able to deal with interference and inhibit task-irrelevant information (see Unsworth, Heitz, & Engle, 2005 for a review). This is the case if irrelevant information for example items of a second list have to be inhibited, so that relevant list items are not interfered with by similar but irrelevant items of another list. Moreover, keeping in mind that WMC is related to attentional control, this means that individuals differ in their ability to focus (e.g. on the relevant list) and maintain attention.

1.2.3.1 … and attentional control

Support for the assumption that individual differences rely on the ability to efficiently control attention, came from results of the dichotic-listening task (Conway, Cowan, & Bunting, 2001) and the anti-saccade task (Kane, Bleckley, Conway, & Enlge, 2001; Larson & Perry, 1999). In the first task, participants were required to monitor a message presented to one ear while ignoring a message presented to the other ear. Most of the participants showed no difficulty monitoring one channel at a time. However, 33 % were interfered by presenting a powerful attentional orienting cue, like the own first name, in the irrelevant ear. This is known as the “cocktail party” effect, which was first reported by Moray (1959). In addition, Conway et al. (2001) showed that 65% of participants classified as low spans showed this effect; whereas only 20% of the high spans reported hearing their names. The results suggest that high spans, in contrast to low spans, are more able to resist an interfering attention capturing cue when it conflicts with task goals (for a review see Unsworth et al., 2005).

The results from the anti-saccade task leads to the same conclusion. The anti-saccade task (Hallet, 1978) requires subjects to fixate on a central cue. After a variable amount of time, a flashing cue appears either to the right or left of fixation. Participants have to shift their attention and gaze at the opposite side of the screen as quickly and accurately as possible. In this experiment the automatic orienting response conflicts with the task goal. Low-span participants made more errors by showing reflexive saccades to the flashing cue than high-span participants (Kane et al., 2001). The findings can be explained by two assumptions. First: Low-spans were particularly deficient in their ability to maintain the task goal in active memory, and second: Even when the task goal was maintained, low spans were slower to implement control and thus resolve the conflict between the automatic orienting response and the task goal. Together with results of several other anti and prosaccade studies (see Unsworth et al., 2005 for a review), it seems quite reasonable to conclude that individuals differ in their ability to control attention. This is also most apparent when active maintenance is needed in conditions of interference.

Transferred to the reading situation, one would expect better comprehension results for high-span readers than for low-span readers. During reading, attention is focused on the current word while maintaining previous word information for comprehension. In line with the assumption that the ability to control attention and maintain information relies on the individual WMC, high-span individuals show better comprehension results than low-span readers (e.g. Miyake, Just, & Carpenter, 1994). Comprehension problems are most likely, if two competing interpretations are possible and have to be remembered until the ambiguous phrase is resolved. Individual differences in comprehension are thus mainly observed for ambiguous sentences (e.g. Caretti, Borella, Cornoldi, & DeBeni, 2009; Engle, Cantor, & Carullo, 1992).

1.3 The link between eye movement control and working memory

The WM is assumed to be the place where the information integration of the currently read words in a sentence appears (Baddeley 1992, Ericson & Kintsch, 1991). Yet, research on the influence of WM on the interword eye-movement behaviour is missing. As described above, research on the functional role of WM during reading relies most exclusively on the reading span (Daneman & Carpenter, 1980), where WM load is manipulated over several sentences. To measure the eye-movements of reading under such complex task demands, it is necessary to link current eye-movement research to theoretical developments in the areas of WM, reasoning, and intelligence.

1.3.1 The Reading Span Task

Complex tasks, as the reading-span task, measure a domain-general construct, which is partly responsible for the performance of a variety of cognitive processes, like executive attention or processing speed (for a review, Unsworth & Spiller, 2010). It was shown to highly correlate with the construction of WM (Oberauer et al., 2000) and has been shown to relate to numerous individual differences such as reading comprehension (e.g., Friedman & Miyake, 2004) as well as psychometric intelligence and age (e.g., Kemper, Crow, & Kemtes, 2004). The reading-span was introduced by Daneman and Carpenter (1980, 1983), who used the task to search for the impact and nature of individual differences in WMC. In the original version, participants read sentences aloud while they were presented one after another on a computer screen. The sentence presentation therefore depended on the individual reading speed. A trial consists of a sequence of sentences. The set sizes increased from three to five sentences, with three trials for each set size. During the task participants were then able to predict the length of a set. That could lead to adjusted reading strategies already at the beginning of a trial. The idea of presenting the set sizes randomly was introduced by Cantor, Engle, and Hamilton (1991). The randomized presentation of set sizes prevented the use of strategies, which come along with the knowledge of set size, and as a consequence of this, lead to a higher variance of results. Early presentations of difficult trials can be frustrating for some participants, especially for older ones. A progressive increase in set size is therefore equally reasonable like a random presentation. Another innovation is to use unrelated items (e.g. Kane et al. 2004) and to print them behind the sentences. This allows much greater control over the nature of the item. In this way digits or spatial cues can be used instead of words, and the influences of different domains can be explored. Engle, Sedek, von Hecker and McIntosh (2005) showed that all these versions of reading-span do not seem to differ in their correlations with other complex WMC tasks and with criterion tasks, such as reading comprehension.

1.3.2 Overview of earlier studies

As clarified in the previous section the reading-span task is an important measure of WMC and very often used in the field of WM research. It has been broadly investigated in more than 1730 studies, but only three of them additionally investigated the eye-movement behavior.

1.3.2.1 Correlation studies

Results of a correlation study by Just and Carpenter (1992) gave evidence for both influences of WM load and WMC on eye movement measures. Increases in task complexity and decreases in reading competence covary with longer fixation durations, more refixations and regressions, and fewer word skippings. The authors manipulated the complexity of a reading task by using object and subject relative sentences. The complex task condition results in longer reading times per word, in contrast to the low complexity condition. The WMC of the participants, defined by the reading span score, had an influence only on the high complexity condition. Just and Carpenter extended the initial observation of Daneman and Carpenter (1980) and described a model of individual differences in reading comprehension. The "capacity limitation model" (Just & Carpenter, 1992) emphasizes the role of limited processing resources in reading comprehension.

Two following studies either showed effects of WM load or capacity. Kennison and Clifton (1995) demonstrated effects of WMC on sentence reading measures. Their participant groups showed significant differences in that low spans show longer total reading times per sentence, more forward fixations, more regressive eye-movements, and longer gaze durations (see also Calvo, 2001). In addition Kennison and Clifton wondered whether low WMC, indexed with reading span would correlate with reduced perceptual span, but they could not support this hypothesis.

1.3.2.2 An experimental study

The only study examining eye movements during the Rspan task was done by Kaakinen and Hyönä (2007). The authors found evidence for memory load influences but failed to replicate the influences of capacity groups of Kennison and Clifton (1995). Their results showed small effects on memory load especially on the target word (sentence-final word), where total fixation times increased with load. The total fixation times in the middle part of a sentence decreased from the no load to the load condition only slightly. The fixational eye-movements did not differ between the capacity groups, so that the distributed processing times within the sentences were the same for both groups. Individual differences were only reported for the use of strategy. High WMC subjects reported more use of semantic encoding strategies in contrast to the low WMC group. Kaakinen and Hyönä (2007) discussed that their results support the knowledge-is-power hypothesis (McNamara, & Scott, 2001), which postulates that some people are more aware of efficient memory encoding strategies than others. Because participants did not differ in there fixation durations the authors propose that the high spans can make use of these strategies without additional processing time. One reason for the unexpected results could be that Kaakinen and Hyönä (2007) looked only at very small set sizes with a maximum of 4 sentences. At these set sizes neither the high spans nor the low span subjects are above their WMC. Both groups have enough capacity available to resolve a task set of 4 sentences. By looking at set sizes above 4, where low capacity groups have to work above their WMC, the groups should show differences in processing time and their fixation durations.

1.4 Summary

Language comprehension, and thereby reading, is a higher-order cognitive activity that involves many processes: Words must be identified, single sentences must be syntactically and semantically analyzed and information from different sentences must be related and integrated into a complete story. Oculomotor processes occur in parallel to this cognitive activity and need to be synchronized to ensure that the relevant input is available.

In this chapter I focused on the fundamental assumptions of working-memory models and models of eye-movement control. Thereby I emphasized that the key focus in both research fields lies on attentional processes. Futhermore, I supposed that this understanding leaves much room for the implementation of working-memory assumptions in models of eye-movement control. The present work does not account for this implementation, but it offers a first step in the correct direction.

The experimental and quasiexperimental manipulation of WM load and WMC opens the opportunity to directly measure the impact for eye-movement patterns.

Until now assumptions of working-memory theories have not been integrated into models of eye guidance. However, several studies have shown the influence of working-memory for sentence-comprehension. In addition, three studies investigated the influence of WM load and capacity on eye-movement patterns, but the results are inconsistent so far.

Research on the functional role of WM during reading relies most exclusively on the reading span task (Daneman & Carpenter, 1980), and although the task was broadly investigated, only one study (Kaakinen & Hyönä, 2007) monitored the eyes during this task. The authors solely used very small set sizes and so failed to replicate the influence of WMC. Thus, it remains ambiguous, for example, whether the probability of refixating the to-be-remembered word or the total time spent on the critical words are predictive of reading span. Similarly, the question is unresolved, whether WM load in reading-span tasks reduces the perceptual span, as expected by proponents of the foveal-load hypothesis.

With the present experiment, I registered eye movements during the reading span task, to shed light on how WM is used for reading. The question asked in this work is: How do WM load and WMC affect lexical access (e.g. frequency effects), memory retrieval (e.g., predictability effects), or spillover and parafoveal-on-foveal effects, that is, eye movements during reading? I expect the results to support theoretical proposals that the perceptual span is dynamically modulated by foveal and possibly parafoveal processing difficulty. In the first step, I am interested in global effects of WM load and capacity (chapter 2). In the following I concentrate on the analysis of parafoveal effects (chapter 3) and the influence of age as additional quasiexperimental manipulation, to enhance the variability of WMC (chapter 4).

Chapter 2

Global effects of working-memory capacity and load

The aim of the present chapter is to examine the impact of WM on eye movements during reading. Thereby a best practised WM task, that is, the reading-span task, is established in the field of eye-movement research. The present experiment represents one of the first attempts to investigate eye movements in this task. Examining eye movements in the reading span task is of great importance for both eye-movement and working-memory research. To know where exactly readers guide their eyes during this task may inform about the use of implicit strategies and may explain WMC differences between individuals.

In the reading span task subjects read a number of sentences while they have to remember the last word of each sentence (target word). The number of target words that have to be remembered increases from sentence to sentence and thus increase the WM load. The number of words correctly recalled at the end is used as an indicator for the individual WMC of a participant. Hence, with the reading span task one can determine how eye movement patterns change for individuals with different WMC, when WM load is increased.

There have already been several attempts to relate WM and reading research, most notably in reading comprehension (e.g. Caretti et al., 2009; Engle et al., 1992). Longer reading times for low WMC subjects were mainly observed for ambiguous sentences. The results can be explained by the compensatory encoding model (Walczyks, 1993, 1995). The model focuses on the influence of WMC during reading. It assumes that low-capacity readers use compensatory mechanisms to overcome word encoding problems and inefficient lexical access. Walczyks differentiates between automatic processes and control processes. According to the author, automatic processes are the result of extensive practice, they put minimal demand on attention and working memory. Whereas, control processes occur for ambiguous sentences, lexical difficulties, and, in general, in difficult texts. Thus, longer sentence-reading times and more look backs in the text are predicted for low but not for high WMC subjects, but only if control mechanisms are necessary, which is the case in ambiguous sentences.

During comprehending ambiguous sentences, low-capacity readers appear to have only the dominant interpretation available, whereas high-capacity readers maintain multiple interpretations of an unresolved lexical ambiguity (e.g. Miyake et al., 1994). These results suggest that readers with a lower WMC might have fewer capacities available for the maintenance of information during sentence comprehension. Thus, they indirectly suggest that individual differences in WMC are reflected by differences in reading behaviour. According to Walczyks predictions low-capacity readers should compensate the maintenance of an incorrect interpretation by regressions to previous text passages, which lead to longer reading times.

The influence of WM on eye-movement measures is inconsistent so far. In an experiment by Just and Carpenter (1992) WM load was manipulated by the task complexity, by using subject-relative (low complexity) versus object-relative (high complexity) sentences. Their results provided evidence for both, influences of WM load; manipulated by task complexity; and WMC; classified by the reading span task. Longer mean reading times per word in object-relative sentences were observed. The effect was strongest for low WMC subjects. Later studies either showed effects of WMC or load. Kennison and Clifton (1995) classified readers in high and low WMC groups on the basis of the reading-span task (Just and Carpenter, 1992). Capacity scores were correlated with eye movement reading measures. Low WMC subjects showed more fixations, as well longer total reading times and gaze durations. The only study examining eye movements during the reading span task was tried out by Kaakinen and Hyönä (2007). The authors found evidence for WM load influences but failed to replicate WMC differences. Total fixation times increased with load, especially on target words. Kaakinen and Hyönä used an adaptive test version, in which the task was aborted when subjects reached their capacity limit.

To clarify the inconsistencies in previous results, the present experiment investigates how subjects with different capacities behave during the reading span task, if they work below, on, or above their WMC. Moreover, the influences of a step by step increase in WM load on eye-movement patterns, is examined. In the reading span task, WM load is progressively increased within a trial. This provides for the opportunity to measure directly the impact of the increase on eye-movement measures. Therefore, I analysed typical measures of eye-movement research for sentence reading times and measures on the target word: total reading time, gaze durations, and percentage of regressions. Maintaining target words during sentence reading should place extra demand on WM. For that reason reading times are expected to increase with the number of target words maintained in WM, that is with WM load. Moreover, the number of forward fixations and regressions should increase with load. Individual differences are expected to occur, particularly in high load conditions where low-capacity subjects work on, or above their WMC limit. The following experiment investigates the impact of working-memory capacity and working-memory load on global eye-movement patterns.

2.1 Method

2.1.1 Subjects

Thirty-three students, native speakers of German with normal or corrected to normal vision, participated in the experiment. They were paid six Euros or received course credit. Data of five subjects was excluded from analysis because they ignored instructions. The 28 remaining subjects were 23.38 years on average (range: 19 to 31 years).

2.1.2 Apparatus

Sentences were presented at a distance of 60 cm centred on a Iiyama Vision Master Pro 514 monitor (1024 x 768 resolution; 21 in.; frame rate 150 Hz; font: regular New Courier 17; visual angle: 0.38° per character). Heads were positioned on a chin rest to minimize head movements. Both eyes were monitored with an EyeLink II system (SR Research, Osgoode, ON, Canada) with a sampling rate of 500 Hz and an instrument spatial resolution of 0.01°. Because of system problems, six participants were monitored with a sampling rate of only 250 Hz. The experimental software was implemented in Matlab (MathWorks, Natick, Massachusetts, USA), using the *Psychophysics* (Brainard, 1997; Pelli, 1997) and *Eyelink* (Cornelissen, Peters, & Palmer, 2002) toolboxes. Data preparation up to the fixation sequence, including blink correction and saccade-detection algorithms, were also implemented in Matlab. The R system (version 2.11.0 R Development Core Team, 2010) under the GNU General Public License (Version 2, June 1991) was used for the calculation of first pass measurements and other parameters, as well as for the statistical computing and plotting. The statistical analysis was based on linear mixed models and was done with the *lmer* and *glmer* program of the *lme4* package (Bates & Maechler, 2010). Other R-packages which were used for plotting and data aggregation were *reshape* (Wickham, 2007) and *lattice* (Sarkar, 2008).

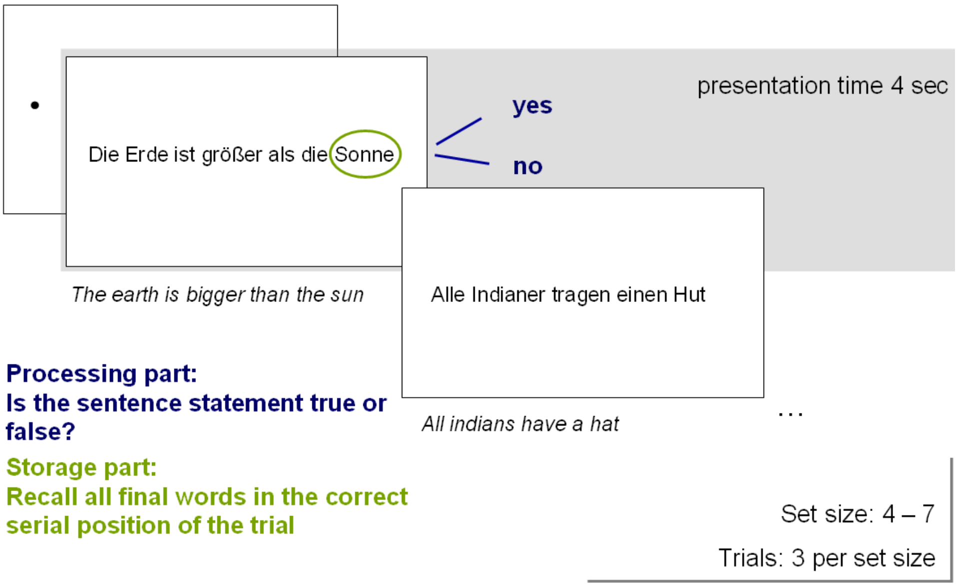

Figure 2.1: Illustration of the reading span task: Sentences were presented one after another centred on screen. Each sentence was presented for four seconds. During that time participants gave a judgement based on the sentence content and memorized the sentence final word. Set sizes increased from four to five sentences, with three trials for each set size.

2.1.3 Material

The sentence material was taken from the German version of the reading span task (Oberauer et al., 2000) and is listed in Appendix A. The sentences vary in length from four to seven words (mean = 5.14). They were presented without punctuation like the example shown in Figure 2.1. Sentences were presented within a set containing four, five, six or seven sentences. The number of sentences within one set denotes the set size. Therefore the serial position of sentences (or targets) was nested under set size.

The task can be conceptualized as consisting of two parts: the processing and the storage part. In the processing part participants read a set of sentences. Each sentence was followed by a judgement based on sentence content. Participants had to judge whether the sentence statement was true or false by pressing the “less than” or the “dash” key, respectively. There were an equal number of sentences with true and false statements. In addition subjects had to remember the last word of each sentence (i.e., the target word). This constituted the storage part of the reading span.

All target words were nouns and none of them was repeated across sentences. Sentence order was constant across subjects.

2.1.4 Procedure

The sentences were presented in black on a white screen. Font was Courier. A calibration started before each trial. All recordings and calibrations were binocular. Minimal head movements were corrected automatically by the Eye-Link II system. To reduce head movements the head was positioned on a chin rest. Subjects were calibrated with a standard nine-point grid for both eyes. After validation of calibration accuracy, a fixation dot appeared on the left side of the centerline on the monitor. Depending on the center location of the following first word, the fixation dot position was determined. If the eye tracker identified a fixation on the fixation spot, the first sentence of a trial was presented. To stay close to the original setup of the reading span task, fixation location was only controlled before the first sentence of each trial. The presentation time of the sentences in a trial was computer-paced. Every 4 seconds a new sentence was presented on screen, independent of whether the processing part was finished or not (i.e., whether a key was pressed or not). At the end of each trial, the request for oral recall of the target words appeared on screen. The answers were recorded by the experimenter. Following two practice trials, with set size three, testing began with two trials consisting of four sentences, followed by three trials each of the set sizes five to seven. The number of sentences within a trail and thus target words, depended on the set size.

Scoring: Following Daneman and Carpenter (1980) a span score was calculated for each participant according to the number of correctly recalled target words. Target words were only regarded as correctly recalled, if they were recalled in the correct serial position. Span score was defined as the value of the largest set size for which the subject was able to recall all target words of at least two thirds of the trials correctly. For example, a person with a span score of five was able to correctly recall all targets of at least two thirds of the trials of set size four and five, but was not able to correctly recall targets of more than one trial of set size six. According to their span score, subjects were divided into four capacity groups, with group numbers reflecting the achieved span score. The distribution of subjects across groups was the following: Capacity group 4, n = 8; group 5, n = 6; group 6, n = 6; and group 7, n = 8.

Blink parameters. In 45 % of the trials a blink occurred causing a loss of measurement. Those data losses were interpolated by using a new tool I developed and of which a detailed description is provided in Appendix B of this thesis.

2.1.5 Data Selection and Statistical Analysis

Forward-saccade criterion. First-pass reading traditionally consists of all initial forward fixations in a sentence. Second-pass reading, however, includes all fixations after a regressive eye movement on those parts of the text that the eye had already passed during the first pass. In the present experiment fixation position was controlled only for the first sentence in a trial. Consequently readers did not start to read with the first word in subsequent sentences, but first jumped to the last word (i.e., the target word), then jumped to the first word to move their eyes across the sentence. Hence, a traditional first pass would include only the target word, after which all other words could only be considered as a regression, and as a result in second-pass reading. Therefore, the definition of first-pass reading was adjusted. First pass reading, as understood in the present experiment, begins with the first fixation on the left side of the sentence with an inter-word forward saccade. For the target words, first pass means the very first fixations on the target, independent of whether there was a prior fixation on the beginning of the sentence or not.

Table 2.1: Number of fixations of various types of fixations. Row 4 = 5 + 6 + 7 + 8; row 8 = 9 + 10 + 11

	N	%
1 N of sentences	1798	100
2 sentences with blinks	819	45.55
3 sentences with interpolated blinks	443	24.76
4 *N of fixations*	13,903	100
5 first word	1367	9.83
6 short/long fixations or amplitudes	2020	14.53
7 forward saccade criterion	2348	16.89
8 *N of valid fixations*	8168	58.75
9 not first pass	2071	25.35
10 multiple fixations	3051	37.35
11 single fixations	3046	37.29

Data selection. Eye-movement data was screened for measurement loss and unacceptable blinks in the data. The top part of Table 2.1 summarizes numbers and percentages of sentences with interpolated or unacceptable blinks. Data of sentences without blinks and with interpolated blinks was reduced to a fixation format using an algorithm for the binocular detection of saccades (Engbert & Kliegl, 2003). Saccades are detected, if their amplitudes are at least half a character space (0.4 deg). Analysis was based on right-eye fixation data. Fixations were assigned to letters. They were excluded when they met the following hierarchical set of criteria: (1) fixations after sentence processing, which means after the judgement of the sentence content (see below for a justification); (2) fixation on first word, (3) fixations shorter than 85 ms, and longer than three standard deviations above the mean cell score; fixations bordered by a saccade amplitude of 25 ore more letters, or by a within-letter saccade; and (4) fixations outside the forward-saccade criterion. The remaining fixations are valid within-sentence fixations. The fixation pattern after sentence processing (see point 1) describes a very unclear picture.

Transformation. Fixation durations as reaction times are normally distributed with a longer tail for long latencies (in Kliegl, Masson, & Richter, 2010; O’Regan, 1990). In order to fulfil the assumption of homoscedasticity data was transformed. Kliegl et al. (2010) discussed two types of transformations and showed that for their priming data speed-transformed reaction times (RTs) were in better agreement with the assumptions of homoscedasticity than log-transformed RTs. Thereby, the type of transformation had no great consequence on group effects, but turned out to be critical for the correlations of mean RT and experimental effects across subjects (in Kliegl et al., 2010). For that reason the lambda coefficient for the Box-Cox power transformation (Box & Cox, 1964) was estimated to decide which transformation is preferable for the present data. Lambda coefficients were 0.45 for sentence processing times and -0.6 for single fixation durations and gaze durations. Thus lambdas were justifying a log transformation.

Analysis. Inferential statistics were based on linear mixed models (LMM, for more details see Baayen, 2008; Kliegl et al. 2010) specifying subjects as random effect, and representing experimental manipulations as fixed effects. A fixed effect is a parameter mean of a linear equation for a certain experimental condition. The random effect represents the covariance of the fixed effect across individuals. Random effects are assumed to be normally distributed with a mean of zero around their fixed effects. For each individual the deviation from the fixed effect is predicted in reliability of the sample mean. That reduces the risk of overfitting the model to unreliable differences between individuals. Moreover, LMM has been shown to suffer substantially less loss of statistical power in unbalanced designs than traditional analysis of variance (Pinheiro & Bates, 2000). Hence, the choice of LMM accounted for the nested design of the reading span, where the serial position of targets and sentences is nested within set size. For the binomial dependent variables (e.g., regressions) the generalized linear mixed model (GLMM) was conducted, with the same random and fixed effect part specification as in the LMM.

2.2 Results

The fixed effects part of all LMM and GLM included the variables SET SIZE, and capacity group (CAPACITY). The factors were coded as linear contrasts. Therefore, nonlinear tendencies of low theoretical interest were ignored, and the monotonic trends were captured with maximal statistical power. Furthermore the set of predictors included comprehension accuracy (COMPREHENSION); which is the measure of accuracy in sentence judgement; and recall accuracy (RECALL), that is whether the target word of a sentence was correctly recalled in its respective serial position. This baseline model was adjusted by further variables, which will be reported in the appropriate place. The report of interactions among the variables will depend on the statistical power for detecting higher-order interactions. Generally, in a preliminary analysis, interaction terms that were not significant were eliminated. All models and their significant effects are listed in Appendix C.

The results of the LMM are interpreted on the basis of *t* values. A *t* value >= |2| approximates a significance level of alpha 0.05 for the fixed effects. This is equivalent to a coefficient magnitude of at least two standard errors. (see Kliegl et al. 2010; also for more details). In addition the estimates of the effect, *b*, and the theoretical range of the estimates, *SE*, are reported. For the GLMER Models the *p* values are additionally reported.

For the illustration of results real-time durations are plottet in the graphics. Graphical illustrations of effects are in line with data patterns of log-transformed and centred data. In the graphical illustrations error bars reflect the mean of the residual errors within one factor level. They are adjusted by the variance of the random effects. For within-subject comparisons, uninformative between-subject variance was removed (see Loftus, & Masson, 1994; Blouin, & Riopelle, 2005). Thus, error bars allow a first interpretation of the differences between the conditions.

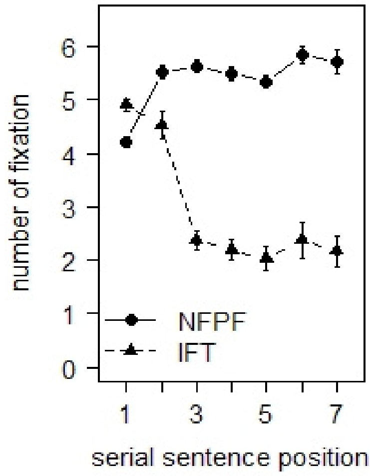

Figure 2.2 Mean number of fixations as function of the serial position of sentences within a trial. The two lines represent the mean number of first-pass fixations (NFPF) on a sentence and the mean initial fixation number for the target words (IFT).

2.2.1 Results of the storage part

The final word in each sentence serves as the target for the storage part of the task. Some global effects on these words are therefore strategy effects, because they are explicitly driven by the task demand. In the following result section, the statistical analysis concentrated on the target words and how WM load and capacity influence the initial fixation number, as well as first (gaze) and second pass fixation durations (total fixation times) on target words. Therefore, a comparison of fixation durations on the target and all words prior the target was also conducted. Moreover, the influence of recall accuracy of target words is considered in the last paragraph of this section.

The strongest effect (shown by all subjects) is that with increasing memory load they tend to start reading with the sentence-final word (see dashed line in Figure **2.2**). The set of predictors for the LMM on the *initial fixations on target* include the linear trend of memory load (LOAD). The LOAD effect was included within each set size and reached significance for all levels: *set size 4:* b = -1.009, SE = 0.117, t = -8.62; *set size 5:* b = -1.082, SE = 0.095, t = -11.44; *set size 6:* b = -0.766, SE = 0.111, t = -6.92; *set size 7:* b = -0.847, SE = 0.098, t = -8.66 (in Table C.1).

This means that within each set size the probability to fixate the target word first increases with memory load, which results in the decrease of the *initial fixation number*. However, this effect is also promoted by the procedure of sentence presentation. And hence can not completely be attributed to memory load. A fixation point occurs only before the first sentence in a trial. Due to this fact subjects are forced to read the first sentence in a trial from left to right. Subjects sustained the reading direction for the second sentence. From the third sentence on, the target word is fixated before the sentence is read as a whole. This is due to the fact that the sentence presentation was computer paced. The eyes of the subjects are still at the end of a sentence, when it is replaced by a new sentence. Hence, the new target is fixated before the eyes are guided to the sentence beginning. The difference in the *initial fixation on target*, between the first two sentences and the sentence positions 3 to 7 is significant with $b = -0.604$, $SE = 0.071$, $t = -8.495$ (in Table C.2). Nevertheless, to start reading with the critical word of the task indicates strategic allocation of processing resources. According to the interpretation Figure 2.2 and the LMM of Table B.2 refer to the serial position (SERPOS) of sentences instead of memory load. To give consideration to the nested design of the task, SERPOS, as LOAD, is assessed within each of 4 to 7 levels of set size.

2.2.1.1 Word position effects

In Figure 2.3 mean fixations durations are plotted as function of word position within a sentence (WORD POSITION), aligned on sentence-final, that is, target words. Fixation durations increase from sentence beginning to the end, with the longest durations for the target words (*total fixation time*: $b = 0.150$, $SE = 0.008$, $t = 19.67$, in Table C.3; *gaze duration*: $b = 0.118$, $SE = 0.006$, $t = 19.15$, in Table C.5). The mean gaze duration for target words was 384 ms ($SD = 285$) and 272 ms ($SD = 163$) for all other words (non targets). Longer duration on sentence-final words is known as sentence wrap-up effect, reflecting integration of information and comprehension of the sentence. In the present study they in addition reflect time for encoding the target word.

The interaction between CAPACITY and WORD POSITION in *total fixation times* was significant with b = -1.821, SE = 0.478, t = -3.81 (in Table C.3). The capacity groups 4 and 5 differ in *total fixation times* significantly from the capacity groups 6 and 7 in their slopes from the word prior to the target words (b = -0.287, SE = 0.067, t = -4.30, in Table C.4). In contrast to the other three groups, the highest capacity group showed significantly slower increases in *total fixation times* also from the word position n-2 to the word prior the target (b = -0.239, SE = 0.068, t = -3.53, in Table C.4). Compared to all other groups, the highest capacity group showed shorter total fixation times for the target word, the word one prior to the target (n-1) and the word at position n-2. Group 6 showed only reduced times on the target word compared to group 4 and 5. The interaction between CAPACITY and WORD POSITION was also observed for *gaze durations* (b = -0.911, SE = 0.385, t = -2.36, in Table C.5). Here the capacity group by word position difference is caused mostly by the lower gaze durations of capacity group 7. Compared to group 6 the highest capacity group shows lower increases of durations from the word prior the target to the target word (b = -0.195, SE = 0.085, t = -2.28) and from the word position n-2

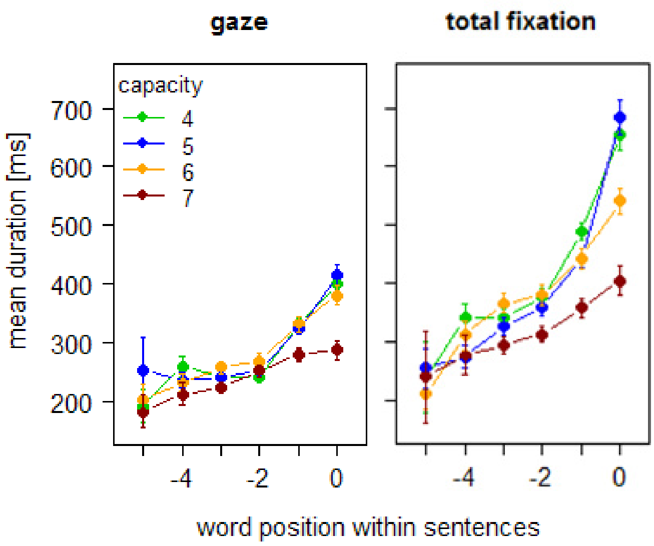

Figure 2.3: WMC differences on mean gaze duration and mean total fixation duration as function of absolute word position aligned at the sentence-final word (target).

to the word prior the target (b = -0.210, SE = 0.085, t = -2.47, in Table C.6). Group dependent variations of fixation durations on target words represent differences in sentence comprehension or a difference in memory encoding of the target. These explanations will be discussed below.

Both WM load and WMC show more pronounced effects on *total fixation times* than on *gaze durations*. This indicates that both factors affected second-pass reading more than first-pass reading. In the left panel of Figure 2.4 the mean duration on *second-pass fixations* is plotted as a function of WMC groups and word position. A similar picture as for gaze and total fixation durations is observed. WMC group 7 differs significantly from other capacity groups in *second-pass fixation durations* on the target (b = -0.101, SE = 0.035, t = -2.90, in Table C.7), but less on the word prior to the target. Moreover, WMC groups differ in the number of *regressions* (b = -0.342, SE = 0.066, t = -5.149) on the target word. This is illustrated in the right panel of Figure 2.4. Hence, differences between WMC groups in *total fixation durations* were a result of both, the number of regressions and the durations of second-pass fixations.

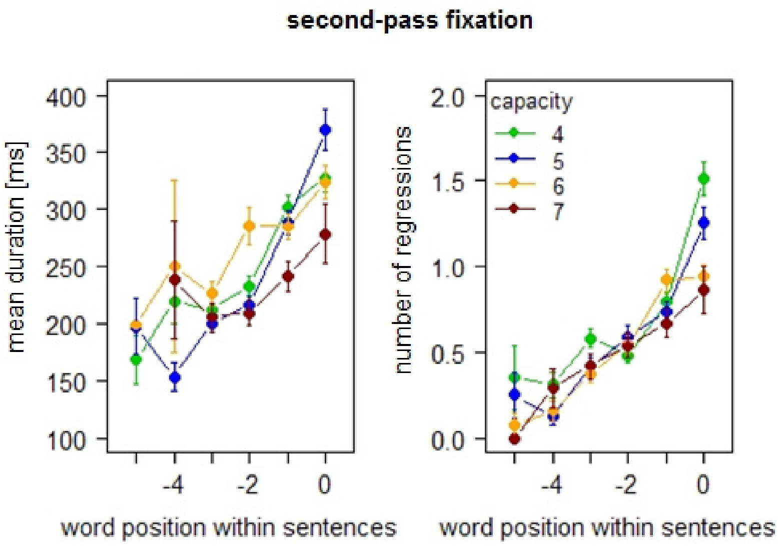

Figure 2.4: WMC differences on mean second-pass fixations durations and on the number of regressions as function of absolute word position aligned at the sentence-final word (target).

Effects of target recall

It was further interesting to see whether the later recall accuracy could be used to predict the previous sentence reading behaviour. The effect of recall accuracy (RECALL), that is whether the target word of a sentence was correctly recalled in its respective serial position, shows a significant influence on the *number of regressions*, as well as for *total fixation times* and *gaze durations*. The main RECALL effect on the *number of regressions* shows significantly more regressions (increase of 7 %) in a sentence, if the target is incorrectly recalled (b = -0.129, SE = 0.036, t = -3.60, in Table C.9). That means, sentences in which the target was later recalled incorrectly include more regressions, longer total fixation times and longer gaze durations. In line with this result, *total fixation times* (b = -0.049, SE = 0.020, t = -2.41, in Table C.3) are significantly longer (20 ms), due to the increase of regressions for sentences with incorrectly recalled target words.

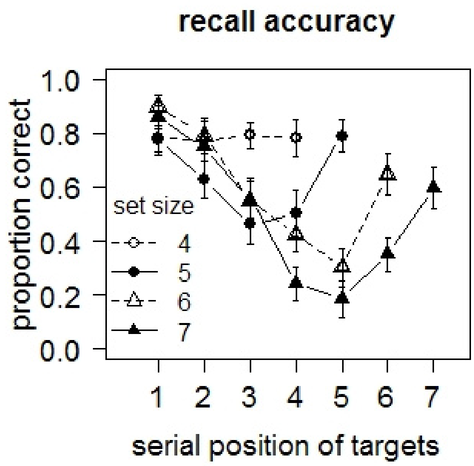

Figure 2.5: Mean proportion correct of target recall for each set size as function of serial position.

Figure 2.5 displays recall accuracy of the target word as a function of its serial position and set size. The curves show clear primacy and the recency effects (especially for higher set sizes), that is, higher accuracies for the target in the first and last list positions, respectively. A GLMM with *recall accuracy* as dependent variable, and SERPOS as one of the fixed effects, shows a significant quadratic trend of serial recall (SERPOS) with $p < .001$ for *set size 5*: $b = 1.593$, $SE = 0.180$; *set size 6*: $b = 1.427$, $SE = 0.178$ and *set size 7*: $b = 2.217$, $SE = 0.189$ (in Table C.10). The exception is recall accuracy of target words in set sizes 4, which does not differ between serial positions. With increasing SET SIZE the accuracy of target recall dramatically decreases ($b = -22.471$, $SE = 2.374$, $p < .001$, in Table C.10) especially for the middle positions. The interaction with capacity group was non significant.

2.2.2 Results of sentence processing

The following section focuses on the influence of WM load and capacity on number of first pass fixations, sentence processing times, total fixation times, gaze durations and regression rates. The last paragraph in this section concentrates on the processing accuracy and its dependency on WM load and capacity.

Figure 2.2 shows the *number of first pass fixations* on the sentences (i.e., non-target words, solid line) as a function of serial position within set size. On average each sentence is considered with 5 to 6 fixations (mean = 5.23, SD = 1.25). The *number of first pass fixations* on a sentence significantly increases linearly with increasing SET SIZE, $b = 1.360$, $SE = 0.648$, $t = 2.10$ (in Table C.11). The mean number of first pass fixations per set size is listed in Table 2.2.

The sentences on the first position within each set size are considered with a reduced *number of first-pass fixations*. The contrast between the first serial position and all other sentence positions is significant with $b = 1.402$, $SE = 0.132$, $t = 10.66$ (in Table C.12). Two explanations are possible for the reduced fixation number of the first sentences. The first explanation follows the subject tendency to fixate the target as fast as possible. At the first sentence subjects are forced to read from left to right. And hence the number of fixations on the first sentence is reduced to an amount that is sufficient for comprehension. The second explanation, however, considers the magnitude of the WM load. During reading the first sentence no target

Table 2.2: Summary: Mean (M) and standard deviation (SD) of eye movement measures for each set size

Eye movement measures		**Set size**			
		4	5	6	7
Sentence processing time (ms)	M	2434	2361	2575	2474
	SD	*504*	*563*	*542*	*572*
Gaze duration (ms) *	M	279	307	321	304
	SD	*88*	*106*	*110*	*95*
Single fixation duration (ms) *	M	231	246	253	253
	SD	*72*	*93*	*77*	*96*
Number of first pass fixation *	M	1.41	1.49	1.58	1.52
	SD	*0.31*	*0.28*	*0.28*	*0.29*
Probability of skipping	M	.20	.23	.19	.22
	SD	*.19*	*.17*	*.16*	*.16*
Probability of regression *	M	.44	.49	.52	.51
	SD	*.23*	*.22*	*.20*	*.19*
Initial fixation on target *	M	3.59	3.11	3.07	2.55
	SD	*2.40*	*2.42*	*2.28*	*1.79*

Note: Inferential statistics are based on linear mixed models specifying participants as random effect. Results are interpreted on the basis of the t value.
* The linear trend of set size was significant ($t > |2|$)

word has to be maintained. Therefore, the fixation number on the first sentence is not reduced, but the number on the other sentences increase in consequence of the WM load. Support for the last explanation comes from the second sentences. Although the reading direction is the same as for the first sentences the number of first pass fixations per sentence significantly increase from the first to the second position ($b = 1.4332$, $SE = 0.141$, $t = 10.18$, in Table C.13). That means the difference between the first and the second sentence position represents the difference between no load and load 1. In addition the number of fixations increases for position 6 and 7 compared to 1 to 5 ($b = 0.450$, $SE = 0.172$, $t = 2.62$, in Table C.12). Transferred to the WM load the level difference occurs from load 4 to load 5. That result is in line with the assumption that the WM has a capacity of about four chunks (Cowan, 2001). Higher memory demands, therefore, result in more fixations.

Moreover, the number of first pass fixations on a sentence significantly increases linearly with increasing WM load. This is true for each of the four set sizes (*set size 4:* b = 0.402, SE = 0.108, t = 3.71; *set size 5:* b = 0.469, SE = 0.087, t = 5.37; *set size 6:* b = 0.261, SE = 0.091, t = 2.87; *set size 7:* b = 0.308, SE = 0.090, t = 3.41; in Table C.11).

Table 2.3: Summary: Mean (M) and standard deviation (SD) of eye movement measures for each capacity groups

Eye movement measures		**Capacity group**			
		4 (n = 10)	5 (n = 6)	6 (n = 7)	7 (n = 6)
Sentence processing time (ms) *	M	2682	2531	2294	2029
	SD	*499*	*493*	*517*	529
Total viewing time (ms)	M	456	451	434	344
	SD	*131*	*122*	*134*	*118*
Gaze duration (ms)	M	307	312	320	266
	SD	*89*	*82*	*129*	*91*
Single fixation duration (ms)	M	243	251	265	224
	SD	*78*	*82*	*107*	*66*
Number of first pass fixation	M	1.19	1.17	1.11	1.13
	SD	*0.70*	*0.69*	*0.67*	*0.69*
Probability of skipping	M	.22	.20	.23	.17
	SD	*.17*	*.15*	*.17*	*.17*
Probability of regression	M	.51	.48	.51	.45
	SD	*.21*	*.20*	*.20*	*.24*
Initial fixation on target	M	3.38	2.79	3.13	3.29
	SD	*2.44*	*2.20*	*2.09*	*2.42*

Note: Inferential statistics are based on linear mixed models specifying participants as random effect. Results are interpreted on the basis of the t value.
* The linear trend of capacity group was significant (t > |2|).

Influences of working-memory capacity

S*entence-processing time* represents the reaction time of the processing part, when subjects pressed a button to decide whether the sentence statement is true or false. Sentence-processing times decreases with increasing WMC ($b = -3.728$, $SE = 1.366$, $t = -2.7$, in Table C.14). Table 2.3 summarizes the mean fixation durations for the different capacity groups.

Effects of working-memory load

In the following I focus on WM load effects on total fixation times, gaze durations and regression rates. The mean *total fixation times* linearly increased from 371 ms ($SD = 89.45$) at memory load 1, to 526 ms ($SE = 121.09$) at memory load 7 with a small drop from load 6 to 7. Figure 2.6 A, displays the mean *total fixation durations* (solid line) aggregated across all set sizes, as function of memory load. The main effect of LOAD, that is the increase of total fixation times with increasing memory load is significant within three of four set sizes: *set size 4:* $b = 0.147$, $SE = 0.051$, $t = 2.87$; *set size 5:* $b = 0.289$, $SE = 0.047$, $t = 6.10$; and *set size 7:* $b = 0.204$, $SE = 0.047$, $t = 4.33$ (in Table C.15). Moreover, the linear trend for SET SIZE, as indication of memory demand, is significant ($b = 1.703$, $SE = 0.639$, $t = 2.67$). That is, total fixation times in general are longer with larger set size. Furthermore, the individual WMC has no effect on *total fixation times*.

Similar results can be found for the *gaze duration* (see Figure 2.6 A, dashed line), with a significant LOAD effect (*set size 4:* $b = 0.108$, $SE = 0.041$, $t = 2.62$; *set size 5:* $b = 0.205$, $SE = 0.038$, $t = 5.39$; *set size 7:* $b = 0.155$, $SE = 0.038$, $t = 4.09$; in Table C.16) and a significant linear trend for factor SET SIZE ($b = 1.511$, $SE = 0.514$, $t = 2.94$).

However, the size of the LOAD effect on *gaze durations* is lower than on *total fixation times.* Gaze durations include all first pass fixations until the eyes left the word for the first time, whereas total fixations include all fixation durations on a word, until sentence processing was finished. Hence, the difference in effect sizes, across the two measures, points to the fact that WM load especially affected second pass reading more strongly.

Effects of WM load on the regression rates support inconsistence to the pattern of results reported so far: The higher the WM load the higher the probability of regression (see Figure 2.6 B). The linear trend of LOAD for regressions is significant with $p < .001$ for *set size 4:* b = 1.023, SE = 0.184; *set size 5:* b = 0.884, SE = 0.168; *set size 6:* b = 0.741, SE = 0.160; and *set size 7:* b = 0.557, SE = 0.161 (in Table C.17). There is a dramatic increase from load 1 to the load condition of 2. The contrast between load 1 versus load 2 is significant (b = 1.888, SE = 0.170, $p < .001$, in Table C.18). This strong increase in the regression rate is due to the presentation difference of the initial sentence. However, the regression rate increases until the load of 4 when it reaches an asymptote. A GLMM with repeated contrast specification shows a significant difference between memory load 1, 2, 3 and 4 versus load 6 and 7 (b = 0.972, SE = 0.207, $p < .001$, in Table C.18). Furthermore, the regression rate significantly increases with SET SIZE (b = 4.765, SE = 2.197, $p < .05$, see Table 2.3 and Table C.17). With increasing set size, regressions become more likely.

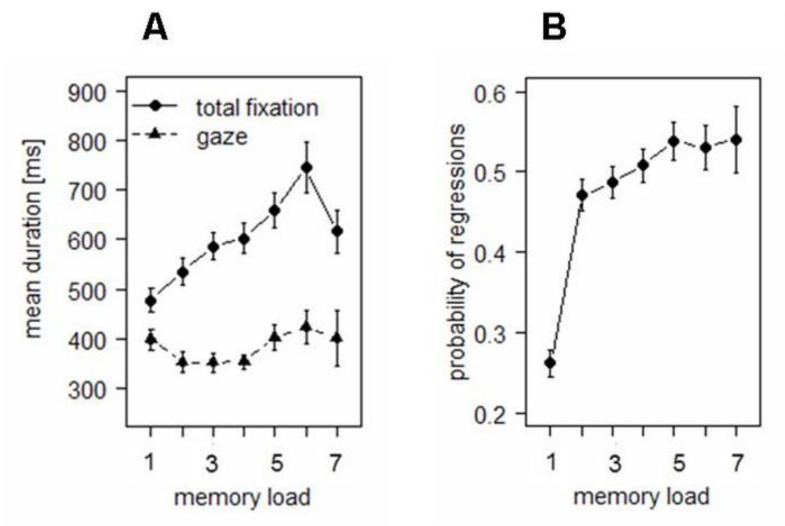

Figure 2.6: A) Mean total fixation duration and gaze duration as function of memory load. B) Probability of regressions as function of memory load.

Effects of working-memory load are modulated by WMC

In addition to the capacity effects, a significant main effect of WM load for the *sentence-processing time* is observed (for the LMM see Table C.14). The effect of LOAD is significant for three out of four set sizes (*set size 4:* b = 0.101, SE = 0.033, t = 3.0; *set size 5*: b = 0.249, SE = 0.018, t = 13.8; *set size 7*: b = 0.132, SE = 0.019, t = 7.0). Figure 2.7 A displays the interaction of WMC and load on mean *sentence processing times* averaged across all set sizes. With increasing WM load the sentence processing times increase. The size of the effect depends on WMC. The smaller the capacity, the more evidently reading times increase. The interaction between LOAD and CAPACITY reaches significance, for two out of four set sizes, that is: *set size 4*: b = -3.297, SE = 1.250, t = -2.6 and *set size 6*: b = -3.766, SE = 1.104, t = -3.2. The results for each set size are plotted in Figure 2.7 B. The influence of WMC on the WM load effect on sentence processing time is also visible as a trend in set size three and five (see Figure 2.7 B).

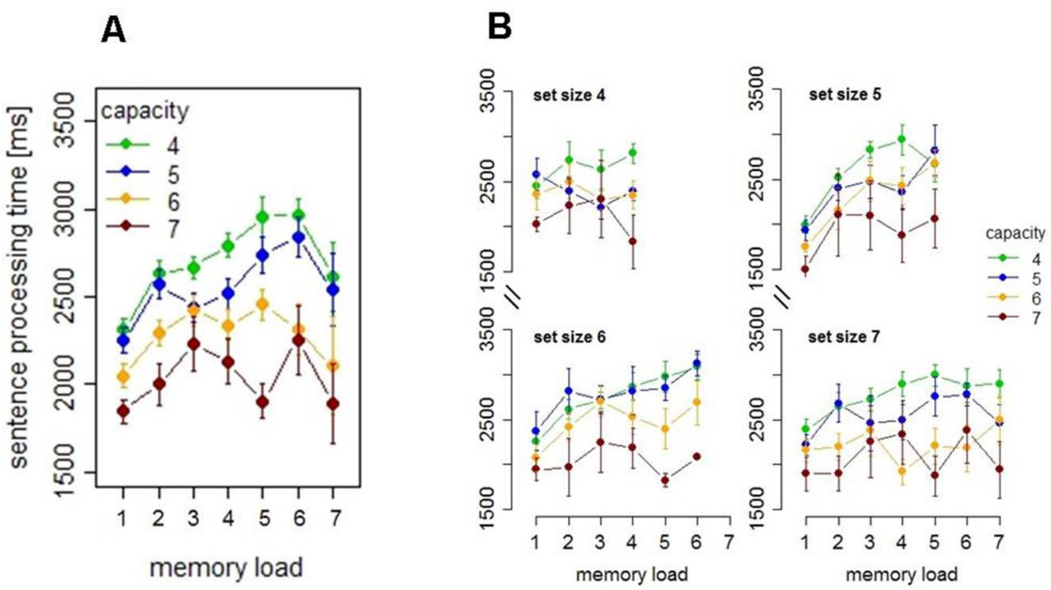

Figure 2.7: Mean sentence processing times as function of memory load. A) Aggregated over all set sizes. Low-capacity groups slow down their reading more as a function of increasing WM load than high-capacity groups. B) Illustration of the effect for each set size. Capacity groups with lower capacity show stronger effects of WM load in all four set sizes. The effect reaches significance only for set size 4 and set size 6.

2.2.2.1 Individual capacity differences are not due to a resource effect

The observed differences between the capacity groups in sentence processing times could be solely due to different amounts of free memory resources. In order to check this hypothesis the effect of a new factor – the number of free memory slots – was examined on the sentence processing times of each capacity group. The factor memory resource (RESOURCE), with 10 factor levels, was computed by subtracting the individual capacity from the serial position of the target word, multiplied by -1. Thus, when the WM load (e.g. 5) of a sentence equalled the individual capacity (5) the number of free memory slots was zero. A further increase of WM load would result in an "overload" and thus a negative value for the number of free memory slots. Figure 2.8 displays the mean sentence processing time of each capacity group as a function of its free WM slots. If the capacity effects on sentence processing time were only due to the capacity resources, the curves of the WMC groups in Figure 2.8 should overlap. They do not. Low-capacity subjects show longer sentence processing times despite the same amount of free memory slots in comparison to high-capacity subjects.

The results of the LMM in Table C.19 confirm this interpretation. The main effect for CAPACITY is significant ($b = -4.402$, $SE = 1.330$, $t = -3.3$), that is as previously shown longer *sentence processing times* for lower in contrast to higher WMC. Moreover, the main effect of RESOURCE is highly significant ($b = -0.034$, $SE = 0.002$, $t = -14.2$). This means that, sentence processing took longer the less memory space was freely available. The interaction between CAPACITY and RESOURCE is significant ($b = 0.645$, $SE = 0.140$, $t = 4.6$). The more exhausted the resource of the low-capacity subjects was, the more time they needed for processing the sentence. This does not apply to the high-capacity subjects, who needed nearly the same processing time irrespective of their resource consumption. The results show that subjects with different WMCs deal differently with their capacity resources. Interestingly the capacity groups 4 and 5 differ significantly from the groups 6 and 7 ($b = 0.135$, $SE = 0.047$, $t = -2.8$). The nested ANOVA contrast (see Table C.20) within the two group pairs is not significant. Neither group 4 shows significant longer sentence processing times in contrast to group 5 (b = -0.039, SE = 0.061, t = -0.6), nor group 6 from group 7 (b = -0.046, SE = 0.072, t = -0.6).

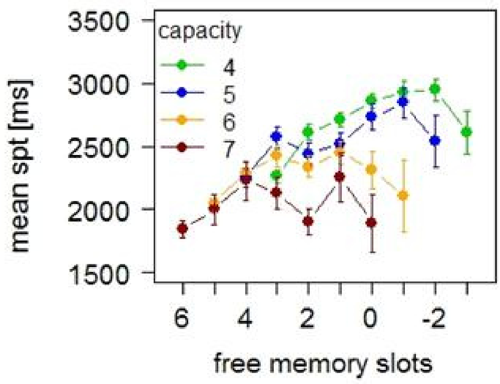

Figure 2.8: Mean sentence processing times (spt) of each capacity group as function of its free working-memory slots.

2.2.2.2 Sentence comprehension depends on working-memory

Understanding of a sentence, measured by the accuracy in judgement of the sentences (defined as COMPREHENSION), influences the intercept of the fixation durations and reading times. For example, *gaze duration* in correctly judged sentences is on average 46.66 ms (SD = 45.03) faster than for incorrectly judged sentences. Correctly judged sentences were read faster, included shorter *gaze durations* (b = -0.090, SE = 0.025, t = -3.53, in Table C.16) and lower *total fixation times* (b = -0.174, SE = 0.032, t = -5.54, in Table C.15). Moreover, *sentence processing times* also decrease significantly for correctly compared to incorrectly judged sentences (b = -0.058, SE = 0.020, t = -2.9, in Table C.14). Interactions between the COMPREHENSION accuracy and LOAD effects are only significant for the global measure of *sentence processing times*. But the direction of interactions are inconsistent across the set sizes (*set size 4*: b = -0.278, SE = 0.067, t = -4.1 and *set size 6*: b = 0.284, SE = 0.076, t = 3.7). The significant interaction between COMPREHENSION and CAPACITY reveals shorter sentence processing times for higher WMC (b = -0.051, SE = 0.017, t = -3.1). Figure 2.9 displays that *sentence* *processing times* increase for incorrectly judged sentences, but only for capacity group 7 and 5. The other two groups show nearly no effect of comprehension accuracy. The ANOVA nested contrast specification between span groups 4 and 5 versus the span groups 6 and 7 is significant (b = -0.137, SE = 0.051, t = -2.71, in Table C.21) and also the interaction of the contrast between the capacity groups 4 and 6 versus 5 and 7 (b = -0.190, SE = 0.036, t = -5.28).

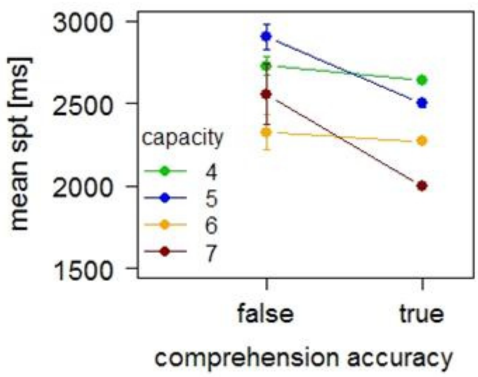

Figure 2.9: WMC differences on mean sentence processing time (spt) for correct and wrong comprehended sentences.

A GLMM with the comprehension accuracy as dependent variable (see Table C.22) shows a significant effect for CAPACITY (b = 43.70, SE = 8.389, p < .001): The higher the capacity, the higher the response accuracy of the comprehension task. The main effect of LOAD is significant within three of four set sizes: set size 4: b = -0.985, SE = 0.391, p < .05; set size 6: b = -0.572, SE = 0.256, p < .05; set size 7: b = -1.061, SE = 0.231, p < .001. The higher the WM load, the more often sentences were incorrectly comprehended. Table 2.4 lists comprehension errors as a function of memory load and capacity group. The interaction between both factors LOAD and CAPACITY is significant for set size 4: b = -68.53, SE = 25.24, p < .01; set size 5: b = 84.77, SE = 20.37, p < .001 and set size 6: b = 41.04, SE = 17.16, p < .05.

Table 2.4 Processing errors in percentage correct for each capacity group and memory load.

	Memory load							
capacity	1	2	3	4	5	6	7	all
4	10.96	22.57	6.25	23.33	32.76	29.06	18.33	18.94
5	10.00	10.53	10.25	12.93	10.48	7.14	25.00	10.99
6	1.48	13.66	6.23	3.82	13.52	15.38	37.21	9.36
7	0	3.26	11.00	1.85	0	4.44	0	3.28
all	6.67	15.01	7.92	12.73	17.93	16.24	22.98	12.43

2.3 Summary of Results and Discussion

The present experiment examined eye movements in the reading span task. The aim was to find evidence for the influence of WM load and capacity on eye-movements during reading. Therefore, WMC was quasi-experimentally controlled by using the reading span task to classify subjects in 4 capacity groups. Moreover, WM load was experimentally controlled by progressively increasing the number of target words within each trial of four possible set sizes. The results showed that both WM load and WMC had an effect on global eye-movement measures during the reading span. In the following, I first focus on results found on the storage part, and then I discuss the results of the processing part, before I give an outlook for following experiments.

2.3.1 Storage

WMC groups did not differ in their fixation durations in general, but the probability of re-fixating target words and the total time spent on targets and the words one prior of the target, was predictive of WMC.

Fixation durations increased across the sentence, with the longest durations on the target words. Longer durations on sentence-final words are known as sentence wrap-up effect (e.g., Just & Carpenter, 1980; Mitchell & Green, 1978; Rayner et al., 1989, 2000), reflecting integration of information and comprehension of the sentence. In the present study, it is reasonable to assume that they additionally reflect time for encoding and storing the target word(s). Friedman and Miyake (2004) interpreted longer reading times for low WMC as a result of limits in resources for the maintenance of information during sentence comprehension. In line with their argument, capacity group differences on the target and the word prior to the target can not solely be explained by longer encoding times for low-capacity subjects. Already, during sentence reading, target words from previous sentences have to be maintained. This alone could result in longer reading times for low-capacity groups and, thus, influence sentence comprehension. Furthermore, a more pronounced capacity effect for total fixation durations in contrast to gaze durations on the target word suggests, that differences in WMC leads to differences in second pass reading. Hence, the observed group differences on target words could reflect (a) a differences in encoding and storing times, (b) differences in comprehension, or (c) a combination of both. The increase of fixation durations within sentences is also in line with results from Kuperman, Dambacher, Nuthmann and Kliegl (2010).

Sentences with target words that were later incorrectly recalled contained a higher rate of regressions and longer sentence processing times, than sentences with correctly recalled target words. In line with Walczyks (1993, 1995) this result suggests sentence-comprehension problems that were caused by the storage part. In the case of word encoding and storage problems or inefficient lexical access, Walczyk’s model predicts longer reading times and regressions to compensate for inadequate comprehension. For the current experiment, longer reading times and higher regression rates were observed for sentences with later incorrectly recalled target words. Thus, more time was used for sentence comprehension, due to processing problems which placed additional demands on working memory. That in consequence hindered memory encoding of the target word.

The results can additionally be explained from an interference perspective (e.g. the interference model by Oberauer & Kliegl, 2001), which assumes that representations maintained for the same time in working memory disturb each other. According to this perspective the encoding or maintenance of target words; and lexical access and word encoding for sentence comprehension disturbed each other. Thus, the target word was incorrectly recalled and sentence comprehension assumed more time and more regressions.

2.3.2 Processing

WM load progressively increased during the task. This enables to directly measure the impact of a step by step increase of WM load on eye-movement reading measures. Sentence-processing times and fixation durations gradually increased with increasing WM load. The load effect replicates the results of Kaakinen and Hyönä (2007). Moreover WM load was related to WMC. The increase of reading times with increasing WM load was the greatest for low capacity groups. The higher the WMC was, the smaller the increase. The influence of both WM load and capacity was previously only shown in a correlation study by Just and Carpenter (1992). Kaakinen and Hyönä, who also measured eye movements during the reading span task, failed to replicate the influences of WMC. This study differs in two ways from the present experiment. First, authors used an adaptive test version, where subjects were not tested with a set size above their span. The second point concerns the sentence presentation. Kaakinen and Hyönä used an experimenter-paced version, where the presentation of sentences depends on individual reading speed. Subjects read aloud and if they finished reading the experimenter pressed a button to display the new sentence. Presentation times in the present experiment were fixed to 4 seconds and the same for all subjects. On average, high capacity subjects needed 2 seconds and low capacity subjects needed 2.8 seconds for sentence processing. The remaining time could be used for memorizing the target words. Low-capacity subjects slowed down their reading more as a function of increasing working-memory load than high-capacity subjects. From a memory performance perspective, longer sentence processing times presumably reduces the time available for ‘pure’ memorization of the target words. Because high-capacity subjects spent less time on the processing task, they had more time for memorizing. Thus, they were able to sequentially work on the processing and storage part of the task. Particularly the longer sentence processing times of low capacity subjects come from longer total fixation times on target words. If the observed group differences on target words are due to differences in encoding and storing times, it is reasonable that low capacity subjects worked on both task components (storage and processing) at once. Thus only the high-capacity group used the presentation time of a sentence effectively by sequentially working on the storage part after processing was finished.

The lower the capacity score and the higher the WM load, the more errors occurred in sentence comprehension. The results are in line with comprehension studies (Miyake et al., 1994), in which high-capacity readers showed better comprehension results than low-capacity readers. The accuracy of sentence comprehension influenced the global reading time of a sentence. Surprisingly, reading times of incorrectly comprehended sentences were longer than correctly understood sentences and not vice versa. Thus, subjects did not strategically skim the processing tasks to increase the time for the target words. The result was in line with the compensatory encoding model by Walczyks (1993, 1995), which predicts longer reading times as a result of inefficient word encoding, and lexical access processes. According to this assumption longer reading times for incorrectly judged sentences mirrors comprehension problems. Importantly, the compensation was not enough.

2.3.2.1 Individual Differences

Group effects, especially in sentence processing times, were not solely due to differences in memory resources. Rather, the WMC groups showed fundamentally different processing times at same amounts of storage space. Thus, the current results can not be interpreted by a resource theory. The WMC groups showed an interesting picture in processing times of incorrectly comprehended sentences. Only group 5 and 7 showed significantly higher processing times for incorrectly compared to correctly judged sentences. Beyond this result, group 6 and 7 (4 and 5, respectively) showed no differences in processing time. The results tentatively suggest different cognitive strategies. Group 5 and 7, for example, tried to answer all sentences correctly, which result in longer reading times for problematic sentences. Group 4 and 6, however, are indiscriminative. Perhaps they are not aware of incorrect comprehension. One could interpret the results with respect to the interference model (Oberauer & Kliegl, 2006). Conceivably the results are due to differences in susceptibility to interference. If we assume a stronger susceptibility of the cognitive system of group 4 and 6 compared to group 5 and 7 the first two groups were maybe not aware of their incorrect sentence judgements. As a consequence of interference, they did not notice the mistake. Therefore, the capacity score differences, for example of group 6 and 7, are not due to a resource difference, but are due rather to differences in the vulnerability of interference. To clarify these assumptions further studies with more narrowly focused questions are necessary.

An unexpected result was that the reading of a sentence tended to start with the sentence final word. All subjects irrespective of their WMC exhibit this behaviour. Given that a fixation was only controlled at task beginning, from the second/ third sentence on, the target was fixated before the eyes were guided to the sentence beginning. As consequence each memory load position (n) 1 to 6 was possibly a mixture of memory load n and n + 1. The result, has no bearing on the current interpretation, but it needs to be controlled if sensitive measures of lexical access are of utmost concern. Therefore, I prevented this strategy in the following experiments by masking the sentence until the initial word was fixated.

2.3.3 Conclusion

The current experiment provided strong evidence for the influence of working-memory processes on eye-movements during the reading span task. These were particularly strong for second-pass reading measures but were also observed in a weaker form in first-pass reading. Furthermore, sentence-processing times increased with increasing WM load, and the increase was greater, the lower the individual WMC. The results clearly divided the WMC groups, with respect to WM load effects. The WMC differences, however, were not a pure result of resource limitations. On the contrary, the groups showed fundamentally different processing times indicative of different cognitive strategies. Moreover, target words of sentences, in which comprehension problems were visible in the eye-movement measures, were frequently not recalled correctly. The result tentatively suggests that additional memory demands, due to sentence comprehension, disturbed the encoding of target words. In general, fixation durations increased across the sentence, with the longest durations on the sentence-final words (i.e., a sentence wrap-up effect), reflecting both comprehension and time for encoding the target word. The WMC is predictive of the probability of refixating targets and the total time spent on the critical words. An unexpected side result was that the reading of a sentence tended to start with the sentence-final word, indicating strategic allocation of processing resources.

Chapter 3

Working memory influences on the perceptual span

The previous chapter focused on the influence of WM on global eye movement measures. Effects of WM load and capacity were shown to affect fixation durations in first and especially second-pass reading. The theoretical question addressed in the present chapter goes a step further by asking how WM load and WMC do affect lexical access during reading.

An established finding in reading research is that word processing is influenced by the word's frequency of occurrence in a given language (see Balota, Pollatsek, & Rayner, 1985; Inhoff, Pollatsek, Posner, & Rayner, 1989; Inhoff & Rayner, 1986; Rayner & Pollatsek, 1987). Word frequency is related to lexical word recognition and to the speed and accuracy with which a word is recognized. Morrison and Ellis (1995) predicted a threshold for word identification. The lower the frequency of a word the higher the threshold and, thus, the longer the time needed for identification, thus for lexical access. In reading first fixation durations and gaze durations are longer for low frequent words (e.g. Inhoff & Rayner, 1986; Rayner & Duffy, 1986). The effect of word frequency was shown for the word that induced those effects (Henderson & Ferreira, 1990, 1993; Raney & Rayner, 1995) and for neighbouring words (Kennison & Clifton, 1995; Kliegl, et al., 2006). The perceptual span is the region of text from which useful information can be extracted during reading (for a review, see Rayner, 1998). The fact that parafoveal word frequency of the not-yet-fixated word showed an effect on current word processing is an indicator that also neighbouring words can fall inside the perceptual span. If the fixated word is short, the word to the right of fixation has a higher probability to fall into the perceptual span than in the case of fixating a long word.

Traditionally the size of the perceptual span was approximated in moving-window studies (e.g., McConkie & Rayner, 1975; Rayner, 1975), where only the text in a ‘window’ around the current gaze position is unmasked. However, influences of the upcoming, not-yet fixated parafoveal word, can be investigated also from normal reading data (Kliegl et al., 2006, for an example). The basic idea is that neighbouring words can only influence fixation durations of the fixated word if they fall into the perceptual span. Thus, reduced influences of neighbouring words are an indication for a reduction of the perceptual span. This was used in the present experimental design to indirectly measure the size of the perceptual span by the influences of parafoveal words on foveal fixation durations.

With respect to what can be processed during a given fixation, additional results suggest a dynamical modulation of the perceptual span, referred to as foveal-load hypothesis (Henderson & Ferreira, 1990; Kennison & Clifton, 1995). Fixation durations on a word decreased when the prior word was easy to process than when the prior word was difficult to process (Henderson & Ferreira, 1990). Reduced reading times on the target after an easy pretarget word are interpreted as the consequence of having more efficiently preprocessed the word in parafoveal vision, hence it must have fallen into the perceptual span. Depending on the difficulty (frequency) of the fixated word in foveal vision, the size of the perceptual span seems to vary. According to this, high foveal-load should result in a smaller perceptual span and fixation durations should be unaffected by the word properties of the upcoming word.

Having to memorize target words from several sentences in the reading span task may have a similar effect of increasing the load during each reading fixation. As a straight-forward prediction, the perceptual span may be restricted only to the fixated word in high memory-load conditions. Given such a reduction of the perceptual span under high memory-load, the neighbouring word should fall outside of the perceptual span and its frequency should not influence the fixation durations of the word in the fovea. In contrast, the frequency of the parafoveal word should affect the fixation durations of the fixated word if the memory-load is low.

In addition it will be investigated how the relation between WM load and the perceptual span depends on WMC. The relationship between the reader's WMC and perceptual processing was suggested by previous research (Fisher & Lefton, 1976; Fisher & Montanary, 1977; Hochberg, 1970; Rayner, 1986; Spragins, Lefton, & Fisher, 1976). In particularly, Rayner (1986) provide evidence that the perceptual span increased with reading skill. In his study, beginning readers showed a smaller perceptual span than adult skilled readers and fourth grade readers. Furthermore, fourth grade readers reduced their perceptual span to the size of that from beginning readers when reading difficult texts. However, individual differences were commonly defined in terms of reading skills. In the present context, a similar relation is assumed but investigating WMC differences between individuals. The reduction of the perceptual span due to memory load is expected to be strongest for readers with low WMC.

With the present experiment I want to further clarify some unresolved questions of the first experiment. As discussed in Chapter 2, individual differences in the reading span task were shown most exclusively for fixation durations on the sentence-final word which was the target in the sentences. Due to the design of the task, subjects tried to read the target first before they started reading the remaining sentence. As a consequence, the WMC differences in fixation durations on the target word were not clearly interpretable. It was discussed earlier that the effects may have been a result of the experimental design and at least not only a result of memory encoding or sentence comprehension.

Therefore, in the following Experiment 2 it was ensured that reading direction was from left to right and could not start at the target word. Participants were forced to start reading at the sentence beginning by masking all words except the sentence-initial one. If the eyes fixated the first word of the sentence the whole sentence was displayed visible. This design provides the possibility to evaluate whether the group differences on target words in the first experiment were due to variations in sentence comprehension or in target encoding and storing time. If sentence comprehension depends on WMC (see also Miyake et al., 1994), the effect of WMC differences from the first experiment should replicate for the current gaze durations on target words. A lack of group effects on target fixation durations would support the interpretation of group-dependent variations in target encoding and storing time. This would be due to the possibility that the target word could be processed in Experiment 1 before sentence reading and would suggest a shift in task-compliance strategy between experiments.

To maximize individual differences in WMC young and old adults were tested. Among other variables (e.g., processing speed, inhibitory control, attentional processes), older adults are typically associated with an age-related WMC reduction (e.g. Park et al., 2002). There age effects were discussed to be associated with a general decline in cognitive functioning (e.g. Craik & Byrd, 1982; Hasher & Zacks, 1988; Park & Schwartz, 2000; Salthouse, 1996). During reading older adults, most notably, show more (e.g. Kemper et al., 2004) and longer fixations (e.g. Laubrock, Kliegl, & Engbert, 2006) than younger adults. Although the effect of age group was included in the LMM of data from Experiment 2, the present chapter globally concentrates on individual differences in WMC on lexical access and does not focus on age groups. How lexical access moreover interacts with age will be discussed in Chapter 4.

In sum, one goal was to replicate the results of the previous chapter. Unresolved questions of the first experiment which were due to the experimental design were supposed to be clarified by the new design where subjects were forced to start reading with the sentence-initial word. Moreover, the experiment was assumed to support theoretical proposals that the perceptual span is dynamically modulated by WM load and WMC. Stronger neighbouring word-frequency effects on current fixation durations were proposed for low rather than high WM load. For WMC, the effect was supposed to be in the same direction: stronger neighbouring word-frequency effects on the fixated word for high-capacity groups rather than for low-capacity groups. If subjects were operating with memory demands that are at or above their WMC, the perceptual span should be reduced to the foveal word only. Foveal WM influences were predicted in the following direction: reduced word-frequency effects for high WMC and low WM load.

3.1 Method

3.1.1 Subjects

A group of 63 students from the University of Potsdam and a secondary school in Potsdam, and a group of 62 older adults participated in the experiment. None of them participated in the previous experiment (see chapter 2). All were native speakers of German with normal or corrected-to-normal vision. They were paid six Euros or received course credit. Data of two young and eight old subjects were excluded from analysis because they ignored instruction. The 61 remaining young subjects were 20.80 years on average (range: 16 to 30 years) and the 54 old subjects were 71.06 years on average (range: 63 to 81 years).

3.1.2 Material and Procedure

The reading span task from chapter 2 was administered to the participants. There was one modification with respect to the experimental design. At initial sentence presentation, all words except the first one were masked with x-strings. As soon as the eyes crossed the last letter of the first word in the sentence, the x-strings were replaced and the whole sentence was displayed visible. The experimental design stayed constant with respect to all other things. Contrary to the oral recall of words in the first experiment not the experimenter but the participant himself wrote the recalled words in an answer sheet.

Table 3.1: Number of fixations for various types of fixations. Row 4 = 5 + 6 + 7 + 8; row 8 = 9 + 10 + 11

	N_{young}	$\%_{young}$	N_{old}	$\%_{old}$
1 N of sentences	3782	100	3348	100
2 sentences with blinks	1563	41.33	2171	64.84
3 sentences with interpolated blinks	1004	26.55	1615	48.24
4 N of fixations	21,888	100	13963	100
5 first word	1197	5.47	768	5.5
6 short/long fixations or amplitudes	2656	12.13	1937	13.87
7 forward saccade criterion	-	-	-	-
8 N of valid fixations	18,035	82.40	11,258	80.63
9 not first pass	7098	39.36	2715	24.12
10 multiple fixations	4412	24.46	2747	24.40
11 single fixations	6525	36.18	5796	51.48

The procedure for the second experiment differed from the former one in three points: (1) Due to the fact, that reading began with the first word of a sentence and reading direction was from left to right the traditional definition of *first-pass reading* was conducted. All fixations on words as they were first encountered (i.e., in contrast to revisiting the word after the eyes had already moved away from the word) were defined as first-pass fixations. (2) Subjects were instructed to *blink* before sentence display. Blinks after sentence display, were interpolated as in the former experiment. The algorithm and its detailed description are provided in Appendix B. An overview about the number of sentences with blinks and both the excluded and remaining fixations is provided in Table 3.1. As additional modification, sentences were presented in random order. Because of program error, for 33 of the young and 25 of the old adults the sentence order was constant across subjects.

To validate the groups on the basis of the reading span score, 56 subjects were tested with three additional tests. The set of tasks involved two memory updating tasks (spatial and verbal) and a spatial coordination task (described in Oberauer et al., 2000, 2003). Oberauer et al. demonstrated that the numerical memory-updating as also the reading-span task, had a high loading on the verbal-numerical factor of working memory, whereas the spatial memory-updating and spatial-coordination had high loadings on a spatial factor of working memory.

During the *numerical memory-updating* task one digit was presented in each of two, four or six frames on the screen. The digits had to be memorized in their corresponding frames. Afterwards, arithmetic operations were presented in randomly selected frames, by which the digit of that frame had to be updated. At the end of a trial the content of selected frames had to be recalled. In the *spatial memory-updating* task the digits were replaced by dots. Each dot was presented in one of 9 possible locations in its frame. Updating was realised with arrows indicating the direction a dot had to be shifted in its frame. In the *spatial-coordination* task a pattern of dots had to be reproduced in an empty grid. Dots were presented one by one in a 10 x 10 grid on screen. Each dot was presented for 1 second. The number of dots increased over trials from three to six. (see Oberauer et al., 2000, 2003, for a description of the tasks).

The additional WMC tests were administered in a single session one week before or after (pseudo-randomly chosen) the day of the reading-span test.

3.2 Results

The percentage of correct scores of each working memory test was transformed in z-scores. Moreover, a mean WMC score was calculated by the mean of z-scores of the two memory-updating tasks and the spatial-coordination task. The grouping on the basis of this mean WMC score replicated the grouping on the basis of the reading span score. The mean WMC score does not enhance the explained variance of the LMM (see Table 3.2 for an example). Hence, it was eliminated from the statistical models.

In addition to the first experiment, the fixed-effects part of the LMM comprise the variables AGE (young vs. old adults) and SENTENCE ORDER (random sentence order vs. fixed sentence order). The results of AGE are provided in the next chapter 4. All models and their significant effects are listed in Appendix D. In the present chapter, before focussing on the influences of WM on lexical access and the perceptual span, I first summarize the results on the storage part of the task, followed by the results of the processing part.

Table 3.2: LMM fitting sentence processing time with mean WMC score

Linear mixed model fit by maximum likelihood			
Random effects:			
Groups	Name	Variance	Std.Dev
subject id	(Intercept)	0.019185	0.13851
Residual		0.046613	0.21590
number of obs: 9306, groups: subject id, 56			
Fixed effects:			
	Estimate	Std. Error	t-value
(Intercept)	7.645749	0.023679	322.9
age	-0.058998	0.055710	-1.1
capacity	-0.008108	0.023746	-0.3
wmc3	-0.079944	0.046509	-1.7
set size	-0.022139	0.002181	-10.2
s4.load	0.003341	0.014555	0.2
s5.load	0.039785	0.011977	3.3
s6.load	0.060666	0.011975	5.1
s7.load	0.051037	0.011663	4.4
capacity * s4.load	0.033828	0.012891	2.6
capacity * s5.load	-0.038341	0.010632	-3.6
capacity * s6.load	-0.036468	0.010627	-3.4
capacity * s7.load	0.015877	0.010438	1.5
wmc3 * s4.load	0.015256	0.020762	0.7
wmc3 * s5.load	-0.002971	0.017319	-0.2
wmc3 * s6.load	0.004519	0.016703	0.3
wmc3 * s7.load	-0.010264	0.016510	-0.6

Note: capacity: WMC groups; WMC3 : mean WMC score of the numerical memory-updating, spatial memory-updating, and spatial coordination task**;** s4: set size 4; s5: set size 5, s6: set size 6, s7: set size (e.g. s4.load: load effect at set size 4)

According to their span scores subjects were divided into four capacity groups with group number reflecting the achieved span score. The distribution of subjects across capacity groups is provided in Table 3.3. More than half of the subjects reached the highest span score and hence were in the capacity group 7. Contrary to expectation, most of the old subjects were very good and reached comparably high span scores.

Table 3.3: Distribution of young and old adults on capacity groups.

	Capacity group				
Age group	4	5	6	7	all
young adults (mean age 23)	5	9	17	30	61
old adults (mean age 73)	4	9	8	33	54
	9	18	25	63	115

3.2.1 Results of the storage part

In this result section, the statistical analysis focuses on how WM load and capacity influence the *initial fixation number, gaze durations,* and *total fixation durations* on target words. With increasing serial position (SERPOS) participants fixated the target word earlier than in the previous trial. In Figure 3.1 the mean *initial fixation number* on target words is plotted as a function of the serial position within each of the four set sizes. This strategy of fixating the target word earlier the later it appeared in the sentence (i.e., targets with higher serial position number) was shown in all set sizes but was significant only for *set size 4* ($b = -0.58$, $SE = 0.128$, $t = -4.561$, in Table D.1). Although not significant in the other set sizes, there was a similar trend for the other three set sizes (as shown in Figure 3.1) suggesting the use of the same strategy to comply with the task and memorize the target words.

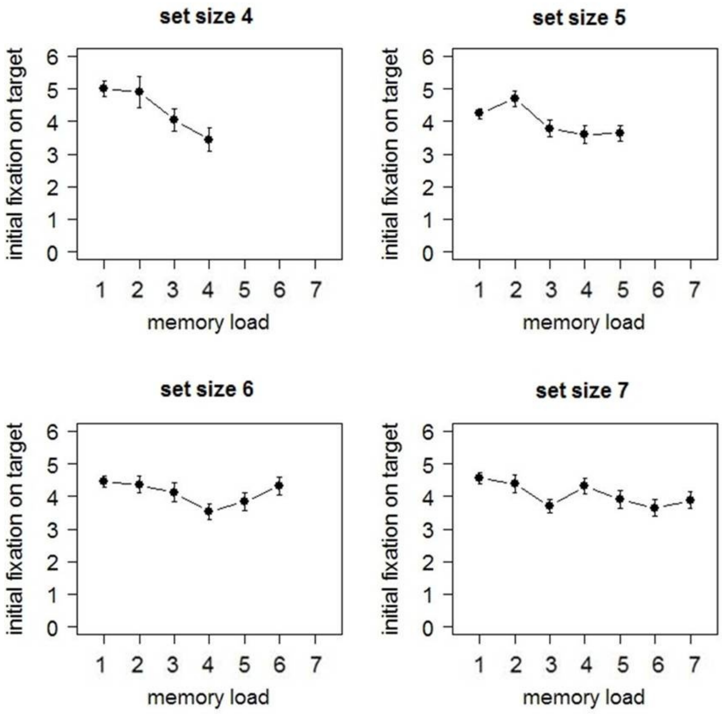

Figure 3.1: Initial fixation on the target for each set size (4-7).

3.2.1.1 Word position effects

The word position effects in Figure 3.2 represent the *total fixation times* of the absolute word position aligned at the sentence-final words (target) aggregated over all sentences of the task. The four lines represent the capacity groups. The left panel displays total fixation times for the fixed and the right panel for the random sentence order, respectively. The total fixation times increase from sentence beginning to the end ($b = 0.083$, $SE = 0.004$, $t = 19.66$, in Table D.2), with the longest duration on sentence-final words. The effect of WORD POSITION is in addition significant for *gaze durations* ($b = 0.009$, $SE = 0.004$, $t = 25.79$, in Table D.3). The mean total fixation times on target words is 499 ms ($SD = 304$) and mean gaze duration is 423 ms ($SD = 276$).

The interaction between WORD POSITION and CAPACITY is only significant for total fixation times ($b = -0.007$, $SE = 0.003$, $t = -2.10$), that means, the lower the capacity group the stronger the increase of total fixation times within a sentence.

The effect is mainly due to the significant longer durations of the lowest capacity group 4. The contrast between capacity group 4 and all other groups was significant with b = -0.153, SE = 0.048, t = -3.2 (in Table D.4). The interaction between capacity group and word position, furthermore, depends on the condition of sentence presentation. The threefold interaction of WORD POSITION, CAPACITY, and SENTENCE ORDER is significant with b = -0.018, SE = 0.007, t = -2.55. In the condition of the fixed sentence order, the lowest capacity group showed the longest total fixation times compared to the other groups (b = -0.025, SE = 0.081, t = -3.08, in Table D.5; see left panel of Figure 3.2). For the condition with random sentence order, however, the highest capacity group showed significantly shorter total fixation times on the target word compared to all other groups (b = -0.032, SE = 0.012, t = -2.8, see right panel of Figure 3.2).

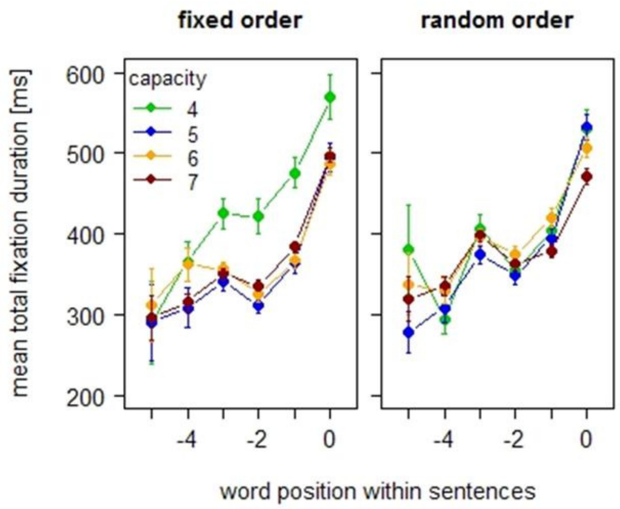

Figure 3.2: WMC differences on mean total fixation duration for the condition of fixed sentence order (left panel) and the condition of random sentence order (right panel).

Moreover, the three-way interaction between WORD POSITION, CAPACITY and SENTENCE ORDER is significant for *gaze durations* (b = -0.027, SE = 0.006, t = -4.59) with the same data patterns as shown for total fixation durations. Longer gaze durations for capacity group 4 in the condition with fixed sentence order and shorter gaze durations for capacity group 7 on target words in the condition with random sentence order. The main effect of CAPACITY, and the interaction between WORD POSITION and CAPACITY respectively, are not significant (see Table D.3).

3.2.1.2 Effects of target recall

The accuracy of target recall (RECALL) influences the *number of regressions* as well as the *total fixation times* on a sentence. The *number of regressions* is significantly higher if the target is incorrectly rather than correctly recalled ($b = -0.173$, $SE = 0.060$, $p < .01$, in Table D.6). In consequence of higher regression rates, also, *total fixation times* increases. Thus, the *total fixation times* for sentences with incorrectly recalled target words were higher compared to sentences with correctly recalled targets ($b = -0.034$, $SE = 0.013$, $t = -2.55$, in Table D.2). In summary, sentences in which the target was later recalled incorrectly included more *regressions* (2 %) and longer *total fixation times* (11 ms). The fixation difference on *gaze durations* is not significant (see Table C.5). Thus, regressions can be an indicator of memory failure.

3.2.2 Results of sentence processing

WM load and capacity are supposed to influence sentence processing that means the *number of first-pass fixations* in a sentence, the *sentence processing times*, *initial fixation number* on target words, *total fixation times*, *regressions* and *gaze durations*. The following section focuses on these variables and their dependency on WMC, on WM load and their interaction, respectively.

On average, sentences were processed 343 ms faster than in the first experiment (2448 ms in the first vs. 2105 ms in the present experiment) with slightly better comprehension accuracy (88 % correct in the first vs. 92 % correct in the present experiment). Each sentence was read with 4 to 5 fixations ($mean = 4.47$, $SD = 1.35$). Contrary to the first experiment, where the *number of first pass fixations* increased with increasing serial position, the number of fixations stayed constant across load conditions for the present experiment. Moreover, the average probability of regressions was dramatically reduced from 46 % (SD = 0.49) in the first to 25 % (SD = 0.43) in the present experiment. The number of fixations and regression rates in the present experiment correspond to those of serial position one in the first experiment (mean number of fixations = 4; regressions: mean = 26 %, SD = 0.49). Thus, the higher fixation and regression rates of the first experiment with beginning of the second sentence in a trial, are a result of the design, which, however, is representative of how the reading span task is usually administered.

3.2.2.1 The role of working-memory capacity

The *sentence processing time* represents the reaction time of the processing part, when subjects pressed a button to decide whether the sentence statement was true or not. The effect of CAPACITY is significant with b = -0.031, SE = 0.013, t = -2.45 (in Table D.7. The time for sentence processing decreased with increasing WMC. Moreover, the CAPACITY effect was significant for the *initial fixation number* on the target word (b = -0.151, SE = 0.075, t = -2.013, in Table D.1). Thus, high capacity readers are faster in sentence processing and are able to fixate the target word earlier than low capacity readers. The mean fixation durations for capacity groups are listed in Table 3.4.

3.2.2.2 Effects of working-memory load

In addition to the capacity effects a significant main effect of WM load for the *sentence processing times* is observed. The effect of LOAD is significant for three set sizes (*set size 5:* b = 0.060, SE = 0.009, t = 6.32; *set size 6:* b = 0.077, SE = 0.009, t = 8.32; *set size 7:* b = 0.091, SE = 0.008, t = 11.79, in Table D.7). With increasing WM load the sentence processing times increase. The linear trend of SET SIZE is significant for *sentence processing times* (b = -0.019, SE = 0.002, t = -11.20). The larger the set size the less time is used for sentence processing. The mean *sentence processing times* per set size are listed in Table 3.5. Moreover, the LOAD effect was more pronounced in higher set sizes.

Table 3.4: Summary: Mean (M) and standard deviation (SD) of eye movement measures for each capacity groups

		Capacity group			
Eye movement measures		4 (n = 9)	5 (n = 18)	6 (n = 25)	7 (n = 63)
Sentence processing time (ms) *	MD	2268	2135	2090	2075
	SD	*565*	*537*	*552*	*509*
Total fixation time (ms) *	MD	436	391	403	390
	SD	*272*	*254*	*251*	*247*
Gaze duration (ms)	MD	317	308	309	302
	SD	*205*	*207*	*202*	*192*
Single fixation duration (ms)	MD	233	232	233	232
	SD	*95*	*115*	*108*	*106*
Number of first pass fixation	MD	4.96	4.56	4.54	4.37
	SD	1.32	1.42	*1.38*	*1.30*
Probability of skipping	MD	.19	.23	.27	.23
	SD	*.39*	*.42*	*.44*	*.42*
Probability of regression	MD	.26	.25	.27	.25
	SD	*.44*	*.43*	*.45*	*.43*
Initial fixation on target *	MD	4.69	4.57	4.31	4.34
	SD	*1.63*	*1.62*	*1.58*	*1.47*

Note: Inferential statistics are based on linear mixed models specifying participants as random effect. Results are interpreted on the basis of the t value.
* The linear trend of capacity group was significant (t > |2|).

Furthermore, the *total fixation times* linearly increased with increasing memory load. The LOAD effect is significant within three set sizes (*set size 4*: b = 0.119, SE = 0.026, t = 4.58; *set size 5*: b = 0.048, SE = 0.021, t = 2.31; *set size 7:* b = 0.068, SE = 0.020, t = 3.33, in Table D.8). For *set size 4* the effect of LOAD significantly interacted with the SENTENCE ORDER (b = -0.199, SE = 0.052, t = -3.82) showing that only in the fixed condition *total fixation times* increased with WM load.

Total fixation times during sentence processing decreased with higher memory demands, which statistically translates in a significant SET SIZE effect (b = -0.013, SE = 0.004, t = -2.83).

Table 3.5: Summary: Mean (M) and standard deviation (SD) of eye movement measures for each set size

Eye movement measures		Set size			
		4	5	6	7
Sentence processing time (ms) *	M	2208	2074	2130	2053
	SD	*559*	*523*	*526*	*520*
Total fixation times (ms) *	M	407	402	394	390
	SD	*263*	*254*	*249*	*246*
Gaze duration (ms) *	M	306	309	303	303
	SD	*198*	*203*	*193*	*196*
Single fixation duration (ms)	M	227	232	232	233
	SD	*104*	*110*	*103*	*110*
Number of first pass fixation	M	4.42	4.39	4.51	4.51
	SD	1.39	1.37	1.36	1.31
Probability of skipping	M	.27	.22	.23	.24
	SD	*.44*	*.42*	*.42*	*.43*
Probability of regression	M	.26	.25	*.25*	.24
	SD	*.44*	*.44*	.43	*.43*
Initial fixation on target	M	4.56	4.21	4.44	4.45
	SD	*1.56*	*1.43*	*1.61*	*1.51*

Note: Inferential statistics are based on linear mixed models specifying participants as random effect. Results are interpreted on the basis of the t value.
* The linear trend of set size was significant (t > |2|)

3.2.2.3 Effects of working-memory load modulated by individual differences in capacity

For *sentence processing times* the interaction of CAPACITY and LOAD is significant for three set sizes (*set size 4:* b = 0.030, SE = 0.010, t = 3.08; *set size 5*: b = -0.019, SE = 0.008, t = -2.53; *set size 6:* b = -0.041, SE = 0.008, t = -5.42, in Table D.7), however, the significant interaction for *set size 4* went into the opposite direction than in the other two set sizes. *Set size 5* and *6*, shows larger LOAD effects for lower compared to higher capacity groups. That means sentence processing times increased with memory load. The increase is the stronger the lower the individual capacity is. Figure 3.3 A displays the interaction of WMC and load on mean sentence processing times averaged across all set sizes.

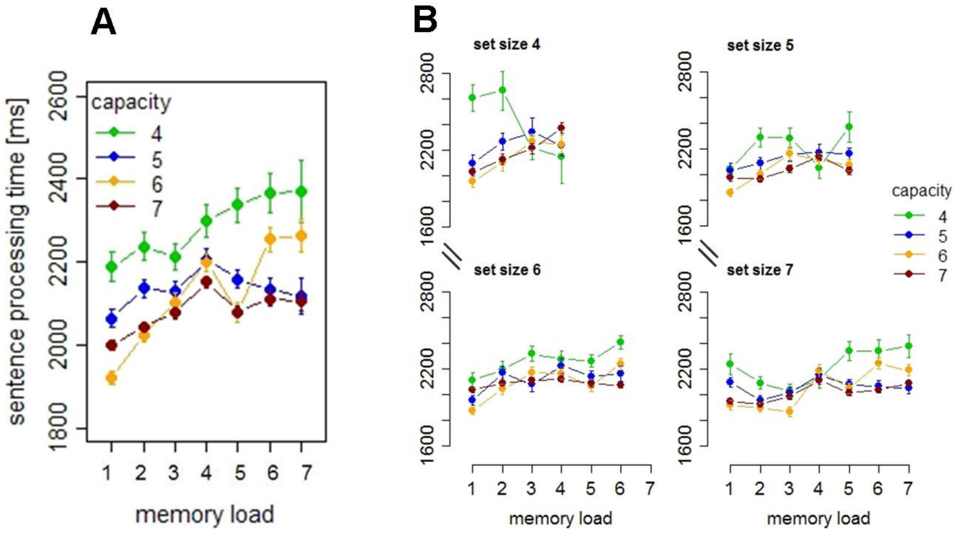

Figure 3.3: Mean sentence processing times as function of memory load. A) Aggregated over all set sizes. Capacity group 4 and 6 slow down their reading more as a function of increasing WM load than the capacity groups 5 and 7. B) Illustration of the effect for each set size. Capacity groups with lower capacity show stronger effect of WM load in three of four set sizes. Only in set size 4 the capacity group 4 decreased sentence processing times with load

3.2.2.4 Are individual capacity differences due to a resource effect?

In order to check if the observed differences between the capacity groups in sentence processing times are a pure result of different amounts of free memory resources, the number of free memory slots was calculated. The factor memory resource (RESOURCE) is computed by subtracting the individual capacity from each serial position, multiplied by -1. If the WM load of a sentence equalled the individual capacity the number of free memory slots is zero. Negative values for the number of free memory slots result indicate an “overload”, when the WM load is above the WMC of the capacity group. If the capacity effects on sentence processing times are only due to the capacity resource, sentence processing times should vary as a function of free memory space but not as a function of capacity groups, that is, the curves of the WMC groups in Figure 3.4 should overlap. As expected by the results of the previous experiment, low-capacity subjects showed longer *sentence processing times* despite the same amount of free memory slots in comparison to high-capacity subjects. The main effect of RESOURCE was significant with b = -0.018, SE = 0.0012, t = -14.97 (in Table D.9). The fewer memory space is freely available the more time was used for sentence processing. Moreover, the interaction between RESOURCE and CAPACITY is significant (b = 0.004, SE = 0.001, t = 4.44). Thus, subjects with different WMCs deal differently with their capacity resources.

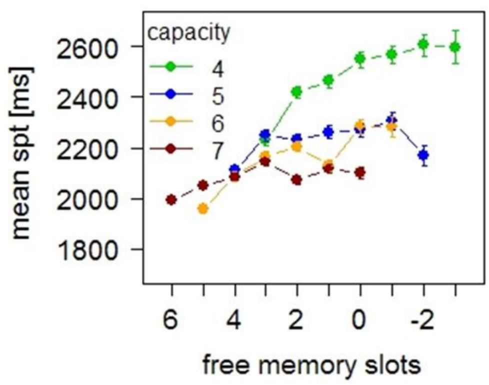

Figure 3.4: WMC differences on mean sentence processing times (spt) as function of the individual free memory slots.

3.2.2.5 Sentence comprehension did not effect reading times

On average 92 % (SD = 0.27) of the sentences were correctly comprehended (judged). Capacity groups showed no significant difference in there comprehension accuracy (see Tables in Appendix D). Thus, contrary to the first experiment incorrectly judged sentences did not result in longer total fixation times, gaze durations or sentence processing times. The sentence comprehension had no effect on all these measures.

3.2.3 WM load and capacity influences parafoveal-on-foveal effects

The following section investigates whether WM load and capacity dynamically modulated the perceptual span as hypothesized according to the foveal load assumption (Henderson & Ferreira, 1990). Thereby, the influence of the word frequency of the fixated word (n, immediacy), the prior word (n-1, lag effect) and the following word (n+1, successor effects) on the current fixation duration were analysed. Immediacy, lag and successor effects of word frequency are examined according to WM load and capacity. The effects were investigated for *gaze durations* (Table D.10) and *single fixation durations* (Table D.11).

3.2.3.1 WM immediacy effects

A well-established frequency effect is replicated for *gaze durations* ($b = -0.101$, $SE = 0.025$, $t = 4.05$). That is, gaze durations were 85 ms shorter for high compared to low frequency words n. The effect was not significant for *single fixation durations*. The interaction between LOAD and word frequency was only significant for *set size 6* for *gaze durations* ($b = 0.084$, $SE = 0.039$, $t = 2.17$) as well as for *single fixation durations* ($b = 0.075$, $SE = 0.035$, $t = 2.13$). A 2 x 3 breakdown of frequency (median split) and load (load 1, 3 and 6) in Figure 3.5 displays reduced immediate word-frequency effects with load for *gaze durations* in the right panel and for *single fixation durations* in the middle panel. Thus, counter to expectation the strongest frequency effects were observed for low load conditions.

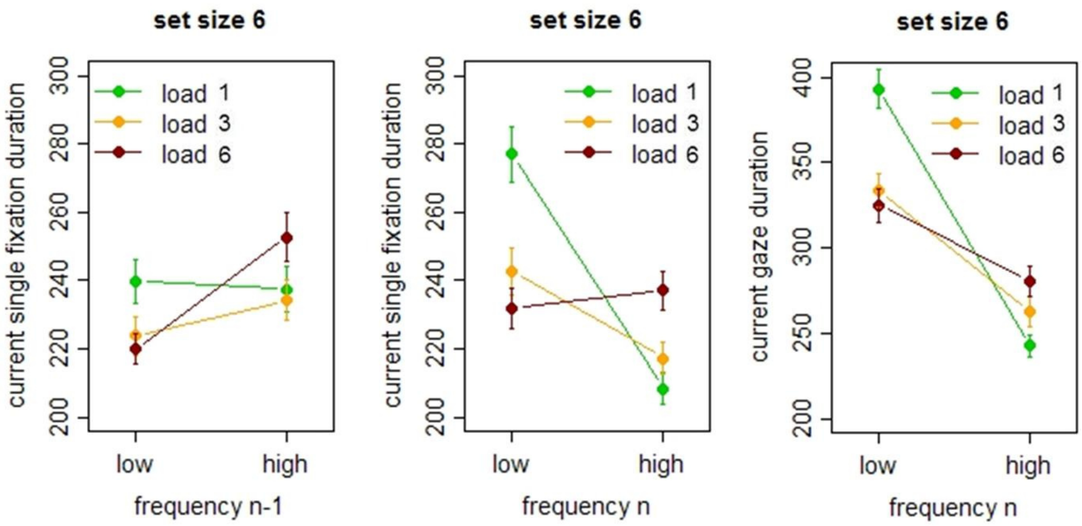

Figure 3.5: WM load differences on frequency effects: Plotted is the mean single fixation duration of the current fixated word n and its dependence on the previously fixated word n-1 (left panel) and its dependence on the foveal word frequency (middle panel). Current gaze durations and its dependence on WM load and the foveal word frequency are plotted in the right panel.

3.2.3.2 WM influences on lag effects

Parafoveal preview of a word is supposed to be reduced by the difficulty of the fixated word (e.g. Henderson & Ferreira, 1990). Processing a low-frequency word should narrow the attentional focus and reduces preview benefit for the upcoming word to the right of fixation. In other words, this should lead to longer fixation durations on a given word n when the previous word n-1 is of low frequency compared to the situation when n-1 is of high frequency. In addition, in the present study it was expected that WM load should reduce parafoveal preview of a word in the same way.

However, such a pattern could not be observed in the present data. The frequency of word n-1 did neither influence current *gaze durations* nor current *single fixation duration.* WM load had also no effect on fixation durations on word n. But the interaction between LOAD and the frequency of word n-1 significantly effected current *single fixation durations*. The interaction was significant for *set size 6*: b = 0.119, SE = 0.042, t = 2.83.

The lag frequency effect changed with load from a slightly positive effect, where word n is longer fixated when word n-1 is of high frequency; to a strong negative frequency effect, where word n is longer fixated after low frequency words n-1 (see left panel of Figure 3.5). This counterintuitive and negative n-1-frequency effect is mainly due to the fixation durations after low frequency words n-1, which reduced with load.

The unexpected lag frequency effect (in the left panel of Figure 3.5), most pronounced for load 5, could be due to spillover effects. High frequency words are often skipped. Thus, the processing of these words (n-1) is assumed to spillover to the subsequent fixation of n. Lexical processing of word n-1 occur in addition to the processing of word n. High memory loads provide additional cognitive demands which result in an over-additive effect on fixation durations on word n if word n-1 was of high frequency because they can reduce the costs of skipping word n-1. Thus, skipping the word n-1 is supposed to increase n-frequency effects. However, including skipping of word n-1 as predictor had no effect for the present data. The interaction between skipping of word n-1 and frequency of word n was neither significant for gaze durations ($b = -0.050$, $SE = 0.031$, $t = -1.64$) nor for single fixation durations ($b = -0.022$, $SE = 0.027$, $t = -0.82$).

3.2.3.3 WM influences on successor effects

Successor effects, as lag effects, were predicted to be influenced by foveal load that is the frequency of the foveal word and additional WM demands, induced by WM load and capacity. Figure 3.6 displays *single fixation duration* on word n as a function of word frequency of word n+1 (median split). Low frequent parafoveal words n+1 generally increased fixation durations on word n (left panel in Figure 3.6). Moreover, in accordance with the foveal load hypothesis, a smaller successor effect was observed for low-frequency foveal words n (left panel). The interaction of frequency of word n and frequency of word n+1 reached significance for *single fixation durations* ($b = -0.066$, $SE = 0.024$, $t = -2.74$).

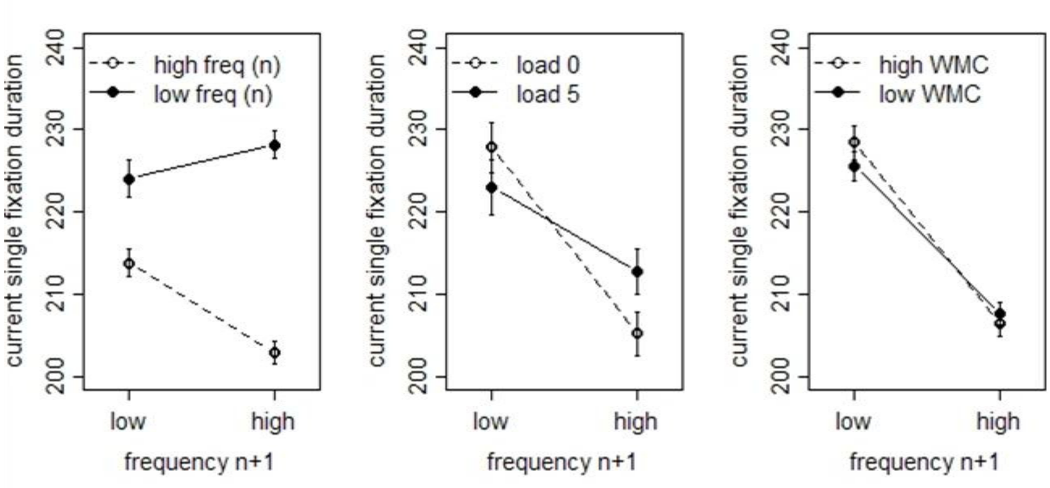

Figure 3.6: Foveal load influences on preprocessing word n+1: Plotted are the mean single fixation durations on the current fixated word n as function of the frequency of word n+1, for the frequency of word n (left panel), the WM load (middle panel), and WMC (right panel).

Moreover, as expected, high memory load reduced the influences of the upcoming word. The interaction between LOAD and the frequency of word n+1 was significant for *set size 5* with the full range of load (b = 0.097, SE = 0.040, t = 2.44). The contrast between load 0 and the highest load within set size was significant for *set size 6* (b = 0.259, SE = 0.119, t = 2.19). Weaker successor effects, as shown for high memory load, suggest a reduction of the perceptual span.

Furthermore, successor frequency effects were weaker for low than high capacity groups, but only marginally so (b = 0.017, SE = 0.008, t = 2.13). The results are expected from the proposition that distributed processing is more focused for high foveal load that is low frequency foveal words, high WM load and low WMC.

3.3 Summary of Results and Discussion

In the present chapter the influence of WM on eye movements was replicated for a modification of the reading span task. Results of the experiment in the first chapter were influenced by the subjects' strategy to start reading at the end of the sentence. In the present experiment subjects were forced to start reading at the beginning by masking all words except the sentence-initial one, and by revealing the remainder of the sentence once the eyes fixated the first word. There were, however, a few modulations by the experimental design. For example, faster reading times and less WMC variability on target fixation-durations were observed.

Moreover, empirical evidence was provided that WM dynamically modulates the perceptual span as proposed by the foveal-load assumption (Henderson & Ferreira, 1990). In the following I first focus on results replicated for global eye movement measures, and their modulation by the experimental design. Then the WM relevance for the perceptual span will be discussed and some critical points will be considered.

3.3.1 Marginal evidence for WMC differences in memory encoding

As in Experiment 1 there were small but consistent influences of WM load and WMC which both affected sentence processing time and fixation durations. Low capacity subjects showed higher increase in sentence processing times and fixation durations with sentence position (i.e., memory load) and set size than high capacity subjects. Hence, they had less time for pure memorizing of the target word(s).

Furthermore, fixation durations increased with word position within a sentence, with the longest duration on target words. The total fixation times on target words however, showed low variability with WMC. The group differences are exclusively based on the lowest capacity group, who not only fixated longer on the targets, but also on words in the middle part of the sentences. Moreover, the longer total fixation durations for low capacity subjects were only observed in the condition of fixed sentence presentation.

Unfortunately, WMC differences were not confirmed by gaze durations. Thus, it seems that the distinct WMC difference on targets total and gaze durations in the first experiment were mainly determined by encoding and storing mechanisms of the target, as a cause of the specifics of that experimental design, which, however, resembled more closely the procedure used in the standard reading span task..

3.3.2 General strategy versus a pure result of design

Sentence processing times, regressions times and number of first-pass fixations were reduced when subjects were forced to start reading at the beginning of a sentence. Their absolute values were compatible with those of the first serial positions in Experiment 1. Therefore, the current results provide clear evidence that in Experiment 1 the higher fixation and regression rates and longer processing times with beginning of the second sentence in a trial were a result of the design.

Independently of the fairly constant number of first-pass fixations, the initial fixation number on the target occurred earlier with an increase of WM load. The subject's strategy to fixate the target as fast as possible was strongest for the lowest set size. High as well as low capacity subjects used this strategy in the condition with lowest memory demand. For higher memory demands (i.e., higher set sizes), the strategy however, missed the conventional level of significance. Possibly, higher WM load induces more variance in the reading profile, leading to a weaker effect overall. Nevertheless, the results provide empirical support for a general strategy subjects used during the reading span task, that is, trying to fixate the target as early as possible.

3.3.3 WM relevance for the perceptual span

WM load and WMC in the reading-span task reduced the perceptual span as expected by the foveal-load hypothesis (Henderson & Ferreira, 1990; i.e., that the perceptual span is modulated by foveal difficulty). This was indirectly measured by the influence of the neighbouring words' frequency. Successor word frequency effects were reduced for low frequent foveal words, in high memory load conditions and for low WMC groups. When the number of target words maintained in WM is low, readers have more capacity to process the next word in parafoveal vision than when a high number of targets have to be maintained. Moreover, also high capacity groups have more capacity to preprocess the upcoming word than low capacity groups.

In the present experiment a well-established frequency effect of the foveal word and of the upcoming word (n+1) in single fixation durations was replicated. Frequent words were longer fixated than infrequent words. In addition, a reversed frequency effect was observed for word n-1. This unexpected effect was more pronounced for high WM load and was caused by reduced fixation durations if the previous word was of low frequency.

WMC showed only marginally influences on successor frequency effects and for WM load the effects were only significant for set size 5 and 6. Nevertheless, as a first attempt the present results provide direct experimental evidence that WM load and WMC dynamically modulate the perceptual span. This result is consistent with Kennison and Clifton (1995) who also failed to find evidence that low WMC acts like high WM load. Thus, WMC appears to be a weaker factor modulating the perceptual span than WM load.

3.3.4 Summary and Outlook

With the present experiment some unresolved questions of the first experiment were answered. A general strategy observed in both experiments was that subjects fixated the target word as fast as possible. The individual WMC differences on target words in the first experiment were caused by additional time of target encoding and storing for low capacity groups. Forcing subjects to start reading with the sentence-initial word reduced sentence processing times. This moreover, increased the time for pure memorizing the target words by on average 343 ms. In addition the results support the view that the perceptual span is dynamically modulated by WM load and WMC. High memory demands (low word frequency and high memory load) and low memory capacities reduced the perceptual span as proposed by Henderson’s and Ferreira’s (1990) foveal load hypothesis.

The report of analysis in the present chapter concentrates on WMC differences and load. As mentioned before old and young adults were tested. The next chapter therefore, focuses on the age dependent results.

Chapter 4

Age differences in global eye-movement measures and in the perceptual span

Reading is a highly practiced and often used skill. Throughout the years older adults become reading experts. Thus, reading seems to be an automated process preserved with age (see Caplan & Waters, 1999). However, age-related restrictions in vision and cognition might result in reading impairments for older adults. Visual restrictions are mainly due to reduced contrast sensitivity and the reduction of central and peripheral acuity with age (for a review see, Fozard & Gordon-Salant, 2001). As mentioned, for several cognitive functions an age-dependent decline was also discussed. Thus, a general reduction in processing speed (Salthouse, 1996), attentional resources (Craik, 1983), inhibitory control (Hasher, Stolzfus, Zacks, & Rypma, 1991), and long-term and WM (Park et al., 2002) were observed. These age-dependent declines in cognition were claimed to be the outcome of a common cause, which Lindenberger and Baltes (1994, 1997) described as aging brain.

Facing these age-dependent declines it seems astonishing that such a complex cognitive skill as reading, which integrates many functions that mostly show declines, is not influenced more. Hence, an interesting question is how old adults compensate for their declines during reading.

A large number of studies investigated age dependencies in eye-movements during reading. Most notably, old adults show more (Kemper et al., 2004; Kemper & Liu, 2007; Kemper & McDowd, 2006; Kliegl, Grabner, Rolfs, & Engbert, 2004; Rayner, Reichle, Stroud, Williams, & Pollatsek, 2006) and longer fixations (Kliegl etl al., 2004; Rayner et al., 2006; Stine-Morrow, Noh, & Shake, 2010). In addition longer saccade amplitudes were observed, combined with higher skipping rates. Moreover, higher skipping rates resulted in more regressions back to previously skipped words compared to young adults (Laubrock, Kliegl, & Engbert, 2006; Rayner et al., 2006). Furthermore, Rayner, Castelhano and Yang (2009) observed a slightly reduced perceptual span for old adults in a moving window experiment, in which only words inside a certain window were unmasked during each fixation. Old compared to young adults showed weaker parafoveal-on-foveal effects of the word right to the fixation. Moreover, reading times of old readers were not reduced when only two words were visible within the window. While young readers showed impaired reading in a two-word window condition (i.e., word n and n+1 visible), old adults' reading times were not reduced when the two words were visible within the window. This result indicates that old adults can read efficiently without processing information beyond the neighbouring word n+1, while young adults seem to use parafoveal information up to word n+2 (i.e., three-word window). Thus, for old adults the perceptual span was reduced in the direction of reading. It was more symmetric around the fixation position compared to a rather asymmetric span size of young adults (see Chapter 1.1.2 for more details about the perceptual span). Rayner et al. (2009) assumed that higher skipping rates let old adults compensate for their slower reading rates and their reduced perceptual span.

However, contrary to Rayner and colleagues Risse and Kliegl (2011) could not find evidence for reduced span sizes for old compared to young adults. If word length of word n+1 was controlled (i.e., always three letters long), young and old adults did not differ in the amount of preview benefit. This was even true for effects of word n+2 on word n. Thus, Risse and Kliegl's results exhibit only very small age differences in the size of the perceptual span during reading (see also Kliegl et al., 2004). This result is in agreement with findings that more global skills and processes, such as verbal knowledge and lexical processing, are relatively constant across the lifespan (e.g. Lima, Hale, & Myerson, 1991; Mayr & Kliegl, 2000; Park et al., 2002; Verheaghen, Cerella, Semenec, Leo, & Bopp, 2002).

Age effects on measurements of eye movements during reading were often associated with declines in working memory. For example, Cerella (1990) showed that age differences increase with higher WM demands through higher task complexity. Furthermore, Kemper and Liu (1997) manipulated WM load in a reading task by varying task complexity via syntactic ambiguity and subject vs. object relative sentences. Old adults compared to young adults showed more regressions and longer fixation times only for high ambiguous, object relative sentences. Therefore, one could assume that the age-dependent increases in fixation durations, regression and skipping rates, are the result of cognitive limitations in the 'aging brain', which correspond to higher cognitive load on WM.

Facing the results of WM influences on age effects in reading, the different results of Rayner et al. and Risse and Kliegl may be attributed to differences in cognitive load evoked by the different text material between studies.

A second line of research relating age effects on measurements of eye movements during reading and WM focuses on individual differences in WMC. In two studies, Kemper and colleagues (Kemtes, & Kemper, 1997; Kemper et al., 2004) monitored different WMC groups of young and old adults during reading garden-path sentences. In these sentences a temporary ambiguity is established and resolved at the end of a sentence. Therefore, two sentence meanings have to be maintained from which one must be inhibited as fast as the ambiguity is resolved. Low WMC readers are associated with reduced abilities to maintain multiple meanings (Miyake et al., 1994) and to inhibit inconsistent or irrelevant meanings (Gunter, Wagner, & Friederici, 2003; Hartman & Hasher, 1991). As old adults show impaired performances in WM tasks (see for example Park et al., 2002), Kemtes and Kemper hypothesized that the group of old adults would show the same resolution problems as young low WMC readers. However, contrary to Kemper et al.'s expectations, old adults were not comparable in general to young low WMC readers. Only old adults with low WMC had problems to answer questions about the ambiguous sentences (Kemtes & Kemper, 1997). Moreover, old readers in general showed more regressions but in a reading time which was comparable to that of young readers (Kemper et al., 2004).

There seems to be a dominant pattern of age-related increases in fixation durations, skipping rates, and especially in the number of regressions. Whether such age effects further vary with task complexity and, thus, with WM demands, and individual WMC is still an open question. The present chapter examines the mediating influence of WM on the relation age on reading measurements by manipulating WM load for different age groups which also vary their WMC (high vs. low capacity).

If age differences in eye movements rely mostly on WMC differences as supposed by some authors (Just & Carpenter, 1993; Kemtes & Kemper, 1997), matched WMC groups of young and old adults should not differ in their eye-movement behaviour. From this perspective, high-WMC young and high-WMC old adults, as groups of young and old readers with low WMC, are expected to show the same gaze pattern. Effects of parafoveal processing such as parafoveal-on-foveal effects should be reduced, mainly for old adults with low-WMC. If differences between young and older adults both with high-WMC will be observed, they can not be attributed to WMC differences alone. Further cognitive processes like inhibitory control must be assumed to be involved to determine such age effects.

The present chapter focuses on the age effects in sentence processing times, fixation durations, and regressions with respect to WMC and WM load. It will be investigated if old adults in general have reduced fixation durations and regression rates, or whether this rather depends on the individuals' WMC. Furthermore, old and young adults are compared in their amount of parafoveal vision.

4.1 Method

Method with subjects, material and procedure as described in chapter 3.

4.2 Results

A typical age effect was found in scores on Lehrl's (Lehrl, 1977) multiple-choice measure of vocabulary and in Wechsler's Digit-Symbol-Test (Wechsler, 1964). Old readers attained a significantly higher vocabulary score, $F(1, 115) = 33.018$, $p < .001$, in comparison to the young group, and a significantly lower score in the Digit-Symbol-Test, $F(1, 115) = 91.525$, $p < .001$. These results support typical findings in age research. Higher vocabulary scores for old adults are expected when considering that old adults are reading experts who already read an enormously large number of words throughout their life. Therefore, for old adults infrequent words and foreign words are also rather common. Lower scores in the Digit-Symbol-Test, however, are an index for reductions in processing speed.

In the following, I report effects guided by the earlier described hypothesis. The complete list of effects included in the respective LMM is listed in Appendix D.

4.2.1 General age dependencies in eye-movement measures

Longer fixation durations and higher regression rates were found for older, compared to younger readers. Old adults *total fixation duration* was on average 28 ms longer than for young readers ($b = 0.095$, $SE = 0.025$, $t = 3.84$, in Table D.2). For gaze and single fixation durations the main effect of age did not reach significance. In comparison to young readers, old readers make 7 % more *regressions* ($b = 0.86$, $SE = 0.29$, $p < .01$, in Table D.6). The interactions of AGE with WMC neither reached the conventional level of significance for total fixation durations, nor for the regression rates.

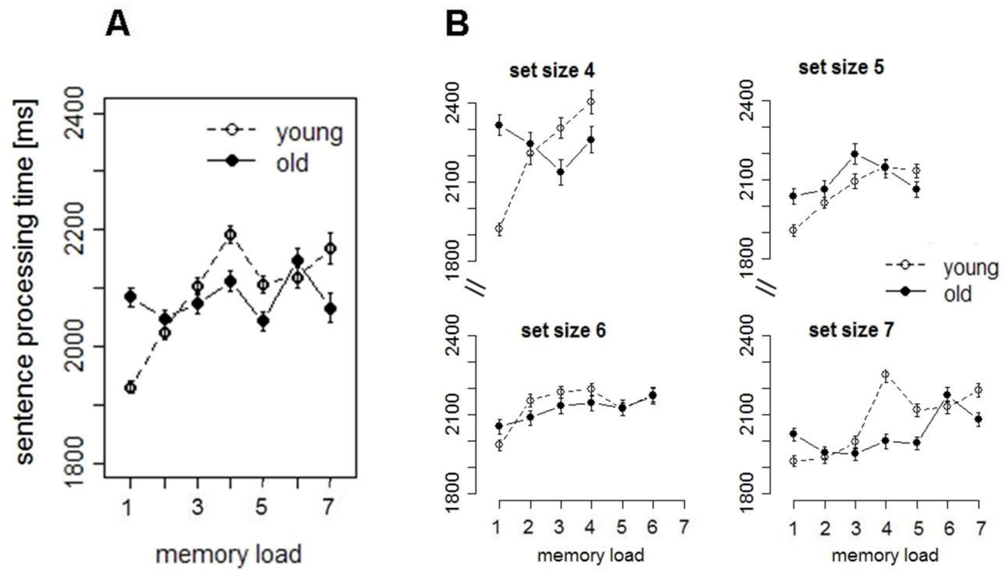

Figure 4.1: Sentence processing times as function of memory load for young and old adults. (A) aggregated over all set sizes; (B) for each set size.

4.2.2 Age dependent influences of WM load

Figure 4.1 A displays the interaction of AGE and LOAD on mean *sentence processing times* aggregated across all set sizes. The sentence processing time is measured as the reaction time in which subjects were able to make a decision on the sentence content irrespective of how many sentences were presented. Sentence-processing times only increase for young adults with WM load, whereas old adults stay fairly constant across load. The graph of young adults can be divided into two parts. In the first part, when up to four sentences were presented, sentence-processing times increased linearly. In the second part, which is for the last three sentences, the increase in sentence-processing time is reduced compared to the first part. Such discontinuity can be an indicator for changes in cognitive functioning. The interaction between AGE and LOAD was significant for *set size 4:* b = -0.096, SE = 0.020, t = -4.83; *set size 5:* b = -0.035, SE = 0.016, t = -2.22; and *set size 7*: b = -0.044; SE = 0.016, t = -2.84 (in Table D.7). In set size 4, old readers actually reduce their reading times with load.

This contra-intuitive reduction in sentence processing times with load for old adults for set size 4 further interacts with the condition of sentence order. The interaction between AGE, LOAD and SENTENCE ORDER was only significant for *set size 4* (b = 0.164, SE = 0.040, t = 4.11, in Table D.9). Old adults sentence-processing time is relatively constant for the condition with fixed sentence order (left panel of Figure 4.2) and decreases before it constantly increases with the random sentence order (right panel of Figure 4.2). The reduction in sentence-processing times with load seems to be most pronounced at the beginning of a trial. Thus, old readers may anticipate the upcoming memory demands which they possibly try to compensate by longer sentence-processing times to better encode the first targets (see also Figure 4.1 A, at memory load 1).

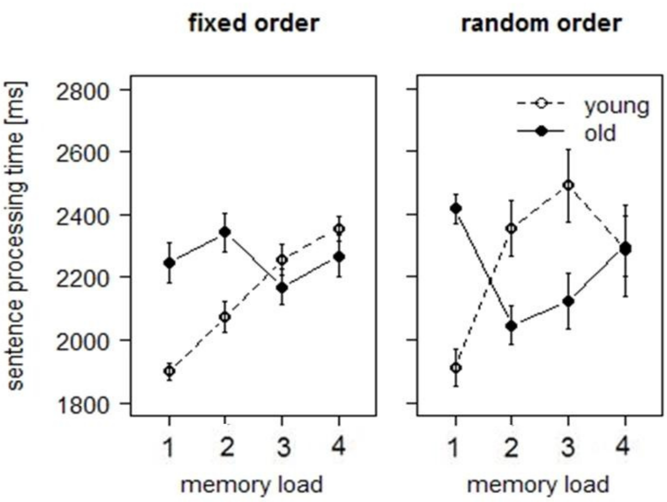

Figure 4.2: Sentence processing time as function of memory load, for both age groups and both sentence orders (fixed, and random).

Young and old readers does not differ in their increase in sentence-processing times, shown in Figure 4.3. Here, the sentence-processing time of young and old subjects is plotted as function of the subject's free memory slots. The number of free memory slots is zero, if the WM load of a sentence equalled the individual capacity. Positive values indicate an "underload", whereas negative values indicate an "overload", (i.e., WM load is above the WMC). The curves for young and old subjects in Figure 4.3 vary as a function of free memory slots and age, but only the memory effect is significant. The curves for old and young readers overlap with one exception: Old adults show longer sentence-processing times at the highest amount of freely available resources (6 memory slots*)* compared to their younger controls. This result statistically translates into a significant interaction between AGE and RESOURCE (b = 0.0119, SE = 0.00196, t = 6.07, in Table D.9). A LMM with nested ANOVA contrast specification for the factor RESOURCE shows a significant interaction between AGE and RESOURCE only for the 6 free memory slots condition (b = 0.161, SE = 0.061, t = 2.65). Thus, matched in the amount of free resources, young and old adults did not differ in their sentence-processing times, except for the maximum number of free slots. Therefore, the age-related increases in processing time with load (shown in Figure 4.1) seems to be due to WMC differences and less to age differences per se. The modulation of age effects by individual WMC is outlined in the following paragraph.

Figure 4.3: Sentence processing time as function of free memory slots, for both age groups**.**

4.2.3 Working-memory capacity and age

Figure 4.4 displays the sentence-processing times comparing only the highest (7) and lowest (4) capacity groups of old and young adults as a function of memory load. Old adults with low WMC show higher mean sentence-processing times and a contra-intuitive reduction in their reading times with increasing memory load from 0 to 3. Sentence-processing times of old adults with high-WMC stayed fairly constant around 2100 ms. Young and old readers both with high WMC do not show significant differences in their sentence processing times. The mean fixation times for capacity groups are listed in Table 4.1 (old adults) and Table 4.2 (young adults). The increase in sentence processing times within the first four load conditions is exclusively shown by young adults. Here the increase is more pronounced for high than for low-WMC groups. Furthermore, the WMC groups of young adults differ in particular for WM load higher than 4. The interaction between Age, Capacity (reduced to group 4 and 7) and Load reached significance for set size 5 ($b = 0.196$, $SE = 0.052$, $t = 3.8$) and set size 6 ($b = 0.309$, $SE = 0.052$, $t = 5.9$, in Table D.12).

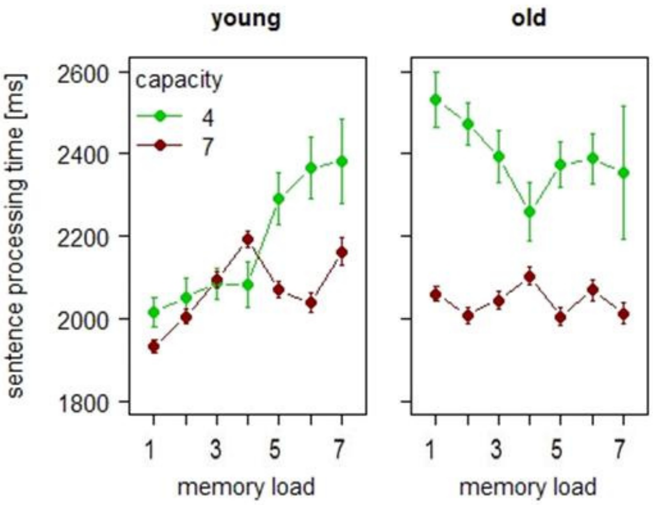

Figure 4.4: Sentence processing times as function of memory load and capacity (only extreme groups 4 and 7), left panel for young and old adults (right panel)

Table 4.1: Summary: Capacity groups old adults

		Capacity groups of old adults			
Eye movement measures		4 (n = 4)	5 (n = 9)	6 (n = 8)	7 (n = 33)
Sentence processing time (ms)	MD	2373	2124	2072	2075
	SD	*547*	*524*	*541*	*519*
Total fixation time (ms)	MD	460	411	424	406
	SD	*271*	*246*	*266*	*250*
Gaze duration (ms)	MD	347	301	296	308
	SD	*205*	*189*	*198*	*186*
Single fixation duration (ms)	MD	256	236	229	240
	SD	*91*	*102*	*104*	*105*
Number of first pass fixation	MD	1.37	1.27	1.29	1.28
	SD	*0.63*	*0.56*	*0.59*	*0.55*
Probability of skipping	MD	.21	.29	.37	.29
	SD	*.41*	*.45*	*.48*	*.45*
Probability of regression	MD	.27	.31	.36	.28
	SD	*.44*	*.46*	*.48*	*.45*
Initial fixation on target	MD	4.95	4.43	4.43	4.18
	SD	*1.76*	*1.63*	*1.42*	*1.50*

Furthermore, WMC-dependent differences in *gaze durations* are only observed for old adults (see Figure 4.5). Low-WMC old adults gaze durations were on average 45 ms longer than those for all other groups and slightly increased with load. However, gaze durations of high-WMC old adults (308 ms) are comparable to that of high-WMC young adults (293 ms). The interaction between AGE, CAPACITY and LOAD reached significance for *set size 6* (b = 0.401, SE = 0.124, t = 3.23, in Table D.13).

Table 4.2: Summary: Capacity groups of young adults

Eye movement measures		Capacity groups of young adults			
		4 (n = 5)	5 (n = 9)	6 (n = 17)	7 (n = 30)
Sentence processing time (ms)	MD	2477	2286	2145	2071
	SD	*637*	*621*	*574*	*510*
Total fixation time (ms)	MD	450	408	408	372
	SD	*323*	*302*	*265*	*243*
Gaze duration (ms)	MD	304	312	312	293
	SD	*214*	*215*	*208*	*193*
Single fixation duration (ms)	MD	230	235	240	222
	SD	*110*	*123*	*123*	*104*
Number of first pass fixation	MD	1.56	1.48	1.43	1.35
	SD	*0.72*	*0.71*	*0.68*	*0.62*
Probability of skipping	MD	.22	.20	.24	.18
	SD	*.42*	*.40*	*.43*	*.38*
Probability of regression	MD	.39	.29	.29	.24
	SD	*.49*	*.46*	*.46*	*.43*
Initial fixation on target	MD	3.73	3.80	4.02	4.37
	SD	*2.40*	*2.31*	*1.99*	*1.69*

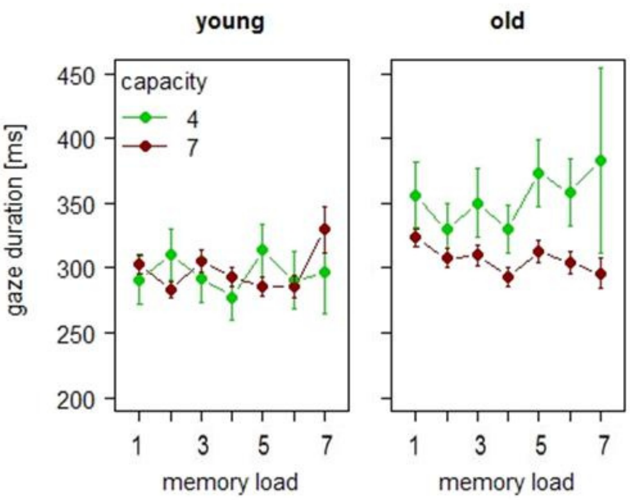

Figure 4.5: Gaze duration as function of memory load and capacity (only extreme groups 4 and 7), left panel for young and right panel for old adults .

4.2.4 AGE dependent parafoveal-on-foveal effects

It was further investigated how age influenced parafovea-on-foveal effects. Therefore, the influences of the frequency of the neighbouring word to the right of fixation on the current fixation duration (i.e., successor effects) were examined. As reported in Chapter 3, low frequent parafoveal words n+1 generally increased single fixation durations on word n (Figure 3.6) compared to the situation when n+1 is of high frequency.

Assuming reduced perceptual spans for older compared to young adults, this effect is expected to be reduced for old adults. However, such a pattern could not be observed in the present data. The interaction of AGE and the frequency of word n+1 neither reached significance for *single fixation durations*, nor for *gaze durations*. The t-values of the effect were less than 1.5 and thus, taken out of the models listed in Appendix C. Furthermore, the interaction between AGE, CAPACITY and the frequency of word n+1 also did not reach significance.

Only the well-established frequency effect of the fixated word n (i.e., immediacy effect) is modulated by age. Figure 4.6 displays *single fixation durations* as a function of current word frequency. The frequency effect of the fixated word is

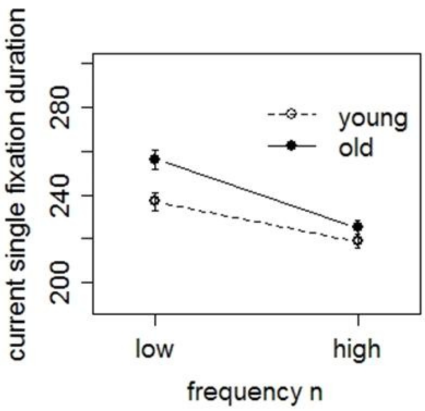

Figure 4.6: Immediate frequency effects of young and old adults. Plotted is the single fixation duration on the fixated word.

slightly stronger for old than for young adults (b = -0.056, SE = 0.014, t = -3.99, in Table D.11). Old adults' single fixation durations are more strongly increased, if word n is of low frequency, than young adults'. For *gaze durations*, this effect was even stronger (b = -0.036, SE = 0.015, t = -2.33, in Table D.10). Thus, old rather than young adults seem to be slowed-down by processing low frequency words.

4.3 Summary of Results and Discussion

The analyses reported in the present chapter focused on age-dependent influences of WM load and capacity on eye-movements during the modified reading span task. There were, however, only a few global modulations by age. With the present study the findings of longer fixation durations (Kliegl et al., 2004; Rayner et al., 2006; Stine-Morrow et al., 2010) and more regressions (Laubrock, et al., 2006; Rayner et al., 2006) of old adults were replicated.

4.3.1 Age effects depend on WMC

The general slowing of old adults is well known for reaction time experiments. In these experiments the absolute age differences increase with increasing task difficulty. This relation is known as age-complexity effect (Cerella, 1990). However, the present results in reaction time (i.e., the sentence-processing time) were contrary to this well-established effect of cognitive aging. Only young adults systematically increased processing time across serial positions of sentences in a trial (especially the first 4 serial positions) reflecting the increase in WM load, whereas old adults allocated a similar amount of time, irrespective of serial position.

Further analysis showed that older and younger adults exhibited similar increases in sentence-processing times when the number of freely available memory slots was taken into account. Thus, the lower the number of free memory slots (resources), the more time was used for sentence processing. This effect was shown by both age groups at the same level, such that the curves of old and young adults overlapped.

Thus, when matched in the amount of freely available resources, old and young adults need the same time for sentence processing. That age effects in lexical processing appear to be much smaller than expected from age-dependent declines in cognitive functioning was already reported by, for example, Laubrock et al. (2006; see also Lima et al., 1991; Mayr and Kliegl 2000; Verheaghen et al. 2002).

The lack of age effects across different levels of resources suggests a reconsideration of age differences with respect to WM load. It is rather possible that the age differences across load are due to WMC differences, and not to WM load per se.

The results of fixation durations support the assumption that age effects rely most exclusively on old adults with low-WMC. Only old adults with low WMC showed increased gaze durations, whereas the high-WMC old and the high-WMC young adults stayed fairly constant across load.

Kemtes and Kemper (1997) already observed that processing of ambiguous sentences was exclusively impaired for low-WMC but not for high-WMC old adults. This dissociation with respect to WMC had not been shown for eye-movements during reading yet, as the focus of typical aging studies was rather on analyzing general age effects. The results of the present study emphasises the importance of differences in WMC, if one aims to understand age-related changes in eye-movements during reading.

4.3.2 Age and the perceptual span

The age by complexity interaction of cognitive ageing predicts an age-related increase in fixation duration with increasing word difficulty, i.e., for words of low frequency. Old adults indeed showed stronger frequency effects of the foveal word, so they more strongly increased single and gaze durations for low frequency words n. Furthermore, following the results of Rayner et al. (2009) one would expect reduced influences of the word right of fixation for old adults. However, the parafoveal-on-foveal influences in the present study missed the conventional level of significance.

4.3.3 Summary

To increase task complexity and to control age and WMC during reading is a promising basic approach for eye-movement research. The present results underline the importance to investigate age-related differences with respect to WMC and WM load. The age comparison supported the notion that reading is a highly automated process, well preserved, for high-WMC old adults with similar overall processing times and gaze durations compared to young readers.

Chapter 5

General Discussion

This dissertation explored the importance of working-memory processes for eye-movements during reading. The comprehension of sentences is linked to different reading skills, and to how working memory can explain individual differences. This is in contrast to the eye-tracking research, where eye-movements are recorded during sentence reading under conditions in which subjects should easily understand the sentences. Thus, the question remains of how reading skills modulate the level of sentence comprehension. The present work bridges the gap between both areas of research, as well as to the research field of working memory by investigating how eye-movement patterns changed for readers with different WMC when WM load was increased.

The *first* central issue addressed in this thesis was the investigation of the impact of working-memory capacity and working-memory load on global eye-movement patterns during the reading-span task. Thus, in a first step I implemented the traditional reading-span test without any manipulation of the design to accommodate eye tracking (Chapter 2). In all sentences, target words were the final words of the sentences. Task demand (i.e., WM load) increased over trials. The subjects' WMC was predictive for the probability of refixating targets and the total time spent on the critical words. Fixation durations increased across the sentence, with the longest durations on the sentence-final words (i.e., a sentence wrap-up effect), reflecting both comprehension and time for encoding the target word. There was also a gradual increase in sentence-processing times related to WMC. Readers with low WMC slowed down their reading as a function of increasing WM load more than readers with high WMC. An unexpected side result was that the number of fixations on the target (sentence-final) words was higher than for all other words and reading of a sentence tended to start with the sentence final word.

The results of the first experiment were influenced by the subjects' strategy to start reading at the end of the sentence. Thus, in the following experiment (Chapter 3 and 4) subjects were forced to start reading at the sentence beginning by masking all words except the sentence-initial one and by revealing the remainder of the sentence once the eyes fixated the first word. Nevertheless, subjects showed the same strategy as in the first experiment by fixating the target word as fast as possible. The attitude of the subjects was therefore independent of the experimental design and thus, indicates strategic allocation of processing resources.

The previous results of Experiment 1, that WM load influenced sentence processing times, were replicated. Load affected fixation durations were also contingent on individual differences in WMC. Readers with high WMC showed only a slight increase with sentence position and set size, and therefore they had more time for pure memorization of the target words.

The *second* question addressed in this thesis was how working memory influences the perceptual span during reading. Thus, it was investigated how working-memory load and working-memory capacity affect lexical access (e.g. frequency effects), and foveal and parafoveal-on-foveal effects (Chapter 3). WM load reduced the perceptual span, as expected by the foveal-load hypothesis (i.e., that the perceptual span is modulated by foveal difficulty). However, WMC had only a marginal effect on the perceptual span. Thus, WMC appears to be a weaker factor in dynamically modulating the perceptual span than WM load.

To maximize individual differences in WMC, I tested young and old adults (Chapter 4). The *third* central issue was to explore if age differences in eye movements rely mostly on WMC differences. The results emphasised the importance of differences in WMC, when one aims to understand age-related changes in eye movements during reading. In particular, old adults with high WMC showed similar overall processing times and gaze durations as young readers. There were, however, a few modulations by age. The well-established effects of longer fixation durations and more regressions of old adults were replicated. Moreover, young and old adults processing times and gaze durations responded differently to WM load. Young adults' durations increased with WM load, whereas the durations of old adults stayed fairly constant. This age effect disappeared when subjects were matched in their amount of freely available resources.

5.1 What is crucial - the sentence or the reader?

The complexity of a text determines the reading behaviour and the time needed for sentence comprehension. Buswell (1922) showed that reading times for words in fiction were shorter than for words in technical literature. It was assumed that complex sentences require more cognitive effort than easy sentences. In the experiments of this thesis the complexity was manipulated independent of the sentences. Not the complexity of the sentences increased but the number of target words a subject had to maintain during reading. Nevertheless, a high number of target words lead to high working-memory load. Thus it also leads to more cognitive effort than in cases when no target word had to be maintained. Therefore, a high number of target words mirrored a high complex sentence or text.

In addition, I expected that the individual reading skill, and thus, the cognitive system and, in particular, the working-memory capacity, would be responsible for differences in eye guidance during reading. In fact, individual constrains (of the reader) as well as the memory demands (of the sentence) determined the reading behaviour. Moreover, the online working-memory load turned out to be more important than the influence of the individual's working-memory capacity. These results will be discussed in the following paragraphs, in reference to the current state of research.

5.2 WMC differences are more than a variation in resources

Consistent with previous studies (e.g., Carpenter & Just, 1992; Engle et al., 1992; Friedman & Miyake, 2004), memory load had a significant impact on processing time devoted to the test materials. The more target words there were which had to be maintained, the more time was needed for sentence processing. This main effect is in agreement with the assumption of a limited pool of resources (e.g. Just & Carpenter, 1992). In accordance to that assumption, a fixed number of resources must be shared for all memory representations and processing tasks. Thus, the more representations (e.g. target words) are maintained in memory, the smaller the amount of resources that is available for each particular one. Several authors (e.g. Just & Carpenter, 1992) assumed that skilled readers (with high WMC) have more resources than less skilled readers (with low WMC). From this perspective one would expect longer reading times for readers with low WMC, which is what I found.

The results strongly suggest that the lower the WMC of a reader the stronger the increase in reading times with load. Differences in reading times stayed, moreover, constant if subjects were matched on their freely available resources. This result stands in contrast to the assumption that the number of resources alone is responsible for differences in WMC. If that was the case, comparable times would be expected after matching. However, subjects still differed with respect to their WMC. These differences may be evoked by basic differences in cognitive processing.

5.3 Individual differences in attentional control

One important factor for the differences between subjects seems to be the selectivity of attentional control and associated mechanisms of inhibition. Hasher and Zacks (1979) gave the impulse to integrate the inhibition theory with the general theory of capacity limitations. Thus, the degree of attentional control defines the efficiency of inhibitory processes. Limited attentional capacity is shared between elements that had to be actively maintained in working-memory. The more elements have to be maintained the less attentional control is available for each particular one. Based on this attentional reduction, inhibitory processes are less efficient. As a consequence, more irrelevant information is allowed to enter working-memory and the relevant information is more susceptible to interference. The cognitive system has to be robust and flexible at the same time. The cognitive system has to be robust and flexible at the same time. Robust enough to inhibit task irrelevant information but flexible enough that new relevant information find its way into memory.

Hasher and Zacks (1979) argued that attentional capacity and thus, inhibitory processes vary both within and between individuals. They applied this assumption of individual differences to the concept of working-memory (Hasher and Zacks, 1988). Moreover, Engle (1996) claimed, that the control of inhibition in working memory is a process constrained by the available attentional resources. Therefore, working-memory capacity is related to attentional control, and individuals differ in their ability to focus and maintain attention.

The inhibition hypothesis received several empirical supports. Subjects differed in their ability to resist task irrelevant information and to deal with interference, dependent on their working-memory capacity (see Unsworth et al., 2005 for a review). Results support the view that individuals with high WMC in contrast to low WMC are more able to resist an interfering attention capturing cue when it conflicts with task goals. Thus, subjects with high WMC reported less than subjects with low WMC to hear their names during inspection of a relevant message (Conway et al. 2001). They better resist giving an automatic orienting response to a flashing cue, which appears on screen and must be ignored (Hallet, 1978). They also better comprehend ambiguous sentences (e.g. Miyake et al., 1994).

The results of the present thesis replicate earlier findings and moreover, provide support for Hasher and Zacks (1979) statement that attentional capacity varies within and between individuals. By increasing working-memory load the sentence-processing times increased within all subjects. However, the increase was stronger the lower the WMC. Matched on their freely available resources fundamental differences between groups preserved.

The inhibition hypothesis could explain why young readers of low and high WMC still show differences in their readings times if the number of resources is taken into account. Given that low working-memory capacity is caused by inhibition deficits, the readers differed in their ability to focus and maintain attention. A computational implementation of the theoretical assumptions of inhibition processes which are linked to attentional control is given with the interference model by Oberauer and Kliegl (2001). Besides the two other parameters, a noise parameter is postulated to explain that some cognitive systems are more susceptible to interference than others. High noise levels in the cognitive systems prevent the focus on task relevant information, which moreover results in longer processing times per se. For the present data one would assume that the noise value is highest for readers with low working-memory capacity. The implementation of the interference model for the present data would be a next step to fortify the given interpretation.

5.4 Attentional reduction with working-memory load

In the reading-span task the working-memory load successively increased from the first to the last sentence in a task by experimentally manipulating the number of target words which were to be maintained during sentence reading. As just described, this results in an increase of sentence-processing times. Furthermore, fixation durations in first-pass reading and particularly in second-pass reading increased with increasing working-memory load. The findings replicate results published by

Kaakinen and Hyönä (2007; see also Just & Carpenter, 1992). In accordance with the previous interpretation, they support the view that also global eye-movement measures were influenced by a reduction in attentional control. The more target words there were which had to be maintained, the less attentional resources were available for the eye-movement control, which moreover resulted into longer fixation times.

Working-memory load furthermore increased within a sentence, which resulted into an increase of fixation durations across the sentence. The longest durations were shown on sentence-final words. Longer fixation durations at the end of a sentence represent a wrap-up effect and reflect integration of information and comprehension of the sentence (e.g., Just & Carpenter, 1980; Mitchell & Green, 1978; Rayner et al., 1989, 2000). In the present study, they additionally reflect the time for encoding and storing the target word. However, the increase of fixation durations within a sentence is irrespective of the target words and replicates findings of Kuperman et al. (2010). Fixation durations increased from word-to-word, although words toward the end of a sentence are likely to be more predictable and of high-frequency, which would lead to reduced fixation times per se (e.g. Boston, Hale, Kliegl, & Vasishth, 2008; Ehrlich & Rayner, 1981; Rayner, 1998; for a review see Kuperman et al. 2010). This contra-intuitive effect was explained by Kuperman et al. to be influenced by language integration and comprehension processes that are not confined to the last word, but rather increases incrementally over several words within a sentence.

A second line of interpretation relates to working memory. For sentence comprehension the reader has to remember all word information until the end of the sentence. Given that the number of words in sentences often grossly exceeds the capacity of the working memory, words have to be bound together in chunks. Baddeley, Hitch, and Allen (2009) postulate that the memory for sentences is enhanced by direct interactions between language knowledge and the phonological loop and by attention demanding binding processes. Thus, it is also possible to assume that increasing fixation durations across the sentence are likewise due to attentional limitations caused by increasing working-memory load.

5.5 High WMC promises youthfulness in reading

Let us go one step back and consider the finding that the absolute working-memory capacity difference increased with increasing task difficulty. Such interaction is known from the aging research as age-complexity effect (Cerella, 1990). Furthermore, older adults were supposed to reflect the reading patterns of young adults with low WMC (Kemper & Liu, 2007). If one assumes that aging is determined by a decline in working-memory capacity (see Park et al., 2002) the age-complexity effect should likewise be determined by WMC differences. However, Kemptes and Kemper (1997) failed to find empirical support for this assumption. In their comprehension study, not all old adults were compatible to young readers with low WMC. In fact, only the old adults with low WMC showed compatible comprehension results and times to young readers with low WMC.

The results of the present thesis shown in Chapter 4 support the findings of Kemptes and Kemper. If subjects were matched on their freely available resources, old adults as a group showed the same reading times as the group of young readers. The lack of age effects across different levels of resources suggests a reconsideration of age differences with respect to WMC. By taking the WMC into account, a much more differentiated picture was observed Fastest mean reading times were shown by readers with high WMC irrespective of age groups, followed by the reading times of young readers with low WMC. The longest mean reading times were due to old adults with low WMC. Moreover, this differentiated picture was observed for various eye-movement measures. Only gaze durations of old adults with low WMC were significantly longer than that of young adults.

The results clear-cut the age groups with respect to their WMC. Counter to the assumption of a general age dependent restriction in sentence comprehension and eye-movement control, reading seems to be an automated process well preserved for old adults, especially those with high WMC (see also Caplan &, Waters 1999, who suggest a WMC independent system for sentence comprehension). Nevertheless, old adults, especially with high WMC, used compensation strategies which become apparent at the task beginning. Contrary to young adults, who systematically increased processing time and gaze durations across serial positions of sentences in a trial, reflecting the concomitant increase in WM load, old adults with high WMC allocated an equal amount of time to each sentence, irrespective of its serial position. Thus, they considered all sentences with a similar amount of time. This reading behaviour suggests that old adults anticipated higher WM load during a task already at its beginning. That made them use more time at the task beginning to thoroughly encode the initial target words and to create highly robust representations.

5.5.1 Counter intuitive age-dependencies of word frequency

The frequency of a word, that is how often it occurs in a given language, influences word processing (see Balota et al., 1985; Inhoff, 1989; Inhoff & Rayner, 1986; Rayner & Pollatsek, 1987). The more frequent a word, the faster and more accurate is its lexical access (Morrison & Ellis, 1995). A prominent effect in eye-movements during reading is that low frequency words induce longer first fixation durations and gaze durations than high frequency words (e.g. Inhoff & Rayner, 1986; Rayner & Duffy, 1986). Old and young adults did not differ in their frequency effect as reported in Kliegl et al. (2004). However, the results of the present thesis revealed a different picture (see Figure 4.6). For old adults the frequency effect of the foveal word turned out to be stronger. Compared to young adults, old adults more strongly increased single and gaze durations for low frequency words n.

These results are counter intuitive, if one assumes that old adults are reading experts who have already read an enormously large number of words throughout their life. Thus, for old adults low frequency words should rather be common. This assumption is supported by the vocabulary scores of my subjects. Old adults attained a significantly higher vocabulary score than the young adults, which is a typical finding in age research. In contrast, the results of the digit symbol test revealed reduced processing speed for the old adults. However, the vocabulary test was conducted without time pressure. Therefore, statements regarding the speed of lexical access are not possible. At this point, I can only provide an initial attempt to formulate an explanation of the observed age-dependent frequency effects. Under the assumption of a general reduction in processing speed, one could assume that the interaction of age-dependent frequency effects is interfered by the general slowing in single fixation durations of old adults. This in turn would lead to the assumption that for old adults, low frequency words did not take longer, but high frequency words more quickly attained lexical access. Such an interpretation would also conform to findings of a reduction in older adults' resilience in response to processing ease in the perceptual span (see Risse & Kliegl, 2010).

5.6 Combining WM models and models of eye-movement control

The key focus of working-memory models as well as eye-movement control models lies on attention processes. The results of the present thesis revealed that increasing working-memory load leads to increases not only in reaction times for comprehending a sentence, but also in fixation durations on the words itself. Moreover, there was clear evidence that working-memory load influenced the pre-processing of the upcoming word and thus, dynamically influences the perceptual span as proposed by Henderson and Ferreira's (1990) foveal load hypothesis. Word properties of the upcoming word modulated fixation durations of the fixated word only in low WM load conditions. Models of eye-movement control could explain the results as following:

(a) Sequential attention shift models (e.g., E-Z Reader: Reichle et al., 2003) assume that a word can only be processed if attention is focused on that word, and attention can only be focused one word at a time. In accordance with these models, in low WM load conditions, in which the upcoming word n+1 was pre-processed, attention was shifted to word n+1 before programming of the saccade to that word was finished. Thus, while the eyes still fixate on word n, attention is already shifted ahead and word n+1 can be pre-processed in parafoveal vision. In contrast, in high WM load conditions processing of word n needs longer due to limitations in attentional resources (WM assumption), and thus, the lexical programming may need as long as programming the saccade to word n+1. As a consequence, there may be no time left to shift attention to word n+1 before moving the eyes and word n+1 will not be pre-processed. (b) Guidance-by-attentional-gradient models (e.g. SWIFT: Engbert et al., 2005; Glenmore: Reilly & Radach, 2003) postulate that attention is distributed across more than one word at a time. In a recent version, the SWIFT-model further implements a dynamical attention span, which changes its size conditional on the foveal processing demand. With respect to the present results, the attention span may extended to the upcoming word(s) in conditions with low WM load and focus on the fixated word only in conditions with high WM load.

In addition, economical conceptions of WM (Cowan, 1999; Engle, 1996, Oberauer, 2002) also postulate a close connection between WM processes and the attention system. If attention is guided to activated information in long-term memory they obtain entrance into working memory. Empirical studies confirmed that only about 4 elements can be actively maintained. In accordance to Oberauer (2002) they are maintained in a region of direct access and in Cowan's model (1999) in a focus of attention. Contrary to Cowan, Oberauer postulates that a focus of attention can only hold one element at a time. Referring to models of working memory, the conception of sequential attention shift models in reading may be similar to the perspective of a restricted focus of attention as Oberauer predicted, and the guidance-by-attentional-gradient models are more like the view of Cowan's focus of attention. However, in explaining the results of eye-movement control, the WM model, as predicted by Cowan, quickly reaches its limits. If one assumes only an activated part of long-term memory and a dynamical focus of attention, one has to assume that all words of a sentence that are already read are in the focus of attention. This representation is not compatible with the view of eye-movement control models and the intention of a perceptual span.

5.6.1 The perceptual span and the focus of attention

Figure 5.1 illustrates different assumptions of WM combined with assumptions of eye-movement control models. In the left upper panel (Figure 5.1 A1), the sequential attention shift view is illustrated with a focus of attention that holds only one element at a time. During sentence reading words are activated in long-term-memory. The region of direct access, as postulated by Oberauer, increases during sentence reading from one element at sentence beginning to 4 elements at the end of a sentence. During reading words are bound together to coherent chunks. Only one word is in the focus of attention. If WM load is high (Figure 5.1 B), and lexical access needs as long as the programming of the saccade, then the focus of attention is always on the fixated word. In contrast, if WM load is low (Figure 5.1 A1) and lexical access is faster than the programming of the saccade, the focus of attention is shifted to the upcoming word and the eyes follow with a delay. The upcoming word enters WM when attention is guided to that word for the first time. Before that, it may however be pre-activated in LTM, for example due to its predictability from the previous sentence context.

Embedding assumptions of WM in guidance-by-attentional-gradient models, one has to assume that the focus of attention is dynamically adjusted to the available attentional resources. In that way the focus of attention is restricted to the fixated word, if that word or additional information occupy all attentional resources (Figure **5.1** B) and it is extended to the upcoming word, if additional resources are available (Figure 5.1 A2). Empirical evidence for a dynamical focus of attention is so far missing in working memory research. If anything, rather the possibility of multiple foci is discussed (see Risse and Oberauer, 2010).

Nevertheless, to restrict the attentional focus only to one word at a time reduces the input of additional information and thus reduces the risk of interference. Computational models of sentence processing have already implemented assumptions of working memory (e.g. Lewis, Vasishth and Dyke, 2006). Among others, computational principles concerning the focus of attention and similarity-based interference were specified. For sentence processing, an extremely limited focus of attention is assumed that reduces the similarity-based interference during the retrieval and encoding of words.

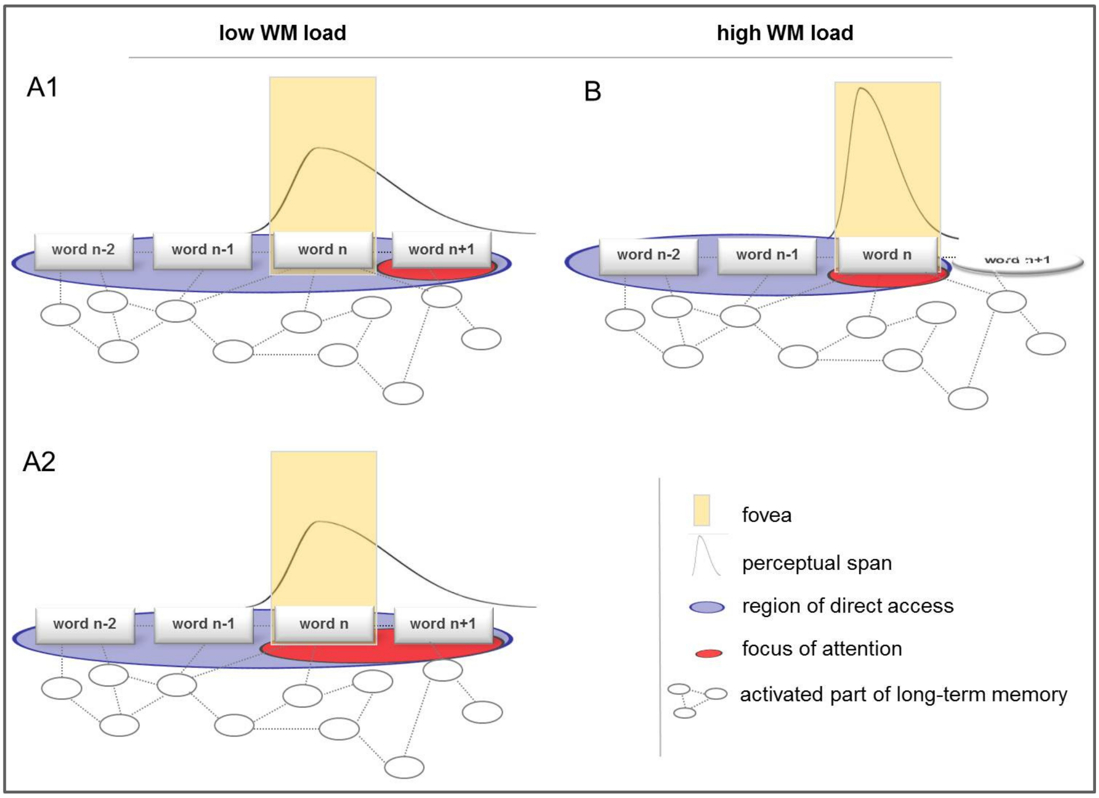

Figure 5.1: Illustration of different assumptions, both from WM models and eye-movement control models. In A1 the sequential attention shift view is illustrated with a focus of attention that holds only one word of a sentence at a time. A2 illustrates the dynamical attention shift view were under low WM load conditions the focus of attention can hold more than one word at a time. In high WM load conditions both, the attention shift view and the dynamical attention shift view, would suggest that attention is focused on one word. The already read words are in the focus of attention and upcoming words can be pre-activated by the activated part of long term memory.

However, in conditions with low working memory demands, there is no reason why attention should not be distributed across a broader region of words as sufficient, if additional attentional resources are freely available. A dynamically adapting focus of attention, therefore, seems to be an interesting starting point and a promising concept in the attempt to integrate working-memory theories and eye-movement models in reading.

5.7 Specifics of the reading span task

A general strategy observed for the reading-span task was that subjects fixated the target word as fast as possible. This lead to specific fixation patterns in the first experiment, were subjects started reading with the sentence final word, which was the target word. Forcing subjects to start reading with the sentence-initial word reduced sentence processing times. Individuals with high WMC read faster and fixated the target word earlier than readers with low WMC. However, individual variance reduced when this intuitive strategy was prohibited. Thus, some of the fixation patterns were task specific, and fixations were influenced by the relevance of the target encoding. To maintain the target words was the pivotal task in this dual-task account. Cognitive relevance plays an important role in dual-task conditions (Meyer & Kieras, 1997; Meyer, Kieras, Lauber, Schumacher, Glass, Zubrigger, 1995). It defines the priority of the task component. In competing situations the task with the highest priority is processed before the other. The results shown in Chapter 3 tentatively suggest that the cognitive relevance of the target word influenced the fixation sequence and the durations. Hence, top-down processes influenced the eye-guidance at the fixation level. The results, therefore, provide preliminary support for the cognitive relevance hypothesis (Henderson, Malcolm, & Schandl, 2009, see also Inhoff et al. 1992) that supposes a top-down processing of eye-movement control which is intention-dependent.

5.8 Outlook: Sentence independent manipulation of working memory load

The present thesis was a first attempt to integrate two fields of research on human cognition that have been largely separated in the past. Clear evidence was given that working-memory has a strong impact on eye guidance during reading. Moreover, the results inspire ideas for subsequent experiments to provide a better understanding of how executive attention processes of WM and eye-movement control are linked together.

In the reading-span task used for this thesis, the target word was the sentence final word. Thus, longer fixation times on target words were shown to reflect both, the sentence wrap-up effect and time for encoding the target word. To disentangle the processes of sentence comprehension and target encoding, one could use *unrelated items*, which are printed to the right of the sentences (e.g. Kane et al. 2004). This allows much greater control over the influence of WM load on sentence comprehension, dependent on individual capacities. Moreover, a sentence independent manipulation of working-memory load provides the opportunity to use items from different domains.

To test the domain specificity of the attentional focus, the *validated WM tests* like memory-updating tests or the operation span, have to be combined with sentence reading. Elements like single letters, words, numbers, or points within a grid could be presented and updated, while a new sentence is presented after each updating step. This procedure ensures a successive increase of working-memory load during each task. Sentence comprehension can be ensured by reading aloud, or by a decision that has to be made on basis of the sentence context.

5.9 Conclusions

The present thesis provided strong evidence for the influence of working-memory processes on eye-movements during reading. Relating to the perceptual span and its dynamical modulation through cognitive demands, the working-memory capacity was shown to be a weaker factor than working-memory load. I therefore conclude that highly automated processes of eye-movement control during reading, such as saccade generation, are performed relatively similar in all brains without strong individual differences. Given that individuals with low working-memory capacity reach there limits more quickly, effects of working-memory load become apparent earlier.

Further, I provided evidence that age-related differences in eye-movements are governed by differences in working-memory capacity. The dissociation of age-effects with respect to WMC had not yet been shown for eye-movements during reading, as the focus of typical aging studies was rather on analysing general age effects. Thus, the present results emphasize the importance of considering differences in WMC if one aims to understand age-related changes in eye-movements during reading.

References

Alvarez, G. A., & Cavanagh, P. (2004). The capacity of visual short-term memory is set both by visual information load and by number of objects. *Psychological Science, 15*, 106-111.

Atkinson, R. C., & Shiffrin, R. M. (1968). Human memory: A proposed system and its control processes. In K. W. Spence & J. T. Spence (Eds.), *The psychology of learning and motivation: Advances in research and theory* (Vol. II, pp. 89-195). New York: Academic Press.

Baayen, R. H. (2008). *Analyzing linguistic data: A practical introduction to statistics using R.* Cambridge: University Press.

Baddeley, A. (1986). *Working Memory*. Oxford: Clarendon Press.

Baddeley, A. (1992). Working memory. *Science, 255*, 556-559.

Baddeley, A. (2000). The episodic buffer: a new component of working memory? *Trends in Cognitive Sciences, 4*(11), 417-423.

Baddeley, A., & Hitch, G. J. (1974). Working memory. In G. H. Bower (Ed.), *The psychology of learning and motivation* (Vol. 8, pp. 47-90). New York: Academic Press.

Baddeley, A., Hitch, G. J., & Allen, R. J. (2009). Working memory and binding in sentence recall. *Journal of Memory and Language, 61*(3), 438-456.

Baddeley, A., Thomas, N., & Buchanan, M. (1975). Word length and the structure of short-term memory. *Journal Of Verbal Learning & Verbal Behavior, 14*, 575-589.

Balota, D., & Rayner, K. (1991). Word recognition processes in foveal and parafoveal vision: The range of influence of lexical variables. In D. Besner & G. Humphreys (Eds.), *Basic processes in reading: Visual word recognition*. Hillsdale, NJ: Erlbaum.

Balota, D. A., Pollatsek, A., & Rayner, K. (1985). The interaction of contextual and parafoveal visual information during eye fixations in reading. *Cognitive Psychology, 17*, 364-390.

Baltes, P. B., & Lindenberger, U. (1997). Emergence of a powerful connection between sensory and cognitive functions across the adult life span: A new window to the study of cognitive aging? *Psychology and Aging, 21*(1), 12-21.

Barrouillet, P., Bernardin, S., & Camos, V. (2004). Time constraints and resource sharing in adults' working memory spans. *Journal of Experimental Psychology: General, 133*(1), 83-100.

Bates, D., & Maechler, M. (2010). *lme4: Linear mixed-effect models using S4 classes. R package version 0.999375-34 [Software]* (R package version 0.995-12. ed.). Vienna, Austria: R Foundation for Statistical Computing.

Blouin, D. C., & Riopelle, A. J. (2005). On confidence intervals for within-subjects designs. *Psychological Methods, 10*(4), 397-412.

Boston, M. F., Hale, J., Kliegl, R., & Vasishth, S. (2008). Parsing costs as predictors of reading difficulty: An evaluation using the Potsdam Sentence Corpus. *Journal of Eye Movement Research, 2*(1), 1-12.

Box, G. E. P., & Cox, D. R. (1964). An analysis of transformations. *Journal of the Royal Statistical Society, 26*(2), 211-243.

Brainard, D. (1997). The psychophysics toolbox. *Spatial Vision, 10*, 433-436.

Braver, T. S., Barch, D. M., & Cohen, H. D. (1999). Cognition and control in schizophrenia: A computational model of dopamine and prefrontal function. *Biological Psychiatry, 46*, 312-328.

Burgess, N., & Hitch, G. J. (1999). Memory for serial order: A network model of the phonological loop and its timing. *Psychological Review, 106*, 551-581.

Buswell, G. (1922). Fundamental reading habits: A study of their development. *Supplementary Educational Monographs, 21*.

Calvo, M. G. (2001). Working memory and inferences: Evidence from eye fixations during reading. *Memory, 9*, 365-381.

Cantor, J., Engle, R. W., & Hamilton, G. (1991). Short-term memory, working memory, and verbal abilities: How do they relate? *Intelligence, 15*(2), 229-246.

Caplan, D., & Waters, G. S. (1999). Verbal working memory and sentence comprehension. *Behavioral and Brain Sciences, 22*, 77-126.

Carpenter, P. A., & Just, M. A. (1992). A capacity theory of comprehension: Individual differences in working memory. *Psychological Review, 99*(1), 122-149.

Carretti, B., Borella, E., Cornoldi, C., & De Beni, R. (2009). Role of working memory in explaining the performance of individuals with specific reading comprehension difficulties: A meta-analysis. *Learning and Individual Differences, 19,* 246-251.

Cerella, J. (1990). Aging and information processing rates in the elderly. In J. E. Birren & K. W. Schaie (Eds.), *Handbook of the psychology of aging* (Vol. 3, pp. 201-221). New York: Academic Press.

Conway, A. R. A., Cowan, N., & Bunting, M. F. (2001). The cocktail party phenomenon revisited: The importance of working memory capacity. *Psychonomic Bulletin & Review, 8*(331-335).

Cornelisson, F. W., Peter, E. M., & Palmer, J. (2002). The eyelink toolbox: Eye tracking with matlab and the psychophysics toolbox. *Behavior Research Methods, Instruments, and Computers, 34*(4), 613-617.

Cowan, N. (1988). Evolving conceptions of memory storage, selective attention, and their mutual constraints within the human information processing system. *Psychological Bulletin, 104*(1), 63-91.

Cowan, N. (1990). Converging evidence about information processing. *Behavioral and Brain Sciences, 13,* 237-238.

Cowan, N. (1999). An embedded-processes model of working-memory. In A. Miyake & P. Shah (Eds.), *Models of working memory* (pp. 62-101). New York: Cambridge University Press.

Cowan, N. (2001). The magical number four in short-term memory: A reconsideration of mental storage capacity. *Behavioral & Brain Sciences, 24*, 87-114.

Cowan, N. (2005). *Working memory capacity*. New York: Psychology Press Taylor & Francis.

Cowan, N. (Ed.). (1995). *Attention and memory: An integrated framework*: Oxford Psychology Series, No. 26. Oxford University Press.

Craik, F. I. M. (1983). On the transfer of information from temporary to permanent memory. *Philosophical Transactions of the Royal Society of London, Series B302*, 341-359.

Craik, F. I. M., & Byrd, M. (1982). Aging and cognitive deficits: The role of attentional resources. In F. I. M. Craik & S. Trehub (Eds.), *Aging and cognitive processes*. New York: Plenum.

Daneman, M., & Carpenter, P. A. (1980). Individual differences in working memory and reading. *Journal Of Verbal Learning And Verbal Behavior, 19,* 450-466.

Daneman, M., & Carpenter, P. A. (1983). Individual differences in integrating information between and within sentences. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 9(4),* 561-584.

Dreher, J.-C., & Burnod, Y. (2002). An integrative theory of the phasic and tonic modes of dopamine modulation in the prefrontal cortex. *Neural Networks, 15,* 583-602.

Drieghe, D., Brysbaert, M., & Desmet, T. (2005). Parafoveal-on-foveal effects on eye movements in text reading: Does an extra space make a difference? *Vision Research, 45*(13), 1693-1706.

Durstewitz, D., & Seamans, J. K. (2002). The computational role of dopamine D1 receptors in working memory. *Neural Networks, 15,* 561-572.

Ehrlich, S. F., & Rayner, K. (1981). Contextual effects on word perception and eye movements during reading. *Journal Of Verbal Learning & Verbal Behavior, 20*(6), 641-655.

Engbert, R., & Kliegl, R. (2003). Microsaccades uncover the orientation of covert attention. *Vision Research, 43,* 1035-1045.

Engbert, R., Nuthmann, A., Richter, E. M., & Kliegl, R. (2005). SWIFT: A dynamical model of saccade generation during reading. *Psychological Review, 112*(4), 777-813.

Engle, R. W. (1996). Working memory and retrieval: An inhibition-resource approach. In *Working memory and human cognition*. New York: Oxford Univ. Press.

Engle, R. W., Cantor, J., & Carullo, J. J. (1992). Individual differences in working memory and comprehension: A test of four hypotheses. *Journal of Experimental Psychology: Learning, Memory and Cognition, 18*(5), 972-992.

Engle, R. W., Sedek, G., von Hecker, U., & McIntosh, D. (2005). *Cognitive limitations in aging and psychopathology.* New York, NY US: Cambridge University Press.

Engle, R. W., Kane, J. M., & Tuholski, S. W. (1999). Individual differences in working memory capacity and what they tell us about controlled attention, general fluid intelligence, and functions of the prefrontal cortex. In A. Miyake & P. Shah (Eds.), *Models of working memory: Mechanisms of active maintenance and executive control* (pp. 102-134). Cambridge: Cambridge University Press.

Ericson, K. A., & Kintsch, W. (1991). Memory in comprehension and problem solving: A long-term working memory. In *(Tech. Rep. ICS No. 91-13)*. Boulder: University of Colorado.

Fisher, D. F., & Lefton, A. L. (1976). Peripheral information extraction: A developmental examination of reading processes. *Journal of Experimental Child Psychology, 21*(1), 77-93.

Fisher, D. F., & Montanary, W. E. (1977). Spatial and contextual factors in beginning reading: Evidence for PSG-CSG complements to developing automaticity? *Memory & Cognition, 5*, 247-251.

Fozard, J. L., & Gordon-Salant, S. (2001). Changes in vision and hearing with aging. In J. E. Birren & K. W. Schaie (Eds.), *Handbook of the psychology of aging* (Vol. 5, pp. 241-266). San Diego: Academic Press.

Friedman, N. P., & Miyake, A. (2004). The reading span test and its predictive power for reading comprehension ability. *Journal of Memory and Language, 51*(6), 634-646.

Garavan, H. (1998). Serial attention within working memory. *Memory Cognition, 26*, 263-276.

Gunter, T. C., Wagner, S., & Friederici, A. D. (2003). Working memory and lexical ambiguity resolution as revealed by ERPs: A difficult case for activation theories. *Journal of Cognitive Neuroscience, 15*(5), 643-657.

Hallett, P. E. (1978). Primary and secondary saccades to goals defined by instructions. *Vision Research, 18*(1279-1296).

Hartman, M., & Hasher, L. (1991). Aging and suppression: Memory for previously relevant information. *Psychology and Aging, 6*(4), 587-594.

Hasher, L., Stoltzfus, E. R., Zacks, R. T., & Rypma, B. (1991). Age and inhibition. *Journal of Experimental Psychology: Learning, Memory and Cognition, 17*(1), 163-169.

Hasher, L., & Zacks, R. T. (1979). Automatic and effortful processes in memory. *Journal of Experimental Psychology: General, 108*(3), 356-388.

Hasher, L., & Zacks, R. T. (1988). Working memory, comprehension, and aging: A review and a new view. In G. H. Bower (Ed.), *The psychology of learning and motivation* (Vol. 22, pp. 193-225). New York: Academic Press.

Henderson, J. M., & Ferreira, F. (1990). Effect of foveal processing difficulty on the perceptual span in reading: Implications for attention and eye movement control. *Journal of experimental Psychogology: Language, Memory, and Cognition, 16*(3), 417-429.

Henderson, J.M., & Ferreira, F. (1993). Eye movement control during reading: Fixation measures reflect foveal but not parafoveal processing difficulty. *Canadian Journal of Experimental Psychology, 47(2),* 201-221.

Henderson, J. M., Malcolm, G. L., & Schandl, C. (2009). Searching in the dark: Cognitive relevance drives attention in real-world scenes. *Psychologic Bulletin & Review, 16*(5), 850-856.

Henson, R. N. A., Norris, D. G., Page, M. P. A., & Baddeley, A. D. (1996). Unchained memory: Error patterns rule out chaining models of immediate serial recall. *Quarterly Journal of Experimental Psychology, 49A*(80-115).

Hochberg, J. (1970). Components of literacy: Speculations and exploratory research. In H. Levin & J. P. Williams (Eds.), *Basic studies in reading*. New York.

Inhoff, A. W. (1989). Parafoveal processing of words and saccade computation during eye fixations in reading. *Journal of Experimental Psychology: Human Perception and Performance, 15*(3), 544-555.

Inhoff, A. W., Eiter, B. M., & Radach, R. (2005). Time course of linguistic information extraction from consecutive words during eye fixation in reading. *Journal of Experimental Psychology: Human Perception and Performance, 31*, 979-995.

Inhoff, A.W., Topolski, R., & Wang, J. (1992). Saccade programming during short duration fixations: An examination of copytyping, letter detection, and reading. *Acta Psychologica, 81(1),* 1-21.

Inhoff, A. W., Pollatsek, A., Posner, M.I., & Rayner, K. (1989). Covert attention and eye movements during reading. *The Quarterly Journal of Experimental Psychology, Section A, 41(1),* 63-89.

Inhoff, A. W., & Rayner, K. (1986). Parafoveal word processing during eye fixations in reading: Effects of word frequency. *Perception & Psychophysics, 40*(6), 431-439.

Just, M. A., & Carpenter, P. A. (1980). A theory of reading: From eye fixations to comprehension. *Psychological Review, 99*(1), 122-149.

Just, M. A., & Carpenter, P. A. (1992). A Capacity theory of comprehension: individual differences in working memory. *Psychological Review, 99*(1), 122-149.

Kaakinen, J. K., & Hyönä, J. (2007). Strategy use in the reading span test: An analysis of eye movements and reported encoding strategies. *Memory, 15*(6), 634-646.

Kane, J. M., Bleckley, M. K., Conway, A. R. A., & Engle, R. W. (2001). A controlled-attention view of working-memory capacity. *Journal of Experimental Psychology: General, 130,* 169-183.

Kane, J. M., Hambrick, D. Z., Tuholski, S. W., Wilhelm, O., Payne, T. W., & Engle, R. W. (2004). The generality of working-memory capacity: A latent-variable approach to verbal and visuo-spatial memory span and reasoning. *Journal of Experimental Psychology: General, 133*, 189-217.

Kemper, S., Crow, A., & Kemtes, K. (2004). Eye-fixation patterns of high- and low-span young and older adults: Down the garden path and back again. *Psychology and Aging, 19*(1), 157-170.

Kemper, S., & Liu, C.-J. (2007). Eye movements of young and older adults during reading. *Psychology and Aging, 22*(1), 84-93.

Kemper, S., & McDowd, J. (2006). Eye movements of young and older adults while reading with distraction. *Psychology and Aging, 21*(1), 32-39.

Kemtes, K., & Kemper, S. (1997). Younger and older adults' on-line processing of syntactically ambiguous sentences. *Psychology and Aging, 12*(2), 362-371.

Kennedy, A. (2000). Parafoveal processing in word recognition. *The Quarterly Journal of Experimental Psychology, 53A*(2), 429-455.

Kennedy, A., & Pynte, J. (2005). Parafoveal-on-foveal effects in normal reading. *Vision Research, 45*, 153-168.

Kennedy, A., Pynte, J., & Ducrot, S. (2002). Parafoveal-on-foveal interactions in word recognition. *The Quarterly Journal of Experimental Psychology, 55A*(4), 1307-1337.

Kennison, S. M., & Clifton, C. J. (1995). Determinants of parafoveal preview benefit in high and low working memory capacity readers: Implications for eye movement control. *Journal of Experimental Psychology: Learning, Memory and Cognition, 21*(1), 68-81.

Kieras, D. E., Meyer, D. E., Mueller, S., & Seymour, T. (1999). Insights into working memory from the perspective of the EPIC architecture for modelling skilled perceptual-motor and cognitive human performance. In A. Miyake & P. Shah (Eds.), *Models of working memory: Mechanisms of active maintenance and executive control* (pp. 183-223). New York: Cambridge University Press.

Kliegl, R., Grabner, E., Rolfs, M., & Engbert, R. (2004). Length, frequency and predictability effects of words on eye movements in reading. *European Journal of Cognitive Psychology, 16*(1/2), 262-284.

Kliegl, R., Masson, M. E. J., & Richter, E. M. (2010). A linear mixed model analysis of masked repetition priming. *Visual Cognition, 18*(5), 655-681.

Kliegl, R., Nuthmann, A., & Engbert, R. (2006). Tracking the mind during reading: The influence of past, present and future words on fixation durations. *Journal of Experimental Psychology, 135*(1), 12-35.

Kuperman, V., Dambacher, M., Nuthmann, A., & Kliegl, R. (2010). The effect of word position on eye-movements in sentence and paragraph reading. *The Quarterly Journal of Experimental Psychology*.

Larson, G. E., & Perry, Z. A. (1999). Visual capture and human error. *Applied Cognitive Psychology, 13*(227-236).

Laubrock, J., Engbert, R., & Kliegl, R. (2004). Microsaccade dynamics during covert attention. *Vision Research, 45*, 721-730.

Laubrock, J., Kliegl, R., & Engbert, R. (2006). SWIFT explorations of age differences in eye movements during reading. *Neuroscience and Biobehavioral Reviews, 30*(6), 872-884.

Lehrl, S. (1977). *Mehrfach-Wortschatz-Test Form B [multiple choice vocabulary test]*. Erlangen, Germany: Perimed.

Lewandowsky, S., Duncan, M., & Brown, G. D. A. (2004). Time does not cause forgetting in short-term serial recall. *Psychonomic Bulletin & Review, 11*, 771-790.

Lewandowsky, S., Geiger, S. M., Morrell, D. B., & Oberauer, K. (2010). Turning simple span into complex span: Time for decay or interference from distractors? *Journal of Experimental Psychology: Learning, Memory and Cognition, 36*(4), 958-978.

Lewis, R. L., Vasishth, S., & Van Dyke, J. A. (2006). Computational principles of working memory in sentence comprehension. *TRENDS in Cognitive Science, 10(10)*, 447-454

Lima, S. D., Hale, S., & Myerson, J. (1991). How general is general slowing? Evidence from the lexical domain. *Psychology and Aging, 6*, 416-425.

Lindenberger, U., & Baltes, P. B. (1994). Sensory functioning and intelligence in old age: A strong connection. *Psychology and Aging, 9*, 339-355.

Logie, R.H. (1995). *Visuo-Spatial Working Memory.* Hove, UK: Lawrence Erlbaum Associates.

Loftus, G. R., & Masson, M. E. J. (1994). Using confidence intervals in within-subject designs. *Psychonomic Bulletin & Review, 1*(4), 476-490.

Matin, E. (1974). Saccadic suppression: A review. *Psychological Bulletin, 81*, 899-917.

Mayr, U., & Kliegl, R. (2000). Complex semantic processing in old age: Does it stay or does it go? *Psychology and Aging, 15*(1), 29-43.

McConkie, G. W., Kerr, P. W., Reddix, M. D., & Zola, D. (1988). Eye movement control during reading: The location of initial eye fixations on words. *Vision Research*, *28*(10), 1107-1118.

McConkie, G. W., & Rayner, K. (1975). The span of the effective stimulus during a fixation in reading. *Perception & Psychophysics, 17*, 578-586.

McNamara, D. S., & Scott, J. L. (2001). Working memory capacity and strategy use. *Memory and Cognition*, *29*(1), 10-17.

Meyer, D.E., & Kieras, D.E. (1997). A computational theory of executive cognitive processes and multiple-task performance: Part 1. Basic mechanisms. *Psychological Review, 104(1), 3-65.*

Meyer, D.E., Kieras, D.E., Lauber, E.J., Schumacher, E.H., Glass, J.M., Zurbriggen, E., et al. (1995). Adaptive executive control: Flexible multi-task performance without pervasive immutable response-selection bottlenecks. *Acta Psychologica, 90*, 163-190.

Mitchell, D. C., & Green, D. W. (1978). The effects of context and content on immediate processing in reading. *Quaterly Journal of Experimental Psychology*, *30*, 609-636.

Miyake, A., Just, M. A., & Carpenter, P. A. (1994). Working memory constraints on the resolution of lexical ambiguity: Maintaining multiple interpretations in neutral contexts. *Journal of Memory and Language, 33*, 175-202.

Moray, N. (1959). Attention in dichotic listening: Affective cues and the influence of instructions. *Quarterly Journal of Experimental Psychology, 11*, 56-60.

Morrison, C. M., & Ellis, A. W. (1995). Roles of word frequency and age of acquisition in word naming and lexical decision. *Journal of Experimental Psychology: Learning, Memory and Cognition*, *21*(1), 116-133.

Morrison, R. E. (1984). Manipulation of stimulus onset delay in reading: Evidence for parallel programming of saccades. *Journal of Experimental Psychology: Human Perception and Performance, 10*(5), 667-682.

Nairne, J. S. (1990). A feature model of immediate memory. *Memory and Cognition, 18*(251-269).

Neath, L. (2000). Modelling the effects of irrelevant speech on memory. *Psychonomic Bulletin & Review, 7*, 403-423.

Oberauer, K. (2002). Access to information in working memory: Exploring the focus of attention. *Journal of Experimental Psychology: Learning, Memory and Cognition, 28*(3), 411-421.

Oberauer, K. (2003). Selective attention to elements in working memory. *Experimental Psychology, 50*(4), 257-269.

Oberauer, K. (2010). Declarative and procedural working memory: Common principles, common capacity limits? *Psychologica Belgica, 2010*(50-3&4), 277-308.

Oberauer, K., & Kliegl, R. (2001). Beyond resources: Formal models of complexity effects and age differences in working memory. *European Journal of Cognitive Psychology, 13*((1/2)), 187-215.

Oberauer, K., & Kliegl, R. (2006). A formal model of capacity limits in working memory.

Oberauer, K., & Lewandowsky, S. (2008). Forgetting in immediate serial recall: Decay, temporal distinctiveness, or interference? *Psychological Review, 115*, 544-576.

Oberauer, K., Süß, H.-M., Schulze, R., Wilhelm, O., & Wittman, W. W. (2000). Working memory capacity - facets of a cognitive ability construct. *Personality and Individual Differences, 29*, 1017-1045.

Oberauer, K., Süß, H.-M., Wilhelm, O., & Wittman, W. W. (2003). The multiple faces of working memory: Storage, processing, supervision, and coordination. *Intelligence, 31*, 167-193.

O'Regan, J. K. (1979). Saccade size control in reading: Evidence for the linguistic control hypothesis. *Perception and Psychophysics, 25*(6), 501-509.

O'Regan, J. K. (1990). Eye movements and reading. *Review of Oculomotor Research, 4*, 395-453.

O'Regan, J. K. (1992). Optimal viewing position in words and the strategy-tactics theory of eye movements in reading. In K. Rayner (Ed.), *Eye movements and visual cognition: Scene perception and reading* (pp. 333-354). New York: Springer.

O'Regan, J. K., & Lévy-Schoen, A. (1987). Eye movement strategy and tactics in word recognition and reading. In M. Coltheart (Ed.), *Attention and performance: Vol.12. The psychology of reading* (pp. 363-383). Hillsdale: NJ: Erlbaum.

Page, M., & Norris, D. (1998). The primacy model. A new model of immediate serial recall. *Psychological Review, 105*, 761-781.

Park, D. C., Lautenschlager, G., Hedden, T., Davidson, N. S., Smith, A. D., & Smith, P. S. (2002). Models of visuospatial and verbal memory across the adult life span. *Psychology and Aging, 17*(299-320).

Park, D. C., & Schwarz, N. (2000). *Cognitive Aging: A Primer*. Philadelphia, PA: Psychol. Press.

Pelli, D. (1997). The videotoolbox software for visual psychophysics: Transforming numbers into movies. *Spatial Vision, 10*, 437-442.

Pinheiro, J., & Bates, D. (2000). *Mixed-effects models in S and S-Plus*. New York: Springer.

Pollatsek, A., Bolozky, S., Well, A. D., & Rayner, K. (1981). Asymmetries in the perceptual span for Israeli readers. *Brain and Language, 14*, 174-180.

Pynte, J., & Kennedy, A. (2006). An influence over eye movements in reading exerted from beyond the level of the word: Evidence form reading English and French. *Vision Research, 46*(22), 3686-3801.

Raney, G. E., & Rayner, K. (1995). Word frequency effects and eye movements during two readings of a text. *Canadian Journal of Experimental Psychology, 49*, 151-172.

Rayner, K. (1975). The perceptual span and peripheral cues in reading. *Cognitive Psychology, 7*(1), 65-81.

Rayner, K. (1986). Eye movements and the perceptual span in beginning and skilled readers. *Journal of Experimental Child Psychology, 41*, 211-236.

Rayner, K. (1998). Eye movements in reading and information processing: 20 years of research. *Psychological Bulletin, 124*(3), 372-422.

Rayner, K., Castelhano, M. S., & Yang, J. (2009). Eye movements and the perceptual span in older and younger readers. *Psychology and Aging, 24*(3), 755-760.

Rayner, K., & Duffy, S. A. (1986). Lexical complexity and fixation times in reading: Effects of word frequency, verb complexity, and lexical ambiguity. *Memory and Cognition, 14*, 191-201.

Rayner, K., Kambe, G., & Duffy, S. A. (2000). The effect of clause wrap-up on eye movements during reading. *The Quarterly Journal of Experimental Psychology, 53A*(4), 1061-1080.

Rayner, K., & Pollatsek, A. (1987). Eye movements in reading: A tutorial review. In M. Coltheart (Ed.), *Attention and performance XII* (pp. 327-362). Hillsdale, NJ: Erlbaum.

Rayner, K., & Pollatsek, A. (1989). *The psychology of reading*. Englewood Cliffs, NJ: Prentice-Hall.

Rayner, K., Reichle, E. D., Stroud, M. J., Williams, C. C., & Pollatsek, A. (2006). The effect of word frequency, word predictability, and font difficulty on the eye movements of young and older readers. *Psychology and Aging, 21*(3), 448-465.

Redick, T. S., Heitz, R. P., & Engle, R. W. (2007). Working memory capacity and inhibition: Cognitive and social consequences. In D. S. Gorfein & C. M. MacLeod (Eds.), *Inhibition in Cognition* (pp. 125-142): American Psychological Association.

Reichle, E. D., Rayner, K., & Pollatsek, A. (2003). The E-Z Reader model of eye-movement control in reading: Comparisons to other models. *Behavioral and Brain Sciences, 26*, 445-526.

Reilly, R., & Radach, R. (2003). Glenmore: An interactive activation model of eye movement control in reading. In J. Hyönä, R. Radach & H. Deubel (Eds.), *The mind's eye: Cognitive and applied aspects of eye movement research* (pp. 429-455). Amsterdam: Elsevier Science.

Richter, E. M., Engbert, R., & Kliegl, R. (2006). Current advances in SWIFT. *Cognitive Systems Research, 7*, 23-33.

Risse, S., & Kliegl, R. (2011). Adult age differences in the perceptual span during reading. *Psychology and Aging, 26*(2), 451-460.

Risse, S., & Oberauer, K. (2010). Selection of objects and tasks in working memory. *The Quarterly Journal of Experimental Psychology, 63*(4), 784-804.

Salthouse, T. A. (1996). The processing-speed theory of adult age differences in cognition. *Psychological Review, 103*(3), 403-428.

Sarkar, D. (2008). *Lattice: Multivariate Data Visualization with R*. New York: Springer.

Schroyens, W., Vitu, F., Brysbaert, M., & d'Idewalle, G. (1999). Eye movement control during reading: Foveal load and parafoveal processing. *The Quarterly Journal of Experimental Psychology, 52A*(4), 1021-1046.

Spragins, A. B., Lefton, L. A., & Fisher, D. F. (1976). Eye movements while reading and searching spatially transformed text: A developmental examination. *Memory & Cognition, 4*, 36-42.

Stine-Morrow, E. A. L., Noh, S. R., & Shake, M. C. (2010). Age differences in the effects of conceptual integration training on resource allocation in sentence processing. *Quarterly Journal of Experimental Psychology, 63*(7), 1430-1455.

Taylor, S. E. (1965). Eye movements in reading: Facts and fallacies. *American Educational Research Journal, 2*, 187-202.

Toms, M., Morris, N., & Ward, D. (1993). Working memory and conditional reasoning. *The Quarterly Journal of Experimental Psychology Section A, 46*(4), 679-699.

Unsworth, N., Heitz, R. P., & Engle, R. W. (2005). Working memory capacity in hot and cold cognition. In R. W. Engle, G. Sedek, U. von Hecker & D. N. McIntosh (Eds.), *Cognitive limitations in aging and psychopathology* (pp. 19-43). Cambridge: Cambridge University Press.

Unsworth, N., & Spillers, G. J. (2010). Working memory capacity: Attention control, secondary memory, or both? A direct test of the dual-component model. *Journal of Memory and Language, 62*, 392-406.

Verhaeghen, P., Cerella, J., Semenec, S. C., Leo, M. A., & Bopp, K. (2002). Cognitive efficiency modes in old age: Performance on sequential and coordinative verbal and visuospatial tasks. *Psychology and Aging, 17*(4), 558-570.

Walczyk, J. J. (1993). Are general resource notions still viable in reading research? *Journal of Educational Psychology, 85*, 127-135.

Walczyk, J. J. (1995). A test of the compensatory-encoding model. *Reading Research Quarterly, 30*, 396-408.

Waters, G. S., & Caplan, D. (1996). The capacity theory of sentence comprehension: Critique of Just and Carpenter. *Psychological Review, 4*, 761-772.

Wechsler, D. (1964). *Der Hamburger Wechsler Intelligenztest für Erwachsene (HAWIE)*. Bern, Switzerland: Huber.

Wickham, H. (2007). Reshaping data with the reshape package. *Journal of Statistical Software, 21*(12), 1-19.

Appendix A

Sentence Material

Sentences of the Reading Span Task from Oberauer et al. (2000) served as stimuli in Chapter 2, 3 and 4. The target words are the last words in the sentences, they are printed in bold. In addition the correctness of sentence statements are shown in brackets.

Die Reichen haben das meiste **Geld**	[true]
Alle Hemden sind aus **Leder**	[false]
Jeder Vogel war ein **Ei**	[true]
In jeder Wolke ist eine **Spinne**	[false]
Alle Hunde fahren gerne **Roller**	[false]
Viele Kinder gehen zur **Schule**	[true]
Manche Häuser sind aus **Holz**	[true]
Alle Menschen haben einen **Vater**	[true]
In der Post kauft man **Schnitzel**	[false]
Der Himmel hat eine **Ecke**	[false]
Die Erde ist größer als die **Sonne**	[false]

Ein Auto verbraucht viel **Kohle**	[false]
Alle Indianer haben einen **Hut**	[false]
Eine Apfelsine ist eine **Frucht**	[true]
Ein Adler hat eine **Flosse**	[false]
Die Banane hat eine gelbe **Schale**	[true]
Salz ist heller als **Pfeffer**	[true]
Ein Mensch hat eine **Nase**	[true]
Jedes Haus hat ein **Dach**	[true]
Licht ist ein **Metall**	[false]
Unter Wasser gibt es kein **Leben**	[false]
Die Sonne paßt in einen **Schrank**	[false]
Blei ist schwerer als **Watte**	[true]
Alle Häuser haben eine **Tür**	[true]
Ein Auto hat einen **Motor**	[true]
Mehrere Sätze ergeben ein **Wort**	[false]
Der Tag beginnt mit dem **Morgen**	[true]
In Frankreich gibt es **Käse**	[true]
Viele Flaschen sind aus **Glas**	[true]
Alle Rosen wachsen im **Kamin**	[false]
In Bonbons ist viel **Zucker**	[true]
Die Blumen blühen im **Winter**	[false]
Ein Papagei hat einen **Schnabel**	[true]
Honig enthält viel **Fett**	[false]
Eine Autobahn fährt immer ans **Meer**	[false]
Eine Pflanze braucht **Licht**	[true]
Im Frühjahr fällt der meiste **Schnee**	[false]
Diebstahl verstößt gegen das **Gesetz**	[true]
Eine Melone ist größer als ein **Apfel**	[true]
Alle Engländer leben in der **Stadt**	[false]
Im Hallenbad spielt man **Fußball**	[false]
Ein Wasserkraftwerk erzeugt **Strom**	[true]

Eine Eiche hat eine **Wurzel**	[true]
Jedes Haus hat einen **Garten**	[false]
Wolken bestehen aus **Kupfer**	[false]
Ein Auto braucht **Benzin**	[true]
Auf dem Mond wächst **Gras**	[false]
Zahlen ergeben das **Alphabet**	[false]
Haare hat man nur auf dem **Kopf**	[false]
Der Bäcker macht **Wurst**	[false]
Auf einem Konzert hört man **Musik**	[true]
Ein Fahrradreifen besteht aus **Gummi**	[true]
Im Kino gibt es eine **Leinwand**	[true]
In Bolivien wächst **Kaffee**	[true]
Bären leben unter der **Erde**	[false]
Eine Heizung spendet **Wärme**	[true]
Jede Katze hat ein **Fell**	[true]
Auf dem Mond lebt ein **Huhn**	[false]
Alle Teller sind aus **Blech**	[false]
Züge fahren auf der **Straße**	[false]
Die Erde hat einen **Mond**	[true]
Im Sommer gefriert der **See**	[false]
Jede Blume ist eine **Tulpe**	[false]
Zum Kochen braucht man ein **Netz**	[false]
Sieben Tage hat die **Woche**	[true]
Mehrere Stufen ergeben eine **Treppe**	[true]
Schuhe sin dimmer aus **Wolle**	[false]
Mit Streichhölzern macht man **Feuer**	[true]

Appendix B

Interpolating the trajectory of the eyes during blinks

Experimental eye movement studies have to cope with the problem of data losses caused by blinks. Eye blinks result in loss of measurement. In long trials, such as normally required for determination of reading span, the number of blink events increases and discarding entire trials with blinks is no longer an option. Hence, the interpolation of blink data is required. Therefore, I developed a new tool, which is characterized by ease of usability and fine-grained interpolation of the data. Different parameter settings allow an adaptation for eye movements during reading as well as during scene perception and visual search tasks. A visualization tool simplifies the setting of parameters.

A blink is a natural brief closing of the eyelids. It can happen as a normal periodic closing reflexively or voluntarily. During a blink each eye typically rotates nasal wards and downward during the closing phase of a blink (Riggs, Kelly, Manning, & Moore, 1987). These eye movements are more rapid than the lid which results in a characteristic data pattern before the eyes are completely closed (see Figure B.2). The positions of the eyes before a blink are mainly inconsistent with the eye position after the blink. To overcome the spatial gap the eyes have to jump from one position to the other. That rapid eye movement with a high velocity is called saccade. Thus, the saccade programming is assumed to be continued during blinking.

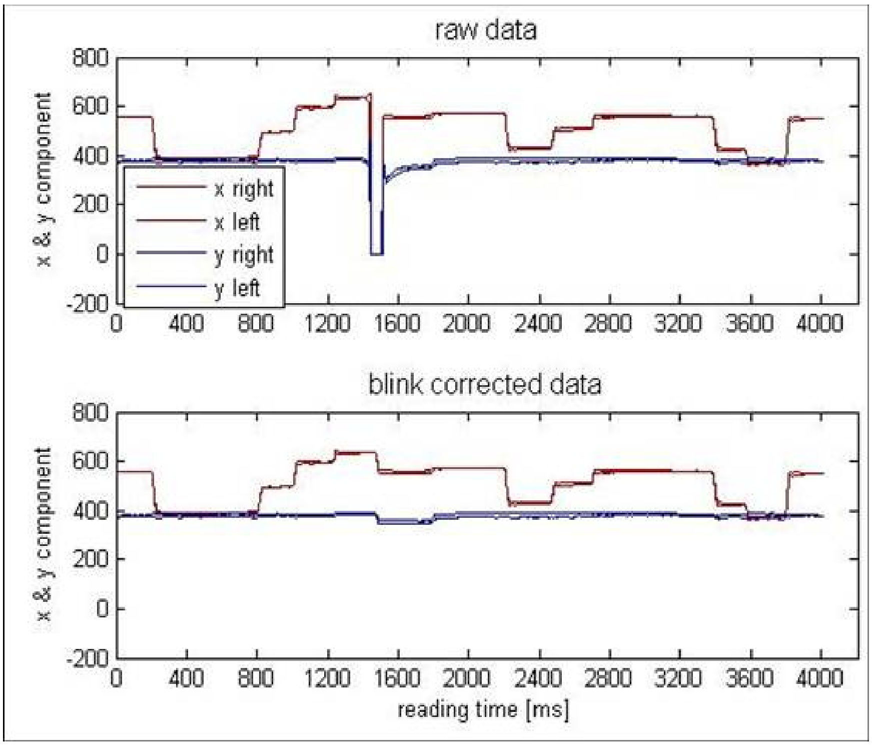

Figure B.1: Vertical and horizontal component of the left and right eye during reading a sentence of the reading span task. The upper panel displays the raw data with a blink between 1400 and 1500 ms. The bottom panel displays the blink corrected data.

B.1 Basic assumption

Figure B.1 displays the eye movements during the four seconds presentation time of one sentence of the reading span task. The upper panel of Figure B.1 shows the raw data with a blink between 1400 and 1500 ms. During the blink the value of the horizontal as well as the value of the vertical component are -1. The bottom part of the figure displays the blink corrected data where a saccade is in place of the blink. The duration and velocity of a saccade depends on the distance the eyes move during a saccade. However, in a first step constant mean saccade amplitudes of 20 ms were implemented to the algorithm. The saccade is assumed to appear in the middle part of the real blink. The real blink is defined as the part where the eyes are completely closed and the values of the horizontal and vertical component are -1. Furthermore, the part of interpolation includes the real blink and the artefacts of the blink, which are caused by the rotation of the eyes during the closing and opening phase of the eye lid. Figure B.2 displays a characteristic blink and its interpolation. In most of the cases the blink time is longer than the mean saccade time. Thus, the remaining blink time before the saccade is assigned to the preceding fixation and the remaining blink time after the saccade is assigned to the subsequent fixation. For that reason the optimal blink region is divided in three parts: The fixation part before the saccade, the saccade part, and the fixation part after the saccade. For all three parts the same linear interpolation was used. In the two fixation parts the slope of the linear function equals zero and a noise value is added to the data points of the function. The noise value is calculated by the variance of the whole sentence excluding the blink part. For the saccade part the noise value equals zero.

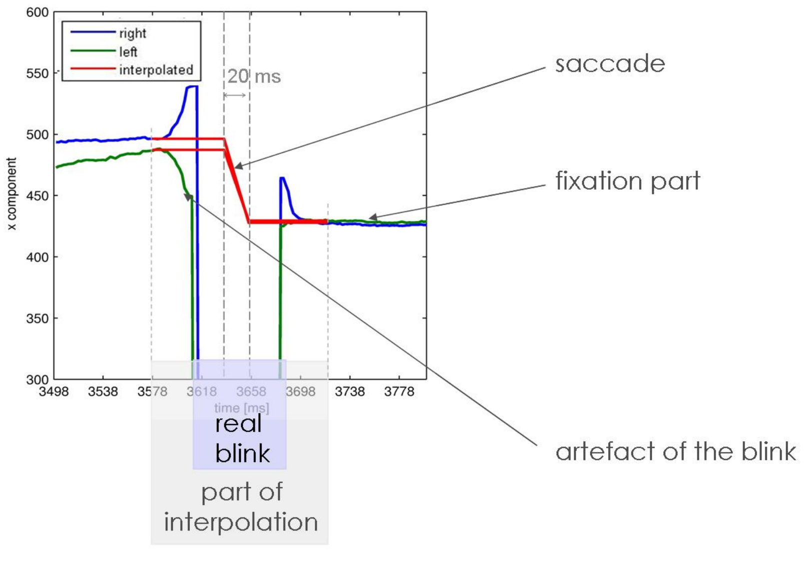

Figure B.2: Characteristic blink pattern: Horizontal component of the left and right eye before and after the blink. The red line represents the interpolated data points.

The slope value of the function depends on the direction of the eye movement and the saccade amplitude. The absolute value of the slope increases with increasing distance between the eye position before and after the blinking. A forward saccade results in a positive slope, whereas a regression results in a negative slope. If the position before and after the blinking is roughly the same, the slope equals zero.

B.2 Exclusion criteria of the method

Not all loss of data is caused by a blink. In the case of a real tracker loss usually only data from one eye are affected. Thus, in such cases the interpolation refers to the existing data of the remaining eye.

Blinks which are longer than 1000 ms or which appear to show a long closing phase which results in a divergence of the eyes as shown in Figure A.3 were not interpolated. Sentences with those blinks or more than two blinks were discarded from analysis. The criteria can be adapted with different parameter settings. It has been shown, that blinks are individually very different. Therefore, I recommend an individual adaptation to each data set.

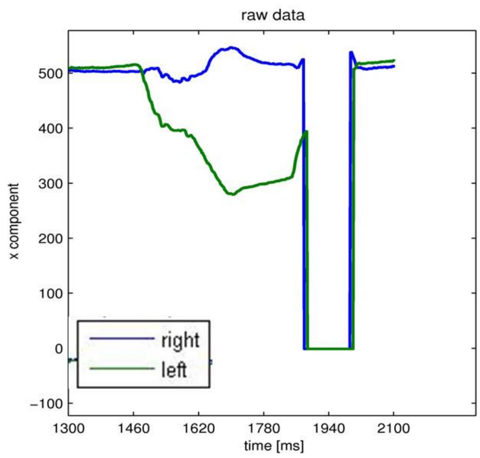

Figure B.3: Horizontal component of the left and right eye and its divergence before the eyes are completely closed within a blink.

B.3 Statistical influence

All interpolated saccades with a visual angle larger than 0.27° were detected by the saccade detection algorithm by Engbert and Kliegl (2003). For reading, saccades with a visual angle greater than 0.38° are relevant. Thus, all relevant saccades were detected. This was valid for following conditions:

```
x resolution of monitor                = 1024
y resolution of monitor                = 768
refresh rate/vertical sync frequency   = 120
font size                              = 18
font                                   = Courier
```

The interpolation of data losses had no influence on the fundamental statistical effects. Figures A.5 and A.6 display that relevant reading measures like the total viewing time and the gaze duration were not affected by the reconstruction method. However, the number of valid data points dramatically increased which moreover increased the power of the test statistics.

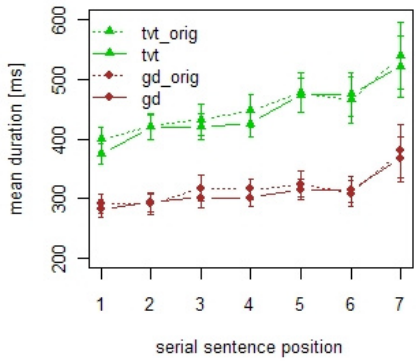

Figure B.4: Total viewing times (tvt) and gaze duration (gd) from the original (orig) data, where entire trials with blinks were discarded and the blink corrected data.

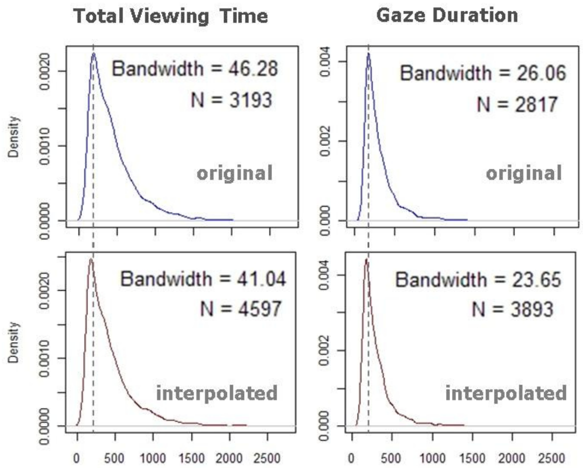

Figure B.5: Density plots of reading measures of the original data (upper panels) and the blink corrected data (bottom panels).

References

Engbert, R. & Kliegl, R. (2003). Binocular coordination in microsaccades. In J. Hyönä, R. Radach, & H. Deubel (Eds.), *The mind's eyes: Cognitive and applied aspects of eye movements.* Oxford: Elsevier, pp. 103-117.

Nuthmann, A., Engbert, R., & Kliegl, R. (in prep.). Not fixating at the line of text comes at a cost. *Unpublished manuscript.*

Riggs, L. A., Kelly, J. P., Manning, K. A., & Moore, R. K. (1987). Blink-related eye movements. *Invest Ophthalmol.Vis.Sci., 28,* 334-342.

Appendix C

Output of lmer-Analysis of the experiment in chapter 2

Table C.1 : Final LMM fitting of initial fixation on target

Linear mixed model fit by maximum likelihood			
Random effects			
Groups	Name	Variance	Std.Dev
subject id	(Intercept)	0.052974	0.23016
Residual		0.473421	0.68806
number of obs: 1043, groups: subject id, 29			
Fixed effects:			
	Estimate	Std. Error	t-value
(Intercept)	1.77754	0.05539	32.09
set size	-4.22395	0.70026	-6.03
capacity	-0.21085	1.50638	-0.14
s4.load	-1.00937	0.11711	-8.62
s5.load	-1.08177	0.09459	-11.44
s6.load	-0.76601	0.11073	-6.92
s7.load	-0.84695	0.09776	-8.66
comprehension	-0.13736	0.06702	-2.05
recall	0.04114	0.04640	0.89
recall * s6.load	-0.47745	0.22350	-2.14

Note: capacity: WMC groups; sentence order (fixed vs. random); s4: set size 4; s5: set size 5, s6: set size 6, s7: set size (e.g. s4.load: load effect at set size 4)

Table C.2: Final LMM fitting of initial fixation on target with repeated contrast specifications for set size and WM load

Linear mixed model fit by maximum likelihood			
Random effects:			
Groups	Name	Variance	Std.Dev
subject id	(Intercept)	0.053268	0.23080
Residual		0.437910	0.66175
number of obs: 1043, groups: subject id, 29			
Fixed effects:			
	Estimate	Std. Error	t-value
(Intercept)	1.74249	0.05605	31.088
set size 4 vs. all	-0.04790	0.07229	-0.663
set size 45 vs. 67	-0.05798	0.05742	-1.010
set size 456 vs. 7	-0.09745	0.05395	-1.806
load 1 vs. all	-0.42575	0.06869	-6.198
load 12 vs. 3to7	-0.60410	0.07111	-8.495
load 1to3 vs. 4to7	-0.04704	0.06889	-0.683
load 1to4 vs. 5to7	-0.14738	0.07491	-1.967
load 1to5 vs. 67	0.14265	0.08963	1.592
load 1to6 vs.7	0.10514	0.12031	0.874
comprehension	-0.11375	0.06431	-1.769

Note: capacity: WMC groups; sentence order (fixed vs. random); s4: set size 4; s5: set size 5, s6: set size 6, s7: set size (e.g. s4.load: load effect at set size 4)

Table C.3: Final LMM fitting of log total fixation time over word position

Linear mixed model fit by maximum likelihood			
Random effects:			
Groups	Name	Variance	Std.Dev
subject id	(Intercept)	0.012010	0.10959
Residual		0.358765	0.59897
number of obs: 3893, groups: subject id, 29			
Fixed effects:			
	Estimate	Std. Error	t-value
(Intercept)	5.734866	0.061062	93.92
set size	0.019062	0.009238	2.06
capacity	1.652739	1.714860	0.96
word position (wp)	0.150032	0.007627	19.67
comprehension (comp)	-0.238543	0.032773	-7.28
recall	-0.049268	0.020429	-2.41
capacity * wp	-1.821137	0.478144	-3.81
capacity * comp	-5.621542	2.198619	-2.56

Note: set size: increased from 4 to 7; capacity: WMC groups 4-7; wp: word position within a sentence; comp: accuracy of sentence comprehension; recall: recall accuracy of the target word

Table C.4: Final LMM fitting of log total fixation time over word position with nested ANOVA contrast specifications

Linear mixed model fit by maximum likelihood			
Random effects:			
Groups	Name	Variance	Std.Dev
subject id	(Intercept)	0.011382	0.10669
Residual		0.356700	0.59724
number of obs: 3893, groups: subject id, 29			
Fixed effects:			
	Estimate	Std. Error	t-value
(Intercept)	5.836053	0.032016	182.28
capacity 45 vs. 67	-0.062081	0.059400	-1.05
capacity 4 vs. 5	0.006099	0.075276	0.08
capacity 6 vs. 7	-0.074288	0.091691	-0.81
word position (wp) 543 vs. 210	0.374164	0.042106	8.89
word position 5 vs. 43	0.322136	0.151019	2.13
word position 4 vs. 3	0.072985	0.049794	1.47
word position 2 vs. 1	0.375693	0.033809	11.11
word position 1 vs. 0	0.337505	0.033301	10.14
comprehension	-0.202637	0.029938	-6.77
recall	-0.057915	0.020102	-2.88
capacity 45vs.67 * wp 543vs.210	-0.108893	0.084246	-1.29
capacity 4vs.5 * wp 543vs.210	0.082784	0.102547	0.81
capacity 6vs.7 * wp 543vs.210	-0.167232	0.133665	-1.25
capacity 45vs.67 * wp 5vs.43	0.013812	0.302074	0.05
capacity 4vs.5 * wp 5vs.43	-0.240739	0.363722	-0.66
capacity 6vs.7 * wp 5vs.43	-0.454483	0.482425	-0.94
capacity 45vs.67 * wp 4vs.3	0.118821	0.099586	1.19
capacity 4vs.5 * wp 4vs.3	0.144005	0.129461	1.11
capacity 6vs.7 * wp 4vs.3	-0.113709	0.151396	-0.75
capacity 45vs.67 * wp 2vs.1	-0.051510	0.084128	-0.61
capacity 4vs.5 * wp 2vs.1	-0.161414	0.105871	-1.52
capacity 6vs.7 * wp 2vs.1	-0.238659	0.067609	-3.53
capacity 45vs.67 * wp 1vs.0	-0.286642	0.066595	-4.30
capacity 4vs.5 * wp 1vs.0	0.143057	0.080603	1.77
capacity 6vs.7 * wp 1vs.0	-0.151675	0.106036	-1.43

Note: set size: increased from 4 to 7; capacity: WMC groups 4-7; wp: word position within a sentence; comprehension: accuracy of sentence comprehension; recall: recall accuracy of the target word

Table C.5: Final LMM fitting of log gaze duration over word position

Linear mixed model fit by maximum likelihood			
Random effects:			
Groups	Name	Variance	Std.Dev
subject id	(Intercept)	0.01824	0.13506
Residual		0.23306	0.48276
number of obs: 3893, groups: subject id, 29			
Fixed effects:			
	Estimate	Std. Error	t-value
(Intercept)	5.490261	0.028993	189.37
set size	0.019165	0.007426	2.58
capacity	0.462892	1.624149	0.29
word position (wp)	0.117724	0.006148	19.15
comprehension	-0.110245	0.024234	-4.55
recall	-0.041171	0.016480	-2.50
capacity * wp	-0.910858	0.385483	-2.36

Note: set size: increased from 4 to 7; capacity: WMC groups 4-7; wp: word position within a sentence; comprehension: accuracy of sentence comprehension; recall: recall accuracy of the target word

Table C.6: Final LMM fitting of log gaze duration with word position of sentences as a predictor and with nested ANOVA contrast specifications

Linear mixed model fit by maximum likelihood			
Random effects:			
Groups	Name	Variance	Std.Dev
subject id	(Intercept)	0.018530	0.13613
Residual		0.231008	0.48063
number of obs: 3893, groups: subject id, 29			
Fixed effects:			
	Estimate	Std. Error	t-value
(Intercept)	5.523136	0.032802	168.38
capacity 45 vs. 67	-0.027443	0.062695	-0.44
capacity 4 vs. 5	0.057535	0.081519	0.71
capacity 6 vs. 7	-0.068931	0.095130	-0.72
word position (wp) 543 vs. 210	0.269924	0.033891	7.96
word position 5 vs. 43	0.207534	0.121556	1.71
word position 4 vs. 3	-0.006605	0.040082	-0.16
word position 2 vs. 1	0.325094	0.027214	11.95
word position 1 vs. 0	0.262163	0.026813	9.78
comprehension	-0.108458	0.024140	-4.49
recall	-0.045616	0.016218	-2.81
capacity 45vs.67 * wp 543vs.210	-0.001317	0.067811	-0.02
capacity 4vs.5 * wp 543vs.210	-0.008853	0.082536	-0.11
capacity 6vs.7 * wp 543vs.210	-0.058320	0.107594	-0.54
capacity 45vs.67 * wp 5vs.43	0.058281	0.243141	0.24
capacity 4vs.5 * wp 5vs.43	-0.345129	0.292740	-1.18
capacity 6vs.7 * wp 5vs.43	0.051791	0.388330	0.13
capacity 45vs.67 * wp 4vs.3	0.089074	0.080163	1.11
capacity 4vs.5 * wp 4vs.3	0.079461	0.104198	0.76
capacity 6vs.7 * wp 4vs.3	-0.067597	0.121878	-0.55
capacity 45vs.67 * wp 2vs.1	-0.182199	0.054420	-3.35
capacity 4vs.5 * wp 2vs.1	-0.018362	0.067710	-0.27
capacity 6vs.7 * wp 2vs.1	-0.210372	0.085222	-2.47
capacity 45vs.67 * wp 1vs.0	-0.174444	0.053621	-3.25
capacity 4vs.5 * wp 1vs.0	0.070951	0.064875	1.09
capacity 6vs.7 * wp 1vs.0	-0.194774	0.085398	-2.28

Note: capacity: WMC groups 4-7; wp: word position within a sentence; comprehension: accuracy of sentence comprehension; recall: recall accuracy of the target word

Table C.7: Final LMM fitting of log second-pass fixation durations with word position within sentences as a predictor and with nested ANOVA contrast specifications

Linear mixed model fit by maximum likelihood			
Random effects:			
Groups	Name	Variance	Std.Dev
subject id	(Intercept)	0.023341	0.15278
Residual		0.129780	0.36025
number of obs: 1178, groups: subject id, 29			
Fixed effects:			
	Estimate	Std. Error	t-value
(Intercept)	5.350391	0.099426	53.81
capacity	-0.034239	0.075688	-0.45
word position (wp) 543 vs. 210	0.491715	0.187768	2.62
word position 5 vs. 43	0.592097	0.742650	0.80
word position 4 vs. 3	0.146348	0.083770	1.75
word position 2 vs. 1	0.311678	0.037541	8.30
word position 1 vs. 0	0.273974	0.037422	7.32
comprehension	-0.061146	0.031499	-1.94
recall	-0.018697	0.022080	-0.85
capacity * wp 543vs.210	0.045498	0.142190	0.32
capacity * wp 5vs.43	0.441384	0.562084	0.79
capacity * wp 4vs.3	-0.001733	0.067157	-0.03
capacity * wp 2vs.1	-0.101107	0.034923	-2.90
capacity * wp 1vs.0	-0.022626	0.033922	-0.67

Note: capacity: WMC groups 4-7; wp: word position within a sentence; comprehension: accuracy of sentence comprehension; recall: recall accuracy of the target word

Table C.8: Final LMM fitting of the number of regressions. With word-position within sentences as a predictor and with nested ANOVA contrast specifications.

Linear mixed model fit by maximum likelihood			
Random effects:			
Groups	Name	Variance	Std.Dev
subject id	(Intercept)	0.027282	0.16517
Residual		0.895008	0.94605
number of obs: 3122, groups: subject id, 29			
Fixed effects:			
	Estimate	Std. Error	t-value
(Intercept)	0.596375	0.050891	11.719
capacity	-0.063088	0.040876	-1.543
word position (wp) 543 vs. 210	0.590306	0.066144	8.925
word position 5 vs. 43	0.184246	0.230685	0.799
word position 4 vs. 3	0.211944	0.078541	2.698
word position 2 vs. 1	0.588486	0.060541	9.721
word position 1 vs. 0	0.636966	0.073923	8.617
comprehension	-0.143907	0.052593	-2.736
recall	-0.128528	0.035667	-3.603
capacity * wp 543vs.210	-0.002946	0.059145	-0.050
capacity * wp 5vs.43	0.106835	0.206268	0.518
capacity * wp 4vs.3	-0.047713	0.070341	-0.678
capacity * wp 2vs.1	-0.210841	0.053829	-3.917
capacity * wp 1vs.0	-0.341679	0.066361	-5.149

Note: capacity: WMC groups 4-7; wp: word position within a sentence; comprehension: accuracy of sentence comprehension; recall: recall accuracy of the target word

Table C.9: Final LMM fitting number of the number of regressions with word-position within sentences as a predictor.

Linear mixed model fit by maximum likelihood			
Random effects:			
Groups	Name	Variance	Std.Dev
subject id	(Intercept)	0.026066	0.16145
Residual		0.908006	0.95289
number of obs: 3122, groups: subject id, 29			
Fixed effects:			
	Estimate	Std. Error	t-value
(Intercept)	0.553936	0.042152	13.142
set size	0.003039	0.016439	0.185
capacity	-1.188303	2.204508	-0.539
word position	0.241090	0.014510	16.616
comprehension	-0.150983	0.052957	-2.851
recall	-0.131043	0.036403	-3.600
capacity * word position	-1.840576	0.918559	-2.004

Note: set size: increased from 4 to 7; capacity: WMC groups 4-7; word position: word position within a sentence; comprehension: accuracy of sentence comprehension; recall: recall accuracy of the target word

Table C.10: Final GLMM fitting of recall accuracy of target words

Generalized linear mixed model fit by the Laplace approximation

Random Effects:

Groups	Name	Variance	Std.Dev
subject id	(Intercept)	0.26990	0.51952

number of obs: 3893, groups: subject id, 29

Fixed Effects:

	Estimate	Std. Error	z-value	Pr(>|z|)	
(Intercept)	0.95189	0.14652	6.497	8.22e-11	***
set size	-22.47062	2.37370	-9.466	< 2e-16	***
capacity	8.59340	6.34133	1.355	0.175372	
s4.load	-0.07017	0.20077	-0.349	0.726731	
s5.load	-0.15346	0.19159	-0.801	0.423146	
s6.load	-1.98652	0.18667	-10.642	< 2e-16	***
s7.load	-2.01921	0.18829	-10.724	< 2e-16	***
comprehension	-0.44286	0.11516	-3.845	0.000120	
poly(s4.load)2	0.28653	0.19708	1.454	0.145991	
poly(s5.load)2	1.59326	0.18005	8.849	< 2e-16	***
poly(s6.load)2	1.42701	0.17818	8.009	1.16e-15	***
poly(s7.load)2	2.21676	0.18864	11.751	< 2e-16	***

Note: Signif. codes: 0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1
capacity: WMC groups; sentence order (fixed vs. random); s4: set size 4; s5: set size 5, s6: set size 6, s7: set size (e.g. s4.load: load effect at set size 4)

Table C.11: Final LMM fitting of the number of first pass fixations

Linear mixed model fit by maximum likelihood			
Random effects:			
Groups	Name	Variance	Std.Dev
subject id	(Intercept)	0.019544	0.13980
Residual		0.405668	0.63692
number of obs: 1043, groups: subject id, 29			
Fixed effects:			
	Estimate	Std. Error	t-value
(Intercept)	2.31941	0.04097	56.62
set size	1.36046	0.64751	2.10
capacity	-1.47284	1.04554	-1.41
s4.load	0.40208	0.10835	3.71
s5.load	0.46901	0.08742	5.37
s6.load	0.26137	0.09118	2.87
s7.load	0.30795	0.09040	3.41
comprehension	-0.14653	0.06191	-2.37
recall	-0.08132	0.04268	-1.91
set size * capacity	-51.38681	20.29133	-2.53

Note: capacity: WMC groups; sentence order (fixed vs. random); s4: set size 4; s5: set size 5, s6: set size 6, s7: set size (e.g. s4.load: load effect at set size 4)

Table C.12: Final LMM fitting of the number of first pass fixations with repeated contrast specifications

Linear mixed model fit by maximum likelihood			
Random effects:			
Groups	Name	Variance	Std.Dev
subject id	(Intercept)	0.26406	0.51387
Residual		1.60566	1.26715
number of obs: 1043, groups: subject id, 29			
Fixed effects:			
	Estimate	Std. Error	t-value
(Intercept)	5.80767	0.11824	49.12
set size 4 vs. all	-0.26461	0.13846	-1.91
set size 45 vs. 67	0.16277	0.10997	1.48
set size 456 vs. 7	-0.15091	0.10336	-1.46
load 1 vs. all	1.40227	0.13157	10.66
load 12 vs. 3to7	0.07611	0.13620	0.56
load 1to3 vs. 4to7	-0.27783	0.13194	-2.11
load 1to4 vs. 5to7	-0.04069	0.14347	-0.28
load 1to5 vs. 67	0.44979	0.17167	2.62
load 1to6 vs. 7	0.06758	0.23043	0.29
comprehension	-0.17085	0.12331	-1.39

Note: capacity: WMC groups; sentence order (fixed vs. random); s4: set size 4; s5: set size 5, s6: set size 6, s7: set size (e.g. s4.load: load effect at set size 4)

Table C.13: Final LMM fitting of the number of first pass fixations per sentence with nested ANOVA contrast specifications

Linear mixed model fit by maximum likelihood			
Random effects:			
Groups	Name	Variance	Std.Dev
subject id	(Intercept)	0.26082	0.5107
Residual		1.61176	1.2696
number of obs: 1043, groups: subject id, 29			
Fixed effects:			
	Estimate	Std. Error	t-value
(Intercept)	5.73056	0.12030	47.63
load 123 vs. 567	0.93378	0.13772	6.78
load 1 vs. 2	1.43328	0.14086	10.18
load 2 vs. 3	-1.60490	0.25059	-6.40
load 3 vs. 4	-1.56989	0.31775	-4.94
load 4 vs. 5	-0.37775	0.32356	-1.17
load 6 vs. 7	-0.21478	0.32313	-0.66
comprehension	-0.13391	0.12505	-1.07
recall	-0.02485	0.08650	-0.29

Note: capacity: WMC groups; sentence order (fixed vs. random); s4: set size 4; s5: set size 5, s6: set size 6, s7: set size (e.g. s4.load: load effect at set size 4)

Table C.14: Final LMM fitting of log sentence processing time

Linear mixed model fit by maximum likelihood			
Random effects:			
Groups	Name	Variance	Std.Dev
subject id	(Intercept)	0.012846	0.11334
Residual		0.052477	0.22908
number of obs: 1043, groups: subject id, 29			
Fixed effects:			
	Estimate	Std. Error	t-value
(Intercept)	7.803895	0.022970	339.8
set size	0.260959	0.245486	1.1
capacity	-3.727634	1.366096	-2.7
s4.load	0.100636	0.033128	3.0
s5.load	0.249128	0.018085	13.8
s6.load	0.040429	0.038089	1.1
s7.load	0.131964	0.018984	7.0
comprehension (comp)	-0.057842	0.020010	-2.9
recall	-0.007179	0.008368	-0.9
capacity * s4.load	-3.297073	1.250762	-2.6
capacity * s6.load	-3.766166	1.103607	-3.4
capacity * comp	-0.050999	0.016717	-3.1
s4.load * comp	-0.277916	0.067434	-4.1
s6.load * comp	0.284173	0.076496	3.7
s7.load * recall	-0.105965	0.038932	-2.7

Note: capacity: WMC groups; sentence order (fixed vs. random); s4: set size 4; s5: set size 5, s6: set size 6, s7: set size (e.g. s4.load: load effect at set size 4)

Table C.15: Final LMM fitting of log total fixation time

Linear mixed model fit by maximum likelihood			
Random effects:			
Groups	Name	Variance	Std.Dev
subject id	(Interce	0.014013	0.11838
Residual	pt)	0.389816	0.62435
number of obs: 3893, groups: subject id, 29			
Fixed effects:			
	Estimate	Std. Error	t-value
(Intercept)	5.96737	0.02746	217.28
set size	1.70321	0.63901	2.67
capacity	-2.70788	1.47971	-1.83
s4.load	0.14704	0.05117	2.87
s5.load	0.28859	0.04728	6.10
s6.load	0.07397	0.04642	1.59
s7.load	0.20416	0.04716	4.33
comprehension	-0.17448	0.03152	-5.54
recall	-0.02736	0.02198	-1.24

Note: capacity: WMC groups; sentence order (fixed vs. random); s4: set size 4; s5: set size 5, s6: set size 6, s7: set size (e.g. s4.load: load effect at set size 4)

Table C.16: Final LMM fitting log gaze duration

Linear mixed model fit by maximum likelihood			
Random effects:			
Groups	Name	Variance	Std.Dev
subject id	(Intercept)	0.018422	0.13573
Residual		0.252224	0.50222
number of obs: 3893, groups: subject id, 29			
Fixed effects:			
	Estimate	Std. Error	t-value
(Intercept)	5.60555	0.02863	195.76
set size	1.51103	0.51420	2.94
capacity	-0.64105	1.59853	-0.40
s4.load	0.10787	0.04117	2.62
s5.load	0.20489	0.03805	5.39
s6.load	0.01514	0.03735	0.41
s7.load	0.15526	0.03795	4.09
comprehension	-0.08955	0.02539	-3.53
recall	-0.02595	0.01772	-1.46

Note: capacity: WMC groups; sentence order (fixed vs. random); s4: set size 4; s5: set size 5, s6: set size 6, s7: set size (e.g. s4.load: load effect at set size 4)

Table C.17: Final GLMM fitting of the probability of regressions

Generalized linear mixed model fit by the Laplace approximation

Random effects:

Groups	Name	Variance	Std.Dev
subject id	(Intercept)	0.061131	0.24725

number of obs: 3567, groups: subject id, 29

Fixed effects:

	Estimate	Std. Error	z-value	Pr(>|z|)	
(Intercept)	-0.15039	0.07134	-2.108	0.035032	*
set size	4.76495	2.19658	2.169	0.030063	*
capacity	-2.42275	3.58311	-0.676	0.498940	
s4.load	1.02270	0.18380	5.564	2.63e-08	***
s5.load	0.88432	0.16795	5.265	1.40e-07	***
s6.load	0.74119	0.16047	4.619	3.86e-06	***
s7.load	0.55657	0.16055	3.467	0.000527	***
comprehension	0.01178	0.10645	0.111	0.911878	
recall	-0.10549	0.07451	-1.416	0.156812	

Note: Signif. codes: 0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1
capacity: WMC groups; sentence order (fixed vs. random); s4: set size 4; s5: set size 5, s6: set size 6, s7: set size (e.g. s4.load: load effect at set size 4)

Table C.18: Final GLMM fitting of the probability of regressions with nested ANOVA contrast specifications

Generalized linear mixed model fit by the Laplace approximation

Random effects:

Groups	Name	Variance	Std.Dev
subject id	(Intercept)	0.069537	0.2637

number of obs: 3567, groups: subject id, 29

Fixed effects:

	Estimate	Std. Error	z-value	Pr(>|z|)	
(Intercept)	-0.12165	0.06642	-1.832	0.0670	.
set size 45 vs. 67	0.12657	0.07719	1.640	0.1011	
set size 4 vs. 5	0.16485	0.11832	1.393	0.1635	
set size 6 vs. 7	-0.20702	0.09057	-2.286	0.0223	*
load 1234 vs. 67	0.97197	0.20715	4.692	2.70e-06	***
load 1 vs. 2	1.88809	0.16997	11.109	< 2e-16	***
load 2 vs. 3	1.38505	0.17167	8.068	7.15e-16	***
load 3 vs. 4	0.77256	0.14431	5.354	8.63e-08	***
load 5 vs. 6	-0.94976	0.22809	-4.164	3.13e-05	***
load 6 vs. 7	-0.35105	0.22975	-1.528	0.1265	

Note: Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Note: capacity: WMC groups; sentence order (fixed vs. random); s4: set size 4; s5: set size 5, s6: set size 6, s7: set size (e.g. s4.load: load effect at set size 4)

Table C.19: Final LMM fitting of log sentence processing time

Linear mixed model fit by maximum likelihood			
Random effects:			
Groups	Name	Variance	Std.Dev
subject id	(Intercept)	0.012813	0.11319
Residual		0.054526	0.23351
number of obs: 3670, groups: subject id, 29			
Fixed effects:			
	Estimate	Std. Error	t-value
(Intercept)	7.845616	0.023222	337.9
set size	-0.660950	0.253335	-2.6
capacity	-4.401721	1.329971	-3.3
resource	-0.034023	0.002404	-14.2
comprehension	-0.026686	0.015323	-1.7
recall	-0.006430	0.008490	-0.8
capacity * resource	0.642980	0.139842	4.6

Note: set size: increased from 4 to 7 capacity: WMC groups 4-7; resource: number of free memory slots; comprehension: accuracy of sentence comprehension; recall: recall accuracy of the target word

Table C.20: Final LMM fitting of log sentence processing times with nested ANOVA contrast specifications

Linear mixed model fit by maximum likelihood			
Random effects:			
Groups	Name	Variance	Std.Dev
subject id	(Intercept)	0.013509	0.11623
Residual		0.053651	0.23163
number of obs: 3670, groups: subject id, 29			
Fixed effects:			
	Estimate	Std. Error	t-value
(Intercept)	7.830e+00	2.475e-02	316.4
set size 45 vs. 67	1.126e-03	8.307e-03	0.1
set size 4 vs. 5	-8.464e-02	1.275e-02	-6.6
set size 6 vs. 7	-4.688e-02	9.944e-03	-4.7
capacity 45 vs. 67	-1.349e-01	4.744e-02	-2.8
capacity 4 vs. 5	-3.932e-02	6.141e-02	-0.6
capacity 6 vs. 7	-4.608e-02	7.226e-02	-0.6
resource	-3.073e-02	2.495e-03	-12.3
comprehension	-2.478e-02	1.522e-02	-1.6
recall	-1.204e-02	8.454e-03	-1.4
resource * capacity 45 vs. 67	1.696e-02	4.687e-03	3.6
resource * capacity 4 vs. 5	1.576e-02	5.769e-03	2.7
resource * capacity 6 vs. 7	1.135e-05	7.387e-03	0.00154

Note: set size: increased from 4 to 7 capacity: WMC groups 4-7; resource: number of free memory slots; comprehension: accuracy of sentence comprehension; recall: recall accuracy of the target word

Table C.21: Final LMM fitting of log sentence processing time with nested ANOVA contrast specifications

Linear mixed model fit by maximum likelihood			
Random effects:			
Groups	Name	Variance	Std.Dev
subject id	(Intercept)	0.014545	0.12060
Residual		0.052547	0.22923
number of obs: 1043, groups: subject id, 29			
Fixed effects:			
	Estimate	Std. Error	t-value
(Intercept)	7.831590	0.025439	307.85
capacity 46 vs. 57	0.017103	0.064825	0.26
capacity 45 vs. 67	-0.136795	0.050504	-2.71
capacity 6 vs. 7	0.023208	0.101041	0.23
load 234 vs. 567	0.049673	0.010029	4.95
load 12 vs. 3	0.342231	0.020055	17.06
load 2 vs. 3	-0.328333	0.024210	-13.56
load 3 vs. 4	-0.020804	0.015080	-1.38
load 5 vs. 6	0.018083	0.021710	0.83
load 6 vs. 7	-0.077610	0.028109	-2.76
comprehension (comp)	-0.070646	0.020655	-3.42
recall	0.001823	0.008451	0.22
capacity 46vs.57 * comp	-0.190053	0.036009	-5.28
capacity 45vs67 * comp	-0.025150	0.041284	-0.61
capacity 6vs.7 * comp	-0.013113	0.082531	-0.16

Note: capacity: WMC groups; sentence order (fixed vs. random); s4: set size 4; s5: set size 5, s6: set size 6, s7: set size (e.g. s4.load: load effect at set size 4)

Table C.22: Final GLMM fitting of comprehension accuracy of the sentences

Generalized linear mixed model fit by the Laplace approximation

Random effects:

Groups	Name	Variance	Std.Dev
subject id	(Intercept)	0.35656	0.59712

number of obs: 3893, groups: subject id, 29

Fixed effects:

	Estimate	Std. Error	z-value	Pr(>|z|)	
(Intercept)	2.667e+00	1.563e-01	17.059	< 2e-16	***
set size	-1.552e+01	4.260e+00	-3.642	0.000271	***
capacity	4.379e+01	8.367e+00	5.234	1.66e-07	***
s4.load	-9.850e-01	3.851e-01	-2.558	0.010537	*
s5.load	-8.761e-02	3.036e-01	-0.289	0.772927	
s6.load	-7.354e-01	2.566e-01	-2.866	0.004159	**
s7.load	-2.194e+00	4.142e-01	-5.296	1.18e-07	***
recall	-5.028e-01	1.181e-01	-4.257	2.08e-05	***
set size * capacity	-1.573e+03	2.728e+02	-5.768	8.03e-09	***
capacity * s4.load	-6.728e+01	2.481e+01	-2.712	0.006690	**
capacity * s5.load	8.049e+01	2.027e+01	3.972	7.14e-05	***
capacity * s6.load	4.026e+01	1.702e+01	2.366	0.018004	*
s7.load * recall	1.510e+00	5.030e-01	3.002	0.002679	**

Note: Signif. codes: 0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Note: capacity: WMC groups; sentence order (fixed vs. random); s4: set size 4; s5: set size 5, s6: set size 6, s7: set size (e.g. s4.load: load effect at set size 4)

Appendix D

Output of lmer-Analysis of the experiment in chapter 3 and 4

Table D.1: Final LMM fitting initial fixation on target

Linear mixed model fit by maximum likelihood			
Random effects:			
Groups	Name	Variance	Std.Dev
subject id	(Intercept)	0.50907	0.71349
Residual		1.79018	1.33798
number of obs: 3545, groups: subject id, 104			
Fixed effects:			
	Estimate	Std. Error	t-value
(Intercept)	5.33781	0.20680	25.812
sentence order (so)	0.11285	0.15357	0.735
age	-0.27035	0.14994	-1.803
capacity	-0.15117	0.07511	-2.013
set size	0.01563	0.02221	0.704
s4.load	-0.58482	0.12822	-4.561
s5.load	-0.14164	0.09642	-1.469
s6.load	0.13144	0.10064	1.306
s7.load	-0.19256	0.10523	-1.830
comprehension	0.21430	0.39207	0.547
recall	0.09976	0.08074	1.236
so * s6.load	-0.49866	0.20130	-2.477
age * s4.load	-0.51126	0.25644	-1.994

Note: age: age groups (young vs. old); capacity: WMC groups 4-7; s4: set size 4; s5: set size 5, s6: set size 6, s7: set size (e.g. s4.load: load effect at set size 4); comprehension: accuracy of sentence comprehension; recall: recall accuracy of the target word

Table D.2: Final LMM fitting of log total fixation time with word position within sentences as one predictor.

Linear mixed model fit by maximum likelihood			
Random effects:			
Groups	Name	Variance	Std.Dev
subject id	(Intercept)	0.012032	0.10969
Residual		0.318794	0.56462
number of obs: 15801, groups: subject id, 104			

Fixed effects:	Estimate	Std. Error	t-value
(Intercept)	5.799898	0.033235	174.51
sentence order (so)	0.033404	0.029854	1.12
age	0.095489	0.024857	3.84
capacity	-0.018712	0.012300	-1.52
set size	-0.009178	0.004374	-2.10
word position (wp)	0.083456	0.004245	19.66
comprehension	-0.057499	0.062607	-0.92
recall	-0.034046	0.013325	-2.55
so * age	0.018410	0.049249	0.37
so * capacity	0.018126	0.024803	0.73
wp * sentence order	-0.017343	0.008491	-2.04
wp * age	-0.012256	0.007180	-1.71
wp * capacity	-0.007349	0.003492	-2.10
wp * sentence order * age	-0.034272	0.014359	-2.39
wp * sentence order * capacity	-0.017778	0.006984	-2.55

Note: so: sentence order within the experiment (serial order vs. random order); age: age groups (young vs. old); capacity: WMC groups 4-7; set size: number of sentences within one trial (varied from 4 to 7); wp: word position within a sentence; comprehension: accuracy of sentence comprehension; recall: recall accuracy of the target word; s4: set size 4; s5: set size 5, s6: set size 6, s7: set size (e.g. s4.load: load effect at set size 4);

Table D.3: Final LMM fitting of log gaze duration with word position within sentences as one predictor.

Linear mixed model fit by maximum likelihood

Random effects:

Groups	Name	Variance	Std.Dev
subject id	(Intercept)	0.017626	0.13276
Residual		0.237001	0.48683

number of obs: 15801, groups: subject id, 104

Fixed effects:

	Estimate	Std. Error	t-value
(Intercept)	5.531862	0.038565	143.44
sentence order (so)	0.028069	0.034216	0.82
age	0.020892	0.027932	0.75
capacity	-0.011987	0.014076	-0.85
set size	-0.006696	0.003772	-1.78
word position (wp)	0.093647	0.003631	25.79
comprehension	-0.050642	0.072820	-0.70
recall	-0.013238	0.015213	-0.87
so * capacity	0.022090	0.028529	0.77
wp * so	0.007626	0.007263	1.05
wp * capacity	-0.003813	0.002986	-1.28
wp * so * capacity	-0.027411	0.005973	-4.59

Note: so: sentence order within the experiment (serial order vs. random order); age: age groups (young vs. old); capacity: WMC groups 4-7; set size: number of sentences within one trial (varied from 4 to 7); wp: word position within a sentence; comprehension: accuracy of sentence comprehension; recall: recall accuracy of the target word; s4: set size 4; s5: set size 5, s6: set size 6, s7: set size (e.g. s4.load: load effect at set size 4);

Table D.4: Final LMM fitting of log total fixation time with word position within sentences as one predictor. Repeated contrast specifications for the WMC groups (capacity).

Linear mixed model fit by maximum likelihood

Random effects:

Groups	Name	Variance	Std.Dev
subject id	(Intercept)	0.010865	0.10423
Residual		0.322097	0.56754

number of obs: 17641, groups: subject id, 115

Fixed effects:

	Estimate	Std. Error	t-value
(Intercept)	5.767620	0.013946	413.6
sentence order (so)	0.004118	0.027012	0.2
age	0.085049	0.022100	3.8
capacity 4 vs. 567	-0.152887	0.048001	-3.2
capacity 45 vs. 67	0.075144	0.036706	2.0
capacity 456 vs. 7	-0.044820	0.028453	-1.6
word position (wp) 543 vs. 210	0.085669	0.004077	21.0
so * capacity 4 vs. 567	0.214324	0.093346	2.3
so * capacity 45 vs. 67	0.009339	0.071162	0.1
so * capacity 456 vs. 7	-0.058815	0.054529	-1.1
wp * capacity 4 vs. 567	0.016691	0.013969	1.2
wp * capacity 45 vs. 67	-0.029442	0.010734	-2.7
wp * capacity 456 vs. 7	-0.001854	0.008418	-0.2

Note: so: sentence order within the experiment (serial order vs. random order); age: age groups (young vs. old); capacity: WMC groups 4-7; wp: word position within a sentence

Table D.5: Final LMM fitting of log total fixation time with word position within sentences as predictor. REPEATED contrast specifications for the condition of fixed sentence order

Linear mixed model fit by maximum likelihood

Random effects:

Groups	Name	Variance	Std.Dev
subject id	(Intercept)	0.014505	0.12044
Residual		0.314171	0.56051

number of obs: 8332, groups: subject id, 59

Fixed effects:

	Estimate	Std. Error	t-value
(Intercept)	5.761397	0.023295	247.32
age	0.096652	0.035097	2.75
capacity 4 vs. 567	-0.249851	0.081040	-3.08
capacity 45 vs. 67	0.073973	0.062654	1.18
capacity 456 vs. 7	-0.040693	0.046042	-0.88
word position (wp)	0.091971	0.006207	14.82
wp * capacity 4 vs. 567	0.002397	0.021557	0.11
wp * capacity 45 vs. 67	-0.029514	0.016390	-1.80
wp * capacity 456 vs. 7	0.027490	0.012315	2.23

Note: so: sentence order within the experiment (serial order vs. random order); age: age groups (young vs. old); capacity: WMC groups 4-7; wp: word position within a sentence

Table D.6: Final GLMM fitting of the probability of regression.

Generalized linear mixed model fit by the Laplace approximation
Random effects:

Groups	Name	Variance	Std.Dev
subject id	(Intercept)	0.19130	0.43738

number of obs: 13861, groups: subject id, 104

Fixed effects:

	Estimate	Std. Error	z-value	Pr(>|z|)	
(Intercept)	-1.24281	0.15508	-8.014	1.11e-15	***
sentence order	0.03662	0.09928	0.369	0.71227	
age	0.86089	0.29099	2.958	0.00309	***
capacity	-0.01414	0.04833	-0.293	0.76986	
set size	-0.01075	0.01925	-0.558	0.57660	
s4.load	0.21302	0.11236	1.896	0.05796	.
s5.load	0.17236	0.09052	1.904	0.05690	.
s6.load	0.26678	0.09025	2.956	0.00312	**
s7.load	0.25093	0.08963	2.800	0.00512	**
comprehension	0.36012	0.29631	1.215	0.22423	
recall	-0.17270	0.06018	-2.870	0.00411	**
age * comprehension	-0.95905	0.58160	-1.649	0.09915	.

Note: Signif. codes: 0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1
capacity: WMC groups; sentence order (fixed vs. random); s4: set size 4; s5: set size 5, s6: set size 6, s7: set size (e.g. s4.load: load effect at set size 4)

Table D.7: Final LMM fitting of log sentence processing times with WM load as predictor.

Linear mixed model fit by maximum likelihood			
Random effects:			
Groups	Name	Variance	Std.Dev
subject id	(Intercept)	0.015275	0.12359
Residual		0.045548	0.21342
number of obs: 15798, groups: subject id, 104			
Fixed effects:			
	Estimate	Std. Error	t-value
(Intercept)	7.650462	0.034599	221.12
sentence order (so)	0.045529	0.030538	1.49
age	0.015423	0.025279	0.61
capacity (cap)	-0.030881	0.012579	-2.45
set size	-0.018585	0.001659	-11.20
s4.load	0.019904	0.011821	1.68
s5.load	0.059860	0.009471	6.32
s6.load	0.077305	0.009290	8.32
s7.load	0.091320	0.007746	11.79
comprehension	-0.024357	0.065460	-0.37
recall	-0.019046	0.013423	-1.42
so * age	0.051524	0.050074	1.03
so * capacity	-0.030715	0.025359	-1.21
so * s4.load	-0.061127	0.020201	-3.03
so * s5.load	-0.039968	0.018723	-2.13
so * s6.load	-0.041518	0.015275	-2.72
age * s4.load	-0.096392	0.019941	-4.83
age * s5.load	-0.035394	0.015937	-2.22
age * s7.load	-0.044129	0.015517	-2.84
capacity * s4.load	0.029694	0.009647	3.08
capacity * s5.load	-0.019329	0.007655	-2.53
capacity * s6.load	-0.041082	0.007576	-5.42
so * age * s4.load	0.164411	0.039993	4.11
so * cap * s4.load	-0.048615	0.015391	-3.16
age * cap * s6.load	0.032647	0.014976	2.20

Note: capacity: WMC groups; sentence order (fixed vs. random); s4: set size 4; s5: set size 5, s6: set size 6, s7: set size (e.g. s4.load: load effect at set size 4)

Table D.8: Final LMM fitting of log total viewing times with WM load as predictor.

Linear mixed model fit by maximum likelihood			
Random effects:			
Groups	Name	Variance	Std.Dev
subject id	(Intercept)	0.011989	0.10949
Residual		0.328784	0.57340
number of obs: 15801, groups: subject id, 104			
Fixed effects:			
	Estimate	Std. Error	t-value
(Intercept)	5.868359	0.032992	177.87
sentence order (so)	0.017642	0.024409	0.72
age	0.093793	0.023880	3.93
capacity (cap)	-0.026188	0.011921	-2.20
set size	-0.012576	0.004451	-2.83
s4.load	0.119240	0.026057	4.58
s5.load	0.048135	0.020867	2.31
s6.load	0.021176	0.020696	1.02
s7.load	0.068083	0.020453	3.33
comprehension	-0.056002	0.062603	-0.89
recall	-0.034484	0.013108	-2.63
so * s4.load	-0.198958	0.052051	-3.82
so * s6.load	0.156533	0.041391	3.78

Note: capacity: WMC groups; sentence order (fixed vs. random); s4: set size 4; s5: set size 5, s6: set size 6, s7: set size (e.g. s4.load: load effect at set size 4)

Table D.9: Final LMM fitting of log sentence processing time with the number free memory slots (resources) as predictor.

Linear mixed model fit by maximum likelihood			
Random effects:			
Groups	Name	Variance	Std.Dev
subject id	(Intercept)	0.015781	0.12562
Residual		0.045758	0.21391
number of obs: 15798, groups: subject id, 104			
Fixed effects:			
	Estimate	Std. Error	t-value
(Intercept)	7.6869442	0.0351951	218.41
sentence order (so)	-0.0009880	0.0272408	-0.04
age	-0.0254059	0.0259911	-0.98
capacity (cap)	-0.0210880	0.0128811	-1.64
set size	-0.0268025	0.0017290	-15.50
resource	-0.0180659	0.0012070	-14.97
comprehension (comp)	-0.0258037	0.0665929	-0.39
recall	-0.0072578	0.0200228	-0.36
resource * so	0.0089860	0.0020466	4.39
resource * cap	0.0042382	0.0009546	4.44
resource * age	0.0119381	0.0019660	6.07
resource * recall	-0.0018314	0.0015372	-1.19
age * recall	-0.0405473	0.0388767	-1.04
resource * age * recall	0.0085313	0.0029974	2.85

Note: so: sentence order within the experiment (serial order vs. random order); age: age groups (young vs. old); cap: WMC groups 4-7; set size: number of sentences within one trial (varied from 4 to 7); resource: number of free memory slots; comprehension: accuracy of sentence comprehension; recall: recall accuracy of the target word

Table D.10: Analysis of lag and successor effects: Final LMM fitting of log gaze duration.

Linear mixed model fit by maximum likelihood			
Random effects:			
Groups	Name	Variance	Std.Dev
subject id	(Intercept)	0.018250	0.13509
word id	(Intercept)	0.030667	0.17512
Residual		0.165897	0.40730
number of obs: 12256, groups: subject id, 104; word id, 95			
Fixed effects:			
	Estimate	Std. Error	t-value
(Intercept)	5.451071	0.049784	109.50
sentence order (so)	0.066083	0.028841	2.29
age	0.005828	0.028112	0.21
capacity	-0.013307	0.014059	-0.95
set size	-0.005581	0.003910	-1.43
s4.load	0.002710	0.023383	0.12
s5.load	-0.011856	0.018866	-0.63
s6.load	0.015210	0.019068	0.80
s7.load	-0.024529	0.017703	-1.39
frequency(n)	-0.101210	0.024972	-4.05
frequency(n-1)	-0.013737	0.016472	-0.83
frequency(n+1)	-0.049509	0.025463	-1.94
1/length(n)	-0.125378	0.016070	-7.80
1/length(n-1)	0.013938	0.016983	0.82
1/length(n+1)	-0.026633	0.022811	-1.17
comprehension	-0.040153	0.073879	-0.54
recall	-0.010580	0.015112	-0.70
so * freq(n+1)	-0.047485	0.020163	-2.36
so * 1/length(n+1)	0.071728	0.021501	3.34
age * freq(n)	-0.035843	0.015362	-2.33
s6.load * freq(n)	0.084314	0.038783	2.17
s6.load * freq(n+1)	-0.096438	0.038343	-2.52
s6.load * 1/length(n-1)	-0.106122	0.038851	-2.73
freq(n) * freq(n+1)	-0.124176	0.029978	-4.14

Note: so: sentence order within the experiment (serial order vs. random order); age: age groups (young vs. old); capacity: WMC groups 4-7; set size: number of sentences within one trial (varied from 4 to 7); frequency: log word frequency of the current word (n) ,the previous word (n-1), and the following word (n+1); length: word length (using the reciprocal value 1/length) of the current word (n), the previous word (n-1), and the following word (n+1); comprehension: accuracy of sentence comprehension; recall: recall accuracy of the target word; s4: set size 4; s5: set size 5, s6: set size 6, s7: set size (e.g. s4.load: load effect at set size 4);

Table D.11: Analysis of successor and lag effects: Final LMM fitting of log single fixation durations.

Linear mixed model fit by maximum likelihood

Random effects:

Groups	Name	Variance	Std.Dev
subject id	(Intercept)	0.0131875	0.11484
word id	(Intercept)	0.0093955	0.09693
Residual		0.0973816	0.31206

number of obs: 9409, groups: subject id, 104; word id, 95

Fixed effects:

	Estimate	Std. Error	t-value
(Intercept)	5.3126014	0.0407080	130.51
sentence order (so)	0.0544686	0.0245630	2.22
age	0.0484644	0.0240201	2.02
capacity	-0.0089237	0.0119850	-0.74
set size	-0.0012867	0.0033709	-0.38
s4.load	0.0179728	0.0203566	0.88
s5.load	-0.0146029	0.0168474	-0.87
s6.load	0.0285152	0.0177631	1.61
s7.load	0.0113795	0.0152757	0.74
frequency(n)	0.0040298	0.0187082	0.22
frequency(n-1)	0.0007324	0.0132102	0.06
frequency(n+1)	-0.0792062	0.0154246	-5.14
1/length(n)	-0.0223618	0.0130843	-1.71
1/length(n-1)	-0.0266435	0.0136194	-1.96
1/length(n+1)	0.0480259	0.0114972	4.18
comprehension	-0.0649777	0.0627423	-1.04
recall	-0.0078893	0.0128073	-0.62
age * freq(n)	-0.0555116	0.0139060	-3.99
capacity * freq(n+1)	0.0172731	0.0080912	2.13
capacity * 1/length(n+1)	-0.0248867	0.0082862	-3.00
set size * freq(n+1)	-0.0132009	0.0067095	-1.97
s5.load * freq(n+1)	0.0967765	0.0396920	2.44
s5.load * 1/length(n+1)	-0.0799959	0.0402918	-1.99
s6.load * freq(n)	0.0745966	0.0350615	2.13
s6.load * freq(n-1)	0.1188511	0.0420675	2.83
s6.load * 1/length(n-1)	-0.1119652	0.0417792	-2.68
freq(n) * freq(n+1)	-0.0661010	0.0241269	-2.74

Note: so: sentence order within the experiment (serial order vs. random order); age: age groups (young vs. old); capacity: WMC groups 4-7; set size: number of sentences within one trial (varied from 4 to 7); frequency: log word frequency of the current word (n) ,the previous word (n-1), and the following word (n+1); length: word length (using the reciprocal value 1/length) of the current word (n), the previous word (n-1), and the following word (n+1); comprehension: accuracy of sentence comprehension; recall: recall accuracy of the target word; s4: set size 4; s5: set size 5, s6: set size 6, s7: set size (e.g. s4.load: load effect at set size 4);

Table D.12: Final LMM fitting of log sentence processing times. WMC groups are reduced to group 4 and 7.

Linear mixed model fit by maximum likelihood			
Random effects:			
Groups	Name	Variance	Std.Dev
subject id	(Intercept)	0.017292	0.13150
Residual		0.048305	0.21978
number of obs: 6870, groups: subject id, 43			
Fixed effects:			
	Estimate	Std. Error	t-value
(Intercept)	7.616600	0.020700	368.0
sentence order (so)	0.063434	0.040257	1.6
age	-0.002154	0.041592	-0.1
capacity (cap)	-0.013916	0.041159	-0.3
s4.load	0.041019	0.015868	2.6
s5.load	0.083792	0.013035	6.4
s6.load	0.108798	0.013281	8.2
s7.load	0.101680	0.011979	8.5
so * s5.load	-0.070597	0.024978	-2.8
so * s6.load	-0.057381	0.024991	-2.3
so * s7.load	-0.059959	0.023777	-2.5
age * s4.load	-0.088541	0.031737	-2.8
age * s5.load	-0.005268	0.025687	-0.2
age * s6.load	0.012059	0.025928	0.5
cap * s5.load	0.080394	0.025727	3.1
cap * s6.load	0.094197	0.025933	3.6
cap * s7.load	0.115818	0.023595	4.9
age * cap * s5.load	0.196451	0.051995	3.8
age * cap * s6.load	0.308763	0.052512	5.9

Note: so: sentence order within the experiment (serial order vs. random order); age: age groups (young vs. old); capacity: WMC groups 4-7; s4: set size 4; s5: set size 5, s6: set size 6, s7: set size (e.g. s4.load: load effect at set size 4)

Table D.13: Final LMM fitting of log gaze durations. Capacity groups are reduced to group 4 and 7.

Linear mixed model fit by maximum likelihood

Random effects:

Groups	Name	Variance	Std.Dev
subject id	(Intercept)	0.018005	0.13418
Residual		0.266672	0.51640

number of obs: 17641, groups: subject id, 115

Fixed effects:

	Estimate	Std. Error	t-value
(Intercept)	5.564058	0.022239	250.19
sentence order (so)	0.036093	0.043203	0.84
age	-0.027021	0.044681	-0.60
capacity (cap)	0.003779	0.044146	0.09
s4.load	0.082636	0.038362	2.15
s5.load	0.009598	0.028779	0.33
s6.load	-0.013516	0.031412	-0.43
s7.load	-0.003685	0.027691	-0.13
so * s4.load	-0.189196	0.072589	-2.61
so * s6.load	0.130507	0.059112	2.21
age * s4.load	-0.201662	0.075142	-2.68
age * s6.load	0.017683	0.061329	0.29
cap * s6.load	0.068764	0.061336	1.12
age * cap * s6.load	0.401481	0.124197	3.23

Note: so: sentence order within the experiment (serial order vs. random order); age: age groups (young vs. old); capacity: WMC groups 4-7; s4: set size 4; s5: set size 5, s6: set size 6, s7: set size (e.g. s4.load: load effect at set size 4)