The dynamics of long-range regressions during sentence reading

INAUGURALDISSERTATION

zur Erlangung des akademischen Grades

Doktor der Naturwissenschaften (Dr. rer. nat.)

durch die Fakultät für Human- und Sozialwissenschaften

der Bergischen Universität Wuppertal

Vorgelegt von

Anne Friede

Wuppertal, im November 2019

The PhD thesis can be quoted as follows:

urn:nbn:de:hbz:468-20210607-115234-4
[http://nbn-resolving.de/urn/resolver.pl?urn=urn%3Anbn%3Ade%3Ahbz%3A468-20210607-115234-4]

DOI: 10.25926/z7dq-9c17
[https://doi.org/10.25926/z7dq-9c17]

Von der Fakultät für Human- und Sozialwissenschaften der Bergischen Universität Wuppertal

als Dissertation im Februar 2021 angenommen.

Erstgutachter: Prof. Dr. Ralph Radach

Zweitgutachterin: Prof. Dr. Nicola Ferdinand

Tag der Disputation: 04.03.2021

Contents

List of figures

List of tables

Danksagung

„What separates those who achieve from those who do not is in direct proportion of one´s ability to ask for help.” (Donald Keough)

Zuerst möchte ich meinen Betreuer Prof. Dr. Ralph Radach danken, der mich für die Leseforschung begeistert hat und der nie gezögert hat, sein Wissen und seine Leidenschaft für Blickbewegungen mit mir zu teilen. Ralph, deine stetigen Ermutigungen und dein Vertrauen in mich als Wissenschaftlerin haben mich und meine Arbeit sehr bereichert. Ich bin sehr dankbar für alle Diskussionen die wir hatten und für deine Geduld beim Auswerten und Interpretieren von Ergebnissen, die auf den ersten Blick chaotisch wirkten. Des Weiteren möchte ich Prof. Dr. Albrecht Inhoff von Herzen danken. Ich habe unglaublich viel von dir gelernt. Ich danke dir für deine Hilfe bei der Datenanalyse, beim Umgang mit R und dafür, dass du meine Fragen immer so schnell beantwortet hast. Deine Rückmeldungen zu meinen Manuskripten haben mir sehr geholfen. Christian Vorstius, ohne dich hätte ich es wahrscheinlich niemals geschafft. Vielen Dank dafür, dass du jedes Drama relativierst und dir mit unglaublicher Geduld Probleme anhörst und sie scheinbar mühelos löst. Ich danke dir für deine Hilfe beim Programmieren der Experimente, beim Auswerten der Daten und für all die kleinen Ratschläge zwischendurch. Ich danke dir, Markus Hofmann, für angenehme Gespräche und hilfreiche Ratschläge. Thomas Alexander und Sebastian Fürth, ihr wart die besten Kollegen, die man sich wünschen kann. Herzlichen Dank für die tolle Zeit, die wir hatten! Ich möchte mich außerdem beim Team der Allgemeinen und Biologischen Psychologie bedanken: André Rölke, Mareike Kleemann, Sabine Hackenberg, Lea Bucher, Ehsan Ansari, Laura Schwalm, sowie bei meinen Praktikantinnen Lara Müller, Lara Steinhoff und Halima Nasri. Ohne euch hätte es nur halb so viel Spaß gemacht. Ein herzlicher Dank gilt den Probanden, die an meinen Experimenten teilgenommen haben und die scheinbar eine halbe Ewigkeit vor dem Eyetracker Sätze lesen mussten. Ganz besonders danke ich den Grundschulen in Duisburg, die bereit waren, meine Arbeit zu unterstützen und den Kindern, die ebenfalls sehr geduldige Leser waren.

Zu guter Letzt möchte ich mich von ganzem Herzen bei meiner Familie bedanken. Danke, dass ihr mir und meinem Weg vertraut habt.

Summary

Regressions are eye movements against the primary direction of reading, used to re-read words and solve comprehension issues. Long-range regressions leave the current word and move backwards to a word further to the left within the present sentence or text segment. The discussion about underlying programming mechanisms for these long regressions has produced two main hypotheses. The *spatial coding hypothesis* assumes that readers use information stored in visuo-spatial memory to direct long-range regressions to their target (Kennedy, Brooks, Flynn, & Prophet, 2003; Kennedy & Murray, 1987; Murray & Kennedy, 1988). As an alternative, the *verbal reconstruction hypothesis*, maintains that linguistic representations are used to reconstruct the location of prior words (Rawson & Miyake, 2002; Inhoff & Weger, 2005; Weger & Inhoff, 2007). Experiment 1 and 2 contributed to this debate, with the aim to disentangle the influence of visuo-spatial memory and reading performance. Experiment 1 showed that, for adult readers, performance in dynamic spatial memory is crucial for the accuracy of long-range regressions, determining the number of highly accurate single shot regressions. In addition, reading comprehension and reading speed affected the effectiveness of corrections, when the initial regression didn´t attain the target, as seen in total number of saccades and regression time. Experiment 2 compared the programming of long regressions between children and adults, finding that for children performance in static spatial memory is the determining factor for accuracy. Again, reading performance was only important for corrections. Overall, in children regressions were less accurate with more searching for targets. In experiment 3, a connection was made between the function of regressions and their programming mechanisms. It was found that difficult words lead to more accurate regressions and a reduced search effort compared to easy targets. This finding demonstrates that the cognitive effort for the processing of a word is closely related to its encoding in the spatial memory, supporting the programming of returning regressions. The question of a reference system for spatial information was addressed in experiment 4. Long-range regressions crossing a sentence boundary turned out to be as accurate as those within the same sentence, when semantic and syntactic properties of two clauses were held constant within a line of text. This provides evidence to suggest that the visual framework of a line is of central importance, providing a more important reference system for word location information compared to the linguistic structure of a sentence.

Zusammenfassung

Blickbewegungen, die während des Lesens gegen die Leserichtung programmiert werden, dienen der Verständnissicherung und werden als Regressionen bezeichnet. Lange Regressionen verlassen dabei das aktuelle Wort, um auf einem Wort weiter links im Satz oder Text zu landen. Die Diskussion um die zugrundeliegenden Mechanismen der Programmierung dieser langen Regressionen hat zwei gegensätzliche Hypothesen hervorgebracht. Die *spatial coding hypothesis* nimmt an, dass die Position von Wörtern im räumlichen Gedächtnis gespeichert wird und dies die benötigten Informationen zur Programmierung von Regressionen liefert. Im Gegensatz dazu geht die *verbal reconstruction hypothesis* davon aus, dass die Position von Wörtern in einem Satz durch linguistisches Wissen, vor allem aus dem verbalen Arbeitsgedächtnis, erschlossen wird. Experiment 1 und 2 der vorliegenden Dissertation tragen zu dieser Diskussion bei, indem sie die Einflüsse des räumlichen Gedächtnisses und der Lesefähigkeit auf die Korrektheit langer Regressionen untersuchen. Die Ergebnisse von Experiment 1 zeigen, dass für erwachsene, geübte Leser die Kapazität des dynamischen räumlichen Arbeitsgedächtnisses ausschlaggebend für die Korrektheit von langen Regressionen ist und damit den Anteil an single shot-Regressionen bestimmt. Leseverständnis und Lesegeschwindigkeit beeinflussen den Korrekturprozess, der notwendig wird, wenn die initiale Regression das Zielwort nicht trifft. Eine hohe Lesekompetenz geht dabei mit einer geringeren Anzahl notwendiger Sakkaden und einer geringeren Regressionszeit einher. In Experiment 2 wurden die Mechanismen der Programmierung von Regressionen zwischen Kindern und Erwachsenen verglichen. Bei Kindern in der 4. Klasse erwies sich das statische räumliche Arbeitsgedächtnis als entscheidender Faktor für die Genauigkeit von Regressionen. Die Lesefähigkeit war, wie auch bei Erwachsenen, bedeutend für das Korrigieren zu kurzer oder zu langer initialer Regressionen. Insgesamt zeigten Kinder weniger genaue Regressionen und mussten häufiger oft nach dem Zielwort suchen. Eine Verbindung zwischen der Funktion von Regressionen und dem zugrundliegenden Mechanismus der Programmierung zeigten die Ergebnisse des Experiments 3. Regressionen zu schwierigen Wörtern waren genauer als Regressionen zu einfachen Wörtern. Damit ist gezeigt, dass der kognitive Aufwand bei der Verarbeitung von Wörtern für die Speicherung der Wortpositionsinformationen im räumlichen Gedächtnis und damit die Genauigkeit der Programmierung von Regressionen von hoher Bedeutung ist. Experiment 4 untersuchte die

Frage nach einem Referenzsystem für die Speicherung räumlicher Informationen über Wortpositionen. Der Vergleich von Regressionen innerhalb eines langen Satzes, der aus zwei Teilsätzen besteht und von Regressionen zwischen zwei kurzen Sätzen, die durch einen Punkt voneinander getrennt sind, zeigte, dass die Genauigkeit langer Regressionen nicht durch das Überqueren einer Satzgrenze beeinflusst wird. Das deutet darauf hin, dass die Zeile als räumliches Kriterium eine wichtige Funktion hat und entscheidender ist, als der Satz als linguistisches Konstrukt.

1 Introduction

1.1 Eye movements in reading

In recent decades, knowledge about eye movement control in reading has steadily increased, thanks to a large amount of data from eye tracking experiments. The recording, visualizing and analyzing of eye movements has contributed to understanding the nature of information processing during reading. While reading, our eyes move over words, sentences and paragraphs of text with very fast saccadic movements. These saccades are in the majority of cases progressive and proceed in the direction of reading (from left to right in German and English). Progressive saccades have an average length of 7-9 letters and take about 20-50 ms. Saccades alternate with fixations, during which the eyes remain relative stable in the current visual axis for about 70 to over 500 ms with means around 200 to 250 ms. Due to saccadic suppression, new visual information is only extracted during fixations (Rayner, 1978; Rayner, 1997; Radach & Kennedy, 2004, for an overview on eye movements in reading). In 15-25 percent of the cases, saccades are programmed against the direction of reading and referred to as regressive saccades or regressions (from right to left in German and English) (Vitu & McConkie, 2000; Vitu, McConkie, & Zola, 1998).

For fluent readers, the typical position of the first (initial) fixation is left to the word center, between the beginning and the middle of a word, often described as the "preferred viewing position" in reading (Rayner, 1979b; Radach & McConkie, 1998; Reilly & O`Regan, 1998; McConkie, Kerr, Reddix, & Zola, 1988). This position is closely related to the optimal viewing position of words. The most effective (and hence optimal) position for word recognition is at the word center, thus a bit right of the preferred viewing position (O'Regan, Lévy-Schoen, Pynte, & Brugaillère, 1984; O`Regan & Lévy-Schoen, 1987; O`Regan & Jacobs, 1992).

The position of the initial fixation on a word is influenced by the launch site of the incoming saccade, corresponding to the previous fixation position. The launch distance is usually measured as the distance (in letter spaces) between this saccade starting point and the beginning of the target word (technically the blank space preceding the target). The further the launch site of an incoming saccade lies to the left of a target word, the further shifted to the left is the landing position within this word, a phenomenon described as the saccadic launch distance error, or saccadic range error (McConkie et al., 1988, replicated by Nuthmann, Engbert, & Kliegl, 2005).

Further determinants of the initial landing position are word length (the longer the word, the more shifted to the right), position within the line of text (shifted to the right for words at the beginning), whether the previous word was fixated or skipped (shifted to the left when the prior word was not fixated) and the number of fixations on the previous word (shifted to the right when the previous word received more than one fixation) (McConkie, Kerr, & Dyre, 1994; McConkie et al., 1988; Radach & Kempe, 1993; Radach & McConkie, 1998). The following figure (figure 1) shows changes of the mean landing position of the initial fixation as a function of both word length and launch distance, in a very clear way. The three columns of line graphs from left to right represent different word lengths (4, 6, and 8 letter words) and the four rows from bottom to top represent different launch sites (-1, -3, -5, and -7 letter spaces left to the blanc space preceding the word). Horizontal axes show the character positions of the word, where zero marks the blank space preceding each word. Vertical axes represent the percentage of initial fixations on that word. As can be seen in figure 1, an additional word length of 2 letters shifts the mean landing position about 0.6 – 0.9 characters to the right. Furthermore, for each character position the launch site shifts to the left, the mean landing site shifts about 0.5 character positions to the left as well. With word length and launch site held constant, the landing position of the initial fixation on a word follows a normal distribution, with a maximum of initial landing positions slightly left of the word center (McConkie et al., 1988).

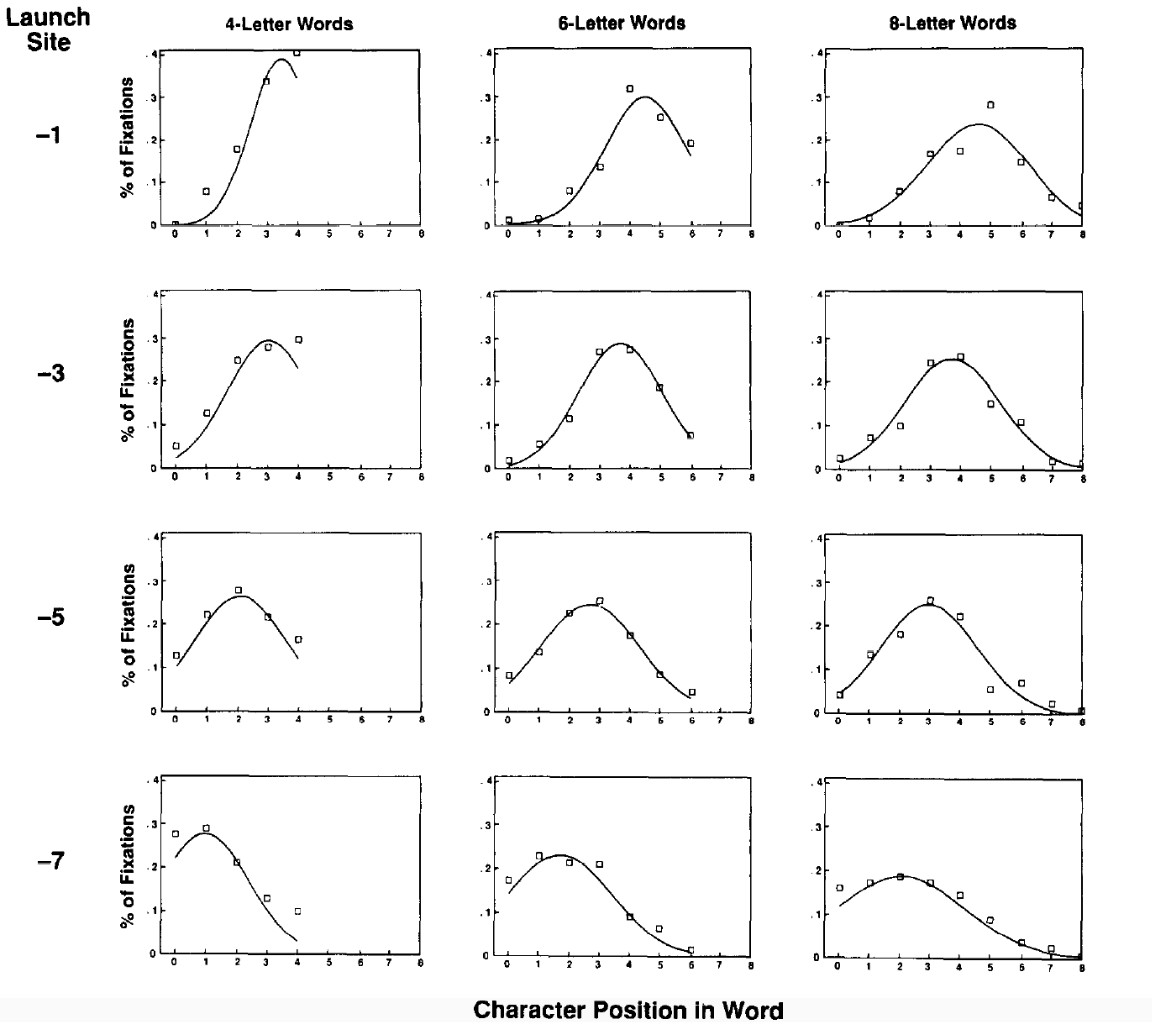


Figure 1. Proportion of initial fixations on words of different length and for different launch sites, taken from McConkie, Kerr, Reddix, & Zola (1988).
Position 0 represents the blank space to the left of the word beginning. Assuming that, negative numbers represent letter positions to the left and positive numbers letter positions to the right of that. Data is based on 43,668 fixations collected from 66 students reading chapters of a novel.

Depending on start and landing position, saccades are either refixation/intraword saccades (start and land within a word) or interword saccades (landing position in another word). Refixations on a word are controlled in two different ways. On the level of visuomotor control, a refixation is most unlikely, when the initial fixation lands near the word center. The more this landing position shifts to the word beginning or the word end, the higher the probability of refixations, yielding a u-shaped refixation curve with maxima at the word's edges and a minimum at the word center (O´Regan, 1984; McConkie, Kerr, Reddix, Zola, & Jacobs, 1989). On a linguistic level, the more difficult a word is (e.g. lower frequency), the higher is the probability of refixations on that word (McConkie et al., 1989; Inhoff & Radach, 1998).

As eye tracking methods and technology have steadily improved over the last decades, current eye movement monitoring is highly accurate and allow analyses of eye movements at the word and even letter level (e.g. Tang, Reilly, & Vorstius, 2012). Word based temporal and spatial eye movement measures reflect oculomotor behavior during reading (see Radach & Kennedy, 2004, for an overview of spatial and temporal eye movement measures). Spatial parameters include saccade amplitude (length of the saccade in letter spaces), fixation probability (relative frequency with which a word is fixated at least once), fixation position (position in the word, where the fixation is located), launch distance (distance between the location of the prior fixation and the beginning or center of the current word) and refixation probability (relative frequency of making at least one additional fixation before leaving the word).

Temporal eye movement measures are, among others, first fixation duration (duration of the first fixation on a word), refixation duration (duration of all additional fixations), gaze duration (duration of all fixations on a word, before leaving the word for the first time), re-reading time (duration of all fixations made on the word after leaving it for the first time) and total reading time (duration of all fixations made on a word) (cf. Radach & Kennedy, 2004). Both, spatial and temporal eye movement measures represent the underlying decision about where and when to move the eyes. The decision about where to move the eyes (spatial aspect of eye movement programming) is widely assumed to be word-based and low-level spatial information are used to guide the saccadic eye movements to target a specific word (McConkie & Zola, 1984; Radach & McConkie, 1998). In contrast, the decision about when to move the eyes (temporal aspect of eye movement programming) is more influenced by cognitive factors and determined by the processing difficulty of the current and the following word(s) (Rayner & Pollatsek, 1981).

The area within which information can be processed successfully during reading is referred to as the perceptual span (Rayner, 1975). The perceptual span is often examined using the so-called moving window technique (McConkie & Rayner, 1975), in which the visible text field can be manipulated during the reading process. The availability of visible letters to the left and right of the current fixation is limited by replacing letters with Xs or other similar or dissimilar letters (see figure 2). To this end, fixation contingent display changes are used, to examine, how many letters have to be visible to achieve a normal reading speed, without a noticeable effect on temporal and spatial eye movement parameters. The perceptual span is typically asymmetric and directed towards the direction of reading. Depending on the distance of the current fixation, various types of information can be extracted and used for parafoveal preprocessing of the reading materials. For successful readers, the span in which information about word length and word boundaries can be extracted extends about 14-15 characters to the right of the current fixation (McConkie & Rayner, 1975). The span in which information about individual letters of a word can be used is slightly shorter and comprises a window of 8-9 letters to the right (Underwood & McConkie, 1985). To the left, the span extends about 3 to 4 letters or up to the beginning of the current word (Rayner, McConkie, & Zola, 1980). Following the foveal-load hypothesis (Henderson & Ferreira, 1990) the perceptual span is influenced by linguistic properties of foveal words. Difficult words lead to a decreased perceptual span in comparison to easy words.

```
XXXX XX a preview sXXXXXX XXXXXX XXXXXXX.
```

Figure 2. Illustration of the moving window technique according to Rayner (1986), taken from Häikiö et al. (2009). Visible are only 12 letter positions (including the blank spaces between words), the rest of the sentence was masked by Xs.

In languages with a reading direction from right to left (e.g. Arabic, Hebrew), the extension of the perceptual span turns around and is longer to the left than to the right side (Pollatsek, Bolozky, Well, & Rayner, 1981). This supports the assumption, that the extent of the perceptual span is a function of attention during the reading process (Pollatsek et al., 1981).

The orientation of the perceptual span can also change dynamically during reading, depending on the direction of the following saccade. Apel, Henderson, & Ferreira (2012) found that regressive saccades are shorter, when, in a masked condition, two words to the left of the starting point of a regression are masked (at least 4 letter positions to the left are visible, to not interrupt the normal perceptual span in reading) and that fixations and gaze durations are longer in this case. This manipulation had no effect on progressive saccades and therefore provides clear evidence that before a regressive saccade is executed, attention is allocated to the left and the perceptual span is prolonged towards the left direction.

1.2 Regressive saccades

Natural reading of sentences and paragraphs includes eye movements against the direction of reading (from right to left in English and German). These “regressive” saccades occur with a frequency between 15 and 25 percent of all saccades depending on materials, instructions etc. There are several possible reasons for the planning and executions of a regression, both on the level of basic visuomotor control and linguistic processing (Inhoff, Weger, & Radach, 2005; Rayner & Pollatsek, 1989; Vitu & McConkie, 2000).

Looking at short regressions, a distinction needs to be made between regressions within and between words. Regressions within words can be triggered for two independent reasons: (1) when the eyes initially land at a letter position at the beginning or end, which is not optimal for word recognition and (2) when linguistic processing of a word turns out to be difficult. Within-word regressions are typically intended to attain the word center share the metric properties described in section 1.1 for progressive saccades (Radach & McConkie, 1998). They are not the topic of the present dissertation thesis.

Oculomotor reasons for relatively short between-word regressions include, as an example, cases where a progressive saccade overshoots a short word. This results in a landing position to the right of the intended word, which is in turn corrected with a regressive saccade. Longer regressions within the same line of text, or even beyond, are usually caused by problems on the cognitive level. This includes, in the majority of cases, either incomplete lexical processing or semantic and/or syntactic comprehension deficits at both sentence and text level. These different scenarios will be discussed in detail in the following sections.

1.2.1 Regressions caused by visuo-motor errors

About a quarter of all regressions are relatively short intra-word regressions (Vitu & McConkie, 2000), which are assumed to correct progressive saccades that have missed the intended target location because of an overshoot in near targets or an undershoot with more distant targets (Taylor, 1971; Reilly, & O'Regan, 1998). Consequently, intra-word regressions are more likely to occur when saccades land at the rightmost letters of the present word, if previous saccades were relatively long and if preceding words were not fixated (Vitu & McConkie, 2000; Vitu, McConkie, & Zola, 1998).

Additionally, errors in the visuo-motor system lead to mislocated landing positions of fixations especially at word boundaries. Experimental studies have shown, that some of the fixations at word beginnings and word ends were intended for adjacent words (Nuthmann, Engbert, & Kliegl, 2005; Nuthmann & Engbert, 2008). The amount of this type of mislocated fixations ranges about 10-30 percent depending on word length (Nuthmann & Engbert, 2008). As a result, more inter-word regressions occur with landing positions of single fixations at the beginning of words (Vitu & McConkie, 2000).

Another cause for regressions based on visuo-motor errors are corrective regressions following return sweeps (Hofmeister, Heller, & Radach, 1999; Parker, Slattery, & Kirkby, 2019). Return sweeps are long goal directed saccades, that move the eyes from the end of one line to the beginning of the next line, when reading a text or a paragraph. These long saccades often undershot the target (the first word on the line) and corrective regressions are necessary to enable a start from the beginning of the line. Between 10 and 53 percent of return sweeps are followed by such corrective regressions, depending on the length of the preceding line (Heller, 1982; longer lines increase the probability of corrective regressions).

1.2.2 Regressions caused by incomplete lexical processing

If a word has not been fully lexically processed, because the word was either skipped in first pass reading or a saccade has left the word before completion of lexical access, a regression has to be programmed to go back to that word in order to complete lexical processing (Vitu, McConkie, & Zola, 1998; Vitu & McConkie, 2000; Engbert, Longtin, & Kliegl, 2002; Engbert, Nuthmann, Richter, & Kliegl, 2005). Consequently, words that are less perceptible are more often the target of regressions. Length and frequency of words modulate regression rates, too, with more regressions targeting long and low-frequency words (Vitu & McConkie, 2000).

In harmony with this assumption, skipped words are more regressed to than unskipped words (Vitu, McConkie, & Zola, 1998; Vitu & McConkie, 2000; White, 2008). Regression rates are influenced by word characteristics of skipped words, regression probability increases with word-length and is higher for low lexical frequency than for high lexical frequency of skipped words (Vitu & McConkie, 2000; White, 2008).

Experiments with words and their orthographic neighbors (words differing in just one letter, which causes a difference in meaning) further support the effect of word frequency on regression probability. In sentences, where a high-frequency neighbor of a word is highly favored, more regressions are programmed back to that word, indicating a false assumption about the meaning of the word in the first place (Perea & Pollatsek, 1998; Pollatsek, Perea, & Binder, 1999; Slattery, 2009). A sentence out of the materials used by Perea & Pollatsek (1998) should serve as an example here: "Everything was clean except for one plate that had egg on it.". "Plate" is the target word and has a higher frequency neighbor, the word "place", who would also fit perfectly into the sentence context. More regressions are programmed to the word "plate" in comparison to the target "spoon" within the same sentence ("Everything was clean except for one spoon that had egg on it), since "spoon" has no higher frequency neighbor.

1.2.3 Regressions caused by comprehension deficits

Another important reason for inter-word regressions are lexical and/or structural ambiguities in sentences, which are raising comprehension difficulties during reading (e.g. garden path sentences). In this case, inter-word regressions are used for re-inspection and integration of the word meaning into the sentence context (Frazier & Rayner, 1982; Liversedge, Paterson, & Pickering, 1998; Mitchell, Shen, Green, & Hodgson, 2008; Levy, Bicknell, Slattery, & Rayner, 2009; Bicknell & Levy, 2011).

For sentences with structural ambiguities, Frazier & Rayner (1982) describe the processing of sentences with the *garden-path theory of sentence comprehension.* According to this hypothesis, readers use only one possible analysis of a sentence at a time, usually the easiest one, even though a further interpretation would be possible. This leads to cases where reanalyzes of sentences are necessary, when the reader is "led down the garden path" (Frazier & Rayner, 1982, p.178). When reading a sentence, readers build a syntactic tree based on their implicit knowledge about sentence phrases. In doing so, they follow the concepts of "minimal attachment" and "late closure", meaning that words are preferably added to the currently read phrase and if an attachment to the phrase is not possible (because the sentence phrase is already complete) the new material is most frequently semantically assigned to the last processed phrase (see Clifton & Staub, 2011, for a summary, including contrary positions to this theory and effects and time course of sentence parsing). A garden path sentence, taken from Frazier & Rayner (1982), can serve as an example: "The girl knew the answer was wrong". Reading this sentence, readers assume that "the answer" is a direct object that refers to the verb "answer" (The girl knew the answer, late closure) and realize later, that it is rather the subject of a complement sentence (the answer was wrong). This kind of reanalysis leads to increased fixation durations and regression probability.

Lexical ambiguities arise from words with more than one meaning. When the first assumed meaning (the meaning with the higher probability of occurrence) of a word does not match the following context, more regressions are programmed back to that word, compared to words with only one meaning (Carpenter & Daneman, 1981; Frazier & Rayner, 1982; Bicknell & Levy, 2013).

The obvious assumption about long-range regressions is that they should aid comprehension. To verify this hypothesis, Schotter, Tran and Rayner (2014) used a trailing mask paradigm where the letters of each word were masked after the word had been read and a progressive saccade had left. This manipulation had a dramatic effect on both semantically easy and difficult (garden path) sentences, suggesting that comprehension decreases globally, if readers are not able to regress and that regressions are an important part of the reading process and support reading comprehension.

The way in which regressions contribute to the understanding of text was further assessed by Booth and Weger (2013). These authors made an attempt to clarify whether regressions can also serve the additional function of facilitate memory for specific words (*deictic pointer hypothesis*). In their experiments, participants read sentences for comprehension. After reading, they were asked questions about the sentence content.

Example: “My shirt is blue and my jacket is green”. Which piece of clothing is blue?

In a first condition, sentences were masked after first pass reading (letters were swapped, but word length and case sensitivity maintained). In this case, participants made fewer and less accurate regressions and more errors in answering the comprehension question. In a further condition, sentences were only presented auditorily (via speakers) as meaningful sentences and written as a random combination of letters (and hence did not make sense) with the same word length. In this condition participants did not make any regressions. In a third condition, target words (words targeted by the questions) were swapped after reading. Participants, who made a regression to the target word after reading the sentence answered the comprehension question incorrectly (responding with the later inserted word). This is strong evidence indicating that readers were more influenced by their second pass reading than by their first one and regressions are used successfully for rereading of words. Nevertheless, long-range regressions might also be used to trigger memory processes, as both processes (rereading and facilitate memory) do not have to be mutually exclusive.

Importantly, the frequency of regressions can be adapted to the demands made on the reader. Expecting difficult questions following the reading materials leads to more regressions compared to easy questions (Wotschak & Kliegl, 2013; Weiss, Kretzschmar, Schlewesky, Bornkessel-Schlesewsky, & Staub, 2018). For example, in Weiss et al. (2018), the sentence “The chef that distracted the waiter sifted the flour onto the counter.” was either followed by the question “Did a chef do something?” (easy question) or the question “Did the waiter distract the chef?” (difficult question). More difficult questions lead to more regressions programmed from the two or three final words of the sentence and to increased re-reading times (Weiss et al., 2018).

The occurrence of regressions has also proven to be a good indicator for processes of comprehension monitoring in reading. Comprehension monitoring can be described as a process of continuous evaluation of how well currently read content is understood and how well it fits the context (Vorstius, Radach, Mayer, & Lonigan, 2013). Comparing sentences with positive (“Erica blushed because she was nervous”) and negative polarity (“Erica blushed although she was confident”) in children (5th grade), Vorstius et al. (2013) found, that negative polarity leads to a significantly increased proportion of inter-word regressions targeting the conjunction of the sentence. Also, increased re-reading time was strongly related to improved comprehension, as indicated by more correct answers to questions on sentence coherence.

In line with these arguments, a similar view on the causes of inter-word regressions has been suggested by Bicknell and Levy (2011). They argue that the likelihood of regressions increases as readers come across words that fit less well with what they have already read and what they think matches the further context. This leads to a so-called *confidence falling*, since the readers trust in their own reading performance is reduced, leading to more re-inspections. This hypothesis has been tested in a text corpus analysis by Bicknell and Levy (2011), using word length (short vs. long words), word frequency (high vs. low frequency), and predictability (high vs. low predictability) of a word (n) and the word to the left of it (n-1) as predictors of regression frequency. Regression rate to word n-1 was influenced by word n-1`s length (more regressions to longer words), frequency (more regressions to high frequency words) and predictability (more regressions to less predictable words). Regression rate was also influenced by predictability of word n (less predictable words n caused more regressions back to word n-1). The authors conclude, that a less predictable word n raises the likelihood that confidence about the correctness of the context will decrease. This happens especially when an unpredictable word n-1 suggests a different (less expected) context, which is mostly the case when these words have a high frequency.

As laid out above, regressive saccades can be elicited for different reasons, including visuo-motor errors, incompletion of lexical processing and comprehension deficits. Comprehension related regressions are saccades (or re-reading saccade patterns) used to remedy integration failure on lexical and syntactic level during reading. In the majority of cases, these are longer regressions (targeting words left the word n-1), which is why they are often referred to as long-range regressions. Obviously, short inter-word regressions may also be used to clarify comprehension problems and the use of long-range regressions as a synonym for comprehension related regressions is not entirely correct, but often used.

No matter what the cause for a regressive saccade is, the eye movement needs to be planned and executed. In the following section the term “programming” refers to the interconnected processes of selecting an intended target word and specifying a saccadic movement aimed at attaining this particular target (see Findlay & Walker, 1999, for a comprehensive theory of saccade programming).

1.3 How regressions are programmed

The underlying programming mechanism of inter-word regressions appears to differ fundamentaly from that of progressive saccades. Regressive saccades do not show an effect of word length and tend to land in the center of the word, irrespective of how long the targeted word is (see figure 3) (Radach & McConkie, 1998). Additionally, an influence of the launch site or whether a word was skipped or not could not be found in regressive saccades (Radach & McConkie, 1998; Vitu, McConkie, Kerr, & O`Regan, 2001), taken as evidence that the programming mechanism of regressions differs from that of progressive saccades. The great accuracy with which regressive saccades attain many target words even outside of the perceptual span raised the question of whether information from visuo-spatial memory is utilized in their programming. The discussion about the importance of visuo-spatial information and visuo-spatial memory for reading already started in the early 1980´s. Several experiments showed that participants could remember the spatial location of words and information quite accurate (see Baccino & Pynte, 1994, as well as Inhoff, Weger, & Radach, 2005, for a summary of experiments and authors). A correlation between spatial memory and remembering and retrieving content of the reading materials was demonstrated: If participants have information about the spatial distribution of a text, the reading content can be reproduced much better (Lovelace & Southall, 1983; Baccino & Pynte, 1994).

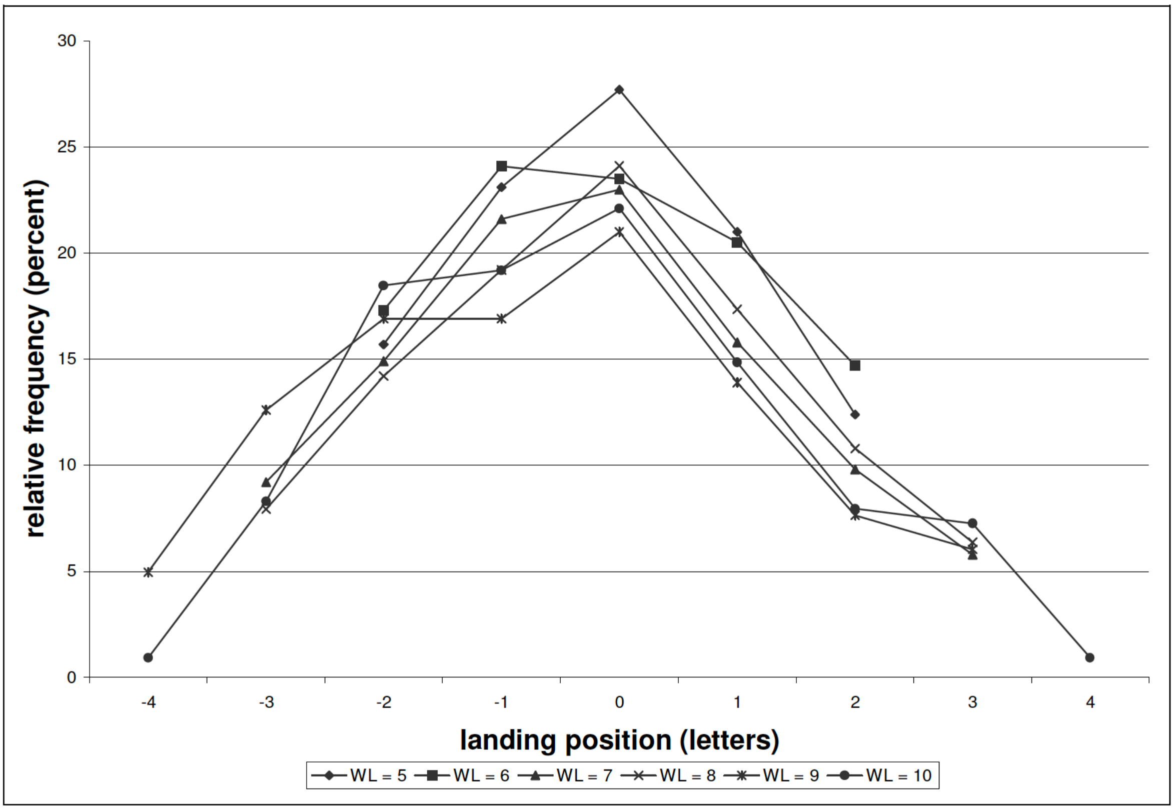


Figure 3. Illustration of relative distribution of landing positions for regressive interword saccades as a function of target word length, taken from Inhoff, Weger, & Radach (2005).
Landing position 0 represents the center of the word.

On text level, Lovelace and Southall (1983) found, that the recall of text content was poorer, when all 12 pages of the text were presented on a continuous sheet with no space between individual pages and no page numbers marked, compared to a presentation mode with extra space between beginning and end of each page and inserted page numbers. Correspondingly, Baccino and Pynte (1994) found a corresponding effect on sentence level. In a decision task, subjects should decide if a particular word was included in a sentence just read (yes/no). Responses were faster and more often correct in the spatial mode (each segment occupied a specific location on the screen, although segments of the sentence were presented one after the other) than in the non-spatial mode (each segment was displayed on the same location in the middle of the screen). Furthermore, the recollection of text can be cued with information about the location of the information (Lovelace & Southall, 1983).

The question arises, how readers program long regressions and how these regressions are guided to targets, further away than the present perceptual span. The spatial coding hypothesis assumes that the boundary of a page or a screen works as a sort of spatially addressed external memory. This position is supported by a number of experiments (Kennedy, Brooks, Flynn, & Prophet, 2003; Kennedy & Murray, 1987; Murray & Kennedy, 1988). Kennedy and Murray (1987) presented single line sentences to their participants. After the end of each sentence a target word was shown. If this target word had been part of the sentence, most participants were able to look back to the word accurately and with only one long-range regressive saccade. This was true, regardless of whether the word was located at the beginning or the end of the sentence, indicating that participants were able to program precise long-range regressions to a given word, irrespective of the distance of the target. Corrective saccades were relatively small (averaging 2.7 character spaces), indicating that readers were actually intending to fixate the target word. These findings led to the conclusion that the spatial coordinates of the sentences are present throughout reading and can be readily retrieved to program long-range regressions.

In a follow-up experiment with children at the age of 9-11 years (Murray & Kennedy, 1988), the authors found that successful developing readers are able to attain target words with only one accurate long-range regression (referred to as single shot regressions). In contrast, less successful readers needed to search for target words with more and shorter regressions and in a backtracking way from the end of the sentence until the target was found. Consequently, they showed less single shot regressions. Reading performance was based on results in the NFER Reading Test BD. Successful readers were those with standard reading scores at least three points above the average of the sample, whereas less successful readers are those with reading scores at least three points beneath the average of the sample. A control group comprised younger readers (two years younger) with a raw score comparable to the score of the less successful readers and the same age based standard score like the more successful readers. Compared to the younger control group, poor readers seem to have noticeable problems in programming regressions, as they showed more backtracks than the control group. This leads to the conclusion that the development of a good visuo-spatial memory seems to be an important precondition for good and fluent reading.

Additional support for the spatial coding hypothesis was also obtained in experiments that manipulated the visual appearance of text (Kennedy et al., 2003). Single line sentences were framed with markers before and after the sentence and a target probe was presented at the very end of the line. The markers build a reference system for visuo-spatial memory during reading and the probe indicated the saccade target in the sentence. In shift conditions, the size of the markers was increased or decreased by two character positions, creating an apparent shift of the sentence relative to the frame without actually altering the physical sentence position. Long-range regression programmed to go back to the target word in the sentence (as indicated by the probe) were influenced by this manipulation. If the sentence frame was shifted to the right, regressions undershot the target by about the distance the markers had been moved (Kennedy et al., 2003). Similar results can be observed in sentence-shift experiments, where instead of added markers, the complete sentence itself was shifted to the right or to the left by one character position (Inhoff, Weger, & Radach, 2005). Mean saccade amplitude was not affected by such line shifts, but remained constant in length over all conditions. The planning of long-range regressions therefore appears to be controlled primarily by knowledge of the spatial location of words out of visuo-spatial memory.

Evidently, spatial information plays an important role in saccade programming. However, as shown in subsequent research, the original spatial coding hypothesis has difficulty accounting for inaccurate regression targeting, for effects of spatial target distance, and effects of linguistic distance.

The role of verbal memory for regression programming has first been pointed out by Rawson and Miyake (2002). They found that the ability to indicate the location of previously read information within a text (page and line) is predicted by verbal memory. Quit surprisingly, performance on visuo-spatial tasks was not correlated with relocation accuracy. These authors claim that participants with higher verbal abilities are better in constructing a text representation and rely on this type of knowledge in re-locating the information in a text, this suggestion will be referred to as the *verbally based reconstruction view*.

However, mechanisms of regression programming may differ between text and sentence level. Therefore, Inhoff and Weger (2005) used sentences with target words located either at the beginning or the end of the sentence, e.g. My mother is younger than my father. Who was born earlier/later?" (left side target/beginning of sentence: mother, right side target/end of sentence: father). Regressions rarely reached the target word and the mean regression error was relatively large (17 letter spaces). It seems that participants programmed one long regression taking the eyes in the general direction of the target followed by a number of small regressions to home in on the target. Both, the size of the regression error and the length of the regression depended on the target location, being larger for distant than for near targets.

In a later experiment, Weger and Inhoff (2007) contrasted effects of spatial distance against effects of linguistic distance (the number of words in-between a denoted regression target and the location from which the regression was launched) to test effects of verbal memory and regression accuracy. To this end, they used sentences presented over two lines, again with target words either at the beginning (distant targets) or end of each line (near targets). It was assumed that linguistic processing demand was low in the first line and high in the second line. Targets were presented over headphones, either when fixating the end of line one (spatial and linguistic distance are identical) or line two (spatial near targets have a far linguistic distance). Readers should then execute a regression to the presented target word. Results showed that both distances influenced regressions from the source location toward the regression target. Regressions were shorter for near compared to distant targets. Importantly, the regression error was larger for distant than for near targets and results were similar for first- and second-line condition. In addition, more corrective regressions were needed for distant targets and this effect was stronger for the second line, thus when verbal memory working load is high. This implied that regressions were controlled by more than one type of knowledge, and Weger and Inhoff proposed a hybrid model in which initial regressions were guided by spatial knowledge while subsequent regressions were guided by verbal/linguistic knowledge.

Mitchell et al. (2008, experiment 2) also pointed out the importance of analyzing the full trajectories of the regression sequence, to see, whether regressive scanpaths are selective for certain linguistic properties. They used sentences with syntactic ambiguity (garden path sentences) either with an early or a late disambiguation area to examine, whether regressions were sent directly back to the area causing the ambiguity. They defined the term regressive sequence, which includes the first regression and all following fixations until an eye movement beyond the disambiguation word is executed. They found (in line with Inhoff and Wegner, 2005), that first regressions were not guided precisely to the target area, but that additional fixations were directed to the area that caused the misanalysis of the sentence. Typically, this takes two to three fixations.

Guérard, Saint-Aubin & Maltais (2012, experiments 2 and 3) showed, however, that an articulatory suppression task during reading of sentences diminished the accuracy of initial regressions. Participants read one-line sentences followed by the auditive presentation of a probe word, after they had crossed the period at the end of the sentence with a saccade. They were instructed to look back to the corresponding target word within the sentence afterwards. Articulating the letters A, B, C, D repeatedly with a duration of 500 ms per letter (articulatory suppression task) when reading the sentences to disrupt verbal memory led to an increased regression error for initial regressions and resulted in a slightly increased number of subsequent corrective regressions to find the target word.

In a following study, based on these results and on the experiments of Inhoff & Weger (2007), Guérard, Saint-Aubin, Maltais, & Lavoie (2014) specified their findings. While participants read sentences presented over two lines, they again performed an articulatory suppression task in 50 percent of the cases. Articulatory suppression increased the initial regression error and the number of corrective saccades, but only for regressions initiated from the second line to the first part of the sentence, not for regressions executed from second line to second part or from first line to first part of the sentence, suggesting, that verbal memory was primarily used to guide initial regressions to relatively distant targets, while another type of memory, likely spatial knowledge, may be used when the regression target is relatively near.

Taken together spatial memory is a promising source of information for the programming of long-range regressions, since targets of these longer regressive saccades are outside the perceptual span in the majority of cases and therefor word location information are not visual perceptible. However, recent studies have shown that spatial memory can not account for all effects found in regression programming, indicating, that a strict acceptance of the spatial coding hypothesis has to be questioned. There is an ongoing discussion about the extent of verbal memory and linguistic knowledge involved in regression programming. Guérard et al. (2012 and 2014) claim, that verbal memory is the primary source of regressions programmed to targets far away from the launch site of a regression and that spatial memory serves, if at all, as source of information for the programming of shorter regressions to near targets. Inhoff et al. (2005 and 2007), on the other hand, sees spatial memory as the primary source for initial regressions and describes the role of verbal memory only in the performance of corrective saccades.

1.4 The role of regressions in computational reading models

Different assumptions about the cause for the programming of regressive saccades are also related to different computational reading models and their presumption about the programming of eye movements. Depending on the theory of attention distribution during reading, a distinction is made between serial (sequential attention shift, SAS models) and parallel (processing gradient, PG) reading models (Radach, Inhoff, & Reilly, 2007).

One of the best-known serial models is the E-Z reader model (Reichle, Pollatsek, & Rayner, 2006; Reichle, Pollatsek, Fisher, & Rayner, 1998). The authors of the model assume that only one word can be processed at a time, and that attention is not shifted to the next word until the lexical processing of the word is completed, leading to a strictly serial mode of word processing. In a recent version of their model (EZ Reader 10, Reichle, Warren, & McConnell, 2009), the authors describe a mechanism to account for short-range context-related regressions. In their model, difficulties in postlexical integration are assumed to be a possible cause for regressions. If difficulties arise in the integration of a word n into the context, the lexical processing of the word n + 1 is stopped, so that the entire processing system loses speed and the attentional spotlight is directed back to the place of difficulties with the programming of an inter-word regression.

If difficulties in integration of the word occur before the programming of a saccade to the next word, attention is also redirected to the origin of processing difficulties within that word and an intra-word regression is programmed. The authors acknowledge the limitations of their account, noting that their “assumptions are not sufficient to explain regressions that result from earlier processing difficulty or problems that are noticed only after some delay”, Reichle, Warren, & McConnell, 2009, p. 10).

Two of the most detailed gradient models are the SWIFT model (Engbert, Longtin, & Kliegl, 2002; Kliegl & Engbert, 2003) and the Glenmore model (Reilly & Radach, 2006). The SWIFT model (Engbert, Longtin, & Kliegl, 2002; Kliegl & Engbert, 2003) is based, in contrast to the E-Z reader model, on the assumption that during reading, attention is spatially distributed and thus the parallel processing of words is possible. During reading, words are perceived within a graded attentional window to the left and right of the current fixation. The authors differentiate between two types of regressions, namely local and global regressions. Local regressions are defined as short regressions within the attentional window and regarded as a special case of refixations. Global regressions are explained through incomplete lexical processing of words before leaving the attentional window. If the lexical processing of a word is not completed during first pass reading of a word, then a regression is programmed back to that word. Note that “global” regressions in this context are still bound to local lexical processing and therefore not necessarily identical to the long-range regressions discussed earlier.

The Glenmore Model (Reilly & Radach, 2006) differs even more from serial models than the SWIFT model. The authors of this model assume that saccades are programmed on the basis of a saliency map, describing the attractiveness of words with the perceptual span. When the information acquisition form a given word is exhausted, the word with the highest saliency value within the perceptual span becomes the goal of the next saccade. Both visual properties of the words as well as linguistic properties such as the frequency of a word are crucial for dynamic changes in word saliency during the course of processing. With this mechanism, the Glenmore model can also account for regressions. If a word to the left of the current fixated word possesses the highest saliency and the activation of word n drops rapidly, an inter-word regression is programmed (Reilly & Radach, 2006).

Even though current computational models of reading are able to describe the programming mechanisms of many short-range inter-word regressions within the model, it is not within their scope to account for the underlying programming mechanisms of comprehension-related regressions. To achieve this goal, such models would have to be coupled with an effective computational model of linguistic processing on the sentence and text level.

1.5 Motivation for the dissertation thesis

Despite the importance of regressions for reading comprehension and their contribution to successful reading, their underlying programming mechanism is still under discussion. It is known that especially long-range regressions are used to re-read words that have caused comprehension difficulties, with the aim of solving problems in understanding und ensure fluent reading. But what is the underlying programming mechanism that enables regressions to attain specific words precisely, even if such targets are located outside the current visual field? As argued above, visuo-spatial memory appears to be a promising source of target location information. However, there is still a need for clarification to what extend the programming relies on spatial memory for previously fixated locations and what other sources of information, like linguistic knowledge and verbal memory, contribute to accurate regression programming.

In their classic experiment, Kennedy and Murray (1988) found that good readers differ from poor readers in their way of programming regressions. Skilled readers show highly effective single shot regressions, attaining the target word with only one accurate movement, or needing at most one small correcting saccade. Less well performing readers seem to have a clear disadvantage here. They need to backtrack from the end of the sentence with more and smaller regressions until the target word is found. This raises the question whether readers with a high accuracy of regression targeting show higher capacities in their spatial working memory than poor readers and if the way in which regressions are programmed is related to reading fluency, reading accuracy and eventually to successful reading comprehension. Another question concerns the role of regression accuracy in the development of reading. Do developing readers show the same way of programming long-range regressions as adults? Or does the skill of accurate regression programming develop with age and/or with the associated level of reading performance?

The central goal of the present dissertation thesis is to contribute to the debate about the role of spatial working memory for the programming of long-range regressions. More specifically, the first aim is to further advance our understanding of how regressions are programmed in adult, skilled readers and to what extent spatial memory is the source of information. Since more than one type of spatial working memory could contribute to regression accuracy, it is important to differentiate between static and dynamic working memory and to assess their relative contributions. In light of the findings of Murray and Kennedy (1988), the effects of a readers' text comprehension and reading fluency skills on regression programming are also examined. This includes a detailed analysis of different types of regressions to see whether the known classification into single shot and backtracking (Murray & Kennedy, 1988) covers all types of regressions or if further types of regressions have to be classified into subgroups and are potentially characteristic for a certain level of performance in either test of spatial memory or reading.

A second aim is to look at children during their reading development. How and in which way visuo-spatial memory contributes to a successful reading development is still largely unknown. Assuming that a good spatial working memory is essential for text orientation and the selective re-inspection of words, a deficit in this area would directly affect reading speed and reading accuracy and thus negatively affecting the development of reading skills. The present dissertation thesis is intended to close this gap by examining children at the age of 9-11 years (grade 4 in the German school system) to compare regression accuracy and regression strategies of children and adults in the light of their spatial working memory capacities and their reading skills.

A further aim targets the relationship between word difficulty and the programming of inter-word regressions as an expression of the memory for word positions in previously read sentences. Comparing easy to difficult words and their representation in memory can shed light on the question, whether spatial representation of word positions in sentences is independent of linguistic properties of words or if processing difficulties lead to less successful retention of target position because of limitation of processing capacity.

Finally, the fourth aim of the dissertation thesis concerns the question whether regressions crossing the sentence boundary (a period) are programmed in a different way compared to regressions within a sentence. It is reasonable to assume that mechanisms for regression programming differ within and across sentence, assuming that reading for meaning requires integration, so that the meaning of individual sentences can be integrated into a common situation model. It might be that the wrap-up of meaning at a sentence boundary interferes with spatial memory for word positions, hampering the accuracy of regressions back into the sentence.

Four eye tracking experiments were conducted. Experiment 1 focused on adult, skilled readers and the way in which they regress to words in already read sentences and their performance in tests of static and dynamic spatial working memory and tests of reading fluency and comprehension, followed by a detailed analysis of strategies that are used to regress to a previously read target. Experiment 2 examined children in fourth grade of school, when children are approximately at the age of 9 to 10 years. This stage of reading development was chosen, because children at that age have usually reached a level of reading competence where they master basic decoding processes in reading (e.g. grapheme to phoneme correspondence), are more sensitive to comprehension problems in reading and are assumed to be familiar with programming long-range regressions because of higher level linguistic reasons. Experiment 3 also examined adult skilled readers and manipulated the difficulty of target words to examine the influence of reading difficulty on word location memory. To go a step beyond sentence reading, Experiment 4 was conducted, to compare regression accuracy in adults between different syntactic types of reading materials to examine the influence of sentence boundaries (the period at the end of the sentence) on the programming of regressions.

1.6 Overview of the experiments

All four experiments included German speaking participants. All adults of experiment 1, 3 and 4 were tested at the Bergische Universität Wuppertal and all children of experiment 2 were examined at primary schools in Duisburg (North Rhine Westphalia). Experiment 1 and experiment 2 are related to the discussion between two existing hypothesis describing long-range regression programming, the spatial coding hypothesis, assuming spatial working memory as the sole source of regression programming and the verbal reconstruction theory, claiming that verbal memory and linguistic skills are also used to reconstruct the position of a previous read word. Experiment 3 examined the influence of word difficulty on the accuracy of regression programming and experiment 4 was conducted to describe the difference and similarities in programming long-range regressions within and between sentences.

1.6.1 Experiment 1: Visuomotor strategies and the role of spatial memory for regressive saccades in reading

Going back to locations left of the current fixation is an essential part of skilled reading, necessary to recover from problems in word processing or to address issues in comprehension. Experiment 1 examined how such regressive eye movements are planned and to what extend they rely on spatial memory for previously fixated locations. The work of experiment 1 included the execution and analysis of regressions to targets that had appeared in a previously read sentence. 48 Skilled adult readers read single line sentences containing a target word close to the beginning or end of the sentence. This manipulation defined a distant (far from the starting point of the regression) and a near (close to the starting point of the regression) target location in the sentence. Participants were asked to go back and check target words for a spelling error, which was present in 50 percent of the trials after reading the sentence.

Importantly, reader's spatial working memory capacity as well as their reading performance were assessed to examine whether regression targeting is modulated by individual differences in spatial and linguistic skills. Both, readers' static and dynamic spatial working memory were examined, since more than one type of spatial working memory could contribute to regression accuracy, the assumption being that higher working memory capacities would predict higher targeting accuracies and more single shot regressions to regression targets, especially for near targets. Performance in reading fluency and reading accuracy were assessed, the assumption being that increased linguistic skills would influence primarily the accuracy of regressions to far targets and, possibly, the occurrence of single shot regressions to these targets. Furthermore, the reading materials used in experiment 1 included a shift manipulation. The location of the previously read sentence was either unchanged or shifted one character to the left (50 percent each) after the sentence was read for the first time but before the regression was executed, using an invisible boundary method (Rayner, 1975). If readers used spatial representations of target locations to direct regressions, then regressions should undershoot regression targets, when the sentence was shifted.

1.6.2 Experiment 2: Long-range regressions in beginning readers

After Experiment 1 focused on the programming mechanisms of long-range regressions in adult, skilled readers, the goal of Experiment 2 was to further advance the understanding of the importance of visuospatial memory for the programming of long-range regressions in children and whether this ability contributes to the development of strong reading skills. Therefore, 78 4th grade students with a mean age of 9;11 years were examined using the same reading materials as in experiment 1, to ensure a high comparability between both groups. Experiment 2 pursued two goals. First, to measure regression accuracy and regression strategies in children and compare them to adults and second, to examine readers' individual capacity of static and dynamic spatial working memory and performance in a test for reading comprehension and reading speed. Experiment 2 therefore determined whether children at the age of 9-11 years show the same regression strategies as skilled readers out of Experiment 1. The analysis of regression strategies in children was intended to show whether different regression strategies develop together or if a certain strategy is dominant over another. Furthermore, regression strategies of children with high and low performance in both reading ability and spatial memory were compared to see whether high performance in one of the tests may lead to a higher proportion of single shot regressions and, in turn, low performance may be related to other strategies.

1.6.3 Experiment 3: The influence of word difficulty on regression accuracy

In experiment 3 adult participants (N=54) were asked to read sentences containing either a difficult or an easy target word. Difficulty was manipulated using lexical frequency and token frequency of initial letter trigrams of nouns. A high frequency of both variables is associated with shorter reading times and hence easier processing. The general procedure followed experiment 1 and 2 and included near and far regression targets, which had to be checked for spelling errors to elicit long-range regressions. Results could potentially turn out in two different ways, depending on capacity considerations: Difficult words might lead to more accurate regressions and less need for correction, because of their higher processing time. The more time is spent on the word, the more elaborate may be the representation of an associated potential target position in memory. On the other hand, easier words could lead to a higher accuracy in regression programming, when more cognitive capacity is available to store information about target position in memory, because of the reduced processing time with easy targets.

1.6.4 Experiment 4: The effect of sentence boundaries on long-range regressions

After experiment 1, 2 and 3 focused on regression accuracy on sentence level, experiment 4 was conducted to go a step beyond sentence level processing, examining long-range regressions traversing a sentence boundary (period). 50 skilled adult readers were tested with reading materials allowing for comparison between intra-sentence and inter-sentence regressions. Two sentences were either separated by a period or combined with the conjunction “und” (*and* in English) to examine whether this leads to differences in regression accuracy. The motivation to use two sentences in one line in comparison to one sentence was to see in which way the end of a sentence (made visible with a period) exclusively effects the information about word locations and accuracy in regression programming, regardless of the information that comes with a line break. The end of a sentence may lead to loss of information about word positions and in this case, regressions should be much more inaccurate for sentences separated by a period. This might be the case especially for far targets, when the sentence boundary has to be crossed with a long regression. Alternatively, the sentence boundary manipulation may end up to have no influence on regression accuracy, suggesting that information about target position is not tied to the linguistic construction of a sentence.

2 Experiment 1: Visuomotor strategies and the role of spatial memory for regressive saccades in reading

2.1 Introduction

Backward-directed saccades can be large, extending across several words or lines, and are linked to processing difficulties and the recovery from processing error (Vitu & McConkie, 2000; Inhoff, Weger, & Radach, 2005; Rayner, Chace, Slattery, & Ashby, 2006; Slattery & Rayner, 2009). But how do readers guide long-range regressions to a previously read word that is relatively far from the current viewing location, and thus outside the range of effective vision? Two accounts have been offered. One, the spatial coding hypothesis, maintains that readers endow the represented linguistic form of an identified word with a spatial tag, and then use this tag, possibly with reference to a visible landmark (such as page margins) to direct long-range regressions to the regression target (Kennedy, Brooks, Flynn, & Prophet, 2003; Kennedy & Murray, 1987; Murray & Kennedy, 1988). The other, the verbal reconstruction hypothesis, maintains that accessible linguistic representations are used to reconstruct the location of prior words in the text (Rawson & Miyake, 2002; Inhoff & Weger, 2005; Weger & Inhoff, 2007; Guérard et al., 2012; Guérard et al., 2014), but the exact way of this effect on regression programming is still vague. Guérard et al. (2012, 2014) differentiate between regressions to near and far targets, assuming that verbal memory primarily influences the programming mechanisms to far targets/targets that have been read a longer time ago. Inhoff et al. (2005, 2007) distinguish initial regressions from additional corrective saccades and describe the influences of verbal memory especially for corrective saccades, executed when the initial regression failed to reach the targeted word. Since Guérard et al. (2014) did not found this effect when analyzing corrective regressions, further examination of both, initial regressions and additional corrective regressions, is essential to specify the actual influence of spatial memory and determine additional effects of linguistic representations.

Just like other processes of reading, the programming of inter-word regressions could be subject to intra-individual differences, anchored in the personal characteristics of the reader. Individual differences in reading texts were described by Hyönä, Lorch and Kaakinen (2002) for adult readers and the importance of analyzing individual scanpaths during reading to capture differences in reading performance was emphasized by von der Malsburg, Kliegl and Vasishth (2015). An early account of individual strategies in regression programming has been provided by Frazier and Rayner (1982) who analyzed sentences with structural ambiguities and found that readers utilized different strategies: They either used a selective reanalysis strategy and regressed directly to the disambiguation area, or they engaged in a forward reanalysis strategy, starting at the beginning of the line to process the whole sentence again, but readers never started at the end of the sentence to backtrack in multiple regressions. A first analysis of individual differences in regression programming related to reading comprehension was published by Murray & Kennedy (1988) who found differences in their participants (children with an age between 9-11 years) according to their reading performance. Good readers usually found the regression target with one single shot regression, regardless of whether the target was near or far away from the starting point of the regression. Poor readers executed significantly less single shot regressions and had to rely on a so called backward scanning strategy, starting at the end of the sentence and working their way through the sentence with small regressive saccades until the target was found.

The current experiment sought to extend this work by determining whether regression targeting is modulated by individual differences in readers' spatial and linguistic skills. Although readers code the (static) spatial location of a target word according to the spatial coding hypothesis, the distance of the eyes to a previously read target word is dynamically changing with the execution of eye movements. Since more than one type of spatial working memory could contribute to regression targeting, the current work examined the influence of readers' static and dynamic spatial working memory on regression accuracy, the assumption being that higher working memory capacities would predict higher targeting accuracies and more single shot regressions to regression targets. A higher accuracy of regressions to far targets would provide evidence for the assumption, that spatial memory primarily guides long-range regressions (Inhoff & Weger, 2005; Kennedy et al., 2003), whereas higher targeting accuracies especially to near targets would support the assumption that verbal memory and linguistic memory but not spatial memory guide regressions to words far away from the starting point of the regression (Guérard et al., 2012; 2014).

Readers' text comprehension skills were also assessed, the assumption being that increased linguistic skills would influence primarily the accuracy of regressions to far targets and, possibly, the occurrence of single shot regressions to these targets. These predictions on accuracy and the use of single shot regressions are in harmony with Murray and Kennedy's (1987) findings, where skilled readers make more single shot regressions than poor readers, with Zabrucky and Commander (1993) who found that good readers were more selective when rereading text whereas poor readers backtracked more randomly and with the assumption of Guérard and collegues (2014), that verbal memory was primarily used to guide longer regressions to far targets.

The present work included two parts, one involved the assessment of readers' spatial and dynamic working memory capacity and of their general comprehension skill, the other involved the reading of one line sentences followed by the execution of regressions to specific target words that had appeared in a previously read sentence. Importantly, sentences reading was followed by comprehension questions targeting both simple and complex sematic relations within sentence to ensure that participants read for meaning and fully understood sentence comprehension as their primary task. In addition, the approach included a shift manipulation so that the location of the previously read sentence was either unchanged when the eyes moved to the probe word or shifted one character to the left. If readers used spatial representations of target locations to direct regressions, then regressions should undershoot regression target when the sentence was shifted, in particular when targets were near and when regressions should be guided by spatial knowledge (Kennedy, Brooks, Flynn, & Prophet, 2003).

Similar to Murray and Kennedy's (1987), participants read one-line sentences that were followed by a probe word that identified a near (close to the probe) or distant (far from the probe) regression target in the sentence. A spell-checking instruction was used to elicit saccades to a specific word, thus approximating the process of going back to check on a certain word during normal reading. Usually, regressions during reading are not under direct control of the experimenter and the goal of regressions is difficult to anticipate. As shown by Inhoff and Weger (2005), the size of regressions to near and far targets within one-line sentences is identical, no matter if regressions occurred naturally or if participants were instructed to regress to a certain target. Weiss and colleagues (2018) showed that difficult sentences lead to more regressions, but not until the final words of the sentence had been reached, again underlining the equivalence of the used procedure to a natural reading situation.

2.2 Methodology

Participants

Forty-eight university students (41 female) volunteered or received experimental course credit for their participation. All had normal or corrected to normal vision and were native German speakers. The mean age was 23 years (range = 18 to 34 years). All students gave written consent prior to participation. Two participants had to be excluded from the data analyses because they did not finish the psychometric assessments for personal reasons.

Psychometric Assessment

Reading Test

Participants' reading skill was assessed with the LGVT 6-12, a test for reading fluency and comprehension (Schneider, Schlagmüller, & Ennemoser, 2007). It is normed for German students up to 12th grade, when students are approximately 18 years old. It requires the reading of text for four minutes. On several locations within the text, three words are presented in parentheses, and participants underline the word that was congruent with the context. Two measures, the number of correctly underlined words and the number of read words, were used to index comprehension and speed, respectively (higher values indicated better comprehension and higher reading speed).

Working memory test battery

Participants' visuo-spatial working memory was assessed with two subtests of the AGTB 5-12 (Hasselhorn et al. 2012), the matrices test and the Corsi block tapping test (Corsi task). The matrices test started with the presentation of a white 4X4 grid pattern. On the first trial, two squares of the grid were shown in black for 2400 ms. The grids then turned back to white, and participants were asked to tap on the grid elements that had been shown in black. After successful completion of a trial, the number of black grid elements was increased by one. The number of black elements remained constant if an error was made, and it was decreased by one – and testing was terminated – when there was more than one consecutive error. Consequently, the test measured an aspect of spatial working memory that represents concurrent (static) spatial locations.

The Corsi task involved the touch-screen presentation of nine spatially separated white boxes on a gray background. A sequence of smileys was then presented one by one (950 ms each) within the boxes. After this, the participant was asked to repeat the *sequence* from memory by tapping the corresponding order of white screen boxes. The progression of set size started with an increase from one to two targets and was continued until the participant made a sequencing error. The computer registered correct and incorrect touch screen sequences. Since the to-be-remembered locations emerged over time and had to be reproduced in their correct order, the test is assumed to index the capacity of dynamic spatial working memory. Together, completion of the reading skill and the spatial working memory assessments required approximately 15 minutes.

The inter-correlation of the four test scores revealed a high positive correlation between the two reading skill components (r = .83) and a moderate correlation between the two spatial working memory tests (r = .33). The correlations between reading speed and the two spatial working memory scores were low (r = .04 and r = -.07) as was the correlation between comprehension and the two spatial working memory test scores (r = .06, r = -.09).

Reading Task

The to-be-read materials consisted of 100 declarative sentences each of which contained a target word or a filler. An additional probe word – that was identical to the target or filler word in the sentence - was positioned five letter spaces to the right of the last word of the sentence (see figure 4). The probe was masked with a sequence of #s while the sentence was read, and unmasked when the eyes crossed the center of the area between sentence ending and the string of #s (the central pixel representing an invisible display change boundary [Rayner, 1975]). Four additional sentences served as practice items.

`Eine sehr schöne **Melodie** wird gerade auf dem alten Klavier vorgespielt. | #######`

A very beautiful **melody** is now being played on the old piano.

`Eine sehr schöne **Melodie** wird gerade auf dem alten Klavier vorgespielt. | Melodie`

Figure 4. Example sentences out of experiment 1.
Note that the probe word at the end of the line was masked with hash marks and only became visible, after participants crossed the invisible boundary. After reading the probe, the corresponding target had to be found in the sentence, to check whether or not it was spelled correctly. Spelling errors were introduced in 50 percent of the cases, simultaneously with the unmasking of the probe. Highlighting of target words and invisible boundary are for demonstration purposes only. Target words were not marked in any way during the experiment.

All sentences were 67-70 characters long and contained 9-14 words. 64 of the 96 experimental sentences contained a target word. All targets (and corresponding probes) were nouns with six or seven constituent letters and either part of the subject or the adverbial complement of the sentence. The word frequency of targets ranged from 5.2 to 93 per million with a mean of 29 (dlexDB; Heister et al., 2011). Each target word was preceded by an adjective with a length of 5 - 8 characters.

Experimental sentences were designed so that the target was located either close to the sentence beginning (far from the probe location) or its ending (near the probe location). If the target word was located at the beginning, it was preceded by two or three words; if it was near the end, it was followed by one or two words. German syntax allows the construction of sentences with changes in word order without altering sentence meaning or grammatical well-formedness, as a subject or adverbial phrase can either be the first element or the penultimate element of a sentence. Both alternatives are regularly used and therefore familiar, although use of the subject at the beginning of a sentence tends to be more common with spoken language. This linguistic property was used to construct sentence pairs with identical meanings in which the identical target occurred near to and also far from the target. Both members of the pair appeared on a list of sentences, so that each of 32 target words appeared near the beginning and near the ending of a sentence.

In addition to the 64 experimental sentences, the list also contained 32 filler sentences with a regression target near the center of the sentence. This was used to counteract the development of location specific target expectancies, as the word that matched the probe could have appeared anywhere in a sentence with equal frequency (except for the very beginning and ending). Figure 5 provides examples of sentences and experimental conditions used in the experiment.

Figure 5. Structure of the sentences of experiment 1.
Note that linguistic components are marked for highlighting the German grammar structure. Sentence 1 represents a sentence with a far target (50 percent of the sentences), which was part of the subject of the sentence. Sentence 2 shows the same sentence frame with the near target position (other 50 percent of the cases). Sentences 3 and 4 present the same sentence frame, but the target was part of the adverbial construction. Each participant saw every target frame in both distance conditions (i.e., either 1+2 or 3+4). Each participant saw 50 percent of all sentences with a target in the adverbial (Klavier) and the other 50percent in the subject (Melodie) part of the sentence. All targets were preceded by an adjective with controlled features (see text for detail).

Apparatus

All to-be-read sentences were presented as a single line of text in black font on a light-grey background. The sentences were centered in the middle of a 21-inch flat-panel monitor (a resolution of 1,680 x 1,050 pixels and a refresh rate of 120Hz). The text was presented in Courier New font, size 15pt and the viewing distance was set to 68cm which resulted in a letter width of 0.3° of visual angle. Eye movements were recorded with an SR EyeLink1000® video-based eye tracking system (SR Research, Toronto, Canada). Viewing was binocular but only the right eye was recorded. The registration is based on infrared light reflection from pupil and cornea at a sampling rate of 2,000 Hz. A three-point calibration and validation routine was performed at the beginning of the experiment and after each question to maintain a spatial resolution of approximately 0.3°. In addition, a drift check was carried out before each sentence to ensure accuracy between calibrations. If drift checks indicated deviations greater than 0.3°, an additional calibration was performed. Display changes were initiated when the right eye crossed an invisible vertical line to the right of the sentence (see above) and completed within less than 11 ms.

Procedure

The experiment started with a description of the experimental procedure followed by the sentence reading task. Participants placed their chin and forehead on a forehead rest so that head movements would be minimized and the eye-monitor distance was approximately 68 cm. This was followed by the calibration of the eye tracker and the subsequent presentation of four practice sentences. During practice and experimental (and filler) sentence reading, participants controlled the on- and offset of each sentence with a button press; they were asked to read at a comfortable pace so that they could answer sentence comprehension questions that would occur after the reading of a subset of sentences. These questions were asked after approximately every sixth sentences, and they queried semantic relations (locations, actors, objects, attributes, conditions, causalities) that were described in the sentence. All questions could be answered with yes or no response, and the answer was provided via pressing of one of two buttons.

Participants were also instructed to look at the masked probe location to the right of the sentence after sentence reading. They were told that this would reveal a probe word corresponding to a target word in the previously read sentence. To entice readers to execute a regression to the target and to make this task meaningful, they were asked to determine whether the target was correctly or incorrectly spelled when re-read. Half of the targets were unchanged and correctly spelled during re-reading, the other half were misspelled by changing a single non-initial letter of the word. In the construction of spelling errors vowel replaced vowel and consonant replaced consonant. Descending, ascending and baseline letter were exchanged by the same category of letters, thus maintaining the original length and word form (e.g. “Melodie” became “Meladie”). These minimal changes were always implemented at near-center locations within words, ensuring that the resulting pronounceable non-words constituted visually non-conspicuous targets for regression saccades.

The saccade to the right of the sentence ending onto the probe word location could also result in the implementation of another display change. On half of the trials, the entire sentence was shifted one letter position to the left (farther away), and no shift occurred on the remaining trials. Target distance (near vs. far), linguistic function of the target (subject vs. adverbial complement) and the factors shift condition (shift vs. no shift) and error (error vs. no error introduced to target word after crossing the boundary) were used to create a total of eight experimental lists. The entire session lasted between 30-45 minutes.

Oculomotor Measures and Data Selection

Eye tracking data were processed using the custom build software suite EyeMap (Tang, Reilly, & Vorstius, 2012). As the experiment sought to examine regression guidance, four regression-related oculomotor measures were extracted and analyzed: (1) initial regression size, consisting of the length of the first saccade (in letter spaces, LS) from the probe toward the target location; (2) the regression error, comprising the distance (in LS) between the initial landing position and the center of the target, (3) the number of regressions that were needed to reach the target word after the probe word was fixated; (4) total regression time, comprising the interval (in ms) between the onset of the first regression toward the target and the fixation of the target.

Data Analyses

Participants' performance on the psychometric assessments was used to define three levels of performance (lower, middle, and upper level) for each of the four test measures (reading comprehension, reading speed, matrices test, Corsi test). The lower level included participants whose performance fell into the lowest quartile (lower 25 percent), the middle level comprised the two center quartiles (the interquartile range), and the upper level consisted of the upper quartile (the upper 25 percent). The number of participants in each group and percentile ranking of participants in each level are provided in Table 1.

Table 1. Participants and percentile ranking by level of test performance for experiment 1.

Level of performance	Reading comprehension		Reading speed		Matrices test		Corsi test	
	N	PR	N	PR	N	PR	N	PR
lower	12	9-31	11	9-34	11	12-31	12	12-42
middle	19	39-72	24	36-74	22	38-82	22	54-86
upper	15	84-100	11	76-100	13	84-88	12	95-98

N= number of participants, PR = percentile ranking

The eye tracking records were examined for missing data, and trials with missing data for one or more of the four oculomotor measures were dropped (1.2 percent). In addition, trials with initial regressions of less than five characters were removed, as they did not move the eyes onto the sentence, and trials with initial regressions of more than 53 LS, as they were extreme outliers. Trials with other outlier values were also removed: total regression times of < 200 ms and >1500 ms, and regression frequencies >8. Combined, this resulted in the exclusion of 6.8 percent of the trials. For statistical analyses, participants' initial regression size values were log transformed to obtain normally distributed values; the distribution of participants' number of regressions and their total regression time data was close to normal and these data were not transformed. Analyses of variance (ANOVAs) with the factors target distance (near vs. far), target shift (present vs. absent), and target order (first vs. second occurrence on the list of to be read sentences) were applied to these three variables.

Subsequent analyses simplified the ANOVA model through the removal of non-significant fixed factors and they added individuals' level of performance for the four test measures.

For each test, a separate ANOVA was calculated, including the simplified ANOVA model plus the fixed factor performance level. Since these levels were ordered (lower, middle, upper), a polynomial contrast (contr.poly) was applied to discern linear trends. In view of the inter-correlation of the four assessment scores, the performance level on a single test was entered into the statistical model which yielded four assessment specific ANOVA models with experimental factors that explained a significant amount of variance in the eye movement data plus the factor assessment level. In view of the high inter-correlation of the two reading skill indices, effects of reading comprehension and reading speed were expected to be quite similar. Correspondingly, matrices and Corsi levels were assumed to yield similar effects on the three oculomotor measures. However, effects of reading skill and of spatial memory were expected to differ.

In contrast to the three oculomotor measures (initial regression size, regression number, and total regression time), the data for the regression error (success of the initial regression) were not normally distributed, as relatively small regression errors were relatively common. Regression strategy analysis revealed, that the size of the regression error is not a meaningful tool to describe the success of a regression, since a common search strategy is to start at the center of the sentence with a relatively small regression error size to both, near and far targets in the first place, but being followed by at least two additional saccades until the target is found, yielding no information about target position stored and retrieved from memory. Consequently, these data were not subjected to an ANOVA. Instead, the accuracy of regressions on individual trials was used to identify underlying regression strategies for successfully attaining the target.

In Murray and Kennedy (1988), readers used two types of strategies to reach the regression target: they either moved the eyes with a single regression onto the target (single shot regressions), or they scanned backward from the sentence end to the sentence beginning to reach the target (backward search). Inspection of the current data revealed a more varied search pattern for the target when it was not attained with a single shot regression.

Specifically, the inspection of eye movement patterns from the probe to the target revealed five different search strategies: (1) *Single shot regressions* which occurred when a single regression positioned the eyes within a target word (19.4 percent); (2) *goal directed regressions* which occurred when initial regressions positioned the eyes more than half way toward the target center (at least 50 percent and at most 150 percent of the distance between the probe and the target), and when this was followed by a single smaller saccade to the target (23.9 percent); (3) *backward searches* which occurred when the initial regression was relatively small, landing in the ending third of the sentence, and when this was followed by at least two additional saccades to the target (29.2 percent); (4) *centered searches* which occurred when the initial regression landed near the center of the sentence, and when this was followed by least two additional saccades toward the target (18.5 percent); (5) *forward (beginning-to-end) searches* occurred when readers executed a long regression to the beginning of the sentence and then executed at least two additional saccades to reach the target (3.1 percent). The eye movements toward the target were haphazard and unclassifiable on the small number of trails (5.7 percent).

In a subsequent step, it was examined whether readers' preferred search strategy for the target was associated with their level of reading skill (comprehension and speed) and/or their spatial working memory (matrices [static], Corsi [dynamic]). For this, participants with an upper or lower level of test performance were selected, and their preferred (modal) search strategy for the target was determined. When a participant used several types of searches with equal frequency in a particular experimental condition, the weaker search strategy was selected (e.g., a tie between one and two step regressions would default to two step regressions [there were only three tied values]). Chi-square tests were then applied to determine whether participants' preferred target search strategy and their level of test performance were associated.

2.3 Results

Regressing to the Target

Table 2 shows the mean size of initial regressions (in character spaces), the mean number of regressions, and mean total search time (in ms) as a function of target distance (near vs. far), target shift (no shift vs. shift), and target order (first vs. second occurance). The analysis of the three sets of data revealed a robust effect of target position, as initial regressions to far targets were larger, $F(1, 352) = 9.1$, $p<.01$, $\omega^2 = .13$, and required more regressions, $F(1, 352) = 69.9$, $p<.001$, $\omega^2 = .17$ and longer total regression times, $F(1, 352) = 54.18$, all $p<.001$, $\omega^2 = .16$ than regressions to near targets. The main effect of target order and its interaction with other factors were negligible in all analyses (all *ps*>.5). Target shifting had virtually no effect on the initial regression size ($F<1$), but regression number and total regression time data showed marginal effects of target shifting, $F(1, 352) = 3.08$, $p<.08$, and $F(1, 352) = 3.05$, $p<.08$, respectively. None of the interactions involving shift approached significance, all $p>.12$. Numerically, a sentence shift barely increased the number of regressions and search time, with .07 more regressions and 15 ms longer total regression times in the shift condition.

Table 2. Means and standard errors for amplitude of initial regressions, total number of regressions and total regression time (in ms) as a function of target position, shift condition and presentation order.

		near				**far**			
		no shift		**shift**		**no shift**		**shift**	
		first	**rep**	**first**	**rep**	**first**	**rep**	**first**	**rep**
Initial Regression	**M**	**30.43**	**28.36**	**28.62**	**28.99**	**37.09**	**33.59**	**37.08**	**35.12**
	SE	1.14	1.00	1.13	1.14	1.33	1.18	1.26	1.17
Number of Regressions	**M**	**2.45**	**2.23**	**2.19**	**2.29**	**3.07**	**3.32**	**3.08**	**3.21**
	SE	0.11	0.09	0.10	0.10	0.11	0.10	0.09	0.11
Regression Time	**M**	**652**	**587**	**602**	**616**	**793**	**794**	**773**	**775**
	SE	25.8	21.3	23.2	22.8	20.8	18.6	19.1	16.1

M= Mean, SE = Standard Error, first = first time of presentation, rep = second time of presentation

The complete absence of a shift effect on initial regression size, the relatively small size of the regression number and total regression time effects, neither of which reached statistical significance, suggests that the influence of the sentence shift on spatial regression targeting was spurious, although some readers could have noted a peripheral brightness change in shift condition. Consequently, neither the factor target order nor the factor target shift was included in subsequent analyses that examined the influence of individual differences in reading skill and spatial working memory on regressions to the target.

Four ANOVAs, using the factors target position and assessment level (for comprehension, reading speed, the matrices test [static spatial working memory], and the Corsi test [dynamic spatial working memory]), revealed similar effects for the two indexes of reading skill, on the one hand, and the two indexes of spatial working memory, on the other hand. Specifically, the level of reading comprehension had virtually no effect on the amplitude of the initial regression ($F<1$), but it influenced the number of regressions (eye movements) to the target, with a higher comprehension level being associated with fewer movements (2.6, 2.7, and 2.9 regressions for upper, middle, and lower levels, respectively), $F(2,360) = 6.36$, $p<.01$, $\omega^2 = .02$ and it also influenced the total regression time data (which are typically a function of the number of executed movements), better comprehension levels being associated with shorter total regression times (673 ms, 697 ms, and 734 ms, for upper, middle, and lower levels), $F(2, 360, = 5.14$, $p<.01$, $\omega^2 = .016$. Participants' level of reading speed had no effect on initial regression size and number of regressions ($F<1$), but faster reading speeds being associated with shorter total regression times (675 ms, 691 ms, and 740 ms, for upper, middle, and lower levels), $F(2,360) = 5.51$, $p<.01$, $\omega^2 = .018$. All interactions of performance level and target position was negligible, with all $p>.4$. Since readers often executed more than one movement to the target, these supplementary analyses suggest that reading skill influenced searches for the target that ensured after the initial regression toward the target had been executed.

In contrast to the effects of reading skill, participants' performance on the matrices test had a profound effect on the size of initial regressions, which decreased in size as test performance increased, $F(2,360) = 7.15$, $p<.001$, $\omega^2 = .027$. The effects of the matrices test on the number of regressions and total regression time were negligible (*p*s>.1). Similarly, increased Corsi test performance was linked to decreased initial regression size (35 LS, 32 LS, and 30 LS, for lower, middle, and upper levels), $F(2,360) = 7.67$, $p<.001$, $\omega^2 = .029$. In addition, Corsi performance modulated the target position effect for the two related measures regression number and total regression time, $F(2,360)= 4.65$, $p<.025$, $\omega^2 = .013$, and $F(2,360) = 3.48$, $p<.05$, $\omega^2 = .01$ respectively. Here, increased Corsi performance influenced primarily the number of regressions to and the total regression time for near targets. The corresponding effect pattern for length of initial regression, number of regression and total regression times is shown in Figure 6. Together, the examination of reading skill and of spatial working memory indicated that higher comprehension and faster reading were associated with more effective searches for the target, which often ensued after an initial regression toward the target had been executed. Conversely, spatial working memory influenced the initial regression toward the sentence target. Similar to effects of reading skill, participants' dynamic spatial working memory also influenced regression number and total search time, but only for near targets.

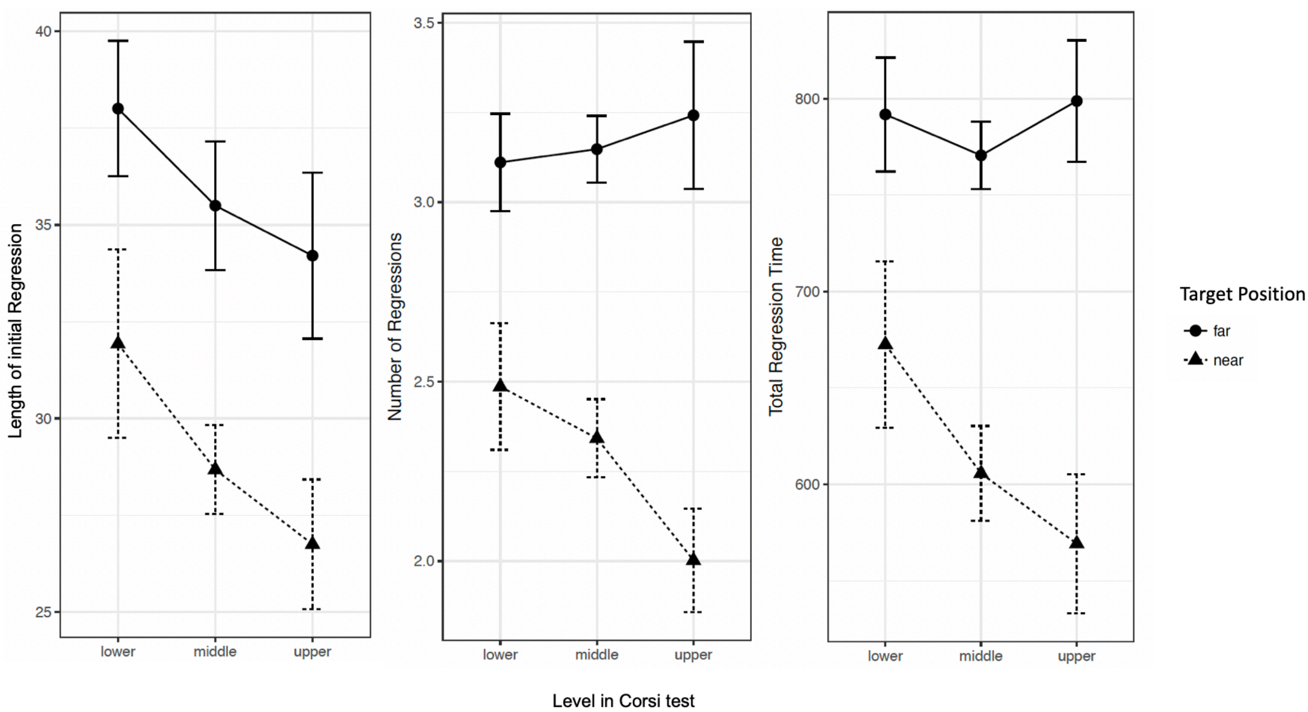


Figure 6. Initial regression amplitude, number of regressions, and total regression time as a function of performance level in the Corsi test measuring dynamic visuo-spatial memory.

Regression Strategies

While the three sets of oculomotor regression data provide quantitative indexes for the reaching of the target, they do not reveal how the target was reached and how they responded when the initial regression did not move the eyes onto the target. Our classification of readers' search for the target showed that the target's position had a profound influence on participants' approach to the target. With near targets, 22 participants used primarily single shot regressions to move the eyes directly onto the target's location, 8 participants preferred goal directed regressions, and 15 participants preferred centered searches. None of the participants preferred a backward search for near targets and just one preferred forward searches. For far targets, none of the participants preferred single shot regressions, 11 preferred goal directed regressions, 22 preferred centered searches, 13 preferred backward searches, and no participant preferred the forward search strategy.

To examine potential influences of participants' reading skill and spatial working memory on target search, Chi^2 tests compared the preferred target search strategy for near and far targets for individuals who scored either in the upper or lower quartile of the four test measures. For near targets, the analyses revealed a robust effect of reading comprehension, $X^2(2) = 7.41$, $p<.05$. Participants with an upper level performance were more likely to use a single shot strategy than participants with a lower level score (9 vs. 3, respectively); participants with a lower level comprehension preferred, by contrast, goal directed regressions (6 vs. 1, respectively). Performance on the Corsi test was also linked to the preferred use of regression strategies, $X^2(2) = 5.05$, $p<.09$. Again, more upper- than lower level performers preferred single shot regressions (8 vs. 3, respectively) for near targets. With this test-based classification, there was no group specific distinction for goal directed searches, but more lower than upper level participants preferred centered searches (7 vs. 2, respectively), see figure 7.

The preferred search for far targets was associated with only one measure, i.e., participants' Corsi performance, $X^2(2) = 6.77$, $p<.05$. More individuals with an upper- than a lower level score used backward searches (6 vs. 1, respectively), whereas fewer individuals with an upper level score used centered searches (3 vs. 9), see figure 8. The backward strategy typically started with an initial regression toward the location that a near target would have had. Had an alternative grammatical sentence construction been used, then this would have been the target's actual location.

Collectively, these analyses show that participants' search for near targets was influenced by their text comprehension and Corsi test performance, as higher performance levels were associated with the preferred use of single shot regressions for near targets. Conversely, relatively low comprehension scores were associated with more than one saccade to near targets. The search for far targets was influenced only by Corsi test performance, which was linked to the preferred use of backward search.

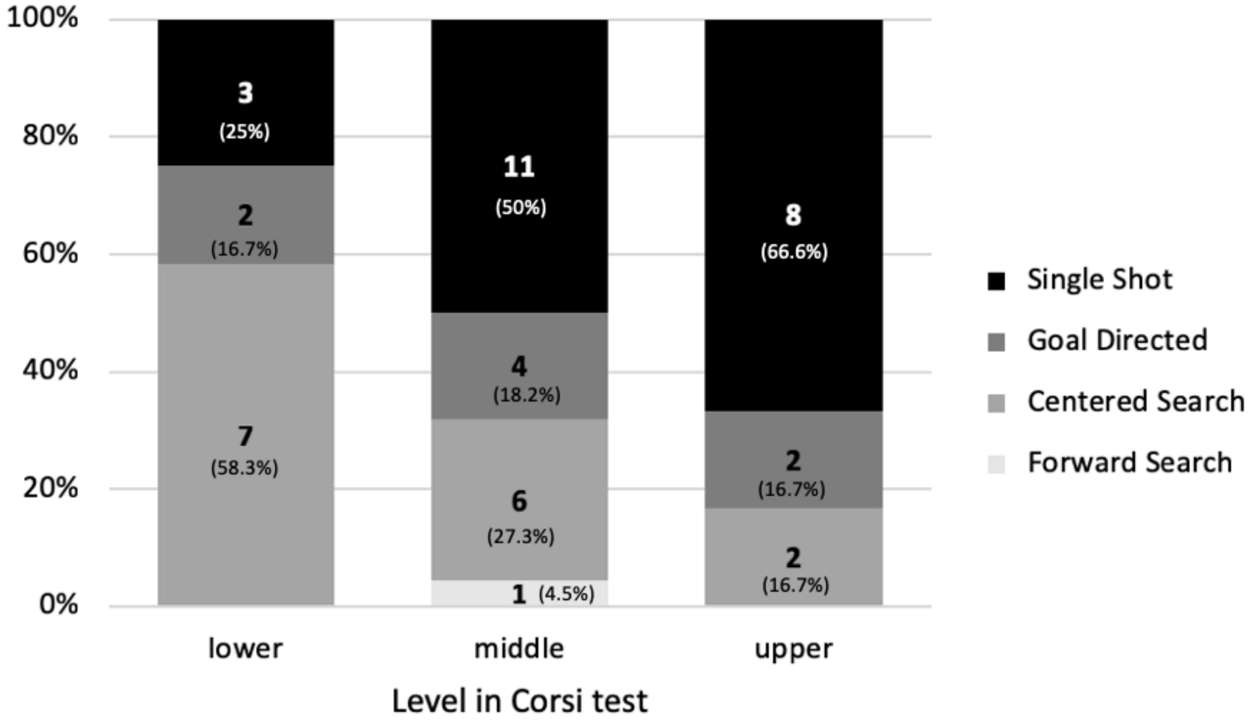


Figure 7. Proportion of preferred strategies (both number of participants and percent) by level of performance in Corsi block tapping test for near targets.

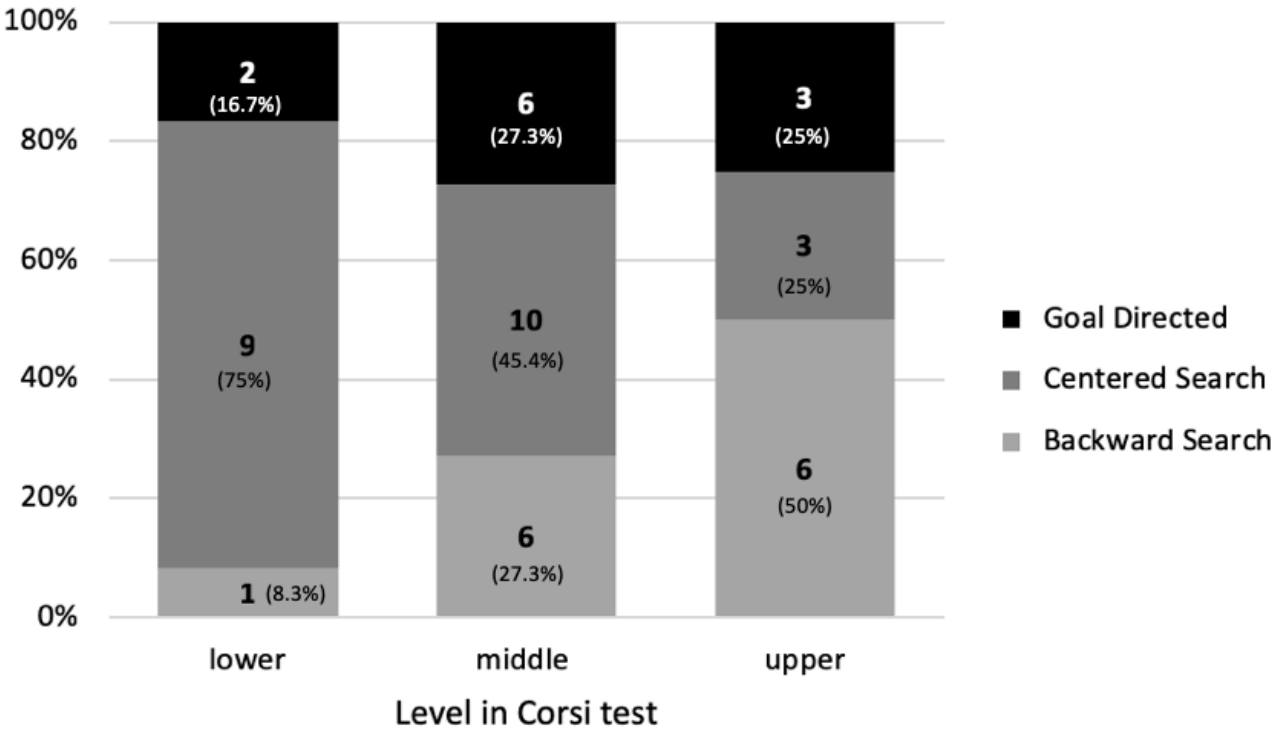


Figure 8. Proportion of preferred strategies (both number of participants and percent) by level of performance in Corsi block tapping test for far targets.

2.4 Discussion

The main goal of the present experiment was to determine whether individual performance of visuo-spatial memory and/or reading skill influences the nature of long-range regressions and whether this is reflected in the occurrence of qualitatively different regression strategies. Long-range regressions in the present experiment were not always precise and did not reach the target directly in somewhat more than 50 percent of the observations. In these cases, one regression is not enough to find a target and approximately 2-3 correcting saccades are needed to home in on the target word. These data are interpreted as evidence against the original spatial coding hypothesis, since regressions in our experiment were not always precise irrespective of target positions similar to those described by Kennedy et al. (2003).

Despite not being precise, regressions are spatially selective, being substantially longer for far as compared to near targets. Overall, target position had a strong influence on all variables. Far targets lead to longer first regressions, but also to more regressions to correct an inaccurate landing position and to a prolonged search time. These results are in line with prior experiments that also found that regressions are selective, but not precise and that the position of the target influences regression accuracy (Inhoff & Weger, 2005; Weger & Inhoff, 2007; Guérard et al., 2012). The fact that long-range regressions adapt to the distance of the target indicates that retrieval of information from memory takes place and that the retrieved information is utilized to enable re-reading.

The present experiment used a saccade-contingent line-shift manipulation to explore a possible involvement of peripheral visual processing in the programming of regressive saccades (Rayner & Pollatsek, 1981). This technique was previously employed by Kennedy, Brooks, Flynn, & Prophet (2003) and by Inhoff, Weger & Radach (2005). These experiments found that a shifting of the current line of text during a saccade at the end of the line (immediately before the planning of the ensuing regression) did not affect primary regression amplitudes. Experiment 1 replicated these findings, confirming that memory and not peripheral visual processing is the primary source of information used for the programming of long-range regressions.

Accuracy of regressions

Performance in the Corsi and matrices tasks substantially affected the amplitude and accuracy of initial regressive saccades. The amplitude of the initial regression decreased with better visuo-spatial working memory performance. Additionally, in near target words, an increased performance in the Corsi task is associated with fewer regressions needed to reach the target and less time until the target is found. This effect disappears with far targets. In contrast, reading comprehension and speed had no influence on the length of the initial regression. However, both measures affected the number of regressions and the time it took to regress back to the target word, irrespective of target position. This advantage for better readers in terms of more efficient saccade correction suggests that more skilled readers can readily utilize linguistic knowledge and/or local linguistic information to correct an incorrect landing position of initial regressions.

On the one hand, such an account is in harmony with the verbal reconstruction hypothesis, that assumes that accessible linguistic representations can be used to reconstruct the location of prior words in the text (Weger & Inhoff, 2007; Guérard et al., 2012; Guérard et al., 2014). On the other hand, our results suggest, that only spatial memory has a direct influence on the programming of initial regressive saccades. One possible interpretation of this pattern could be that spatial memory is the source of target location information for regression programming irrespective of target location. In near targets this is reflected through fewer regressions and less time used to reach the target word. With far targets, good reading comprehension and reading speed could have compensated for poor spatial memory information, explaining the disappearance of the spatial memory effect for far targets.

This line of reasoning contradicts Guerard et al´s (2012) assumption that verbal memory is the dominant source for regressions to words that have been read a longer time ago, but goes hand in hand with Weger and Inhoff's (2007) conclusion that "initial regressions are primarily guided by spatial memory and rely little on linguistic knowledge, whereas corrective regressions are guided to a much larger extent by linguistic knowledge" (*p.* 1304). However, there is a way to reconcile the present results with those found by Guérard et al. (2014). These authors used a verbal interference task (repeating irrelevant letters) during reading, known as articulatory suppression. They found that the initial regression accuracy was significantly impaired by articulatory suppression and suggested that verbal memory played an important role during regressions. Spatial memory could have an indirect influence on regression programming in this case, considering, that the central executive (which is limited [see Baddeley, 2010]) is continuously overloaded by the articulatory suppression task, which makes it impossible to simultaneously storage spatial information into visuo-spatial memory. Through this indirect mechanism, spatial memory could be the origin of the impaired regression accuracy even though the articulatory suppression task originally targets verbal memory.

Regression strategies

In their seminal analyses of long-range regressions Murray and Kennedy (1988) classified their observations into two categories, a single shot and a backtracking strategy. The idea behind this distinction was that when participants had access to sufficient information, the target word could be attained with a single regression saccade, while the absence of such information necessitates a laborious search process. Careful inspection of the present data indicated that this simple dichotomy may not provide a sufficient description of re-reading variability in natural continuous reading. Indeed, additional regression strategies with a high probability of occurrence were found.

The centered search strategy, starting with a regression landing in the center region of the line, followed by correcting saccades either to the beginning or the end of the sentence, was used mainly by lower level performers in Corsi task for both, near and far targets. This suggests that the centered search strategy may be used when participants have no real idea where the target is located. This is a very interesting finding, as had been expected that poor spatial memory would be associated with backtracking as suggested by Kennedy et al. (1988; 2003), but with better performance in Corsi test, the proportion of centered search was reduced and backtracking became the dominant strategy if the target could not be reached with accurate regressions.

The backward search strategy occurred quite frequently for far targets and was used by readers with advanced dynamic spatial memory performance (upper performer in Corsi test), assuming, that backtracking is often utilized when readers expect an incorrect location of the target, e.g. a position at the end of the sentence, when the target is actually at the beginning. Far targets are in general more error-prone. Readers with better spatial memory tend to make mistakes in the form of sending the initial regression to the end of the sentence assuming a near target instead of a far target. The following process of correction is laborious, as a substantial deviation has to be compensated, often with a large number of extra saccades. Consequently, more intelligent errors made by readers with better spatial memory can lead to a paradoxical advantage for less able readers who tend to use the more economical centered search strategy.

The present regression strategy analysis revealed that re-reading with the aim of attaining a specific word occurs in more than one way. In accordance with Kennedy and Murray (1988) populations of single shot regressions and backward scanning were found. Unlike Frazier and Rayner (1982) the strategy of forward scanning was observed relatively rare, although occurring in a few observations. The phenomenon of centered search has to my knowledge not been described before. This can be seen as a major innovative contribution of the present work, especially in conjunction with the functional relation to spatial memory described above. Murray and Kennedy (1988) linked the use of backward scanning to children with poor reading abilities, whereas the results of experiment 1 did not show a correlation between the use of backward scanning and reading competence but instead spatial memory capacities. This could be interpreted as another hint for a possible indirect mechanism of spatial memory. It seems likely that readers with deficits or decoding and/or comprehension need to invest more cognitive resources into linguistic moment-to-moment processing or the word and/or sentence level (e.g. Vorstius, Radach, Mayer, & Lonigan, 2013). As a result, not enough cognitive capacity could be available to store spatial information about word positions in memory and regression targeting becomes invalid. The precise interplay of an indirect mechanism of spatial memory and reading difficulties will be an interesting topic for experiment 2 and experiment 3. If the accuracy of long-distance regressions would be reconstructed from linguistic knowledge, more single shot regressions in participants with superior comprehension skills, especially with far targets should have occurred, which was clearly not the case. The conclusion from the results are that word location is primarily not reconstructed from temporal order and item memory as suggested, among others, by Fischer (1999), at least not in continuous reading. In this context, it is interesting to note that the proportion of single shots is significantly higher for near than for far targets irrespective of performance in any of our tests. This is consistent with other authors (Inhoff & Weger, 2005; Weger & Inhoff, 2007; Guérard et al., 2012) and compatible with the assumption that the number of words between the target word and the starting point of the regression might/can influence regression accuracy.

3 Experiment 2: Long-range regressions in beginning readers

3.1 Introduction

In contrast to the extensive empirical evidence on eye movements during reading in adults, studies describing the development of eye movement behavior in reading acquisition are less frequent. On a fairly general level, developing reading is characterized by a higher number of fixations with longer fixation durations, resulting in longer gaze durations and longer reading times per word. The increased number of fixations is necessarily associated with shorter saccades (Taylor, Frackenpohl, & Petee, 1960; McConkie, Zola, Grimes, Kerr, Bryant, & Wolff, 1991; Blythe, Liversedge, Joseph, White, & Rayner, 2009; Joseph, Liversedge, Blythe, White, & Rayner, 2009; Huestegge, Radach, Corbic, & Huestegge, 2009; Blythe, Häikö, Bertam, Liversedge, & Hyönä, 2011).

Early work on reading development indicated that the number and the duration of fixations decreases with increasing reading skills over the first six years of school (Buswell, 1992; Taylor, Frackenpohl, & Petee, 1960). Recent experiments elaborate on these results, focusing on the dynamics of word-level fixation patterns. Results indicate, that number of fixations, fixation durations and word viewing times decrease and saccade length increases with reading development and level of reading (McConkie et al., 1991; Blythe, Liversedge, Joseph, White, Findlay, & Rayner, 2006; Blythe et al., 2009; Joseph et al., 2009; Huestegge et al., 2009; Blythe, Häikö, Bertram, Liversedge, & Hyönä, 2010; Vorstius, Radach, Mayer, & Lonigan, 2013; Vorstius, Radach, & Lonigan, 2014). According to Blythe & Joseph (2011), eye movement behavior of children in England reaches the level of adult, skilled readers at about the age of eleven.

Children have already developed a typical asymmetrical perceptual span when they are at the second grade of school (Rayner, 1986; Häikiö, Bertram, Hyönä, & Niemi, 2009). However, the size of their span is significantly smaller than the size of the perceptual span in adult readers. The use of parafoveal information increases with reading performance (Marx, Hutzler, Schuster, & Hawelka, 2016), reducing the effort of word identification when a word is actually fixated (Häikiö, Betram, & Hyönä, 2010; Häikiö et al., 2009). On average, the length of the perceptual span reaches the level of adults in the sixth grade of school, when children are approximately at the age of eleven (Häikiö et al., 2009).

Work on developing readers has shown that children make many more regressions compared to skilled adult readers (Blythe, Liversedge, Joseph, White, & Rayner, 2009; Blythe, Häikiö, Bertam, Liversedge, & Hyönä, 2011; Buswell, 1922; Joseph, Liversedge, Blythe, White, & Rayner, 2009; McConkie, Zola, Grimes, Kerr, Bryant, & Wolff, 1991; Rayner, 1986). Unfortunately, only a few of these developmental studies distinguished between intra- and inter-word regressions. McConkie et al. (1991) compared U.S. developmental data from first to fifth grade and reported that the number of regressions within a word (intra-word regressions) decreases with age, while inter-word regressions actually become more frequent with increasing reading ability. Vorstius, Radach, and Lonigan (2014) also reported an increase of inter-word regressions in children from first to fifth grade in their U.S. sample. This highlights the importance of long-range regressions (reaching at least one word to the left of the currently fixated word [N-1]) for more skilled readers, who need less refixations, have shorter first pass reading times, and are able to selectively go back to words causing comprehension problems if needed (Vorstius, Radach, Mayer, & Lonigan, 2013; Vorstius, Radach, & Lonigan, 2014).

Spatial memory was the first candidate to be considered as the functional basis for re-inspections during reading (see Kennedy, Brooks, Flynn, & Prophet, 2003 for an overview), based on experiments showing very well targeted long-range regressions (Kennedy & Murray, 1987; Murray & Kennedy, 1988; Kennedy et al., 2003), and resulting in the *spatial coding hypothesis*. According to this view the printed page or the display screen could act as an external memory, providing reference for spatial orientation.

However, the spatial coding hypothesis as the sole explanation for regression programming has difficulties to account for effects of spatial and linguistic distance on regression accuracy. Regressions to near targets are usually more accurate than regressions to far targets. Increasing linguistic distance (number of words between starting point of the regression and target word) can lead to a higher number of corrective saccades, when one single regression alone is not sufficient for finding the target word (Inhoff & Weger, 2005; Weger & Inhoff, 2007; Guérard, Saint-Aubin & Maltais, 2013; Guérard, Saint-Aubin, Maltais, & Lavoie, 2014). Such findings suggest that information needed for the programming of long-range regressions is not exclusively provided by spatial memory but that further influencing mechanisms are at work. According to the *verbal reconstruction hypothesis,* the position of a word may be ascertained from information out of verbal memory and the knowledge about linguistic positions in sentences (Rawson & Miyake, 2002; Weger & Inhoff, 2007; Guérard, Saint-Aubin, & Maltais, 2013; Guérard, Saint-Aubin, Maltais, & Lavoie, 2014).

Experiment 1 examined the contribution of visuospatial memory and reading performance on regression accuracy in adult readers (see chapter 2). Single-line sentences with targets either at the beginning or at the end of a sentence were presented. Participants had to check target words for spelling errors (inserted after reading the sentence) to evoke long-range regressions to near and far positions relative to the location from which the regression was launched. Individual performances in visuospatial memory were examined using Corsi block tapping task (Corsi) and matrices task (matrices), while literacy was assessed with a test capturing reading accuracy and fluency.

It was found that regressions were spatially selective, being larger for far than for near targets, and target position influenced regression accuracy, with far regression targets being more error-prone than near regression targets. Spatial memory performance had a substantial effect on regression programming, with high level performer in Corsi showing the largest proportion of accurate regressions. Reading performance was only crucial for correcting processes, necessary when the first regression did not reach the intended target. When the target was not attained with single shot or goal directed regressions, readers executed alternative strategies of re-reading. The *centered search strategy* started with an initial regression to the center of the sentence, followed by regressions either to the beginning or the end of the sentence, depending on the position of the target. This strategy was mostly used by participants with low performance in Corsi, suggesting that this strategy is used when spatial memory provides no useful information about the position of the target. The *backward scanning strategy* started with a regression to the end of the current sentence, followed by backtracking regressions until the target was found. This strategy was mostly used for far targets by participants with high performance in Corsi, indicating that this strategy is used when, despite good spatial memory, an incorrect target position is assumed.

The ability to program regressions in children (9-11 years), and the contribution of reading abilities on regression accuracy, was first investigated by Murray and Kennedy (1988). They found that good readers were perfectly capable to program long-range regressions to near and far targets in a previously read sentence. Poor readers could usually not program single shot regressions and had to search for the target with more and shorter regressions. They also compared the results of both good and poor readers with a group of younger readers who had the same *absolute* reading ability (raw score in NFER Reading Test BD) as the poor readers and the same *relative* reading ability (age based standard score in NFER Reading Test BD) as the good readers. Regression patterns in the younger group showed considerably more similarities to the good than to the poor readers, as the younger group was better at programming accurate long-range regressions compared to poor readers. This led to the conclusion that strategies of re-inspecting words with regressions not only differ because of a difference in the absolute reading ability, but may indicate a deficit in poor readers with respect to finding required information on a page. Therefore, the ability to program accurate long-range regressions appears to play an important role in the development of skilled reading performance (Kennedy & Murray, 1987; Murray & Kennedy, 1988).

In a recent experiment, published by Tiffin-Richards and Schroeder (2018), the authors found, that although adults show fewer regressions than children in general, the proportion of regressions initiated at sentence boundaries was greater in adults than in children, “suggesting that the use of sentence endings to initiate regression increases with reading experience and proficiency” (Tiffin-Richards & Schroeder, 2018, p. 1059) and underlines the importance of long-range regressions for skilled reading.

Taken together, regressions play a crucial role for fluent reading, support reading comprehension, and are of vital importance for reading development. The type of contribution of visuospatial memory on that ability still remains unclear, but there is ample evidence for an influence of visuospatial memory on the programming of long-range regressions, suggesting a direct influence of visuospatial memory on reading development.

The goal of the present experiment is to further advance our understanding of the importance of visuospatial memory for the programming of long-range regressions in children and whether and how this ability contributes to the development of good reading skills. German students in 4th grade of school were examined, with students at an age of approximately 10 years. The same sentence paradigm as reported for experiment 1 was used, with content designed to be also appropriate for children in 4th grade. This includes one-line sentences with near and far regression target words, indicated by a probe word at the end of the sentence, to examine how precise children are able to program long-range regressions. Reading of sentences was accompanied by comprehension questions, to make sure that participants read for understanding as their main task. In addition, children´s static and dynamic spatial working memory as well as reading performance were assessed.

The present experiment addresses two main research questions. First, what is the accuracy of regressions in developing readers and can spatial memory be verified as the major source of regression programming in children, as found for adult readers in experiment 1? Regressions being longer for far than for near targets would represent spatial selective regression programming and would support the hypothesis, that memory processes influence regression programming already at an early stage in reading. Spatial memory as the source of programming information for regressions should be evident in an influence of individual performance in spatial memory tests on the length of the initial regression and accuracy of regression programming, especially for near targets. Assuming that reading performance is the source of effective correcting processes, a high level of reading comprehension should lead to a reduced number of regressions and less time to attain target words.

Experiment 1 provides the opportunity to compare results of children with those obtained with adults. A larger influence of target distance and the length of the initial regression saccade in adults would suggest that the accuracy of regressions increases with age and reading proficiency. Better performance with regard to number of regressions and total regression time in adults would lead to the conclusion that increased reading skill helps to correct a misplaced first regression and can therefore serve as a secondary source for efficient regression strategies.

Second, the use of individual regression strategies was examined, focusing on whether children at the age of 9-11 years are already comparable to strategies known from adults (experiment 1). In addition, the question arises whether one strategy is dominant over another in the development of reading or if different regression strategies can develop together in parallel and how children deal with regressions not attaining the targeted word. Consequently, regression strategies of children with high and low performance in both, reading ability and spatial memory was conducted to see, whether high performance in one of the tests leads to a higher proportion of single shot regressions and, in turn, low performance reflects certain other strategies, as suggested by Murray and Kennedy (1988) and as evident in experiment 1.

3.2 Methodology

Participants

A total of 78 (38 female) 4th grade students from three different primary schools in Duisburg (North Rhine-Westphalia, Germany) took part in this experiment. The average age was 9;11 years (max = 11;5, min = 9;2). All participants were native speakers of German, although 23 (29,5 percent) of them grew up bilingually. All participants had normal or corrected-to-normal vision. Parents or legal guardians were informed about the purpose of the experiment and gave written consent. Child assent was also collected before the testing session started. All children participated unpaid, but were given candy at the end of the session to show appreciation.

Materials

Reading task

The same materials as described in detail in experiment 1 was used for experiment 2. The materials consisted of 100 declarative sentences including topics suitable both for children and for adults. Each sentence included one target word, either at the beginning (far target) or at the end of the sentences (near target, in reference to the starting point of the regression). Target words were in all cases nouns with a length of 6 or 7 characters. They were preceded by an adjective of 5-8 characters length and were either part of the subject or the adverbial complement of the sentence. All target words were of medium frequency with a mean frequency of 29 counts per million (M = 29.21, dlex DB, Heister et al., 2011). In total, sentences were 67-70 characters long and consisted of 9-14 words. The special structure of German grammar was used to swap parts of the sentences without changing the meaning of the sentences to present each target once at a location near to the probe location (sentence end) and once as a target far to the probe region (sentence beginning) (see experiment 1 for details). Thirty-two sentences were used as fillers with targets around the center of the sentences to avoid any training effects for a certain position.

Psychometric Tests

All participants were examined with a test of reading performance, the Salzburger Reading Screening for first to fourth graders (Salzburger Lese-Screening für die Klassenstufen 1-4, SLS 1-4, Mayringer & Wimmer, 2012) and a test of visuo-spatial working memory, taken from the working memory test battery for children at the age of 5-12 years (Arbeitsgedächtnistestbatterie für Kinder von 5-12 Jahren, AGTB 5-12, Hasselhorn et al., 2012)

Reading Test

The Salzburger Reading Screening (SLS 1-4; Mayringer & Wimmer, 2012) is a common assessment of reading fluency for German readers, combining speed with a measure of accuracy. Within a three minute time period, children have to read as many sentences as possible, to decide whether the content of the sentence is right or wrong, and to circle the tick (if the sentence is right) or the cross (if the sentence is wrong) at the end of the sentence. They are instructed to do this as fast as possible in three minutes. The SLS provides norm data for boys and girls separately as well as regardless of gender. There are 4 parallel test forms.

Working memory test battery

The working memory test battery (AGTB 5-12; Hasselhorn et al. 2012) is normed for children aged five to twelve and is suitable for measuring the functions of working memory in three areas: central executive, phonological memory, and visuo-spatial memory (VSM). As the focus of the experiment lies mainly on the influences of the visuo-spatial memory, only the subtests *matrices task* and *corsi block tapping task* were used. Both tests were performed consecutively on a touch screen.

The visual part of the VSM is assessed using the matrices task. A 4X4 grid of white boxes is initially presented. This is followed by a number of boxes shown in black for a predetermined time, starting with two boxes for 2400ms. As a consequence of correct responses, the number of highlighted boxes and presentation duration increases. In case of an incorrect response, the number of black boxes is maintained and decreased in case of further errors.

The spatial part of the VSM is captured by the Corsi block-tapping task. Nine spatially separated white boxes are initially presented on a background of gray boxes. A sequence of smileys is subsequently shown one by one in the gray boxes, starting with only 2 smileys. Participants are asked to reconstruct the sequence and tap the gray boxes in the order of the previously seen smileys. Again, the number of consecutively presented smileys in a trial increases if the previous sequence is reported correctly, but maintained and eventually reduced in case of incorrect responses.

TNO Stereotest

The TNO is designed as a screening test for deficits in binocular vision. It is appropriate for children and consists of seven plates with figures which must be viewed with red/green spectacles. The first three plates allow for screening of stereoscopic vision in general. The remaining three plates are used to examine stereoacuity with binocular disparities between 480 and 15 seconds of arc. All children had good stereoscopic vision.

General Procedure

Reading tests were administered in the classroom and took approximately 15 minutes to complete. For eye movement recordings, each child was picked up individually by the experimenter from the classroom and guided to the testing room in the school. Children were seated in front of the screen and could place their chin on a chin-rest and their forehead against a forehead-rest to avoid head movements. Instructions were provided orally in child-appropriate language, via pre-recorded texts played back to the participants and supplemented by the experimenter, if necessary. Children were instructed to read the sentences silently and for full understanding. The first four sentences were practiced together with the experimenter to make sure that each child completely understood the task. Sentences were presented individually on a screen. At the end of each sentence a masked probe word became visible after crossing an invisible boundary at the sentence end (see experiment 1, for a more detailed description). Participants were instructed to go back to the identical word in the current sentence and to decide, whether this target word was correct or had a misspelling in it. Children had knowledge that the misspellings in target words appeared after they had read the sentence for the first time. Spelling errors were inserted to make the task meaningful for the children.

Randomly and on average after six sentences a question related to the content of the sentence had to be answered. All questions could be answered with yes or no and by pressing one of two assigned buttons. After half of the sentences a 5-minute break was taken. When all eye movement measurements had been completed, the AGTB 5-12 was administered, which took about 10 minutes. In total, a test session lasted about 45 -50 min. All children were guided back to the classroom after the session. Results of reading and memory tests were reported in written form to parents, if requested, and to the head- and class teachers if parents had given consent.

Apparatus

Eye movements were recorded using an EyeLink 1000 video-based eye tracker (SR-Research, Toronto, Canada) with a sampling rate of 2000Hz. Viewing was binocular, but only the movements of the right eye were recorded. Single line sentences were presented in black on a light-gray background on a 21-inch monitor with a resolution of 1680 x 1050 and a refresh rate of 120Hz. Type font was Courier New and font size was 15 pt. At a viewing distance of 68 cm for each participant, letters had an extent of 0.3° visual angle. At the beginning of the experiment and after each question a three-point calibration, followed by validation, ensured a spatial resolution of less than 0.3° visual angle.

Data selection

Raw eye movement data were analyzed using the Eye Movement Data Analyzer (EyeMAP, Tang, Reilly, Vorstius, 2012) as well as SPSS Statistics 24. All eye movement data were inspected visually to detect any possible problems and practice trials were excluded from the data. One participant had to be excluded because he/she could not finish working memory tests due to personal reasons. Two participants were excluded due to major calibration problems and another two participants, because they read only less than half of the sentences due to very low reading skills. The remaining sample consisted of 75 participants.

As means of analyses, three variables for regression accuracy were determined. Namely, (1) the length of the initial regression, (2) the total number of saccades needed to detect the target word, as well as (3) the time it took until the target was found. The ability to adjust the length of the initial regression to the distance of the target reflects the availability of information about the target location in memory. In addition, the number of saccades and time until attaining the target reflect the effort necessary for correcting inaccurate or completely misguided long distance regressions.

Participants' performance on the standardized assessments (reading ability and spatial working memory) was used to define three levels of performance, referred to as lower, center, and upper, for each of the four tests. Participants whose performance fell into the lowest quartile (lower 25 percent) formed the lower level, the interquartile range (the two center quartiles) comprised the middle level, and the upper level comprised the upper quartile (the upper 25 percent). Table 3 provides the number of participants and percentile ranking of participants in each group.

Table 3. Participants and percentile ranking by level of test performance for experiment 2.

Level of performance	Reading comprehension		Matrices test		Corsi test	
	N	LQ	N	PR	N	PR
lower	16	66-95	17	4-27	16	4-12
middle	36	94-133	34	14-76	38	12-79
upper	21	115-145	22	66-97	19	62-95

N= number of participants, PR = percentile ranking, LQ = reading quotient, SLS only provides a reading quotient: values are comparable to intelligence quotient; overlaps in percentile ranking are due to different age levels.

Data analysis

Analysis of data from the present experiment closely followed the procedure of experiment 1. Trials with missing data were excluded (2.6 percent) and trials with initial regressions of less than 5 LS, regression frequencies of > 14 and total regression time of > 4000 were removed, resulting in the exclusion of 2.1 percent of the data. As in experiment 1, length of initial regressions was log-transformed to better fit a normal distribution, but number of regressions and regression time were not transformed as their distribution were close to normal.

All effects on regression accuracy were analyzed using ANOVA models using the R environment for statistical computing (R Core Team 2014; version 1.0.136). In a first step the three variables of regression accuracy were analyzed separately and nonsignificant fixed factors were removed from the model. In a second step, the assessment data from spatial memory and reading tasks was entered into simplified ANOVA with only position as fixed factor and individual level of performance in each test. (see experiment 1 for a detailed description).

The classification of regressions follows the procedure described for experiment 1. (see chapter 2.2). Every regression that immediately attains the target counts as *single shot* regression. *Goal directed regressions* are regressions with a slight under- (at least 50 percent of the distance) or overshoot (at most 150 percent of the distance) followed by only one corrective saccade. *Backward searches* describe a searching process starting in the last third of the sentences and needing at least two corrective saccades to attain the target word. *Centered searches* describe a pattern where the first regression is targeted towards the center of the sentences (middle third), followed by at least two corrective saccades. Finally, *Forward searches* start in the first third of the sentences and also include at least two corrective saccades.

During visual inspection, no regression strategies differing from those identified in the adult sample from experiment 1 were found. All trials that could not be assigned to any of the patterns were counted as unclassifiable.

As in experiment 1, the preferred search strategy for every participant was determined in association with level of performance in each of the tests. Chi2 tests were then used to determine an association between preferred target search strategy and level of test performance.

Differences in regression strategies between children and adults were compared using a one-way repeated measures ANOVA and independent samples t-tests, because sample sizes are not identical between experiment 1 and experiment 2.

3.3 Results

Regression Accuracy

Target position had a strong influence on all three dependent variables. The length of the first regression was significantly longer for far than for near targets, $F(1, 525) = 19.038$, $p<.001$, $\omega^2 = .102$, and number of regressions and total regression time increased with distance of the target. When targets are far away, more regressions are needed, $F(1, 525) = 77.184$, $p<.001$, $\omega^2 = .184$, and regression time is longer, $F(1, 525) = 68.062$, $p<.001$, $\omega^2 = .144$. The effect of shift was negligibly for all variables with *p*s>.3 and no interaction between shift condition and target position could be found (all *p*s>.4). Target order had a significant effect on the length of the initial regression, $F(1, 525) = 5.138$, $p<.05$, $\omega^2 = .007$, with initial regressions being shorter for targets occurring for the second time (but in another position within the sentence). The interaction between target position and target order reached significance for number of regressions, $F(1, 525) = 4.293$, $p<.05$, $\omega^2 = .004$, and showed a marginal effect for total regression time, $F(1, 525) = 3.762$, $p=.06$, $\omega^2 = .004$. Second occurrence of the target in a near position leads to fewer additional regressions and shorter regression times. Table 4 provides data of all regression measures as function of target position, target shift and target order.

Table 4. Means and standard errors for amplitude of initial regressions (in letter spaces), total number of regressions and total regression time (in ms) as a function of target position, shift condition and presentation order.

		near				far			
		no shift		shift		no shift		shift	
		first	rep	first	rep	first	rep	first	rep
Initial Regression	M	**27.72**	**27.79**	**28.13**	**27.75**	**36.32**	**33.25**	**34.19**	**32.82**
	SE	0.96	0.95	1	1.14	0.98	1.05	0.92	1.05
Number of Regressions	M	**2.97**	**2.84**	**3.12**	**2.79**	**4.35**	**4.34**	**4.32**	**4.45**
	SE	0.12	0.12	0.14	0.13	0.15	0.15	0.14	0.15
Regression Time	M	**942.8**	**913.4**	**987.2**	**915**	**1296.5**	**1273.3**	**1286.3**	**1320.9**
	SE	32.5	30.6	38.5	39.2	40.3	39.1	38.7	48.2

M = mean, SE = standard error, first = first time of presentation, rep = second time of presentation

The factors target position and participants` reading performance (SLS) and dynamic (Corsi test) and static (matrices test) spatial memory were included individually to ANOVAs for each of the dependent variables, length of initial regression, number of regressions and total regression time, respectively (see table 5 for means and standard errors of each variable).

Table 5. Means and standard errors (in parenthesis) for amplitude of initial regressions (in letter spaces), total number of regressions and total regression time (in ms) as a function of level of performance in reading and static (matrices) and dynamic (Corsi) spatial memory.

		Length of initial Regression		Number of Regressions		Regression Time	
		near	far	near	far	near	far
Reading	lower	26.3 (1.3)	34.9 (1.6)	3.1 (0.2)	4.7 (0.2)	1052 (53)	1432 (46)
	middle	27.5 (1.3)	34.2 (1.2)	3 (0.1)	4.5 (0.2)	947 (41)	1347 (57)
	upper	29.6 (1.6)	33.8 (1.4)	2.9 (0.2)	4 (0.2)	896 (41)	1158 (41)
Matrices test	lower	32.1 (1.9)	38.3 (1.7)	3.6 (0.2)	4.1 (0.2)	1137 (62)	1266 (62)
	middle	27 (1)	34 (1)	2.9 (0.1)	4.5 (0.2)	939 (28)	1357 (47)
	upper	24.8 (1.2)	30.7 (1.5)	2.6 (0.1)	4.5 (0.2)	818 (37)	1286 (68)
Corsi test	lower	27.5 (1.2)	35.8 (1.2)	3.2 (0.2)	4.4 (0.3)	1012 (46)	1341 (83)
	middle	28.9 (1.4)	34.6 (1.3)	3.1 (0.1)	4.4 (0.2)	989 (43)	1300 (46)
	upper	25.7 (1.3)	32.4 (1.1)	2.6 (0.1)	4.5 (0.2)	861 (34)	1327 (50)

Performance in reading had no effect on the length of the initial regressions ($p>.01$), but significantly influenced the number of regressions, $F(2, 533) = 4.636$, $p<.05$, $\omega^2 = .009$, and the total regression time, $F(2, 533) = 14.882$, $p<.001$, $\omega^2 = .037$. The number of regressions and the total regression time are associated with correcting processes following an over- or undershoot of initial regressions. The better the reading performance, the more effective the correction of initial regressions, represented by fewer additional regressions and less time to find the target. An additional interaction effect between target position and number of regressions, $F(2, 533) = 4.474$, $p<.05$, $\omega^2 = .009$, and total regression time, $F(2, 533)=3.084$, $p<.05$, $\omega^2 = .006$, emphasized the importance of reading comprehension for the correction of wrong landing positions of regressions aimed to far targets (see figure 9).

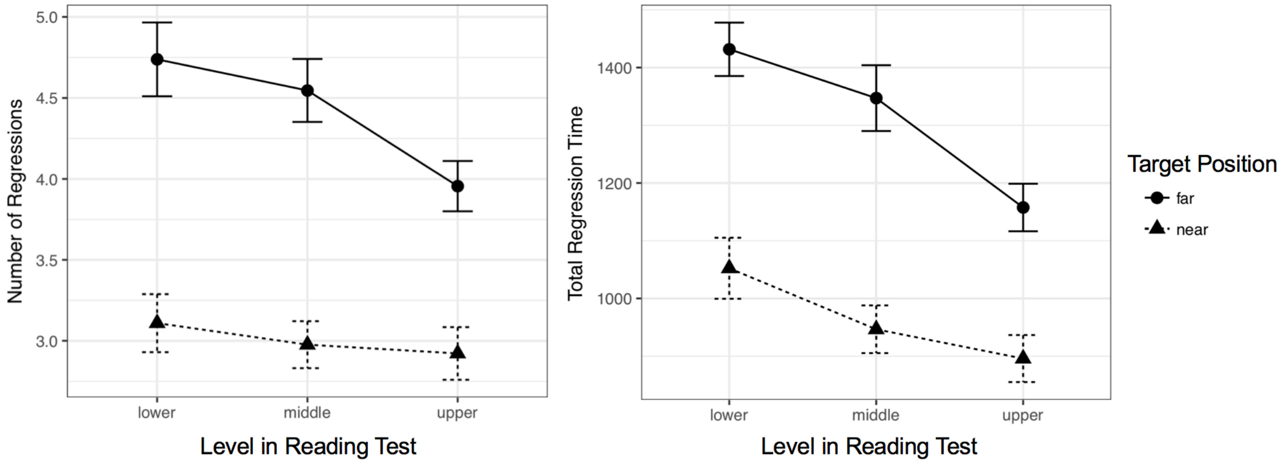


Figure 9. Mean number of regressions and total regression time as a function of target position (near and far) and level of performance (lower, middle, upper) in the reading test.

Performance in both spatial memory tests influenced the length of the initial regression. A better performance in Corsi, $F(2, 533) = 4.299$, $p<.05$, $\omega^2 = .01$, and matrices test, $F(2, 533) = 25.874$, $p<.001$, $\omega^2 = .073$, was associated with shorter initial regressions (see figure 10).

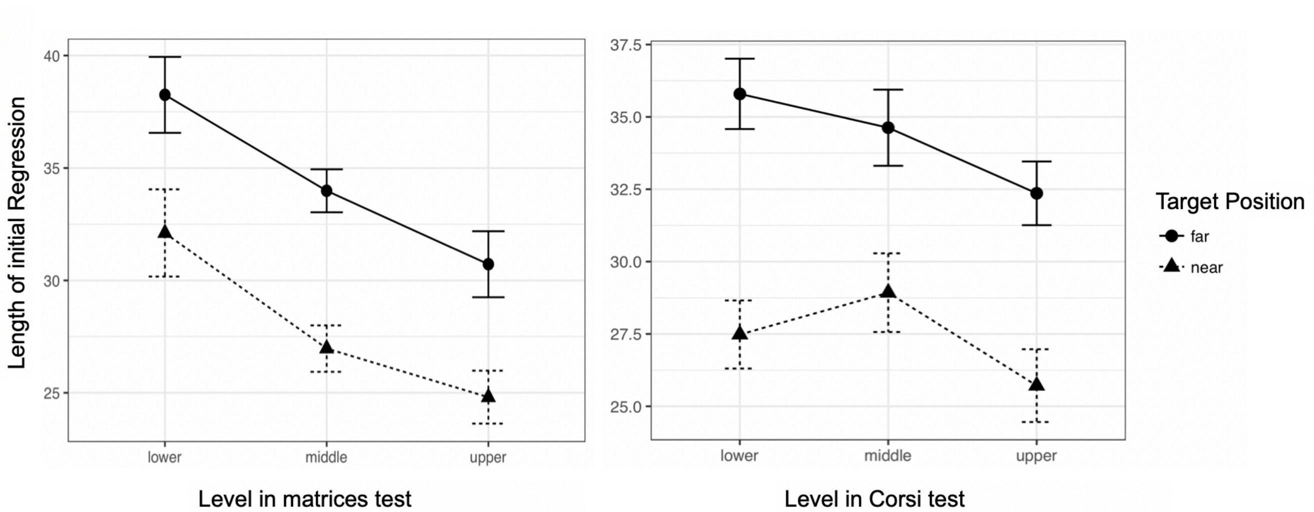


Figure 10. Mean length of initial regression as a function of target position (near and far) and level of performance (lower, middle, upper) in matrices (left) and Corsi test (right).

Level in matrices test also influenced the total regression time, indicated by a main effect of matrices test on the time it took, to reach the target word $F(2, 533) = 6.185$, $p<.01$, $\omega^2 = .014$, indicating that a higher level of performance is associated with less time needed. Interaction effects between level of performance in matrices and target position revealed that both number of regressions, $F(2, 533) = 13.171$, $p<.001$, $\omega^2 = .031$, and regression time, $F(2, 533) = 10.436$, $p<.001$, $\omega^2 = .025$, are reduced with higher performance in matrices test, especially for near targets (see figure 11).

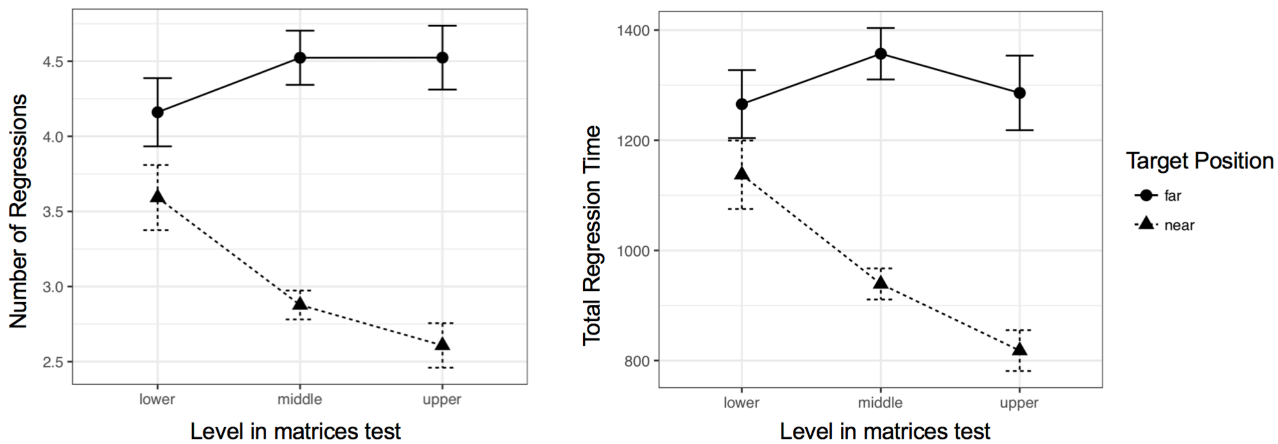


Figure 11. Mean number of regressions and total regression time as a function of target position (near and far) and level of performance (lower, middle, upper) in static spatial memory (matrices test).

Children vs Adults

In this section, the data collected with children in experiment 2 are compared to those obtained in experiment 1 with adult readers. Results for both samples are provided in table 6. An ANOVA with the factors age and position of the target revealed that the length of the initial regression did not significantly differ between children and adults ($p=.068$), but shows a tendency that initial regressions of children are slightly shorter than those of adults. Both variables associated with correcting processes show significant differences between children and adults. Children needed substantially more regressions, $F(1, 232) = 56.80$, $p<.001$, $\omega^2 = .123$, and more time, $F(1, 232) = 173.49$, $p<.001$, $\omega^2 = .311$, to reach the target word. An additional interaction effect of target position and number of regressions, $F(1, 232) = 8.835$, $p<.001$, $\omega^2 = .017$, and regression time, $F(1, 232) = 11.35$, $p<.001$, $\omega^2 = .019$, indicates that this effect is pronounced with far targets.

Table 6. Means and standard errors for amplitude of initial regressions (in letter spaces), total number of regressions and total regression time (in ms) as a function of age (children vs. adults).

		near		far	
		Kids	**Adults**	**Kids**	**Adults**
Initial Regression	**M**	**27.77**	**28.9**	**34.34**	**35.66**
	SE	0.83	0.95	0.79	1.01
Number of Regressions	**M**	**3.0**	**2.37**	**4.43**	**3.25**
	SE	0.09	0.1	0.12	0.08
Regression Time	**M**	**960.9**	**628.9**	**1315.2**	**809.9**
	SE	26.6	26.6	32.4	19.4

M = mean, SE = standard error

Regression strategies

The analysis of regression strategies provided detailed insight into the way children regress to target words and how they respond to initial regressions not attaining the target. 97.4 percent of all trials were successful and targets were eventually hit by a regression. The few remaining trials were apparently aborted too early with a button press before the target had been found. 17.5 percent of all executed regressions were single shot regressions. In this case, only one long regression was necessary to get to the target. The proportion of goal directed regressions was 16.8 percent, representing a total of 34.3 percent very well targeted long-range regressions. The next frequent strategies are centered search with 27.4 percent and backward scanning with 28.7 percent, while forward scanning accounts for only 4.9 percent. Another 4.8 percent of the cases could not be classified into one of the defined strategies.

Chi² tests compared the preferred use of strategies with respect to lower and upper performance in test assessments. As it turned out, only the matrices test revealed a significant difference in strategy use $X^2(4) = 14.262$, $p<.01$ (cf. figure 12). More upper (10) than lower performers (5) used single shot regressions as their main strategy. Further, more lower (19) than upper performers (5) preferred centered search, whereas more upper performers (17) relied on backward search compared to lower performer (8). For near targets, again, reading comprehension showed no influence on the preferred use of strategies (p=.6). In contrast, static spatial memory showed a significant influence with higher performance in the matrices test leading to more single shot regressions (see figure 12). Search strategies were also influenced by level of performance in matrices. The use of backward search increased from lower (1) to upper level performance (4) at the expense of the centered search strategy, which decreased from lower (10) to upper (1) level. Level of performance in Corsi test, measuring dynamic spatial memory, had no influence on regression strategies in near targets.

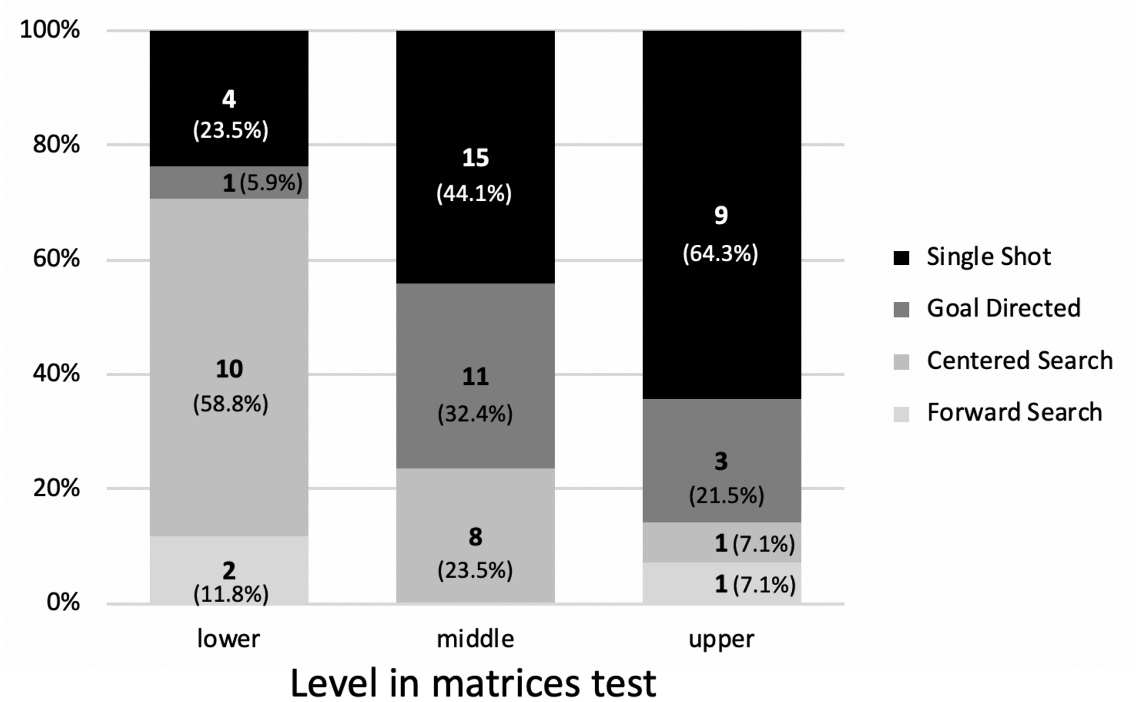


Figure 12. Proportion of preferred strategies (both number of participants and percent) by level of performance in matrices test for near targets in children.

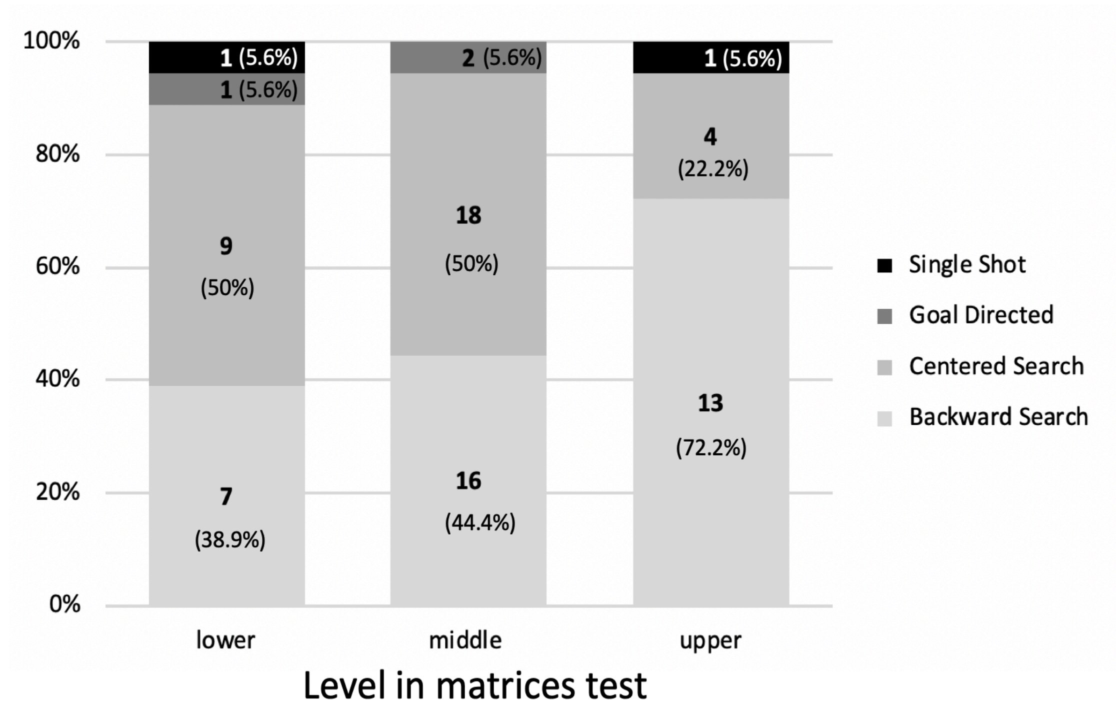


Figure 13. Proportion of preferred strategies (both number of participants and percent) by level of performance in matrices test for far targets in children.

For far targets, none of the test assessments showed a significant influence on the use of regression strategies, but there was a tendency that performance in matrices test influences search strategies for far targets, with a marginal effect of $X^2(3) = 7$, p=.072. Upper performer tended to use more backward scanning (13 vs. 7), whereas lower performers appeared to rely more often on centered search (9 vs. 4, see figure 13).

Children vs Adults

The comparison of strategies between children and adults is based on means of used strategies for either all targets or near and far target words separated, because sample size of children and adults was not equal and Chi^2 tests could not be used. A repeated measures ANOVA revealed an interaction effect between Strategies (Single Shot, Goal Directed, Centered Search, Backward Scanning and Forward Scanning) and Age (Children vs. Adults), $F(3.096, 362.201) = 11.786$, $p \leq .001$ and an interaction effect between Strategies, Age and Position of Target $F(2.803, 327.980) = 5.503$, $p \leq .001$. For both, degrees of freedom were corrected using Greenhouse-Geiser estimates of sphericity.

Overall, children had less single shot $t(117) = 4.013$ $p \leq .001$, $r = .35$, and less goal directed regressions $t(117) = 5.677$, $p \leq .001$, $r = .46$ than adults. They had a dramatically higher proportion of backward searches $t(117) = -3.588$, $p \leq .001$, $r = .31$ but significantly less centered search $t(117) = 2.682$, $p \leq .01$, $r = .24$, compared to adult readers.

With near targets, children showed less single shots $t(73.8) = 3.773$, $p \leq .001$, $r = .4$ but about the same number of goal directed regressions. They exhibited less centered search $t(81.9) = 2.066$, $p \leq .05$, $r = .22$, but significantly more backward scanning $t(116.6) = 5.723$, $p \leq .001$, $r = .47$. In far targets the amount of goal directed regressions was about twice as high in adults $t(69.2) = 5.243$, $p \leq .001$, $r = .53$. Children engaged less in centered search $t(112.1) = 2.609$, $p \leq .01$, $r = .24$, but more in backward scanning $t(117) = -2.333$, $p \leq .05$, $r = .21$.

Similar tendencies emerged when comparing low and high performance in visuospatial memory tests in children and in adults. High performers had more single shot and accurate regressions. Their main search strategy is backward scanning followed by centered search and forward scanning. Low performer had less single shot and accurate regressions and their main search strategy is centered search, followed by backward scanning and forward scanning. In children, these effects were observed based on performance in static spatial memory (matrices test), whereas in adults the same pattern was driven by dynamic spatial memory (Corsi test).

Test correlation

Performance in matrices task was significantly related to performance in the Corsi block tapping task, $r = .46$, $p<.01$. Participants with a high level of performance in Corsi also tended to have superior performance in the matrices task. There is also a significant relationship between performance in visuo-spatial tasks and reading test. High performers in Corsi, $r = .28$, $p<.05$ as well as high performers in the matrices task, $r = .38$, $p<.01$ tend to be those who are also successful in the reading task.

The correlation between performance in visuo-spatial memory and reading ability, which was found in children, was not present in the adult sample.

3.4 Discussion

The present experiment examined how children plan and execute long-range regressions while reading. Target position strongly influenced the length of the first regression programmed to go back into the sentence. Initial regressions were substantially longer for far than for near targets, indicating spatial selectivity and suggesting that spatial information retrieved from memory is used to guide regressions towards certain words. At the same time, the results indicate that distance also affects the accuracy of long-range regressions and that far targets are more error-prone than near targets (as seen in experiment 1 and Guérard et al., 2014; Weger & Inhoff, 2007; Inhoff & Weger, 2005). Far targets need more regressions and more time to be located and eventually fixated, mostly related to the fact that inaccurate landing positions of initial regressions needed to be corrected.

The influence of spatial memory

The results show a clear influence of static spatial memory on regression accuracy and regression strategies. High performers in matrices task show significantly more single shot and goal directed regressions than low performers, whereas low performers engaged in a lot more searching for targets. Low performers mostly started the searching process in the center of the sentence, followed by saccades towards the sentence end or the beginning. This differs from the way high performers search for targets, mostly with a regression to the end of the sentence to backtrack until the target is attained. Moreover, individual differences in spatial test influenced regression strategies, high performer in Corsi and matrices test showed shorter initial regressions. A main effect of matrices test performance on number of regressions and total regression time emphasizes the importance of static spatial memory for the effectiveness of regression programming in children, especially for near words.

Reading skill influenced the way a misguided first regression was corrected. Highly skilled level readers needed fewer regressions and less time to correct an incorrect initial landing position compared to less skilled readers. These results clearly demonstrate that spatial memory impacts the way children program long-range regressions. Participants with good spatial memory are more likely to program single shot regressions. In contrast, poorer performance in spatial memory is associated with inaccurate initial regressions, leading to effortful and time-consuming correction that interrupts the reading process.

The results outlined above are seen as direct evidence for a functional role of spatial memory in reading. This view goes hand in hand with findings suggesting that visuospatial memory is impaired in children with poor reading abilities and that individuals with dyslexia are characterized by deficits in tasks examining visuospatial skills or visuospatial short-term memory (Bacon, Parmentier, & Barr, 2013; Winner et al., 2001; Lipowska, Czaplewska, & Wysocka, 2011; Menghini, Finzi, Carlesimo, & Vicari, 2011; Eden, Wood, & Stein, 2003).

On the other hand, there are also several studies reporting no clear link between poor reading performance or dyslexia and visuospatial memory. Here, children with dyslexia show intact visuospatial short-term memory despite their low reading performance (Del Guidice et al., 2000; Kibby, 2009; Nation et al., 1999; Rawson & Miyake, 2002; Von Karolyi, 2001; Von Karolyi, Winner, Gray, & Sherman, 2003). Most of the apparent discrepancy can be reconciled when considering different processing stages in reading, specifically problems in decoding (the accuracy or fluency of reading aloud) vs. problems with comprehension (the adequacy of understanding text, Hulme & Snowling, 2016). Since long-range regressions are used for re-reading and repairing comprehension difficulties in the majority of cases (Gregg & Inhoff, 2016; Booth & Weger, 2013; Slattery & Rayner, 2009), only reading difficulties at this stage should be related to spatial memory deficits. Any deficiencies that occur in reading processing prior to this stage (e.g. problems with grapheme to phoneme correspondence) may well occur independently of spatial memory.

Differences and similarities between Children and Adults

Very similar results were obtained up in experiment 1 with adult participants. In both samples, initial regressions were longer for far than for near targets, and far targets needed more additional saccades and more time to be reached. Differences between children and adults emerged in the two dependent variables associated with correcting processes, executed following initial regressions with inaccurate landing positions. In such cases, children needed more regressions and more time to attain target words compared to adults. For children, performance in static spatial memory (matrices test) modulates searching processes to near targets, whereas for adults dynamic spatial memory (Corsi test), assumes this role, in both cases indicated by fewer saccades and less searching time for high level performers.

For both children and adults, very similar relations between high and low performance in visuo-spatial memory tests and the use of search strategies emerged. When high performers missed the target, they most often used the backward scanning strategy. This can be interpreted as the result of an incorrect assumption about the position of the target (in contrast to having no idea at all). The position of far targets is more error-prone, possibly inducing a bias for longer initial regressions, resulting in frequent backtracking saccades to find the actual position of the target.

Low performers relied more often on a centered search strategy. The use of this strategy suggests that low performers, did often not mark the target word with a spatial tag during reading and therefore had no idea where it could be. Starting at a central position is quite economical, since a large part of the sentence is brought into central vision, which may help to decide where further correcting saccades should be directed. This might also be the reason why, seen globally, adults used centered search more often than children.

At this point two questions arise: Why did adults execute more single shot regressions than children and why did regression accuracy depend on performance in matrices test in children and on Corsi test in adults? One possible answer to the first question is suggested by research indicating that that spatial memory is developing until 14-15 years of age (Gathercole, Pickering, Ambridge, & Wearing, 2004). These authors reported that the predominantly spatial tasks block recall and mazes memory (comparable to the tests used in experiment 1 and 2) had a particularly late developmental asymptote compared to a visual patterns test. In the present work the sample of children had an average age of 9;11 years, an age when neither visual nor spatial memory are fully developed. Taking this view, the proportion of single shot regressions during reading appears to be a consequence of general spatial memory development.

An alternative possibility could be that linguistic and spatial processing rely to some extent on shared capacity in working memory. If this is the case, the more extensive use of mental resources for decoding and comprehension may limit the availability of such resources for spatial tagging of word positions. Such a mechanism is likely to operate within the limits of spatial memory development outlined above. Its role can be studied in research varying the local difficulty of linguistic processing during sentence or text reading (see experiment 3).

Back to the second question raised above: Why is the matrices test more sensitive for children and the Corsi test more sensitive for adults? It is known from decades of research that young readers show a dramatically inflated number of fixations with longer fixation durations, resulting in longer gaze durations and longer reading times per word (McConkie et al., 1991; Blythe, Liversedge, Joseph, White, Findlay, & Rayner, 2006; Joseph et al. 2009; Huestegge, Radach, Corbic, & Huestegge, 2009; Blythe, Häikö, Bertam, Liversedge, & Hyönä, 2011). They show more refixations on a word and shorter saccade amplitudes, reflecting a small-unit decoding strategy. This reflects a relatively static reading process with a high demand on local cognitive processing.

With more skilled reading, eye movement behavior changes: words can be read in a very short time and often with only one fixation representing a more efficient whole-word recognition of words (e.g. Rau, Moeller, & Landerl, 2014; Taylor & Perfetti, 2016; Conrad & Deacon, 2016). In addition, the use of parafoveal information increases with reading performance (Marx, Hutzler, Schuster, & Hawelka, 2016), reducing the effort of word identification when a word is subsequently fixated (Häikiö, Betram, & Hyönä, 2010; Häikiö, Bertram, Hyönä, & Niemi, 2009). All this reflects a change from a more localized, static reading strategy in beginning reading to a much more dynamic reading strategy in skilled readers. This fundamental shift may explain why different spatial memory tasks are particularly sensitive at certain stages of reading development.

The effect of repetition

In the present experiment target words were presented twice to each participant, once at a near and once at a far position, to examine word position effects. In adults, the repetition of words had no effect on regression accuracy. In children, the second presentation of targets led to shorter initial regressions and to fewer correcting regressions, when the second presentation was in a near position. This effect can be interpreted as another hint for an indirect spatial memory mechanism. In children, the second presentation of a word may lead to a more efficient linguistic processing of that word (cf. Radach, Günther, & Huestegge, 2012), leaving more capacity for the storage of spatial information about the target position. This may in turn result in a more effective way of attaining the target word, an effect that no longer manifests itself with increasing reading skill.

4 Experiment 3: The influence of word difficulty on regression accuracy

4.1 Introduction

Interword regressions, which go back further as to the word immediately left to the currently fixated word, ensure reading for understanding (Booth & Weger, 2013; Schotter, Tran, & Rayner, 2014) and help solving comprehension problems (Frazier & Rayner, 1982; Altmann, Garnham, & Dennis, 1992; Mitchell, Shen, Green, & Hodgson, 2008; Rayner, Pollatsek, Ashby, & Clifton, 2012; Slattery & Rayner, 2009). Consequently, words, that are difficult to identify are more likely the target of inter-word regressions compared to easy words (Vitu & McConkie, 2000; Kliegl, Grabner, Rolfs, & Engbert, 2004). On the lexical level, processing difficulty of a word is determined by the frequency of that word in the language (Vitu & McConkie, 2000; Inhoff & Rayner, 1986). Words with low frequency are generally fixated longer than words with higher frequency (Kliegl, Nuthmann, & Engbert, 2006; Kliegl et al., 2004; Schilling, Rayner, & Chumbley, 1998; Rayner, Sereno, & Raney, 1996; Henderson & Ferreira, 1990; Inhoff & Rayner, 1986), resulting in longer gaze durations (Henderson & Ferreira, 1990; Pollatsek, Juhasz, Machacek, & Rayner, 2008; Schuster, Hawelka, Hutzler, Kronbichler, & Richlan, 2016), overall longer processing times and less acquisition of parafoveal information (Henderson & Ferreira, 1990). Additionally, low frequency words are skipped less often (Inhoff & Rayner, 1986; White, 2008; Schuster et al., 2016), but when skipped, they are more likely the target of regressions than high frequency words (White, 2008; Vitu & McConkie, 2000).

In single word recognition experiments (lexical decision and single word naming), words with higher frequency led to shorter reaction times (Gardner, Rothkopf, Lapan, & Lafferty, 1987; Grainger, 1990; Brysbaert, 1996; Gerhand & Barry, 1998; Schilling, Rayner, & Chumbley, 1998; Brysbaert, Lange, & Van Wijnendaele, 2010), supporting the hypothesis that representations of words with high frequency are easier and hence more quickly retrieved from lexical memory, because of their higher exposure in language (Coltheart, Rastle, Perry, Langdon, & Ziegler, 2001; Reichle & Perfetti, 2003; Vanyukov, Warren, Wheeler, & Reichle, 2012).

On the other hand, in recognition tasks challenging episodic memory (recognition of words [yes/no] that had appeared in an earlier presented list of words) low frequency words achieve better results, as indicated in a lower rate of false alarms and a higher rate of hits than high frequency words, which is described as the word-frequency mirror effect (Glanzer & Adams, 1985; Glanzer & Adams, 1990; Glanzer, Adams, Iverson, & Kim, 1993; Mulligan, 2001; Heathcote, Ditton, & Mitchell, 2006).

The processing times of words are also influenced by the frequency of the beginning letters of a word, which is referred to as orthographic regularity. Manipulating the frequency of initial letter trigrams and quadrigrams within three categories (low, medium and high orthographic regularity), Radach, Inhoff, and Heller (2004) found that first fixation durations and gaze durations increase for less regular word beginnings. It is generally assumed that orthographic regularity of words also affects the programming of eye movements on a visuo-motor level. Saccade amplitudes are shorter for words with irregular beginnings and the landing site distribution is shifted to the left with less regular letter trigrams and quadrigrams accordingly (Radach et al., 2004; White & Liversedge, 2004; Reichle, Rayner, & Pollatsek, 2003; Vonk, Radach, & van Rijn, 2000; Hyönä, 1995).

In the present experiment, frequency and orthographic regularity of target words were varied within two categories (high vs. low) to test for an indirect linguistic influence on spatial memory and the accuracy of the programming of long-range regressions. The existence of a substantial direct effect of spatial memory on the programming of long-range regressions was confirmed with the findings of experiments 1 and 2. The performance of a person in tests of spatial memory was a good predictor for regression accuracy in sentence reading. A higher performance resulted in a higher number of regressions attaining a target word precisely without any need of correction (single shot regressions), whereas lower performance led to searching for the target word with multiple saccades. On a more detailed level, search strategies differed with performance in spatial memory tests. A higher level of performance yielded an increasing use of backward scanning regression strategy (starting at the end of the sentence and scanning back until reaching the target word). In contrast, a lower level of performance led to a higher use of the centered search strategy (starting in the center of the sentence to bring as much information as possible into central vision to decide in which direction the scanning process should continue). Performance in reading tests did not show similar correlations. High performers in reading tests were able to correct an incorrect landing position regression faster and with les effort, but did not execute more accurate initial regressions.

A first presumption of the existence of an additional indirect mechanism of spatial memory comes from considerations regarding the interpretation of our results from experiment 1 and 2 in relation to Guérard et al. (2012 and 2014). These authors suggested that the impact of a verbal interference task on regression accuracy implies some kind of involvement of verbal memory in the programming of regressions. A possible interpretation among these lines would be that due to the utilization of the verbal memory, information regarding target locations can´t be stored in spatial memory, because of an overload in the central executive (see Baddeley, 2010, for the proposition that the central executive is attentional limited). Another hint comes from the results of experiment 2. A comparison of regression accuracy of children and adults indicated that adults showed significantly more single shot regressions than children. In addition to the direct influence of the spatial memory (which is developing in its efficiency with age), a more indirect mechanism could come to play here. With increasing reading skills, less cognitive capacity is used for decoding processes during reading, leaving more cognitive capacity for the storage of spatial information about the position of words during reading (see Blythe, 2014, for a summary of the literature considering eye movement differences in beginning and skilled readers as well as Vorstius, Radach, & Lonigan, 2014, for a detailed description of changes in eye movements from first to fifth grade).

Based on the above considerations, experiment 3 was designed to test for an indirect mechanism of spatial memory on regression accuracy. To achieve this goal, a set of sentences was developed so that each sentence included targets that are either difficult or easy to process. Two opposing hypotheses appear viable. On the one hand, the position of difficult words might be located particularly well during the planning of a regression, because a lot of processing time had been spent on this position within the sentence, creating ample opportunity to store word position information in visuo-spatial working memory. This would lead to the programming of more accurate initial regressions and to a higher rate of single shot regressions for difficult than for easy words.

On the other hand, following the idea about an indirect influence of spatial memory on regression programming, regressing to a more difficult word might be problematic, as less resources for spatial memory encoding were available while reading the word. In this case, one would expect that difficult words are targeted by less accurate initial regressions, needing more effort to correct erroneous landing positions and more time to reach the target word. Consequently, there should be a higher likelihood of search strategies, perhaps in the use of more centered search for difficult targets.

As in experiment 1 and 2, targets were presented both near the sentence beginning and near the sentence end (and hence far vs. near relative to the starting point of regressions). Participants read the sentence for understanding, which was ensured by asking questions randomly (but frequently) after the sentence had been read. Again, the execution of long-range regressions was elicited by the use of an error-finding task. After the sentence had been read, participants were asked to check for spelling errors, which were inserted into target words in 50 percent of the cases after the participant crossed an invisible boundary at the end of the sentence, a procedure well known for its efficiency from experiment 1 and 2.

4.2 Methodology

Participants

Fifty-four university students (46 female) with a mean age of 22;1 years (min = 17;10 years, max= 30;10 years) participated in this experiment. They gave written consent prior the experiment and received course credits for participation. One participant needed to be excluded, because assessments could not be completed for personal reasons.

Materials

The materials comprised 100 declarative sentences with a length between 69 and 79 letter spaces and contained 9-14 words. Every sentence contained one target word. Target words were nouns with a length of 6-7 letters and were either part of the subject or the adverbial complement, to avoid any training effects for a special grammatical element. Each target word was preceded by an adjective, which was 5 - 8 characters long. 64 out of the 96 experimental sentences were designed so that the target word was located either close to the beginning (far target) or near the end of one sentence (near target). Target words at the beginning of a sentence were preceded by two or three words, whereas those located at the end of a sentence were followed by one or two words. During the course of the experiment each target word was presented twice to each participant, once at the beginning position and once at the end position, following the procedure established in experiment 1 and experiment 2 (see chapters 2.2 and 3.2). The following figure (figure 14) provides examples of the sentences and experimental conditions used in the present experiment. Again, the remaining 32 sentences were used as filler sentences with target words near the center of the sentences to avoid any training effects for specific target locations.

Word frequency and orthographic regularity were used to define word difficulty. Word frequency is defined as number of word types per million based on the dlex DB word corpus (Heister et al., 2011). High frequency was set to ≥150 per million and low frequency to ≥1 ≤3 words per million. Orthographic regularity is defined for a given word type as the number of types with the same word length and the same initial letter trigram, again based on the dlex DB database (Heister et al., 2011) and was set to >10 for high and to <5 for low orthographic regularity. Word length was set to 6-7 letters, to avoid that frequency effects be confounded with word length (e.g. Kliegl et al., 2004).

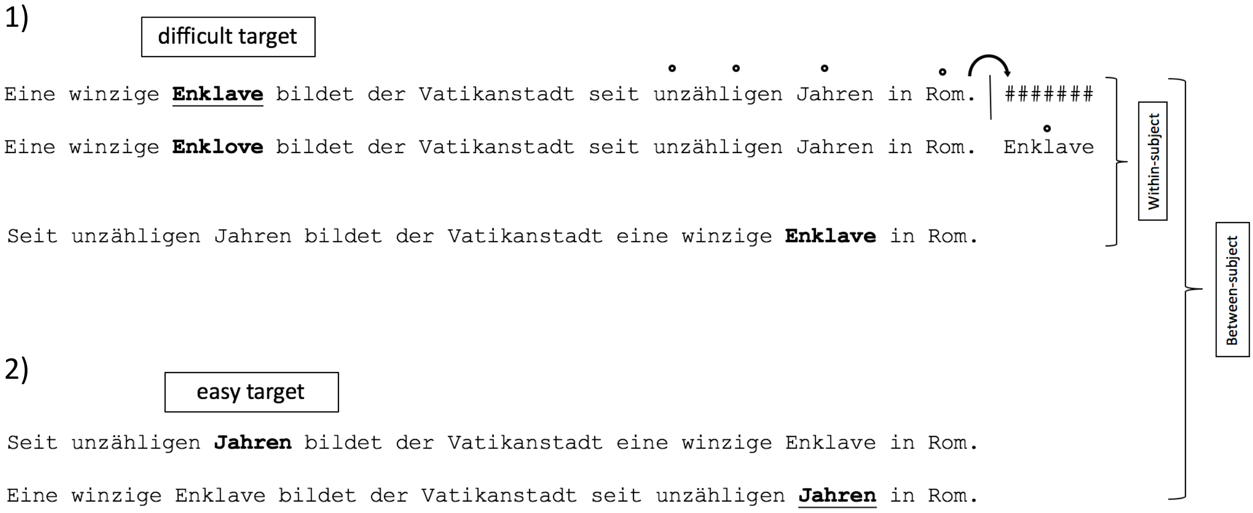


Figure 14. Example sentences out of experiment 3.
Target words are not marked in any way during the experiment. Condition 1 represents sentences with a difficult word, each as far (first two sentences) and as a near target (third sentence) (50 percent of the cases). Condition 2 represents the same sentence frame, but with the easy word being the target word (other 50 percent of the cases), please see text for details. Condition 1 illustrates the invisible boundary. When crossing this boundary with a saccade, the probe word at the end of the sentence gets visible, which was masked through hashmarks beforehand.

Retrieval of high-frequency 6 to 7 letter long nouns (≥150 p. M.) in the dlex DB word corpus resulted in a total of 74 hits. These candidate words were sorted in descending order of orthographic regularity and the 52 words that had an orthographic regularity of> 10 were selected. For the difficult words, a normalized frequency of ≥1 ≤3 and an orthographic regularity of <5 were specified. For 6-7 letter long nouns, this resulted in an item pool of 128 words. These two lists were used to construct the sentence materials for the experiment.

Procedure

Participants sat in front of the computer. They read the instructions of on the computer screen and were instructed by the experimenter additionally if needed. Out of the 100 sentences the first four served as practice items to make sure that all instructions were completely understood. Sentences were presented in Courier New with a font size of 15 pt in black on a light grey background on a 21-inch flat-panel monitor with a resolution of 1,680 x 1,050 pixels and a refresh rate of 120Hz. Viewing distance was set to 68cm, so that every letter occupied a width of 0.3° of visual angle. Eye movements were recorded using an SR EyeLink1000® video-based eye tracking system (SR Research, Toronto, Canada) and a refresh rate of 2000Hz. A chinrest and forehead rest minimized head movements. At the beginning and at several points during the experiment a three-point calibration, followed by a validation routine, was performed. A drift check before the start of each sentence ensured accuracy between calibrations. Participants were instructed to read the following sentences silently and for good comprehension. After they read the sentence for the first time, a probe word at the end of the sentence became visible. At his point and the corresponding word in the sentence had to be checked for a spelling mistake (inserted after the sentence had been read for the first time), followed by pressing the appropriate button for correct or wrong spelling. Target distance (near vs. far position), word difficulty (easy vs. difficult word) and error (error vs. no error introduced to target word after crossing the boundary) were used to create a total of eight experimental lists.

To ensure that participants read for understanding, comprehension questions were asked after approximately every sixth sentence, which had to be answered orally. Answers were noted by the experimenter. The session lasted about 30-45 minutes, depending on the reading speed of the individual participant.

Data selection

Saccades and fixations (recorded during the reading task) were classified online using the EyeLink software. Raw data output from the recording system was further processed using the software suite EyeMap (Tang, Reilly, & Vorstius, 2012) that generated a wide range of oculomotor indices for each word of the sentence. Three measures of regression performance are reported: First, the length of the initial regression back into the sentence, second, the number of regressions that were needed to attain the target word after boundary crossing and third, the interval in between the onset of the first regression into the sentence and the fixation of the target (total regression time).

Values of all parameters had to be log-transformed for statistical analyses to better fit normal distributions. Figures and tables report non-transformed values for ease of interpretation.

To remove outliers, the following exclusion criteria were used: for length of initial regression outliers with log values of <2.0 (0.6 percent), for number of regressions outliers with log values of > 2.2 (0.8 percent), and for total regression time outliers with log values of < 4.5 and > 7.7 (0.8 percent) were excluded.

Data analysis

In a first step, the three regression measures from the reading task were analyzed separately to identify significant predictors of readers' regressions. The second step comprised a qualitative selection of regression strategies (types) and an examination of regression accuracy as a function of regression strategies. Linear mixed models (LMM), as implemented in the lmer function from the lme4 package (Bates, Mächler, Bolker, & Walker, 2015) in the R environment for statistical computing (R Core Team 2014; version 1.0.136), were used to analyze the three regressions measures. The LMMs included the fixed effects position (near vs. far) and difficulty (difficult vs. easy) and interaction of these factors. The LMM started with a maximal random effect structure but this structure was simplified as maximal models failed to converge.

Specifically, maximum likelihood comparison revealed that only the variance components of participant and item intercepts and participants' slope for the position effect improved the model. The fixed effects of position and difficulty and their interaction improved the prediction of the statistical model and were consequently not removed.

Regression strategies

The classification of regression patterns from the probe to the target word into different categories followed the procedure described and successfully used in experiment 1 and 2 (see chapters 2.2 and 3.2). The categories are (1) *single shot regressions*, (2) *goal directed regressions*, (3) *backward searches,* (4) *centered searches* and (5) *forward searches*. For a detailed description of classification principles please see the methods section of experiment 1. Participants used single shot regressions in 22.2 percent of all cases, goal directed searches in 18.2 percent, centered searches in 29.6 percent, backward searches in 18 percent and forward searches in 5.1 percent. In 0.6 percent of all observations targets were never reached by a regression, hence button presses were executed before attaining the target. In 6.4 of the trails regressive searching behavior could not be captured within the classification and remained unclassified.

Eye movement measures

To determine the processing time of difficult and easy target words, temporal eye movements representing early, intermediate and late linguistic processing were used (see Radach & Kennedy, 2013 and Payne & Stine-Morrow, 2012, for overviews). For each target, that was fixated at least once, the early measures of temporal measures, first fixation duration and single fixation duration and are measured. The single fixation duration specifies the duration of a single fixation on a word, when this word is fixated exactly once. If a word is fixated by more than one fixation, the duration of the first fixation on that word is reported (the first fixation duration). These two first-pass measures are generally assumed to reflect orthographic and early lexical processing (Hyona, Lorch Jr., & Rinck, 2003; Inhoff & Radach, 1998). The sum of all fixations on a word, until this word is left for the first time, is reported as gaze duration and, as an intermediate measure, reflects later stages of word processing such as lexical access (Radach & Kennedy, 2013). Additionally, the later-pass measures re-reading time and total viewing time are reported. The re-reading time describes the duration of all fixations made on the word after leaving it for the first time and total viewing duration is the sum of gaze duration and re-reading time and reflects the duration of all fixations made on a word (see Radach & Kennedy, 2004, for an overview of spatial and temporal eye movement measures). These later measures are associated with comprehension issues on sentence and passage level and the integration of the meaning into previously read materials (Radach & Kennedy, 2013).

304 (9.6 percent) of all target words were skipped and never fixated. Of these, 78 words were classified as difficult and 226 were classified as easy words. Again, all temporal eye movement measures were log-transformed for a better fit of normal distribution. Exclusion criteria were log-values of >4.3 and <6.5 (1.2 percent) for single fixation duration, >4.0 and <6.2 (1.3 percent) for first fixation duration, >4.3 and <6.7 (1.3 percent) for gaze duration, >5.5 (0.5 percent) for total viewing duration and >5.0 (1.4 percent) for re-reading time.

4.3 Results

Effects of word difficulty

Word difficulty was manipulated using word frequency and orthographic regularity of words. Difficult words had lower frequencies (≤3 per million) and lower regularity of initial letter trigrams (<5 per million) compared to easy words with a frequency of more than 150 per million and an initial trigram frequency of more than 10 per million. This manipulation leads to distinctive differences in word processing, as demonstrated by temporal eye movement measures. Easy and difficult words differed significantly in single- and first fixation duration, gaze duration, re-reading time, and total viewing time (see table 7 and figures 15-18). Evidently, difficult words need to be processed with substantially more effort during all stages of word processing and are more likely the target of regressions (hence the longer re-reading time). These results confirm that the distinction between easy and difficult words is effective, allowing for subsequent analyses of regression accuracy on the basis of word difficulty.

Table 7. Means and standard error for basic temporal eye movement parameters for easy and difficult words and results of the linear mixed model analysis (LMM).

	M (SE)		Estimate	SE	/t/-value
	Easy	Difficult			
Single fixation duration	216 (1.9)	249 (2.6)	0.138	0.018	**7.78*** **
First fixation duration	185 (5.3)	212 (4.3)	0.123	0.034	**3.608*** **
Gaze duration	232 (2.4)	289 (3.3)	0.2	0.023	**8.596*** **
Total viewing duration	804 (6.7)	921 (8.5)	0.126	0.02	**6.448*** **
Re-reading time	575 (6)	632 (7)	0.09	0.02	**4.474*** **

*** p < .001, M = Mean, SE = Standard Error, Easy = easy words, Difficult = difficult words

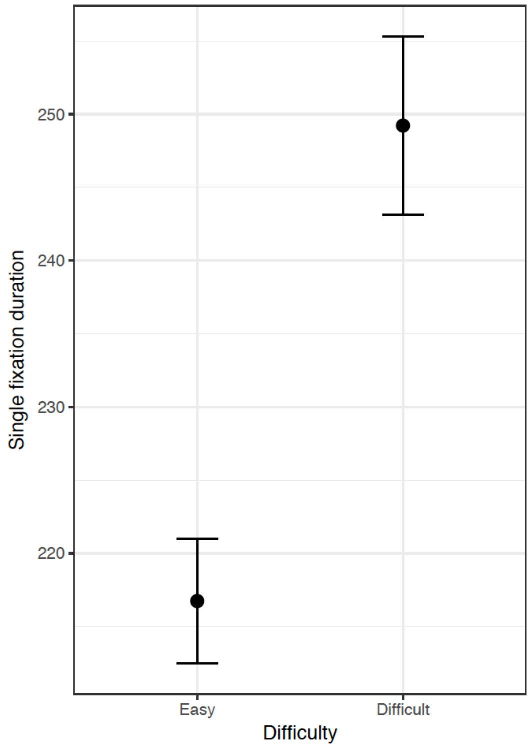


Figure 15. Means of single fixation duration (in ms) as a function of word difficulty.

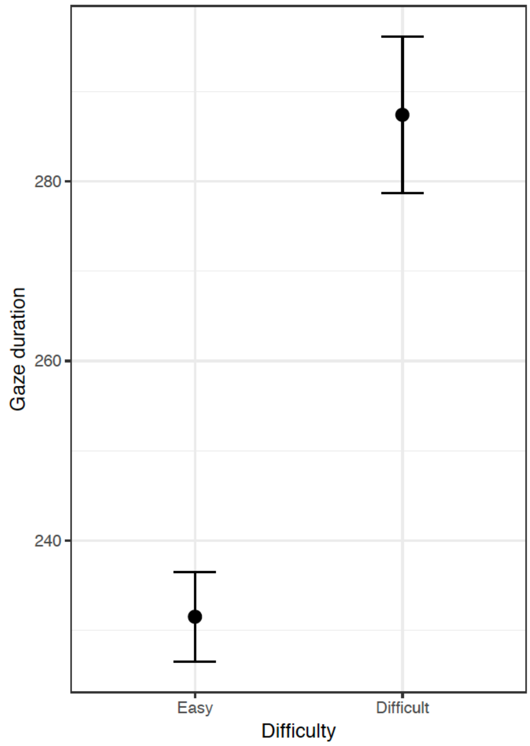


Figure 16. Means of gaze duration (in ms) as a function of word difficulty.

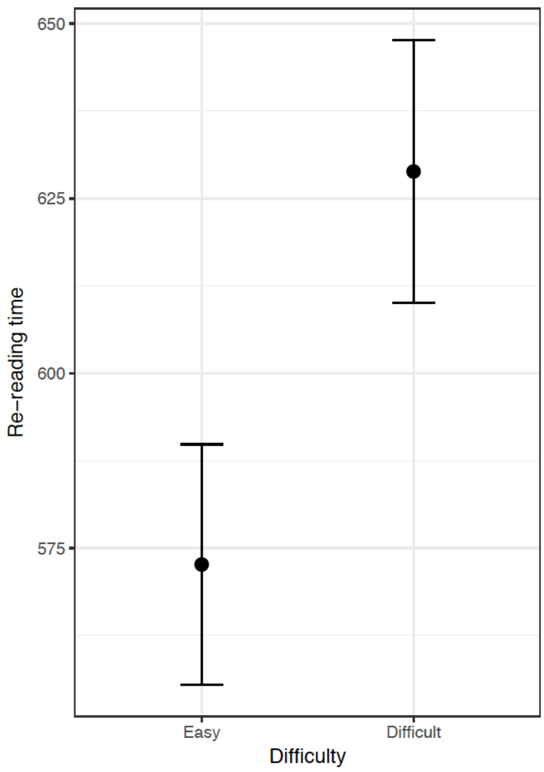


Figure 17. Means of re-reading time (in ms) as a function of word difficulty

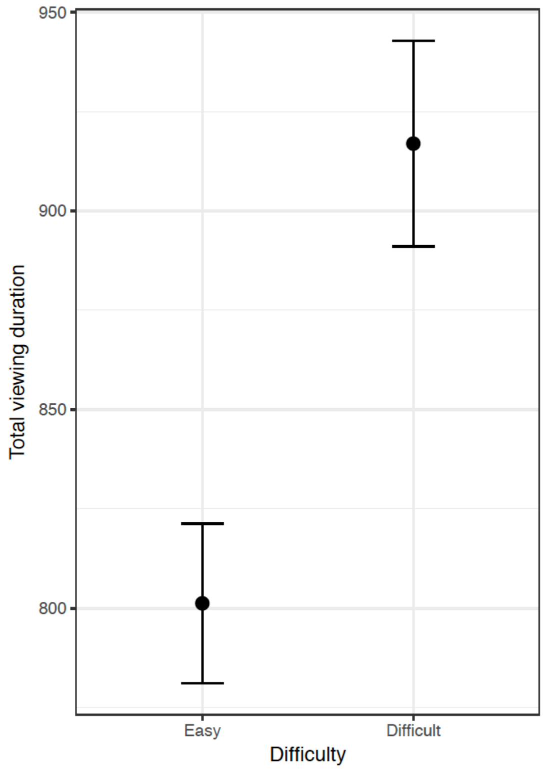


Figure 18. Means of total viewing duration (in ms) as a function of word difficulty.

Effects of word difficulty on Regression Measures

The effects of target word difficulty on length of initial regression, number of regressions and total regression time was examined using linear mixed models separately for the specified dependent variables (table 8 provides means and standard deviations for all dependent variables). The initial regression was significantly longer for far (M = 35.5, SE = 0.3) than for near (M = 26.9, SE = 0.3) targets, confirming spatial selectivity for target position, (b = 0.29, SE = 0.029, t = 10.066, p<.001).

An interaction effect between position and difficulty revealed, that spatial selectivity is especially pronounced for difficult words (see figure 18) (b = 0.34, SE = 0.0482, t = 7.065, p<.001), leading to a significant larger difference in regression length between far and near targets for difficult (14 LS) than for easy target words (3.6 LS) (see figure 19).

Correction processes in cases when initial regressions did not land on the target were generally influenced by word difficulty (see figures 20 and 21). Difficult words resulted in fewer regressions (b = 0.199, SE = 0.025, t = 7.921, p<.001) and less total regression time (b = 0.175, SE = 0.022, t = 7.926, p<.001) to eventually attain the target word.

Both variables interact with target position. For both variables, the differences are significantly larger for near targets (number of regressions: 0.7, b = 0.197, SE = 0.054, t = 3.631, p<.001, total regression time: 152 ms, b = 0.182, SE = 0.046, t = 3.965, p<.001) than for far targets (number of regressions: 0.3, total regression time: 61 ms).

Target position had a main effect, too. Near targets generally required less regressions (M = 2.2, SE = 0.04, b = 0.5, SE = 0.047, t = 10.744, p<.001) and less total time (M = 586 ms, SE = 8.1, b = 0.396, SE = 0.038, t = 10.387, p<.001) than far targets (number of regressions: M = 3.4, SE = 0.03, total regression time: M = 806.8 ms, SE = 7.1) to be reached.

The highly significant effect of word difficulty (caused by differences in word frequency and orthographic regularity), reflected in basic temporal eye movement parameters, led to significant effects on programming of initial regressions and in correcting wrong landing positions, hence influenced the representation of word positions in memory, especially for near targets.

Table 8. Means and standard error for amplitude of initial regressions (in letter spaces), total number of regressions and total regression time (in ms) as a function of word difficulty (easy vs. difficult) and target position (near vs. far).

		Near		Far	
		Easy	**Difficult**	**Easy**	**Difficult**
Length of initial Regression	*M*	29.6	24	33.2	38
	SE	0.4	0.4	0.5	0.5
Number of Regressions	*M*	2.6	1.9	3.5	3.2
	SE	0.06	0.05	0.05	0.05
Total Regression Time	*M*	660	508	835	774
	SE	12	10	10	10

M = Mean, SE = Standard Error, Length of initial Regression is given in letter spaces and total regression time in ms.

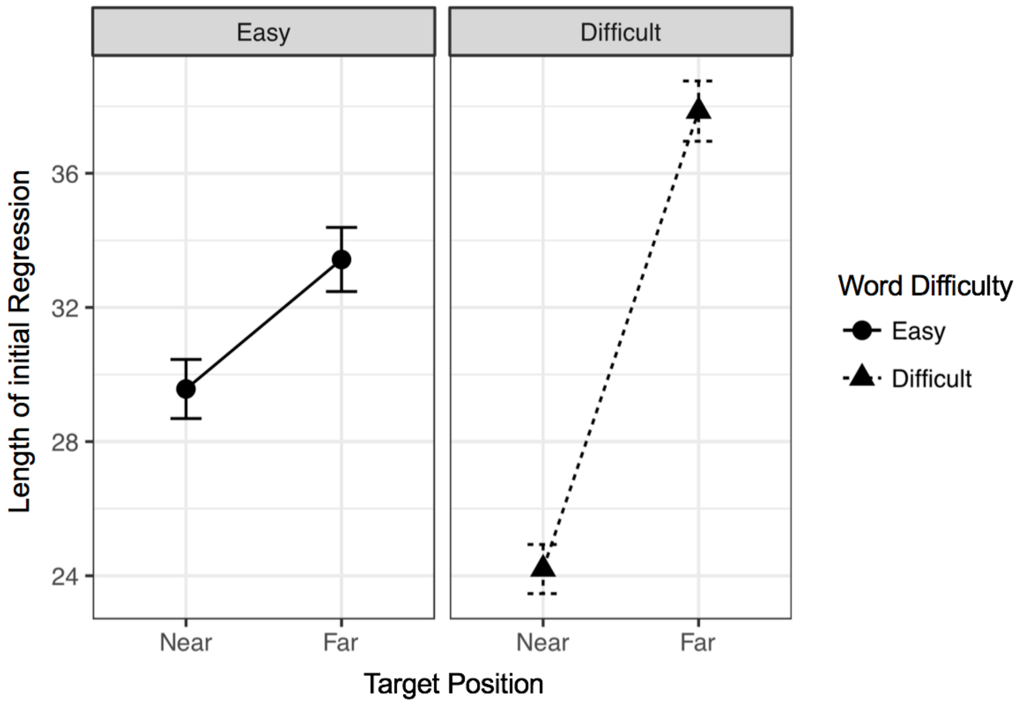


Figure 19. Means of length of initial regressions as a function of word difficulty and target position.

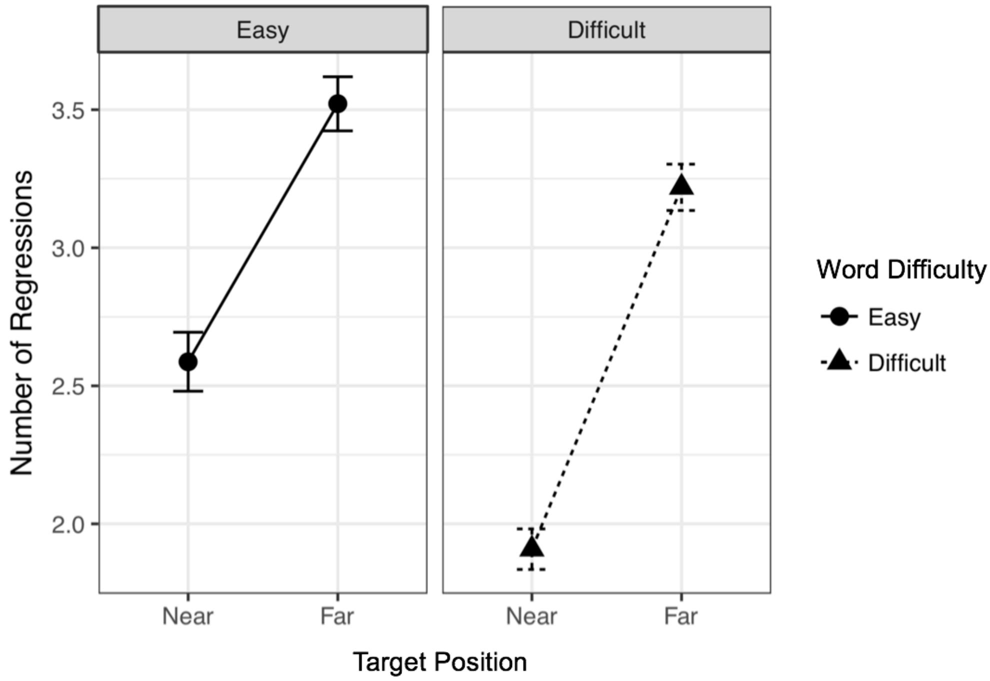


Figure 20. Means of number of regressions as a function of word difficulty and target position.

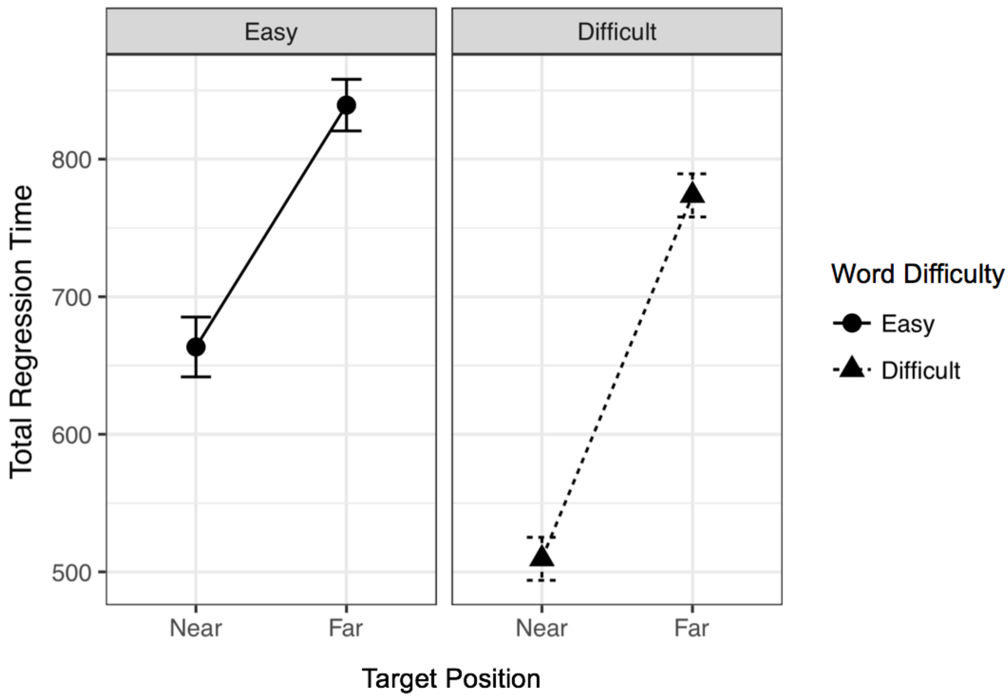


Figure 21. Means of total regression time as a function of word difficulty and target position.

Regression Strategies

To test the influence of word difficulty on regression strategies, a repeated measures ANOVA with the within-subject factors regression strategy (single shot, accurate regression, centered search, backward scanning and forward scanning) and word difficulty (easy vs difficult) was carried out. Degrees of freedom were corrected using Greenhouse Geiser estimates of sphericity ($X^2(20) = 226.928$, $p \leq .001$, $\varepsilon = .487$). Word difficulty had a significant influence on the used strategies, $F(4.223, 219.581) = 23.637$, $p < .001$. Paired t-tests revealed, that difficult words were targeted by a higher number of single shot regressions ($M = 8.38$, $SD = 3.8$) than easy words ($M = 4.94$, $SD = 3.3$), $t(52) = -5.953$, $p \leq .001$, $r = .64$. The amount of goal directed regressions also increased, but did not reach significance ($t(52) = -1.633$, $p = .109$). Easy words were more likely reached by search strategies, represented by a significantly higher amount of centered search (easy: $M = 10.11$, $SD = 5$, difficult: $M = 7.6$, $SD = 4.1$) $t(52) = 3.545$, $p \leq .001$, $r = .44$, backward search (easy: $M = 6.68$, $SD = 4.3$, difficult: $M = 4.1$, $SD = 2.8$), $t(52) = 5.478$, $p \leq .001$, $r = .6$ and forward search (easy: $M = 1.87$, $SD = 2.1$, difficult: $M = 1.2$, $SD = 1.2$), $t(52) = 2.359$, $p \leq .05$, $r = .31$ regression strategies in easy target words.

Additional repeated measures ANOVAs tested for differences in the use of strategies for near and far targets. For near targets, word difficulty influenced the preferred use of regressions, $F(3.866, 201.030) = 26.818$, $p \leq .001$, see figure 22. Degrees of freedom were again corrected using Greenhouse Geiser estimates of sphericity ($X^2(20) = 285.519$, $p \leq .001$, $\varepsilon = .354$). Subsequent t-tests indicated that for near targets the number of single shot regressions dramatically increased with difficult targets, $t(52) = -6.187$, $p \leq .001$, $r = .65$ (easy: $M = 4.3$, $SD = 3.2$, difficult: $M = 7.47$, $SD = 3.5$). The resulting lower proportion of search strategies is reflected in a significantly reduced use of centered search regressions toward difficult (M = 2.02, SD = 2.1) compared to easy targets ($M = 4.2$, $SD = 3.2$), $t(52) = 5.544$, $p \leq .001$, r = .61, whereas the proportion of backward searches remained very similar, $t(52) = .237$, $p = .814$.

With far targets, the ANOVA again revealed a significant influence of word difficulty on strategy, $F(3.218, 167.315) = 14.651$, $p \leq .001$, also with corrected degrees of freedom ($X^2(20) = 327.054$, $p \leq .001$, $\varepsilon = .389$), see figure 23. A higher proportion of accurate regressions was reflected in a significant higher number of goal directed regressions in difficult ($M = 3.0$, $SD = 2.4$) than in easy targets ($M = 2.2$, $SD = 2.2$), $t(52) = -2.473$, $p \leq .05$, $r = .32$, while the number of single shot regressions remained constant $t(52) = -1.816$, $p = .075$. Furthermore, the number of backward searches increased with easy targets ($M = 5.7$, $SD = 3.9$, difficult targets: $M = 3.2$, $SD = 2.3$), $t(52) = 5.770$, $p \leq .001$, $r = .62$, but the number of centered searches was not different between easy and difficult targets $t(52) = 0.766$, $p = .447$.

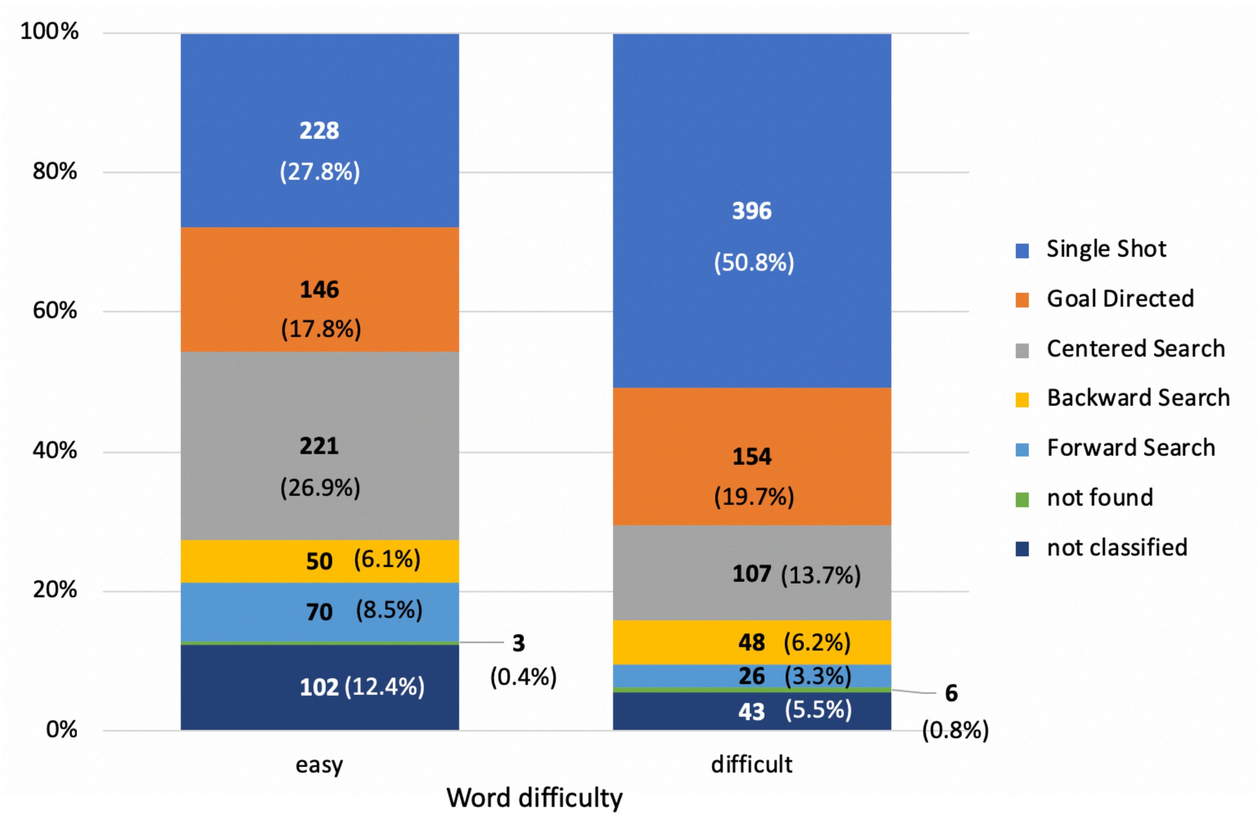


Figure 22. Distribution of regression strategies (both number of cases and percentage) as a function of word difficulty for near targets.
Data for easy words include 820 executed regressions toward targets, while results for difficult words are based on 753 regressions aimed at target words.

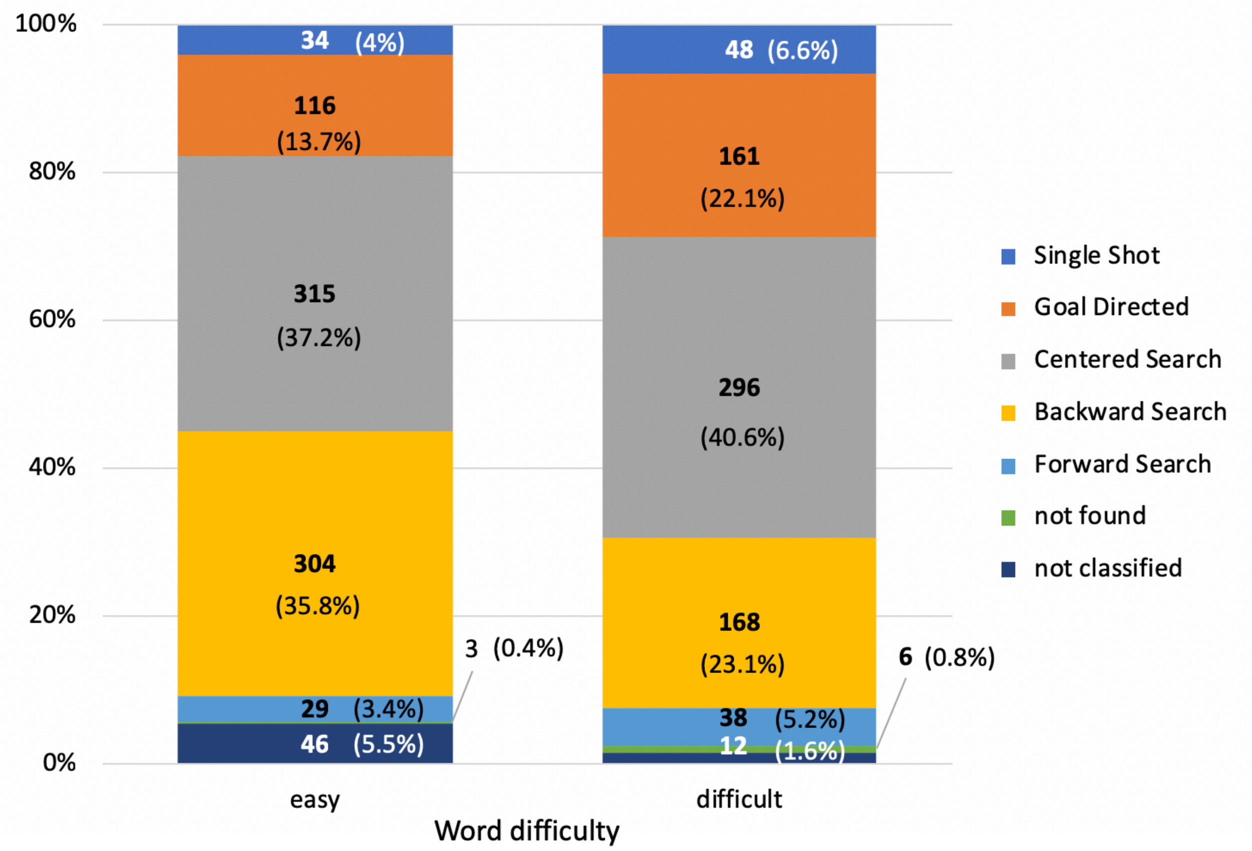


Figure 23. Distribution of regression strategies as a function of word difficulty for far targets.
Data for easy words include 847 regressions toward targets, while results for difficult words are based on 729 regressions aimed at target words.

4.4 Discussion

In this experiment, the difficulty of target words was manipulated to examine the influence of word processing difficulty on the ability to store information about word positions in memory and use this information to program accurate long-range regressions. Difficult words had a lower lexical frequency as well as a lower frequency of their initial letter trigrams. As expected, this caused large differences in temporal eye movement measures reflecting early-, intermediate- and later-pass reading. Difficult words substantially increased the precision of long-range regressions, being more often targeted by accurate regressions. Consequently, less pronounced search processes took place, represented in fewer regressions and less regression time to attain difficult targets in comparison to easy targets.

The findings of this experiment again confirmed an effect of target distance: initial regressions are longer for far than for near targets, representing spatial selectivity of long-range regressions and the retrieval of word location information from spatial memory. At the same time, far targets are more error prone, represented by a higher number of regressions and a prolonged total regression time. These results are in line with the findings from experiment 1 and experiment 2 and earlier findings from Inhoff & Weger (2005), Weger & Inhoff (2007) and Guérard and colleagues (2012).

The effect of processing difficulty

The effect of spatial selectivity of initial regressions is substantially more pronounced for difficult targets. Apparently, the ability to adapt the length of the initial regression to the distance of the target word increases with the processing time spend on a word, which is in turn based on an increasing word difficulty. Far and difficult targets lead to longer regressions than far and easy targets, whereas near and difficult targets evoke shorter regressions than near and easy targets, leading to more regressions attaining the word without any need of landing position correction. These findings are the first to indicate that the complexity of single words affects the way how spatial word location information is stored in visuo-spatial working memory and utilized for the programming of regressions.

One possible explanation for this mechanism could be that difficult words are better represented in memory because of their higher processing time. This assumption goes hand in hand with classical findings about the remembering of easy and difficult materials. Prior work on the *desirable difficulty effect* suggests that difficult learning conditions, which engage the learner more deeply and slow learning speed, can have a positive effect on memory retention and transfer after training (Bjork, 1994; Bjork & Bjork, 2014). This effect can also be transferred to reading. When reading, gaps in text coherence, which are making reading more difficult and challenge readers to consider inferences and draw conclusions, are advantageous for learning, at least for readers with adequate background knowledge (Kintsch, 1994). For competent readers, texts with insufficient coherence can activate prior knowledge, helping to anchor new information into memory, whereas readers with little background knowledge benefit most from well-defined text coherence (Fulmer, D'Mello, Strain, & Graesser, 2015; McNamara, Kintsch, Songer, & Kintsch, 1996; McNamara, 2001). General legibility of materials may produce a comparable effect. Diemand-Yauman, Oppenheimer, & Vaughan (2011) found that a hard-to-read font improved the retention of information from texts compared to an easy-to-read font. The deeper processing of the font, while creating disfluency in reading, was associated with better retention of information. A similar effect has also been suggested for a font that was specifically created for students to enhance memory processes and better remembering notes (*sans forgetica*, https://sansforgetica.rmit). In the context of the present experiment it could also be the case that location information is more easily retrieved for difficult words because of deeper processing during the inflated prior processing time.

Rather than being a product of word-level processing, the regression advantage for more difficult words might also originate on the level of sentence processing. It may be that difficult words are preferentially stored in spatial memory because it is more likely that comprehension difficulties arise in relation to these words. Consequently, regressions need to be programmed more frequently back to such locations to clarify any issues. In contrast, easy words leave less explicit traces in spatial memory, because they are not likely associated with comprehension issues and necessary repairs. In harmony with this view, prior work has shown that sentences with lexical and syntactical ambiguities (Mitchel et al., 2008; Rayner & Duffy, 1986; Frazier & Rayner, 1982) provoke more regressions from the disambiguating sentence region back to an ambiguous target word.

A final piece of evidence supporting this view comes from work on the "inhibition of return" (IOR) phenomenon, where the return of attention to a target that has been attended before leads to an increased latency compared to a target that has not been attended previously (Posner & Cohen, 1984). This effect has also been applied to reading, based on longer fixation durations prior to regressions back to previously fixated words in comparison to previously skipped words (Rayner, Juhasz, Ashby, & Clifton, 2003). As recently shown by Eskenazi & Folk (2017), this IOR-like effect is virtually absent for regressive saccades caused by comprehension difficulties in the case of ambiguous words. This result points to the possibility that readers could have a sense that difficult words may lead to uncertainties in understanding and that retaining information on their location may benefit the resolution of later comprehension difficulties.

Indirect mechanism of spatial memory on regression programming

As discussed earlier, the idea of an indirect mechanism assumes that when linguistic processing is particularly difficult, this may draw resources form a common pool, reducing the available capacity for the storage of spatial information. At first sight, the results of experiment 3 appear to strongly contradict any such mechanism, because regressions towards difficult words enjoyed a benefit rather than suffering in terms of accuracy. However, in defense of the indirect mechanism hypotheses, it could be argued that the task used here was not challenging enough to provoke competition for limited cognitive resources. To address this objection, a follow-up experiment should use more complex sentence materials, e.g. containing syntactic ambiguity of sentences, in turn to causing misleading interpretations and comprehension difficulties related to certain target words. These target words should again be easy or difficult, to examine whether complex sentence structure hinders or facilitates effective memory traces for difficult words.

Finally, according to Diemand-Yauman, Oppenheimer, & Vaughan (2011) the influence of disfluency caused by processing difficulty on the ability of retention from memory could follow a U-shaped curve, with the best effects on retention, when fonts are neither too difficult nor to easy. A similar pattern may emerge for the storage and availability for spatial information of word localization during reading, if local and global (indirect) mechanisms act together. A moderate level of difficulty (as used in the present experiment) could lead to the most effective storage of spatial location information, whereas reading materials that are either too difficult or too easy might result in insufficient storage of spatial information about word positions. In any case, experiment 3 has opened an interesting new avenue of research that should be explored more fully in subsequent work.

5 Experiment 4: The effect of sentence boundaries on long-range regressions

5.1 Introduction

In all previous experiments of the present dissertation thesis (experiments 1-3) sentences occupied exactly one line of the display, resulting in an identical spatial position of target words within sentence and line. While this design provided a highly successful framework for the analysis of regressions and strategies of re-reading, it necessarily confounded position within sentence and position on the line of text. To address this confound, a design was needed than can disentangle the linguistic aspect of distance from the spatial distance within an array of horizontally arranges letter strings.

Evidence for a strong influence of sentence boundaries on eye movement control comes from research about wrap-up effects, beginning with the now classical work by Just and Carpenter (1980). In their reading model, reading and understanding of more than one sentence consists of a recurring circle of five major stages: get next input, encoding and lexical access, case role assignment, inter-clause integration, and sentence wrap-up. The process of wrap-up includes checking for any inconsistencies and misunderstandings, assigning case roles that have not yet been assigned at this point and constructing further inter-clause relations. The last word of a sentence is central for this process, because it is the location, where ambiguities are often solved and the period at the sentence end signals unmistakably the end of a thought. Several eye movement experiments showed that punctuated words (followed by a period or comma) received longer gaze durations, compared to unpunctuated words (Just & Carpenter, 1980; Rayner, Sereno, Morris, Schmauder, & Clifton, 1989; Hill & Murray, 2000; Rayner, Kambe, & Duffy, 2000; Kennedy et al., 2003; Hiratoni, Frazier & Rayner, 2006; Payne & Stine-Morrow, 2012), longer single fixation durations (Pynte & Kennedy, 2007; Payne & Stine-Morrow, 2012), more refixations, more regressions (Rayner, Kambe, & Duffy, 2000; Payne & Stine-Morrow, 2012) and longer go-past times[1] (Hiratoni, Frazier, & Rayner, 2006; Payne & Stine-Morrow, 2012).

[1] The term go-past time is also referred to as regression path duration and includes the duration of all fixations made on a word together with the duration of all refixations made on earlier words until the word is left to the right for the first time.

These findings have been interpreted in two different ways. First, the *semantic integration hypothesis* explains the additional time spent on the last word before the punctuation sign in terms of a process of integration into the memory representation of previously read text, updating the accruing discourse model und re-checking the integrity of comprehension. As a result, the mental workspace is cleared after leaving the sentence and processing resources are available for new information (Just & Carpenter, 1980; Rayner et al., 2000).

An alternative way of interpreting wrap-up effects was offered by the *dwell-time hypothesis,* driven by findings in some studies that increasing the complexity of semantic or syntactic structures did not lead to longer wrap-up times. There was also no interaction between punctuation (a word followed by a punctuation mark or not) and mental workload generated by the reading materials (Hill & Murray, 2000; Hiratoni, Frayzer, & Rayner, 2007). According to the hypothesis, prolonged gaze durations on words before punctuation marks could have their cause in low-level mechanisms, representing prosodic features like a short pause after a comma, normally produced during reading aloud, but transferred to silent reading mode (implicit prosody).

So far, the question if and how punctuation marks can change the programming of regressions was only addressed in one study by Kennedy et al. (2003). In their experiment (experiment 1), they used sentences containing two clauses either separated by a full stop (period) or concatenated by the conjunction “and”. After each sentence, a question was asked that contained a pronoun either referring to the first or to the second sentence. Sentences with a full stop contained a blank space of three letters length after the period to ensure the same sentence length in both conditions. Regression programming was not influenced by this manipulation, the probability of making a long regression out of the question into clause one was equal in both conditions and length and the accuracy of regressions did not differ between conditions.

Evidence on the influence of the line as a spatial attribute is quite limited and indirect. Experiments examining regression accuracy in multiple line experiments showed that the number of lines appeared to influence regression programming. The effect, that regressions to targets in the first part of a sentence are less precise than to targets in the second part of the sentence no longer exists, when regressions are executed from the second line of a two-line experiment (Guérard, Saint-Aubin, Maltais, & Lavoie, 2014). In contrast, Inhoff and Weger

(2007) found that in their two-line experiment the effect of spatial selectivity for targets in the first vs. last region of the sentence was similar for both lines, but that correcting regressions (when needed) were more accurate in the first line than in the second line. Another finding illuminating the importance of word position within the line for regression programming was reported by Mitchell, Shen, Green, & Hodgson (2008). They compared the frequency of regressions when a word is placed on the last position of a line (late linebreak) compared to the same word at the beginning of a second line (early linebreak) and found that the word placed at a line-final position evokes four times as many regressions. Mitchell et al. concluded that not the word itself is crucial for the decision to program a regression but its position at the end of the line. Unfortunately, in all three experiments, the used materials included one sentence presented over two lines, rather than two or more sentences. Therefore, it is impossible to distinguish whether layout effects and/or linguistic factors were responsible for the observed results.

In the present experiment, long-range regressions within one sentence were compared to regressions across a sentence boundary within a similar spatial layout. Participants read either one long or two short sentences, providing the same amount of semantic information, but different syntactic conditions. Items consisted either of two clauses connected with the conjunction "und" (and-condition) or two short sentences separated by a period (period-condition). After reading each item, a probe word was shown at the sentence end, indicating either a near (within the sentence in both conditions) or a far target word (within the sentence in the and-condition or in the previous sentence in the period-condition). The task was to check for a subsequently introduced spelling mistake, a procedure well established for evoking long-range regressions (see experiment 1, 2 and 3 of the present dissertation thesis). Sentences were followed by comprehension questions at random points in the experiment (about every sixth sentence), to ensure reading for understanding.

Based on the semantic integration hypothesis (Rayner, Kambe, & Duffy, 2000) it can be expected that regressions to far targets are more error-prone and less accurate in the period condition. With updating the discourse model and checking for understanding of all propositions, spatial information should be cleared from working memory, since understanding of the sentence is completed (Just & Carpenter, 1980). This should also diminish the memory trace for word position, resulting in reduced regression accuracy.

Consequently, items with punctuation marks should also result in a higher proportion of search strategies.

In contrast, the dwell time hypothesis (Hirotani, Frazier, & Rayner, 2006) would not predict an interaction between the two syntactic conditions. More specifically, near and far targets may be represented in spatial memory based on physical distance, irrespective of whether a sentence boundary has to be crossed or not. The use of different regression strategies should also not be influenced by the conditions in experiment 4. Such a result would clearly contradict the semantic integration hypothesis, suggesting that linguistic processing of sentence boundaries does not interfere with the control of long-range regressions.

5.2 Methodology

Participants

Data of 50 students (42 female) was analyzed in this experiment. Mean age was 22;1 years (Min = 17;10 years; Max = 30;3 years). All participants were native German speaker and had normal or corrected to normal vision. They received course credits for participation.

Materials

The experiment consisted of 100 one-line items displayed individually on the screen. The to-be-read materials could either be a sentence consisting of two clauses, connected by “und” (and-condition) or two sentences, separated by a period (period-condition). Sentences for the and- and period-condition were always identical, with the difference only in punctuation, as displayed in figure 24, 1). In the period-condition (= two sentences separated by a period) the second sentence started with a personal pronoun, either “er”, “sie” or “es” (english: “he, she, it”, two sentences started either with “der” or with “man”, see figure 24, 1b). In the and-condition (= two clauses connected with an “und”) the personal pronoun was omitted and replaced by an “und”, which doesn´t affect the grammatical well-formedness of the sentences in any way. This leads to either identical length of both conditions (sentences with “er” [he] and “es” [it]) or sentences with a length prolonged by only one letter space (sentences with “sie” [she]). Altogether, the length of one line (either one long or two short sentences) ranged from 68 to 80 letter spaces and items contained ten to 14 words.

Every item contained a target word, either at the beginning or at the end of the line. Targets were nouns of mid-frequency, ranging from 5.4 to 97.6 per million, with a mean of 39.9 (dlex DB, Heister et al., 2011) and were either six or seven letters long. Each target was preceded by an adjective with a length between five and eight letters.

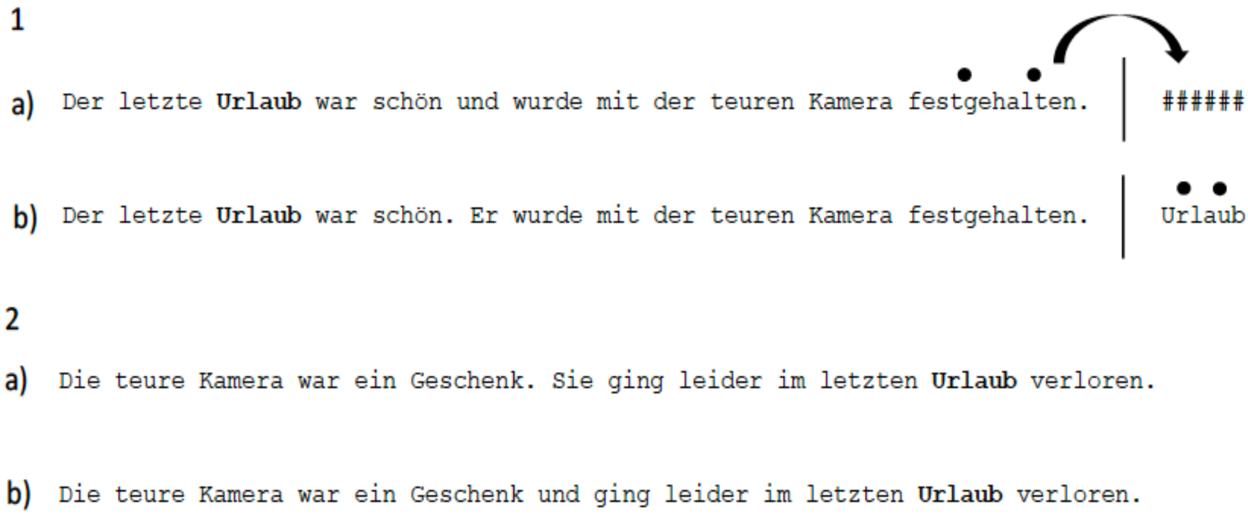


Figure 24. Sentence materials out of experiment 4.
Condition 1 represents sentence frames with far targets (far from the point of execution of the regression). Two clauses are either connected with the conjunction “and” as shown in 1a) or separated by a period, as shown in 1b). Condition 2 represents sentence frames with a near target, again the two clauses are either connected by the conjunction “and” (2a)) or separated by a period (2b)). Target words are not marked in any way during the course of the experiment. Condition 1 illustrates the invisible boundary. When crossing this boundary with a saccade, the probe word at the end of the sentence gets visible, which was masked through hashmarks beforehand.

Target words are presented twice to each participant, once at a beginning position (first third of the item, but never the first word) and once at an end position (last third of the item, but never the last word) In addition, one of each target is presented in the and-condition and one in the period-condition. As seen in figure 24, a participant read either 1a) and 2a) or 1b) and 2b).

32 out of the 100 items had their target in the middle region (middle third), to avoid any training effect for a certain position. Therefore, targets could appear at any position in the item, except as the first or last word.

Errors were introduced into the target words in 50 percent of the items after the sentence had been read, with crossing the invisible boundary, only to make the task meaningful for the participants and evoke long-range regressions back into the line to check whether the target has now a spelling error or not, a procedure well established in previous experiments (see Experiments 1-3, chapters 2.2, 3.2 and 4.2).

Altogether eight experimental lists were constructed with the factors punctuation (period vs “und”), target distance (near vs far), and error condition (error vs no error in target word).

Procedure

The procedure in experiment 4 was the same as in experiment 3, please see the methods section of experiment 3 (chapter 4.2) for details.

Data selection and analysis

The oculomotor measures for regression accuracy were again the length of the initial regression, the number of regressions, that were needed to attain the target word and the total regression time, which represents the time from initiation of the initial regression until the target was found. The successful programming of regressions was further specified by the use of several regression strategies. Classification of regression strategies followed the procedure used in experiment 1, 2 and 3, dividing trials into *single shot regressions, goal directed regressions, centered searches, backward searches,* and *forward searches* and are described in detail in the methods section of experiment 1 [chapter 2.2].

Saccades and fixations were recorded during the experiment using the SR EyeLink1000® video-based eye tracking system (SR Research, Toronto, Canada). Output from the camera system was classified and visualized using the software suite EyeMap (Tang, Reilly, & Vorstius, 2012) the output of which was imported into SPSS. During this procedure data were also inspected visually to detect any possible problems and data from practice trials were eliminated from further analyses. As in experiment 3 all three parameters had to be log-transformed for statistical analyses to better fit normal distributions. To remove outliers, the following exclusion criteria were used: for length of initial regression outliers with log values of <2.0 (0.8 percent), for number of regressions outliers with log values of > 2.5 (1.2 percent), and for total regression time outliers with log values of < 4.5 (1.3 percent) were excluded.

Statistical inferences about effects on length of first regression, number of regression and regression time were based on linear mixed models (LMMs), specifying participants and items (=sentences) as random effects. All effects were estimated using the lmer function from the lme4 package (Bates, Mächler, & Bolker, 2015) in the R environment for statistical computing (R Core Team 2014; version 1.0.136). Figures and tables presented use non-transformed values for ease of interpretation.

Initial models included the fixed effects position and punctuation and the two-way interactions between position and punctuation and a maximal random effect structure for subjects and items. Models were simplified through the removal of fixed effects that did not improve the model, and simplifying the random effect structure that did not diminish the quality of the model.

For random effects of all three accuracy variables, maximum likelihood comparison revealed, that only the position effect was needed for the random factors subjects and items and that the effects of punctuation did not improve the model. For fixed effects punctuation and its interactions (punctuation and position) were removed. This had no significant influence on the model. Inclusion of punctuation and its interaction with position did not improve the model, yielding a best fitting model with only position as a main effect.

Eye movement measures

Differences in processing times of words which mark either the end of the first clause (and-condition) or the first sentence (period-condition) were examined using several eye movement measures reflecting early, intermediate and later pass reading processes (see Radach & Kennedy, 2004 and Radach & Kennedy, 2013 for overviews). For early pass reading initial fixation duration was measured, reflecting the length of the first fixation on a word. Intermediate reading was reflected by gaze duration (the summed duration of all fixations on a word) and proportion of refixations on a word (frequency of additional fixations after the first fixation of a word), both before the word was left to the right for the first time. Based on the assumption, that the wrap-up effect reflects understanding processes and assigning case roles, differences between the conditions should show up most clearly in later pass reading measures. Go past time, proportion of regressions, number of gazes, and total viewing time are used to measure differences on that reading stage. Go past time is the summed duration of all fixations made on a word until it is left to the right, including all refixations made on earlier words. The proportion of regressions reflects the frequency of making regressions out of the word and the number of gazes counts the number of passes needed to process a word. The total viewing time is the duration of all fixations made on a word. The processing of a challenging word can lead to spill-over effects, describing the influences of processing difficulty of a word N on the processing time of the next word N+1 (Rayner & Duffy, 1986). To capture these effects, the following variables are also reported: next fixation duration (duration of the first fixation on the next word), gaze duration N+1 (gaze duration of the next word), and gaze duration next region (summed gaze durations of the next two words [N1 + N2]).

Of all sentence or clause final words 414 (13.2 percent) were not fixated and had to be excluded from analysis. All eye movement measures had to be log-transformed to achieve a normal distribution of the data. Outliers with log-values of >4 and <6.5 (1 percent) for initial fixation duration, of >4 and <7 (0.7 percent) for gaze duration, of >5.5 and <8 (3.5 percent) for go past time, of >4 and <7 (0.9 percent) for total viewing duration, of >4.5 and < 6.5 (1.6 percent) next fixation duration, of >4 and <6.5 (0.9 percent) for duration of N+1 and of >4 and <7 (0.4 percent) for duration for the next region (N1 + N2).

5.3 Results

Wrap-Up Effects based on the punctuation condition

To examine the effect of wrap-up processes, eye movement measures concerning the last word of the first clause before the “und” (in the and-condition) and the last word of the first sentence before the period (in the period-condition) were compared. A summary of basic eye movement parameters based on word means is provided in table 9. It was tested with LMMs whether the temporal parameters initial fixation duration, gaze duration, total viewing duration and go-past time significantly differed for the last word in the two conditions. None of the parameters reached significance level, with all *p*s>.05 (see table 9), revealing no difference between words in the and- and the period-condition in temporal parameters. Furthermore, the non-temporal parameters number of gazes, refixation probability and number of regressions out of the critical word were estimated. Words followed by a period did not receive a higher number of gazes compared to words followed by an “und” (p>.05).

A significant difference between both conditions was found in the number of refixations in the first gaze (b = 0.048, SE = 0.016, t = 3.092, $p \leq .01$). Words followed by a period received significantly more fixations than words followed by an “und” (see figure 25). The number of regressions out of the critical word did also differ between both conditions. Significantly more regressions were programmed out of a word followed by a period (b = 0.055, SE = 0.012, t = 4.424, $p \leq .001$) than out of a word followed by an “und” (see figure 26). To examine spill-over effects we looked at the word to the right of the critical word (N+1) and the duration of the next fixation, after the last saccade was leaving the critical (last) word. Both variables revealed no difference between the two conditions (*p*s>.05). Since word N+1 is never longer than 3 letters and hence not fixated in 55 percent of all cases, the gaze duration of the next region was additionally examined. This region included the two next words following the critical word (N+1 and N+2). The gaze duration of that region is significantly longer in the period than in the and-condition (b = 0.043, SE = 0.018, t = 2.354, $p \leq .05$), indicating a spill-over effect of wrap-up processes on the following region (see figure 27).

Table 9. Means and standard errors (in parenthesis) for punctuation-condition and results of the linear mixed model analysis for basic eye movement parameters.

	M (SE)		Estimate	SE	/t/-value
	And	Period			
Initial fixation duration	223 (2.1)	221 (2.2)	-0.024	0.017	-1.449
Gaze duration	272 (3.5)	283 (4.1)	0.005	0.018	0.25
Go past time	740 (38)	767 (27)	0.039	0.052	0.759
Total viewing duration	358 (5.1)	358 (5.2)	-0.02	0.022	-0.903
Number of gazes	1.1 (0.01)	1.1 (0.01)	0.003	0.01	0.326
Proportion of refixation	0.2 (0.01)	0.3 (0.01)	0.048	0.016	**3.092****
Proportion of regressions	0.08 (0.006)	0.13 (0.008)	0.055	0.012	**4.424*****
Next fixation duration	225 (4.5)	226 (4.5)	0.008	0.016	0.514
Gaze duration N+1	189 (2.4)	204 (2.5)	0.016	0.025	0.622
Gaze duration next region	298 (3.4)	315 (3.6)	0.043	0.018	**2.354***

M=mean, SE = standard error, *= p≤.05, **= p≤.01, ***= p≤.001.

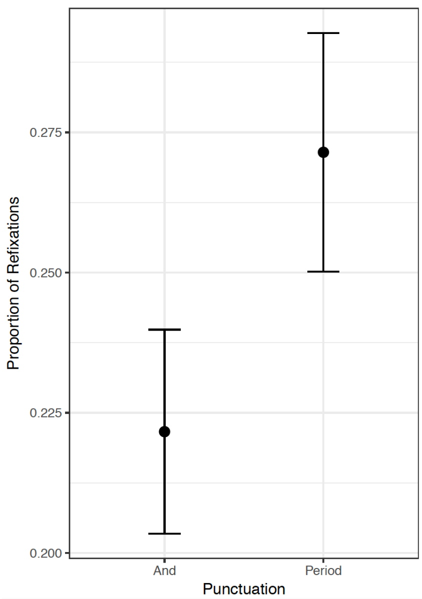


Figure 25. Proportion of refixations as a function of punctuation.

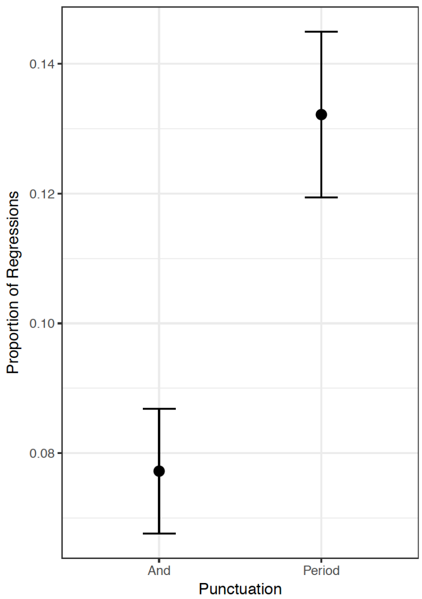


Figure 26. Proportion of regressions as a function of punctuation.

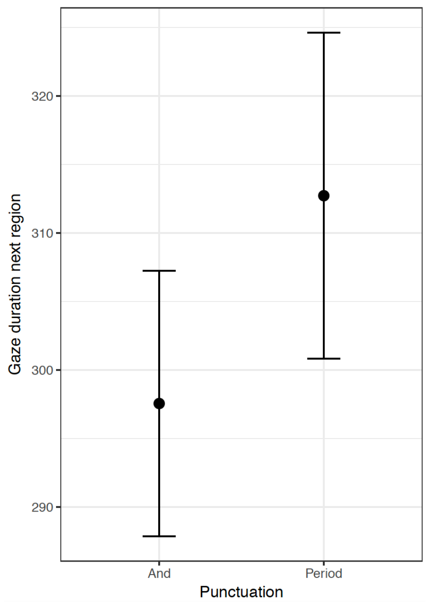


Figure 27. Gaze duration of the next two words following N as a function of punctuation.

Accuracy measures

Table 10 provides means and standard deviations in both conditions for all dependent variables. The accuracy of regression programming was examined regarding the length of the first regression. The length of the primary regression was strongly influenced by target distance and longer for far (*M* = 37.1, *SE* = 0.4) than for near (*M* = 27.1, *SE* = 0.3) targets (*b* = 0.325, *SE* = 0.035, *t* = 9.263, *p*<.001). The mean difference between a regression directed to a far than to a near target word amounts 10 letter spaces (LS), a difference comparable to the results obtained in experiment 1 and experiment 3. However, differences in punctuation led to no significant differences in the length of the primary regression (*b* = 0.002, *SE* = 0.014, *t* = 0.19) with nearly identical values for the period- (*M* = 32.0, *SE* = 0.4) and the and- (*M* = 32.3, *SE* = 0.4) condition (see figure 28). An interaction effect between position and punctuation could not be found (*b* = -0.03, *SE* = 0.028, *t* = -1.11).

Correcting processes, necessary in cases when initial regressions are not attaining the target word are indicated by the number of regressions and the time it takes to find the target word (total regression time). The number of regressions used to correct a landing position of a first regression not hitting the target was not significantly influenced by the manipulation of punctuation. There is a tendency of an effect reflecting a more laborious correcting process when two sentences are separated by a period compared to the and-condition (*b* = 0.03, *SE* = 0.017, *t* = 1.9, *p*=.0564). Position and punctuation showed no interaction (*b* = -0.013, *SE* = 0.034, *t* = -0.394), but number of regressions differed highly significant with target position (*b* = 0.536, *SE* = 0.04, *t* = 13.13, *p*<.001). Participants needed approximately 1.2 regression more to reach a far target (*M* = 3.5, *SE* = 0.04) than a near target (*M* = 2.3, *SE* = 0.04) if the initial regression did not attain the target word immediately (figure 29).

The total time it took to find the target word was not influenced by punctuation condition (*b* = 0.03, *SE* = 0.023, *t* = 1.451), but was significantly larger for far (*M* = 867 ms, *SE* = 8.6) than for near (*M* = 598 ms, *SE* = 8.9) targets (*b* = 0.426, *SE* = 0.03, *t* = 13.583, *p*<.001), with an average additional searching time of 269 ms for far targets (see figure 1). An interaction between position and punctuation did not emerge (*b* = -0.02, *SE* = 0.033, *t* = -0.738), see figure 30.

Altogether, the presence vs. absence of punctuation did not influence the adaption of initial regression amplitude to target distance. The correction of misguided initial regressions (number of regressions and total regression time) was also not significantly affected by this manipulation. However, comparison of number of regressions between both conditions suggests a tendency for somewhat more laborious correcting processes for targets in the period condition.

Table 10. Means and standard errors of length of initial regression, regression error, number of regressions and total regression time for target position and punctuation condition.

		Total		Near		Far	
		And	**Period**	**And**	**Period**	**And**	**Period**
Length of initial Regression	*M*	32.3	32.0	27.1	27.1	37.5	36.8
	SE	(0.4)	(0.4)	(0.5)	(0.4)	(0.5)	(0.5)
Number of Regressions	*M*	2.88	2.94	2.26	2.3	3.5	3.6
	SE	0.04	0.04	(0.06)	(0.05)	(0.05)	(0.06)
Total Regression Time	*M*	727.4	736.5	592.5	602.4	861.7	872.2
	SE	9.4	9.5	13.1	12.1	11.5	12.9

M = Mean, SE = Standard Error, length of initial regression is given in letter spaces and total regression time in ms.

Regression Strategies

The classification of regressions into five categories is based on the earlier work described in the chapters on experiment 1 and 2. As before, regressions were divided into two relatively accurate and three searching strategies. The strategy of single shot regressions was observed in 20.9 percent of all cases, where participants attained the target word with one accurate regression. Another 18.3 percent of the cases were classified as accurate regressions, needing only one small correction to attain the target word. Taken together, 39.2 percent of all cases represent very well and accurate programmed regressions.

Another 54.1 percent of all regressions were assigned to searching strategies. Most of them represented cases of centered search, starting in the center of the sentence (32 percent), followed by backward scanning regressions (15.7 percent). Only a small proportion of the observations consisted of forward scanning regressions (6.4 percent), starting at the beginning of the line and scanning from start to end. Another 5.7 percent of all cases could not be accounted for with the regression strategies described above and remained unclassified. Finally, 1.2 percent of all targets were excluded from analyses because participants pressed the button before looking at the target word.

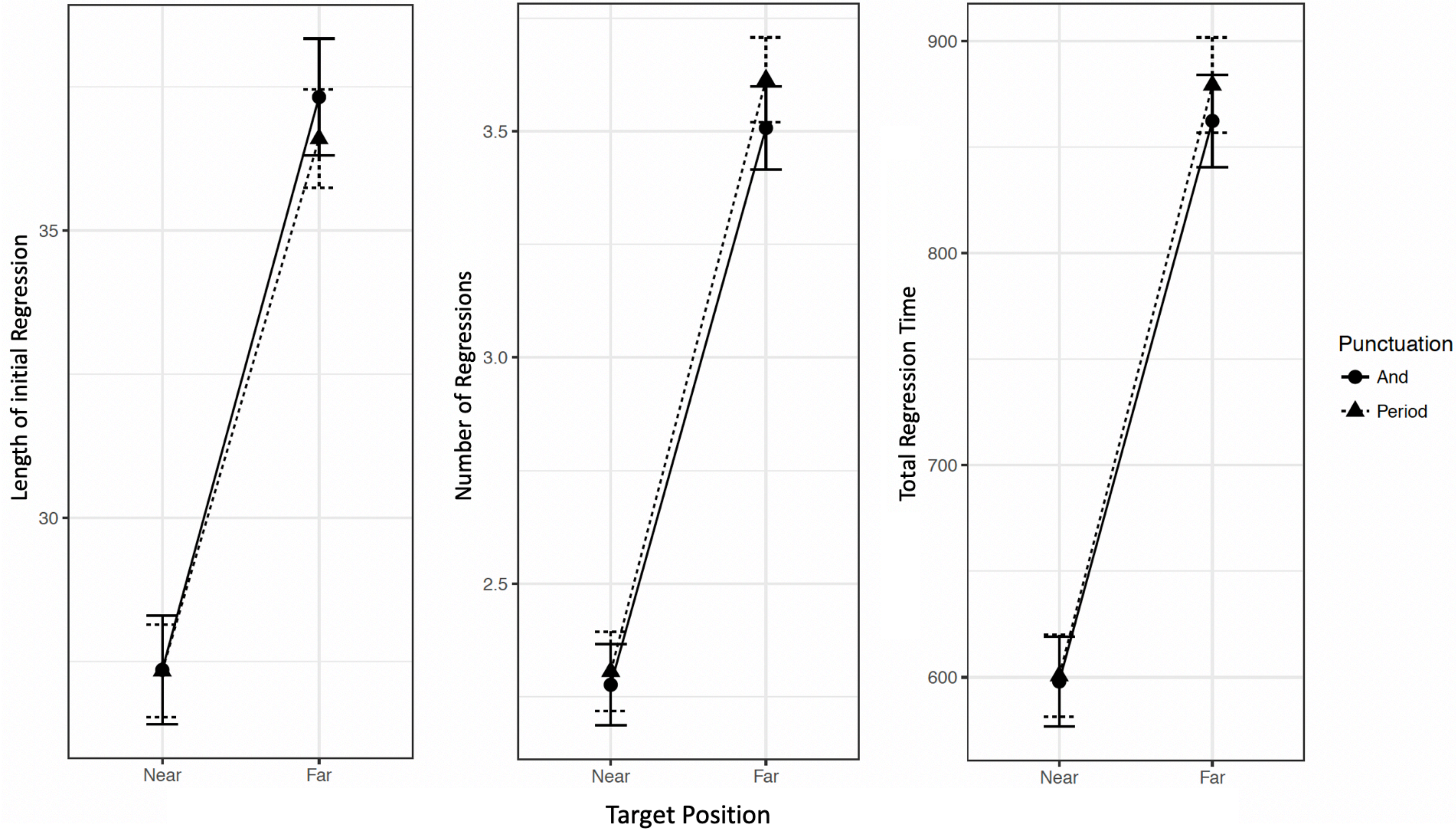


Figure 28. Length of the initial regression as a function of punctuation condition (“period” vs. “und”) and position (near vs. far target word).

Figure 29. Number of regressions needed as a function of punctuation condition (“period” vs. “und”) and position (near vs. far target word).

Figure 30. Total regression time as a function of punctuation condition (“period” vs. “und”) and position (near vs. far target word).

A repeated measures ANOVA was carried out with the factors strategy and punctuation manipulation, with results indicating that the type of strategy used to attain the target was significantly affected by the punctuation manipulation $F(3.425, 167.801) = 2.653$, $p \leq .05$. Mauchly´s Test indicated that the assumption of sphericity had been violated, $X^2(20) = 148.073$, $p \leq .001$, therefore degrees of freedom were corrected using Greenhouse Geiser estimates of sphericity ($\varepsilon = .571$). A subsequent comparison with paired t-tests showed, that the proportion of regressions with centered search strategy was significantly higher in the " period-condition ($M = 10.8$, $SD = 4.4$) than in the and-condition ($M = 9.2$, $SD = 4.4$), $t(49) = -2.818$, $p \leq .05$, r = .37. All other regression strategies did not differ significantly between the two conditions (all *ps*>.05), see figure 31.

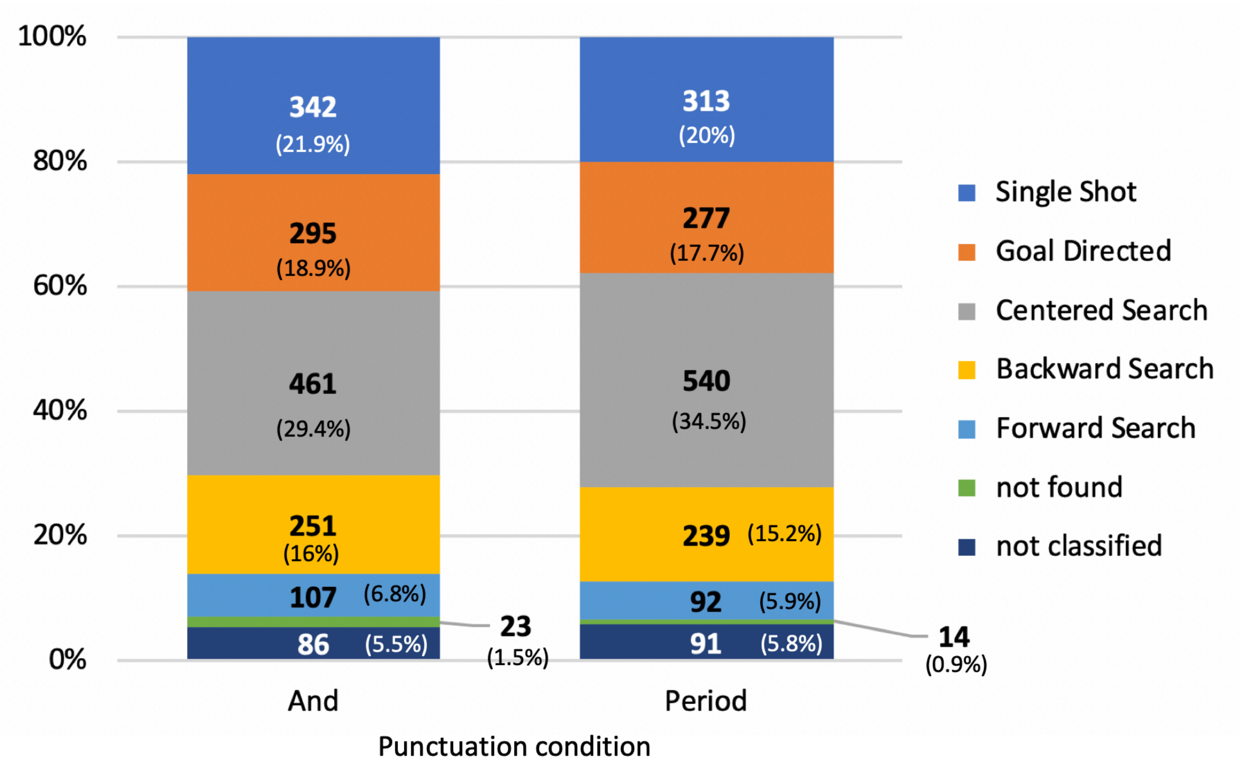


Figure 31. Distribution of regression strategies (both absolute number of regressions and percent) as a function of punctuation condition ("und" vs. period) for all targets.
The and-condition includes 1565 regressions to target words, while the period-condition includes 1566 observations

It is reasonable to assume that the punctuation manipulation should particularly influence far targets, because these targets are located in the first sentence which is separated from the second sentence with a period in the period-condition. Consequently, an ANOVA was computed to test for significant differences in regression programming for far targets only. The repeated measures ANOVA revealed that the difference in regression strategies used for different punctuation conditions did not reach the level of significance $F(3.431, 144.102) = 2.329$, $p=.069$ (see figure 32), suggesting that the effects found in the regression strategy analysis were overall effects and not specific for far targets. However, the data include a tendency towards a slightly more pronounced use of centered search in the period-condition.

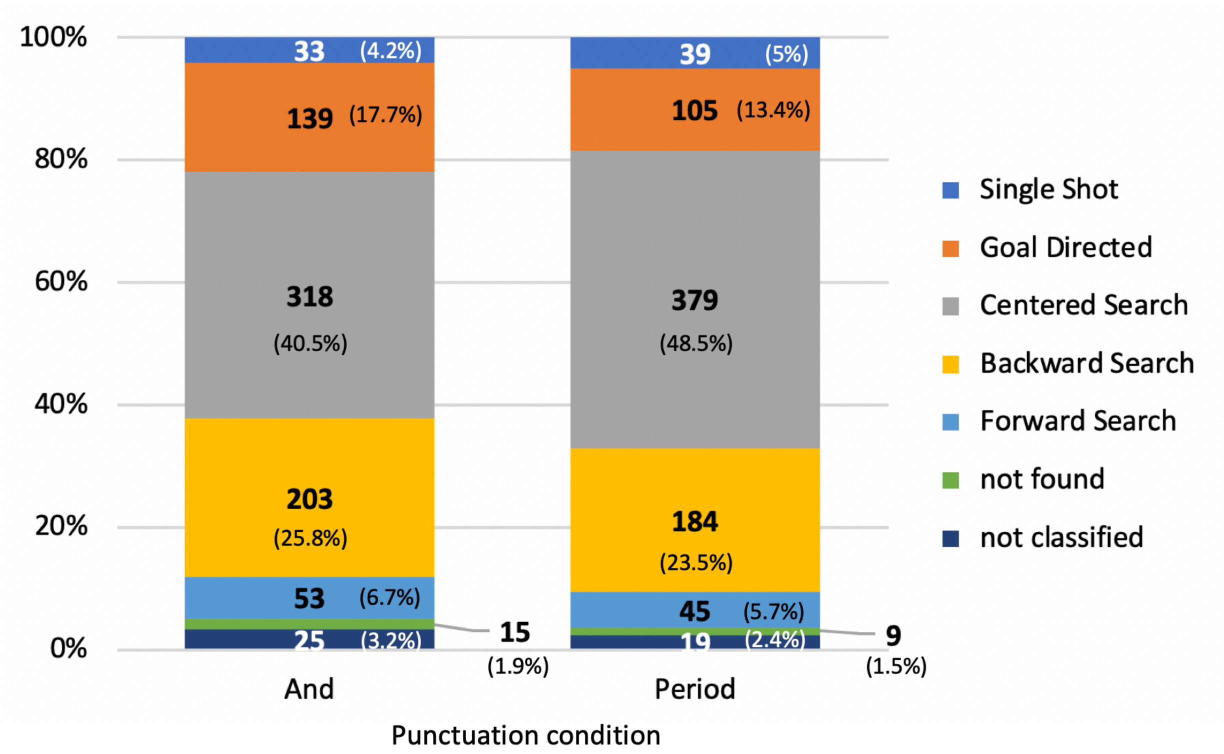


Figure 32. Distribution of regression strategies (both absolute number of regressions and percent) as a function of punctuation condition ("und" vs. period) for far targets.
And-condition includes 786 executed regressions to target words and period-condition includes 780 executed regressions to target words

5.4 Discussion

The present experiment was designed to examine whether long-range regressions across a sentence boundary differ from regressions within a sentence. Similar to prior work on sentence wrap-up effects, materials had identical semantic meanings and word order. Inserted was either a full stop (period) to create two short sentences, or an “und” to create one long sentence. This manipulation lead to considerable wrap-up effects, but did not influence the accuracy of regression programming. The strategy analysis indicated that in cases where a search was necessary, the type of searching strategy was moderately influenced by the manipulation of sentence boundaries.

Effects of punctuation manipulation

The number of refixations and the number of regressions were higher in the period (word N sentence-final) than in the and-condition (word N clause-final), revealing that more fixations were needed and more regressions were programmed out of words followed by a full stop (period) than by an “und” (cf. Rayner, Kambe, & Duffy, 2000; Payne & Stine-Morrow, 2012). The duration of the first fixation, the gaze duration and the total viewing duration were not affected by this manipulation, resulting in no significant differences between both conditions in early and late temporal eye movement measures. This comes somewhat unexpected, considering the multitude of experiments describing these effects (Just & Carpenter, 1980; Rayner, Sereno, Morris, Schmauder, & Clifton, 1989; Hill & Murray, 2000; Rayner, Kambe, & Duffy, 2000; Kennedy et al., 2003; Hiratoni, Frazier, & Rayner, 2006; Pynte & Kennedy, 2007).

However, looking the previous studies more closely, it turns out that most of the experiments compared sentence or clause final words to sentence internal words, resulting in a comparison of words being part of different sentence phrases and having completely different stress patterns during sentence reading. Comparing clause-final and sentence-final words, Payne & Stine-Morrow (2012) also found no difference on word N in temporal eye movement measures. However, in a comparison of both of these boundary conditions vs. sentence-internal words, wrap-up effects were shown to emerge.

One possibility to account for the lack of a difference in viewing time in the present experiment could be in terms of the dwell-time hypothesis (Hirotani, Frazier, & Rayner, 2006). According to this view, viewing times on words N (in both, sentence final and clause final positions) are strongly influenced by a transfer of prosody from oral to silent reading. Given the similar prosody in both present conditions, a difference in viewing times on word N appears unlikely under such a hypothesis. Moreover, even in the framework of the semantic integration hypothesis (Rayner, Kambe, & Duffy, 2000), the lack of a difference in viewing times may not be all that surprising. Both item types in experiment 4 had exactly the same length, and, even more important, semantic and syntactic complexity until word N was reached. This may be taken to suggest that there was an equal amount of immediate linguistic processing demands for word N at the sentence vs. clause ending.

It is very interesting that the proportion of refixations and regressions on word N is increased when this word is followed by a punctuation mark. Due to the necessary latency for the execution of refixations, these movements are triggered early during the cause of the initial fixation, either when the word appears difficult to process or when the eyes land on a non-optimal letter position (Deubel, O'Regan, & Radach, 2000; McConkie et al., 1989). It may well be that the punctuation mark serves as an early visual trigger that additional local processing might be necessary, while the amount of actual processing needed does not exceed the demands imposed by a clause boundary. In the case of the increased rate of regressions originating from word N in the punctuation condition, an interpretation appears even more straightforward: It makes perfect sense that the sentence boundary indicated by the punctuation mark triggered some regressions and re-reading. The fact that fixation durations before regressions are significantly reduced (Kliegl, Nuthmann, & Engbert, 2006) is likely to counteract any increased fixation durations at the sentence boundary, contributing to the observed non-difference in viewing times.

Looking at the region following the critical word (word N) after the end of the clause/sentence (N+1 + N+2), gaze duration is significantly longer in the period- compared to the and- condition. This result indicates that the expected increase in viewing times is not missing, but materialized with a delay, as it is also typical for lagged effects of high-demand lexical processing (Rayner & Duffy, 1986; Kliegl, Nuthmann, & Engbert, 2006). In addition, any immediate increase in the mental workload for word N is also likely to diminish the parafoveal preprocessing of words to the right (Henderson & Ferreira, 1990), again inflating gaze duration when these words are fixated and adding to the total, spatially distributed sentence end effect. Taken together, the present pattern of data indicates that the sentence boundary significantly affected eye movements during initial sentence reading, presumably in terms of sentence wrap-up effects that materialized within a region extending from word N to word N+2.

Regression Accuracy

Experiment 4 found that the accuracy of long-range regressions to near and far target words is not influenced by the existence of a sentence boundary within a line of text. This finding goes hand in hand with the results of Kennedy et al. (2003) who similarly reported that the precision of long-range regressions is not modulated by the type of concatenation (period or and) between two clauses (their experiment 1).

The present study can go one step further in considering the strategies used to search for targets when an accurate regression saccade cannot be executed. Here, the significant result is an increase from 29.5 to 34.5 percent in the use of the centered search strategy when a sentence boundary is present (see table 31). The analysis of individual differences reported for experiment 1 suggested that the centered search strategy is presumably preferred in the absence of any knowledge on target location. Applying this idea to the present experiment would lead to the expectation of a difference between near and far targets. If there is any effect of the sentence boundary on clearing visual-spatial memory, the resulting absence of target location knowledge should occur much more frequently in the case of far targets. However, as table 32 indicates, the difference between conditions in the use of centered search is no longer significant when looking only at far target items. Why then, is the centered search strategy used more often when a sentence boundary is present? One plausible speculation could be, that the presence of the sentence boundary may invite a larger number of initial regressions towards the center of the line, simply due to the presence of a clearly demarcated visual landmark, presumably also with high saliency in visual-spatial memory.

Line vs. sentence effect

The lack of an effect of punctuation marks on the accuracy of regressions in the present experiment appears to suggest that for visuo-spatial memory the line is a more crucial construct than the sentence. Support for this assumption comes from the “line effect”, described by Mitchell et al. (2008), arguing that word position within a sentence has a more crucial influence on regression frequency than word meaning. That could mean that clearing out working memory at a sentence final position does not include visuo-spatial information, because this information is stored independently of the sentence unit for the whole line. This would explain why the present effects were rather small. However, since this are only first hints of explanation, further experiments are needed to examine a possible line effect for the programming of long-range regressions. A follow-up study should examine sentences over two lines. The first sentence should either extend over the first line and the first half of the second line or the first line should form a closed sentence, resulting in a new sentence in the first half of the second line. This could evoke much larger wrap-up effects and increase the possibility to distinguish between line and sentence effects.

6 General Discussion

6.1 Summary of the results

Regressions are an integral part of the reading process, necessary to ensure optimal word recognition and reading for understanding. The underlying programming mechanisms for long-range regressions have been under debate, with two hypotheses facing each other. The *spatial coding hypotheses* considers spatial memory as the fundamental mechanism, whereas the *verbal reconstruction hypothesis* regards verbal memory and the reconstruction of word positions from linguistic knowledge as the key basis for the programming of long-range regressions.

Experiment 1 sought to contribute to this discussion by examining the performance of adult skilled readers in the execution of long-range regressions to far- and near-distance targets within sentences on one full line of text. Importantly, participants also completed standardized assessments of static and dynamic spatial memory as well as reading speed and reading comprehension so that inter-individual differences could be considered.

Initial regressions were larger for far than for near targets, representing spatial selectivity, while number of regressions and total regression time increased for far targets, showing the susceptibility to regression errors with increasing target distance. Spatial memory performance influenced the programming of initial regressions, whereas reading performance only guided the correction of over- or undershooting regressions, but not the programming of initial regressions.

Spatial memory performance had a profound influence on the programming of long-range regressions. A high level of performance in a dynamic spatial memory task (Corsi block tapping test) was associated with more accurate regressions (single shot as well as goal directed regressions). Lower level performers in this test needed more regressions and more time to reach near targets, an effect that disappeared with far target words. Good reading skills could have been used in the case of far targets to compensate for poor spatial memory information.

Long-range regressions in developing readers were examined in experiment 2. Using the same sentence materials as with the adult sample in experiment 1 allowed for direct comparison of regression programming between children and adults. Again, long-range regressions were spatially selective; children showed longer regressions to far than to near targets, confirming the retrieval of information about target word locations from memory. Again, far targets led to a higher number of regressions and a prolonged total regression time, validating the error-proneness associated with far targets.

Spatial memory influenced the programming of the initial regressions whereas reading performance affected the correction of misguided initial regressions, comparable to the results found in experiment 1 for adults. A high performance in the static memory test (matrices test) leaded to an increased number of highly accurate single shot regressions, especially with near targets. Furthermore, a high performance in the matrices test was associated with less additional regressions and less regression time to attain near targets, emphasizing the importance of spatial memory performance for the near periphery. Reading performance influenced the correction process especially for far targets. Presumably, the erroneous retrieval of location information of far targets was compensated to some extent with a high level of reading performance. With far targets, an effective correction process was associated with less regressions and less time for high level performer in both, reading speed and reading comprehension.

Developing readers differed from adult readers in their number of highly accurate regressions. Adult readers had a higher proportion of single shot and goal directed regressions. Moreover, they also differed in how they searched for target words that could not be attained accurately. Children preferred the use of a backward scanning strategy, whereas adult readers more often relied on the centered search strategy. Adult readers showed a generally more effective way of correcting for incorrect landing positions of initial regressions, needing fewer regressions and less time compared to children. Both adults and developing readers showed no effect in the shift condition. Shifting the line for one letter position to the left after the sentence was read but before executing a regression did not affect the length of the initial regression. This clearly confirms the use of information from memory and not from the current peripheral vision during the preparation of regression saccades.

The results of experiment 3 demonstrated that the difficulty of a word can modulate the way in which the position of a word is stored in spatial memory. As expected, words with lower lexical frequency and orthographic regularity were read with more cognitive effort as reflected in longer word viewing times. Subsequently, participants needed fewer regression saccades and less time to re-read these more difficult target words and the amplitude of initial regressions was better adapted to target distance. Regression strategy analyses showed that difficult targets were more often attained with single shot regressions. Accordingly, easier target word more often necessitated the use of search strategies, with a preference for the centered search strategy. These results indicate that when readers invest more effort into processing a word, this leads to a better representation of the word's location in spatial memory.

In experiment 4 the programming of long-range regressions within and between sentences were compared. Two clauses were either connected with the conjunction “und” to create one long sentence or separated by a period to create two short sentences, both presented as one line of text. The meaning and the syntactic structure within the two clauses were kept identical in both conditions. Results demonstrated that the need to cross a sentence boundary to reach far targets, did neither affect the accuracy of the initial regression, nor the number of regressions or the total regression time. Even the proportion of single shot and accurate regressions was nearly identical between both conditions. The fact that the centered search strategy was used slightly more often in the period-condition was most likely the result of the period being a visually salient landmark and hence an attractive regression goal.

6.2 Linking results

The results of experiment 1 and 2 showed very clearly that the programming accuracy of long-range regressions depends on the level of performance in visuo-spatial memory. Reading performance, on the other hand, only influenced the effectiveness of correcting processes in cases when the initial regression did not attain the target word. Both results support the hypothesis that the ability to program accurate long-range regressions depends primarily on the performance of spatial memory. However, spatial working memory is limited in capacity and therefore particularly vulnerable to disruption. Visuo-spatial working memory span is typically limited to about three to four items (Baddeley, 2003). Sentences used in experiment 1 and 2 included between 9 and 14 words, clearly exceeding this span. Consequently, far targets should be particularly difficult to regress back to, as the first part of the sentence may already be deleted from memory. On the other hand, the capacity to remember a string of words is up to 15 words, if words build a meaningful sentence (Baddeley, 2010). If reconstruction of word location from linguistic knowledge would be possible, better performance in reading comprehension should lead to more single shot regressions for far targets. This was clearly not the case in the present work.

A hybrid model of regression programming was first suggested by Weger and Inhoff (2007). They proposed that spatial memory guides the programming of initial regressions and that subsequent corrective saccades are guided by linguistic knowledge. Our findings support this view very clearly. They also speak against both classical hypotheses described in some detail in the introduction. The smaller number of accurate regressions (both single shot and goal directed regressions) and the more laborious correcting process in far targets argue against the *spatial coding hypothesis*, according to which all positions in a sentence are represented equally well (Kennedy et al., 2003). The non-existence of any influence of reading performance on the programming of the initial regression and on the proportion of single shot regressions, especially for far targets, argues equally strongly against the *verbal reconstruction hypotheses.* Essentially, these results rule out the idea that word location information is directly reconstructed from linguistic knowledge. Instead of being the one or the other, both mechanisms seem to work efficiently together, with spatial memory determining the programming of initial long-range regressions and linguistic information subsequently modulating correction and guiding the homing in on the target (Weger & Inhoff, 2007).

However, the results of experiment 3 suggest that spatial memory does not work independently of the linguistic features of the reading materials. Linguistic difficulty clearly influenced the way in which information about word positions is stored and made accessible for the programming of regressions. Location information about words that received longer processing times during reading are easier to access for accurate regression programming. These results show a close connection between the function of regressions in the process of reading and their underlying programming mechanisms. Long-range regressions are mainly used for the re-reading of words and text segments that had caused local comprehension difficulty and insecurity about global text understanding (Weger & Inhoff, 2013; Bicknell & Levy, 2011). Assuming that difficult words are more frequently the source of comprehension problems (and given that the capacity of the visuo-spatial memory is limited), it makes perfect sense to store the location of difficult words for potential use in the service of successful re-reading.

Experiment 3 was planned with a reasonable alternative hypothesis in mind. This hypothesis about an indirect effect of spatial memory on regression programming had been derived from the comparison of the results from experiment 1 and 2 and the results of Guérard et al. (2012, 2014). Guérard and colleagues found (experiment 2 and 3 in their 2012 paper), that when participants articulate the letters A, B, C, D (articulatory suppression task) during their experiment, the accuracy of regressions was reduced. This leads to the conclusion that verbal memory, and not spatial memory, is crucial for the programming of long-range regressions. Experiment 1 and 2 of the present dissertation thesis showed exactly the opposite.

It appears possible that due to an overload of the central executive during the articulatory suppression task in Guérard et al.´s experiments, no information about word positions could be stored into visuo-spatial memory. In their following experiment, Guérard et al. (2014) found that the effect was present particularly for targets far away from the starting point of regressions. This in turn goes hand in hand with the present findings that far targets are especially error-prone and spatial memory performance influenced accuracy and time for regressions to near targets in both adults and children. For far targets, reading comprehension compensated for poor spatial memory information. If indeed spatial memory for far targets is more error-prone because of limited capacity, it is not surprising that distant words are more susceptible to articulatory suppression, which severely limits the capacity of the central executive. It may be possible that an indirect influence of spatial memory only comes into play with additional cognitive load not directly related to the lexical processing of the reading materials. After all, Guérard et al. used a task that has nothing to do with the process of natural reading.

The key role of spatial memory discussed so far raises the question about a reference system for spatial coordinates of sentences, which can be utilized for the programming of accurate long-range regressions to prior words/text segments (Kennedy et al., 2003; Inhoff et al., 2005). Experiment 4 asked whether such a reference system should be thought of in terms of sentences or simply the current line of text. To this end, one-line items were compared that had two identical clauses either connected with the conjunction “und” or separated by a period. Following the semantic integration hypothesis of wrap-up processes, the last word of a sentence is used to ensure comprehension and program regressions when needed (Just & Carpenter, 1980; Rayner et al., 2000). Crossing this word position could go along with deleting information about word location information, because comprehension is validated and programming of regressions becomes much less likely, resulting in less accurate regressions when crossing sentence boundaries. The results of experiment 4 clearly refuted this hypothesis. Crossing the end of a sentence within a line did not affect the accuracy of the initial regression programmed form the end of the line to an earlier position within the line. This can be taken as evidence that the linguistic framework of a sentence has less weight for the programming of accurate long-range regressions than the spatial framework of a line.

However, experiment 4 only examined the programming of regressions within one single line. Re-Reading on the text level typically involves the relocation of information across multiple lines or even pages. Indeed, examining the relocation of information in a text (12 pages with 30 lines on each page), Rawson and Miyake (2002) found no correlation between relocation accuracy and visuo-spatial ability, but between relocation accuracy and verbal memory. They concluded that spatial memory is not crucial for the relocation of information within extended text, especially with respect to previous pages. It is reasonable to assume that, given the capacity limitations of spatial memory, it can only serve as the base for long-range regressions with a limited spatial and temporal range. Outside of this range, the relocation of information will have to rely on attempted reconstruction from text memory and/or employ searching strategies like those described in the present dissertation thesis. In this case, verbal memory and reading comprehension are likely to account for individual differences in relocation and re-reading.

The detailed examination of regression saccade accuracy and the proposed classification of regressions strategies provide a novel framework for the analysis of differences arising either from reader's performance or from characteristics of reading materials. One important consideration with respect to accuracy was to consider two subgroups, single shot regressions (attaining the target word with only one highly accurate regression) and goal directed regressions (an initial regression programmed into the direction of the target and only one additional saccade that attains the target). The successful programming of accurate regressions was different for near and far targets. Near targets were more often attained by single shot regressions, while far targets were more often attained with one primary regression and one corrective saccade.

The higher number of two-step regressions to far targets is certainly based on the well-known tendency to undershoot far targets whenever a saccade is executed (e.g. Becker, 1989; Deubel, Wolf, & Hauske, 1982). In the present work this basic effect was modulated by individual differences in spatial memory.

As could be expected, High-performers in terms of spatial memory showed more single shots for near targets and more goal directed regressions to far targets, reflecting their better knowledge of previously encoded word positions. At the same time, difficult target words lead to more single shots regressions in near targets and more goal directed regressions to far targets compared to easy words.

It appears reasonably to argue that that additional saccades needed to attain far targets (although knowledge about the target location presumably exists) reflects a reversed saccadic launch distance error. As a general rule, the further left a progressive saccade is launched, the further shifted to the left is the landing position within the target word (McConkie et al., 1988). This could also apply to long-range regressions. The further to the right the target word is located, the greater could be the probability of an undershot. Speaking against this hypothesis is the missing launch site effect for regressions, described by Radach & McConkie (1998), and by Vitu et al. (2001). The landing position of regressions within a word is usually not influenced by the launch site of the regression.

Looking at the present data, the following picture emerges in both Experiment 1 (for adults) and in Experiment 2 (for children). The landing position of the first fixation on a target word (following a progressive saccade) showed the classic saccade launch distance effect. In contrast, landing positions of regressions attaining the target word (irrespective of whether additional saccades were executed on the way towards the target) deviated from this pattern and showed landing positions around the center of the word (please see appendix for figures 33-36). Thus, there is the possibility that an undershot in far targets causes regressions too short to reach the target, making the programming of an additional regression necessary. These additional regressions will not show the launch distance effect since they are attaining the aimed word.

On the other hand, it can be questioned whether goal directed regressions reflect rather a search strategy than an accurate strategy. It might be possible, that the first regression was actually a regression starting a centered search strategy, directed to the center of the sentence and that the second regression happened to hit the target and following this further corrective regressions were not necessary. This possibility can´t be entirely excluded and may have occurred in some cases.

In addition to the two categories of accurate regressions, three categories of search strategies were defined in the present work: a centered search, a backward search and a forward search category. Earlier experiments described a dichotomy of strategies consisting of accurate regressions and some way of searching for target words. Frazier and Rayner (1982) argued in their seminal study on “garden path” sentences that adult readers either program regressions directly to the location causing understanding problems (selective reanalysis strategy) or that they start at the beginning of the sentence to search for the critical region (forward reanalysis strategy). None of their participants started to search from the end of the sentence. These two strategies appear to make sense in an experiment where readers can expect the occurrence of syntactic ambiguities, a scenario that appears quite remote from normal reading.

Murray and Kennedy (1988) reported that their participants either found the target with one highly accurate regression or started to backtrack from the end of the sentence. Looking at their work from today’s perspective, it appears possible that their data analysis strategy (e.g. using a small target region for accurate saccades) and discarding a large number of trials is partially responsible for their two-category view.

In all four experiments included in the present dissertation thesis, a simple dichotomy between accurate regressions and a rest category of searching never provided an appropriate account of the data. Instead, for all samples a careful visual inspection (based on a replay of viewing patterns in real time or slow motion; Tang, Reilly, & Vorstius, 2012) yielded a distinction into three search categories. Looking at these categories, one could argue that some centered searches are essentially the same as backward searches, just starting a bit more towards the sentence center. However, a strong argument for the need to distinguish between these two strategies arises from the modulation of searching process by spatial memory performance. The results of experiment 1 and 2 showed quite clearly that low-level performers in spatial memory needed to rely much more often on searching instead of accurate regressions. More importantly, less successful participants more often used the centered search strategy, while high performing participants more often relied on backward searches. Altogether, the detailed description and classification of regression strategies is an innovative approach that makes it possible to understand how participants search for and attain target words during sentence reading.

6.3 Strengths and limitations

All four experiments of the present dissertation thesis were designed to approximate natural reading as closely as possible. Despite this, the selection of regression targets was determined by the experimenter (induced by probe words at the end of the sentence) and not by the reader, as is typical for natural reading. However, preselecting the target has the advantage of having precise knowledge about the goal of the regression and hence being able to decide whether a regression has attained the intended goal or not. It is important to note that there is empirical evidence suggesting that the metrical properties of regressions in a setting similar to the present work are comparable to regression behavior in natural reading. Inhoff and Weger (2005) found that regressions in their experiments were identical in amplitude and accuracy, no matter if they were triggered by the experimenter (regression target known) or occurred naturally (regression target triggered indirectly via comprehension questions).

The decision to use the last word on the line as the starting point for the regression is supported by the fact that during normal reading most of the comprehension related long-range regressions are not programmed before the end of the sentence has been reached (Weiss et al., 2018). It is further backed by experiments suggesting that the ability to program selective regressions from the end of the sentence characterizes skilled and proficient reading (Tiffin-Richards and Schroeder, 2018). It appears therefore that the present methodology represents a reasonable approximation to a scenario typical for natural reading.

Nevertheless, it might be that the programming of regressions in natural reading is different in some detail from the situation described in experiments one to four. In all experiments, regressions did not really originate from the last word of the sentence, but from the last word of the line, located a bit to the right of the sentence final word. However, considering the results of experiment 4 and those of Mitchel et al. (2008), showing that the position of a word on the line is more important for regressions than the position within a sentence, this problem appears not critical.

The present reading materials had several advantages, but also limitations. The sentences were constructed so that each participant read each target word once as a far target and once as a near target. This was achieved by swapping subject and adverbial complement of the sentences, which is possible in the German language and does not change the meaning or the well-formedness of the sentences. This allowed for a within-subject comparison of the exact same target in exactly the same sentence context, which, to my knowledge, was not yet achieved in any previous study on regressions and re-reading.

This maximally balanced design had one minimal side effect, however. Since each critical word was read as a near and as a far target, it received two entries into spatial memory and there is a possibility, that the first entry had an impact on the way the target location was stored and assessed during the second pass. As the influence of the order of presentation is not significant for adults (see experiment 1), this should, if at all, only be a minimal influence.

Furthermore, only nouns were used as target words and other word categories, such as verbs and adjectives, were not. It is possible that the programming of regressions to these word classes somewhat differs from nouns. There are good arguments for the possibility of such differences. As an example, work on aphasia found a double dissociation between verb and noun retrieval in patients following brain damage (Damasio & Tranel, 1993; Daniele, Giustolisi, Silveri, Colosimo, & Gainotti, 1994). This was later confirmed for healthy participants using event-related potentials (Pulvermüller, Preissl, Lutzenberger, & Birbaumer, 1996; Pulvermüller, Lutzenberger, & Preissl, 1999), event-related functional MRI (Shapiro, Moo, & Caramazza, 2006) and PET-experiments (Perani et al., 1999). According to these results, lexical processing of nouns is localized in anterior temporal regions, while lexical processing of verbs is localized in left frontal regions. To examine possible influences of different word classes on regressions and re-reading, follow up studies using a similar methodological approach should be beneficial.

Another potentially limiting factor of the present work might be that experiments 1, 3 and 4 relied on convenience samples of adult readers, most of them students at the University of Wuppertal, with a high proportion of female participants. As there is no debate on gender differences in the reading literature on adults, this factor is probably negligible. For children, in contrast, gender differences during reading development are well established. Girls outperform boys throughout elementary school in both reading fluency and comprehension, and reading disabilities are also more frequent in boys (e.g. Lewis, Hitch, & Walker, 1994; Flynn & Rhabar, 1994; Rutter et al., 2004; Logan & Johnston, 2009). Therefore, it was important to keep the gender distribution in experiment 2 as equal as possible.

Another limitation might be that in experiment 1 and 2 performance in verbal working memory tests was not examined, limiting the extent to which the present results can be compared with by Guérard et al., 2012 and 2014. However, it is well established that verbal working memory is closely related to reading comprehension (see Carretti, Borella, Cornoldi, & De Beni 2009 for a meta-analysis), which was carefully assessed for adults and children.

6.4 Conclusion

The present dissertation thesis contributes to the scientific discussion about long-range regressions during reading. The most important findings are the following.

1) Skilled readers can be remarkably accurate in regression programming during sentence reading and attain target words precise in nearly 50 percent of the cases. Search processes take the form of either backward, centered or forward searches.
2) Individual differences in spatial memory substantially determine how the amplitude of an initial regression can be adapted to the distance to the intended target. Reading skill does not influence the programming of initial regressions aimed towards specific words. Instead, it modulates the efficiency of correcting processes, when the initial regression did not attain the target word.
3) Children are less likely to attain a target with an accurate regression and therefore rely more often on search strategies. If this is the case, the regression strategies employed by children are remarkably similar to those found in adults.
4) The difficulty of words influences the way in which location information about these words is stored and accessed for regression programming. The cognitive effort required to process an unfamiliar word does not compromise the formation of location memory. Instead, regressions to difficult words are programmed with higher accuracy compared to easier words.
5) Regressions crossing a sentence boundary are as accurate as regressions within a sentence, at least when one-line items are considered.

Taken together, the results support the view that the programming of regression is best accounted for by a hybrid model. Visuo-spatial memory appears to take up a primary role, whereas reading skill serves as an important base for correction, if needed. The necessity to store word location information can be modulated by word difficulty and the accuracy of a spatial representation can be adapted to the demands of linguistic processing. The visual framework of a line rather than the structure of a sentence provides the primary reference system for the storage of word location information.

7 References

Altmann, G. T. M., Garnham, A., & Dennis, Y. (1992). Avoiding the garden path: Eye movements in context. *Journal of Memory and Language*, 31(5), 685–712.

Apel, J. K., Henderson, J. M., & Ferreira, F. (2012). Targeting regressions: Do readers pay attention to the left? *Psychonomic Bulletin & Review*, 19, 1108-1113.

Baccino, T. & Pynte, J. (1994). Spatial coding and discourse models during text reading. *Language & Cognitive Processes,* 9, 143-155.

Bacon, A. M., Parmentier, F. B., & Barr, P. (2013). Visuospatial memory in dyslexia: evidence for strategic deficits. *Memory*, 21(2), 189-209.

Baddeley, A. (2010). Working memory. *Current Biology*, 20(4), 136-140.

Bates, D., Maechler, M., Bolker, B., & Walker, S. (2015). Fitting Linear Mixed-Effects Models Using lme4. *Journal of Statistical Software*, 67(1), 1-48.

Becker, W. (1989). Metrics. In: R. H. Wurtz, M. E. Goldberg (Eds.), *The Neurobiology of Saccadic Eye Movements* (pp 13-61)*,* Amsterdam: Elsevier.

Bicknell, K. & Levy, R. (2011). Why readers regress to previous words: A statistical analysis. In L. Carlson, C. Hölscher, & T. Shipley (Eds.), *Proceedings of the 33rd Annual Conference of the Cognitive Science Society.* Austin, TX: Cognitive Science Society.

Bjork, R. A. (1994). Memory and Metamemory Considerations in the Training of Human Beings. In J. Metcalfe and A. P. Shimamura (Eds.), *Metacognition: Knowing About Knowing* (pp. 185-205), Cambridge, MA: MIT Press.

Bjork, E. L. & Bjork, R. A. (2014). Making Things Hard ON Yourself, But In a Good Way: Creating Desirable Difficulties to Enhance Learning. In M. A. Gernsbacher, J. R. Pomerantz (Eds.), *Psychology and the Real World: Essays Illustrating Fundamental Contributions to Society (2nd edn)* (pp. 59-68), New York: Worth.

Blythe, H. I. (2014). Developmental Changes in Eye Movements and Visual Information Encoding Associated with Learning to Read. *Current Directions in Psychological Science*, 23(3), 201-207.

Blythe, H. I. & Joseph, H. S. S. L. (2011). Children's eye movements during reading. In Liversedge, S. P., Gilchrist, I. D., Everling, S. (Eds.), *The Oxford handbook of eye movements* (pp. 643–662). Oxford, England: Oxford University Press.

Blythe H. I., Häikiö T., Bertram, R., Liversedge S. P., & Hyönä, J. (2011). Reading disappearing text: Why do children refixate words? *Vision Research*, 51, 84-92.

Blythe, H. I., Liversedge, S. P., Joseph, H. S. S. L., White, S. J., Findlay, J. M., & Rayner, K. (2006). The binocular co-ordination of eye movements during reading in children and adults. *Vision Research*, 46, 3898-3908.

Blythe, H. I., Liversedge, S. P., Joseph, H. S. S. L., White, S. J., & Rayner, K. (2009). The uptake of visual information during fixations in reading in children and adults. *Vision Research*, 49, 1583-1591.

Booth, R. W. & Weger, U. W. (2013). The function of regressions in reading: Backward eye movements allow rereading. *Memory & Cognition*, 41, 82-97.

Boston, M. F., Hale, J. T., Vasishth, S., & Kliegl, R. (2011). Parallel processing and sentence comprehension difficulty. *Language and Cognitive Processes*, 26(3), 301-349.

Brysbaert, M., Lange, M., & Van Wijnendaele, I. (2010). The effects of age-of-acquisition adn frequency-of-occurance in visual word recognition: Further evidence from the Dutch language. *European Journal of Cognitive Psychology*, 12, 65-85.

Buswell, G. T. (1922). *Fundamental reading habits: A experiment of their development*. Chicago: University of Chicago Press.

Carpenter, P. A. & Danemann, M. (1981). Lexical retrieval and error recovery in reading: A model based on eye fixations. *Journal of Verbal Learning and Verbal Behavior*, 20, 137-190.

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(2), 246-251.

Clifton, C. & Staub, A. (2011). Syntactic influences on eye movements during reading. In S. P. Liversedge, I. Gilchrist, S. Everling (Eds.), *The Oxford Handbook of Eye Movements* (pp. 895-910). Oxford, United Kingdom: Oxford University Press.

Coltheart, M., Rastle, K., Perry, C., Langdon, R., & Ziegler, J. (2001). DRC: a dual route cascaded model of vidual word recognition and reading aloud. *Psychological Review*, 108(1), 204-256.

Conrad, N. J. & Deacon, S. H. (2016). Children´s Orthographic Knowledge and Their Word Reading Skill: Testing Bidirectional Relations. *Scientific Experiments of Reading*, 20(4), 339-347.

Damasion, A. R. & Tranel, D. (1993). Nouns and verbs are retrieved with differently distributed neural systems. *Proceedings of the National Academy of Sciences of the United States of America,* 90(11), 4957-4960.

Daniele, A., Giustolisi, L., Silveri, M. C., Colosimo, C., & Gainotti, G. (1994). Evidence for a possible neuroanatomical basis for lexical processing of nouns and verbs. *Neuropsychologia,* 32(11), 1325-1341.

Del Giudice, E., Trojano, L., Fragassi, N.A., Posteraro, S., Crisanti, A.F., Tanzarella, P., Marino, A., & Grossi, D. (2000). Spatial cognition in children. II. Visuospatial and constructional skills in developmental reading disability. *Brain & Development*, 22, 368-372.

Deubel H., Wolf W., & Hauske G. (1982). Corrective saccades: effect of shifting the saccade goal. *Vision Research,* 22(3), 353–364.

Diemand-Yauman, C., Oppenheimer, D. M., & Vaughan, E. B. (2011). Fortune favors the Bold (and the Italics): Effects of disfluency on educational outcomes. Cognition, 118(1), 111-115

Eden, G. F., Wood, F. B., & Stein, J. F. (2003). Clock Drawing in Developmental Dyslexia. *Journal of Learning Disabilities*, 36, 216-228.

Engbert, R., Longtin, A., & Kliegl, R. (2002). A dynamical model of saccade generation in reading based on spatially distributed lexical processing. *Vision Research*, 42, 621-636.

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.

Eskenazi, M. A. & Folk, J., R. (2017). Regressions during reading: The cost depends on the cause. *Psychonomic Bulletin & Review*, 24(4), 1211-1216.

Fischer, M. H. (1999). Memory for word locations in reading. *Memory*, 7, 79-116.

Flynn, J. M. & Rahbar, M. H. (1994). Prevalence of reading failure in boys compared with girls. *International Journal of Psychology*, 31(1), 66-71.

Frazier, L. & Rayner, K. (1982). Making and Correcting errors during Sentence Comprehension: Eye Movements in the Analysis of Structurally Ambiguous Sentences. *Cognitive Psychology*, 14, 178-210.

Fulmer, S. M., D'Mello, S. K., Strain, A., & Graesser, A. C. (2015). Interest-based text preference moderates the effect of text difficulty on engagement and learning. *Contemporary Educational Psychology,* 41, 98-110.

Gardner, M. K., Rothkopf, E. Z., Lapan, R., & Lafferty, T. (1987). The word frequency effect in lexical decision: Finding a frequency-based component. *Memory & Cognition*, 15(1), 24-28.

Gathercole, S. E., Pickering, S. J., Ambridge, B., & Wearing, H. (2004). The Structure of Working Memory From 4 tp 15 Years of Age. *Developmental Psychology*, 40(2), 177-190.

Gerhand, S. & Barry, C. (1998). Word frequency effects in oral reading are not merely age- of-acquisition effects in disguise. *Journal of Experimental Psychology: Learning, Memory, and Cognition,* 24, 267-283.

Glanzer, M. & Adams, J. K. (1985). The mirror effect in recognition memory. *Memory & Cognition*, 13, 8-20.

Glanzer, M. & Adams, J. K. (1990). The mirror effect in recognition memory: Data and theory. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, 16, 5-16.

Glanzer, M., Adams, J. K., Iverson, G. J., & Kim, K. (1993). The regularities of recognition memory. *Psychological Review, 100*(3), 546-567.

Grainger, J. (1990). Word frequency and neighborhood frequency effects in lexical decision and naming. *Journal of Memory and Language,* 29(2), 228-244.

Gregg, J. & Inhoff, A. W. (2016). Misperception of Orthographic Neighbors During Silent and oral Reading, *Journal of Experimental Psychology: Human Perception and Performance*, 42(6), 799 – 820.

Guérard, K., Saint-Aubin J., & Maltais, M. (2012). The role of verbal memory in regressions during reading. *Memory & Cognition*, 41, 122-136.

Guérard, K., Saint-Aubin, J., Maltais, M., & Lavoie, H. (2014). The role of verbal memory in regressions during reading is modulated by the target word`s recency in memory. *Memory & Cognition*, 42, 1155-1170.

Häikiö T., Bertram R., & Hyönä J. (2010). Development of parafoveal processing within and across words in reading: evidence from the boundary paradigm. *Quarterly Journal of Experimental Psychology*, 63, 1982–1998.

Häikiö, T., Bertram, R., Hyönä, J., & Niemi, P. (2009). Development of the letter identity span in reading: Evidence from the eye movement moving window paradigm. *Journal of Experimental Child Psychology*, 102, 167-181.

Hasselhorn, M., Schumann-Hengsteler, R., Gronauer, J., Grube, D., Mähler, C., Schmid, I., Seitz-Stein, K., & Zoelch, C. (2012). *Arbeitsgedächtnistesbatterie für Kinder von 5-12 Jahren (AGTB 5-12)*. Göttingen: Hogrefe.

Heathcote, A., Ditton, E., & Mitchell, K. (2006). Word frequency and word likeness mirror effects in episodic recognition memory. *Memory & Cognition*, 34(4), 826-838.

Heister, J., Würzner, K.-M., Bubenzer, J., Pohl, E., Hanneforth, T., Geyken, A., & Kliegl, R. (2011). dlexDB - eine lexikalische Datenbank für die psychologische und linguistische Forschung. *Psychologische Rundschau*, 62(1), 10-20.

Heller, D. (1982). Eye movements in reading. In R. Groner & P. Fraisse (Eds.), *Cognition and eye movements* (pp. 139-154). Berlin, Germany: Deutscher Verlag der Wissenschaften.

Henderson, J. M. & Ferreira, F. (1990). Effects of Foveal Processing Difficulty on the Perceptual Span in Reading: Implications for Attention and Eye Movement Control. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 16(3), 417-429.

Hill, R. L. & Murray, W. S. (2000). Commas and spaces: Effects of punctuation on eye movements and sentence parsing. In A. Kennedy, R. Radach, D. Heller, & J. Pynte (Eds.), *Reading as a perceptual process* (pp. 565-589). Amsterdam, Netherlands: North-Holland/Elsevier Science Publishers.

Hirotani, M., Frazier, L., & Rayner, K. (2006). Punctuation and intonation effects on clause and sentence wrap-up: Evidence from eye movements. *Journal of Memory and Language*, 54(3), 425-443.

Hofmeister J., Heller D., & Radach R. (1999). The Return Sweep in Reading. In: W. Becker, H. Deubel, & T. Mergner (Eds.), *Current Oculomotor Research* (pp.349-357). Springer, Boston: MA.

Huestegge, L., Radach, R., Corbic, C., & Huestegge, S. M. (2009). Oculomotor and linguistic determinants of reading development: A longitudinal experiment. *Vision Research*, 49, 2948-2959.

Hyönä, J. (1995). An Eye Movement Analysis of Topic-Shift Effect During Repeated Reading. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 21(5), 1365-1373.

Hyönä, J. & Olson, R. K. (1995). Eye fixation patterns among dyslexic and normal readers: effects of word-length and word-frequency. *Journal of Experimental Psychology: Learning Memory and Cognition,* 21, 1430-1440.

Hyönä, J., Lorch, R. F., & Kaakinen, J. K. (2002). Individual Differences in Reading to Summarize expository Text: Evidence From Eye Fixation Patterns. *Journal of Educational Psychology*, 94(1), 44-55.

Hyönä, J., Lorch, R. F., & Rinck, M. (2003). Eye movement measures for experimenting global text processing. In R. Radach, J. Hyönä, H. Deubel (Eds.), *The mind´s eyes: Cognitive and applied aspects of eye movement research* (pp. 313-334). Amsterdam: Elsevier.

Hulme, C. & Snowling, M. J. (2016). Reading disorders and dyslexia. *Current Opinion in Pediatrics*, 28(6), 731-735.

Inhoff, A. W. & Radach, R. (1998). Definition and Computation of Oculomotor Measures in the Experiment of Cognitive Processes. In G. Underwood (Ed.), *Eye guidance in reading and scene perception* (pp. 29-53). Oxford, England: Elsevier Science Ltd.

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

Inhoff, A. W., Weger, U. W., & Radach, R. (2005). Sources of Information for the Programming of Short- and Long-Range Regressions during Reading. In G. Underwood (Ed.), *Cognitive Processes in Eye Guidance* (pp. 33-52). Oxford, England: Oxford University Press.

Inhoff, A. W. & Weger, U. W. (2005). Memory for word location during reading: Eye movements to previously read words are spatially selective but not precise. *Memory & Cognition*, 33(3), 447-461.

Joseph, H. S. S. L., Liversedge, S. P., Blythe, H. I., White, S. J., & Rayner, K. (2009). Word length and landing position effects during reading in children and adults. *Vision Research*, 49, 2078-2086.

Just, M. A. & Carpenter, P. A. (1980). A theory of reading: From eye fixations to comprehension. *Psychological Review*, 87, 329-354.

Kennedy, A. & Murray, W. S. (1987). Spatial coordinates and reading: Comments on Monk (1985). *Quarterly Journal of Experimental Psychology*, 39A, 649-656.

Kennedy, A., Brooks, R., Flynn, L.-A., & Prophet, C. (2003). The reader´s spatial code. In J. Hyona, R. Radach, & H. Deubel (Eds.), *The mind`s eye: Cognitive and applied aspects of eye movement research* (pp. 193-212). Amsterdam, The Netherlands: North Holland.

Kibby, M. Y. (2009). Memory Functioning in Developmental Dyslexia: An Analysis Using Two Clinical Memory Measures. *Archives of Clinical Neuropsychology*, 24, 245-254.

Kintsch, W. (1994). Text comprehension, memory, and learning. *American Psychologist*, 49(4), 294-303.

Kintsch, W. & van Dijk, T. A. (1978). Toward a model of text comprehension and production. *Psychological Review*, 85(5), 363-394.

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., 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: General,* 135, 13-35.

Kwon, M. Y., Legge, G. E., & Dubbels, B. R. (2007). Developmental Changes in the Visual Span for Reading. *Vision Research*, 47(22), 2889-2900.

Levy, R., Bicknell, K., Slattery, T., & Rayner, K. (2009). Eye movement evidence that readers maintain and act on uncertainty about past linguistic input. *Proceedings of the National Academy of Sciences*, 106(50), 21086-21090.

Lewis, C., Hitch, G. J., & Walker, P. (1994). The Prevalence of Specific Arithmetic Difficulties and Specific Reading Difficulties in 9- to 10-year-old Boys and Girls. *The Journal of Child Psychology and Psychiatry,* 35(2), 283-292.

Lipowska, M., Czaplewska, E., & Wysocka, A. (2011). Visuospatial deficits of dyslexic children. *Medical Science Monitor*, 17(4), 216-221.

Liversedge, S. P., Paterson, K. B., & Pickering, M. J. (1998). Eye movements and measures of reading time. In G. Underwood (Ed.), *Eye guidance in reading and scene perception* (pp. 55-75). Oxford, England: Elsevier Science Ltd.

Logan, S. & Johnston, R. (2009). Gender differences in reading ability and attitudes: examining where these differences lie. *Journal of Research in Reading*, 32(2), 199-214.

Lovelace, E. A. & Southall, S. D. (1983). Memory for words in prose and their locations on the page. *Memory & Cognition*, 11(5), 429-434.

Mayringer, H. & Wimmer, H. (2012). *Salzburger Lese-Screening für die Klassenstufen 1-4.* Verlag Hans Huber, Hogrefe AG: Bern.

Marx, C., Hutzler, F., Schuster, S., & Hawelka, A. (2016). On the Development of Parafoveal Preprocessing: Evidence from the Incremental Boundary Paradigm. *Frontiers in Psychology*, 7, 1-13.

McConkie, G. W., Kerr, P. W., & Dyre, B. P. (1994). What are 'normal' eye movements during reading: Toward a mathematical description. In J. Ygge & G. Lennerstrand (Eds.), *Eye movements in reading* (pp. 315-328). Oxford, England: Pergamon.

McConkie, G. W. & Rayner, K. (1975). The span of the effective stimulus during a fixation in reading. *Perception & Psychophysics*, 17(6), 578-586.

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

McConkie, G. W., Kerr, P. W., Reddix, M. D., Zola, D., & Jacobs, A. M. (1989). Eye movement control during reading: II. Frequency of refixating a word. *Perception & Psychophysics*, 46, 245–253.

McConkie, G. W. & Zola, D. (1984). Eye movement control during reading: the effects of word units. In W. Prinz & A. T. Sanders (Eds.), *Cognition and Motor Processes* (pp. 63-74). Berlin: Springer.

McConkie, G. W., Zola, D., Grimes, J., Kerr, P. W., Bryant, R. B., & Wolff, P. M. (1991). Children's eye movements during reading. In: J. F. Stein (Ed.), *Vision and visual dyslexia* (pp. 251-262). London, England: Macmillan Press.

Menghini, D., Finzi, A., Carlesimo, G. A., & Vicari, S. (2011). Working memory impairment in children with developmental dyslexia: is it just a phonological deficit? *Developmental Neuropsychology,* 36, 199–213.

Mitchell, D. C., Shen, X., Green, M. J., & Hodgson, T. (2008). Accounting for regressive eye-movements in models of sentence processing: A reappraisal of the Selective Reanalysis hypothesis. *Journal of Memory and Language*, 59, 266-293.

McNamara, D. S. (2001). Reading both high-coherence and low-coherence texts: effects of text sequence and prior knowledge. *Canadian Journal of Experimental Psychology*, 55(1), 51-62.

McNamara, D. S., Kintsch, E., Songer, N. B., & Kintsch, W. (1996). Are good texts always better? Interactions of text coherence, background knowledge, and levels of understanding in learning from text. *Cognition and Instruction*, 14(1), 1-43.

Mulligan, N. W. (2001). Word frequency and memory: Effects on absolute versus relative order memory and on item memory versus order memory. *Memory & Cognition*, 29(7), 977-985.

Murray, W. S. & Kennedy, A. (1988). Spatial coding in the processing of anaphor by good and poor readers. Evidence from eye movement analyses. *Quarterly Journal of Experimental Psychology*, 40A, 693-718.

Nation, K., Adams, J. W., Bowyer-Crane, C. A., & Snowling, M. J. (1999). Working Memory Deficits in Poor Comprehenders Reflect Underlying Language Impairments. *Journal of Experimental Child Psychology*, 73, 139-158.

O'Regan, J. K. (1984). How the eye scans isolated words. In A. G. Gale & F. Johnson (Eds.), *Theoretical and applied aspects of eye movement research* (pp. 159–168). Amsterdam, The Netherlands: North-Holland.

O'Regan, J. K. & Jacobs, A. M. (1992). The optimal viewing position effect in word recognition: A challenge to current theory. *Journal of Experimental Psychology: Human Perception and Performance,* 18(1), 185-197.

O'Regan, J. K. & Levy-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.

O'Regan, J. K., Lévy-Schoen, A., Pynte, J., & Brugaillère, B. (1984). Convenient fixation location within isolated words of different length and structure. *Journal of Experimental Psychology: Human Perception and performance*. 10(2), 250–257.

Parker, A., Slattery, T., & Kirkby, J. A. (2019). Return-sweep saccades during reading in adults and children. *Vision Research*, 155, 35-43.

Payne, B. R. & Stine-Morrow, E. A. L. (2012). Aging, Parafoveal Preview, and Semantic Integration in Sentence Processing: Testing the Cognitive Workload of Wrap-up. *Psychology and Aging*, 27(3), 638-649.

Payne, B. R. & Stine-Morrow, E. A. L. (2014). Adult Age Differences in Wrap-Up During Sentence Comprehension: Evidence From Ex-Gaussian Distributional Analyses of Reading Time. *Psychology and Aging*, 29(2), 213-228.

Perani, D., Cappa, S. F., Schnur, T., Tettamanti, M., Collina, S., Rosa, M. M., & Fazio, F. (1999). The neural correlates of verb and noun processing: A PET experiment. *Brain*, 122(12,1), 2337-2344.

Perea, M. & Pollatsek, A. (1998). The Effects of Neighborhood Frequency in Reading and Lexical Decision. *Journal of Experimental Psychology: Human Perception and Performance*, 24(3), 767-779.

Pollatsek, A., Bolozky, S., Well, A. D., & Rayner, K. (1981). Asymmetries in the perceptual span for Israeli readers. *Brain and Language*, 14(1), 174-180.

Pollatsek, A., Juhasz, B. J., Reichle, E. D., Machacek, D., & Rayner, K. (2008). Immediate and delayed effects of word frequency and word length on eye movements in reading: A reversed delayed effect of word length. *Journal of Experimental Psychology: Human Perception and Performance,* 34(3), 726-750.

Pollatsek, A., Perea, M., & Binder, K. S. (1999). The effects of „neighborhood size" in reading and lexical decision. *Journal of Experimental Psychology: Human Perception and Performance*, 25(4), 1142-1158.

Posner, M. I. & Cohen, Y. (1984). Components of visual orienting. *Attention and Performance X: Control of Language Processes*, 32, 531–556.

Pulvermüller, F., Lutzenberger, W., & Preissl, H. (1999). Nouns and Verbs in the Intact Brain: Evidence from Event-related Potentials and High-frequency Cortical Responses. *Cerebral Cortex,* 9(5), 497-506.

Pulvermüller, F., Preissl, H., Lutzenberger, W., & Birbaumer, N. (1996). Brain Rhythms of Language: Nouns Versus Verbs. *European Journal of Neurosciences*, 8(5), 937-941.

Pynte, J. & Kennedy, A. (2007). The influence of punctuation and word class on distributed processing in normal reading. *Vision research*, 47, 1215-1227.

Radach, R. (1996). *Blickbewegungen beim Lesen. Psychologische Aspekte der Determination von Fixationspositionen. (Eye movements in reading. Psychological aspects of fixation position control).* Münster/New York: Waxmann.

Radach, R., Günther, T., & Huestegge, L. (2012). Blickbewegungen beim Lesen, Leseentwicklung und Legasthenie*.* *Lernen und Lernstörungen*, 1(3), 185-204.

Radach, R., Inhoff, A. W., & Heller, D. (2004). Orthographic regularity gradually modulates saccade amplitudes in reading. *Journal of Cognitive Psychology*, 16, 27 – 51.

Radach, R. & McConkie, G.W. (1998). Determinants of fixation positions in words during reading. In G. Underwood (Ed.) *Eye guidance in Reading and Scene Perception* (pp. 77-124). Amsterdam, The Netherlands: Elsevier Science Ltd.

Radach, R. & Kempe, V. (1993). An individual analysis of fixation positions in reading. In G. d'Ydevalle, & J. van Rensbergen. (Eds.). *Perception and Cognition. Advances in eye movement research* (pp. 213-225). Amsterdam, The Netherlands: Elsevier.

Radach, R. & Kennedy, A. (2004). Theoretical perspectives on eye movements in reading: Past controversies, current issues, and an agenda for future research. *European Journal of Cognitive Psychology*, 16(1/2), 3-26.

Radach, R. & Kennedy, A. (2013). Eye movements in reading: Some theoretical context. *The Quarterly Journal of Experimental Psychology*, 66(3), 429-452.

Radach, R. & Vorstius, C. (2011). Spatial coding and visual processing for long-range regressions in reading. European conference on eye movements, Marseille, France.

Rau, A. K., Moeller, K., & Landerl, K. (2014). The Transition from Sublexical to Lexical Processing in a Consistent Orthography: An eye-Tracking Experiment. *Scientific Experiments of Reading*, 18(3), 224-233.

Rayner, K. (1975). The Perceptual Span and Peripheral Cues in Reading. *Cognitive Psychology*, 7, 65-81.

Rayner, K. (1978). Eye movements in reading and information processing. *Psychological ´Bulletin*, 85(3), 618-660.

Rayner K. (1979a) Eye Movements in Reading: Eye Guidance and Integration. In: Kolers P.A., Wrolstad M.E., Bouma H. (eds) *Processing of Visible Language.* Nato Conference Series, Vol 13. Springer, Boston, MA

Rayner, K. (1979b). Eye guidance in reading: Fixation locations within words. *Perception*, 8, 21-30.

Rayner, K. (1986). Eye movements and the perceptual span in beginning and skilled readers. *Journal of Experimental Child Psychology*, 41(2), 211-236.

Rayner, K. (1997). Understanding eye movements in reading. *Scientific Experiments of Reading*, 1, 301-323.

Rayner, K. (1998). Eye Movements in reading and Information Processing: 20 Years of Research. *Psychological Bulletin*, 124(3), 372-422.

Rayner, K. (2009). Eye movements and attention in reading, scene perception and visual search. *The Quarterly Journal of Experimental Psychology*, 62(8), 1457-1506.

Rayner, K., Chace, K., H., Slattery, T. J., & Ashby, J. (2006). Eye Movements as Reflections of Comprehension Processes in reading. *Scientific Experiments of Reading*, 10(3), 241-255.

Rayner, K. & Duffy, S. A. (1986). Lexical complexity and fixation times in reading: Effects of word frequency, verb complexity, and lexical ambiquity. *Memory & Cognition*, 14(3), 191-201.

Rayner, K., Juhasz, B., Ashby, J., & Clifton, C. (2003). Inhibition of saccade return in reading. *Vision Research*, 43(9), 1027-1034.

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. a, Human Experimental Psychology*. 53, 1061-1080.

Rayner, K., McConkie, G. W., & Zola, D. (1980). Integrating information across eye movements. *Cognitive Psychology*, 12(2), 206-226.

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

Rayner, K., Pollatsek, A., Ashby, J., & Clifton, C. (2012). *The Psychology of Reading*. New York: Psychology Press.

Rayner, K., Sereno, S. C., & Raney, G. E. (1996). Eye Movement Control in Reading: A Comparison of Two Types of Models. *Journal of Experimental Psychology Human Perception & Performance*, 22(5), 1188-1200.

Rayner K., Sereno S. C., Morris R. K., Schmauder A. R., & Clifton C. (1989). Eye movements and on-line language. *Language and Cognitive Processes*, 4, 3-4.

Rawson K. A. & Miyake, A. (2002). Does relocation in text depend on verbal or visuospatial abilities? An individual-differences analysis. *Psychonomic Bulletin & Review*, 9 (4), 801-806.

Reichle, E. D. & Perfetti, C. A. (2003). Morphology in Word Identification: A Word-Experience Model That Accounts for Morpheme Frequency Effects. *Scientific Experiments of Reading*, 7(3), 219-237.

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(4), 445-476.

Reilly, R. G. & O'Regan, J. K. (1998). Eye movement control during reading: A simulation of some word-targeting strategies. *Vision Research*, 38, 303–317.

Rutter, M., Caspi, A., Fergusson, D., Horwood, J., Goodman, R., Maughan, B., Moffitt, T. E., Meltzer, H., & Carroll, J. (2004). Sex Differences in Developmental Reading Disability New Findings From 4 Epidemiological Experiments. *Special Communication*, 291(16), 1931-2040.

Schneider, W., Schlagmüller, M., & Ennemoser, M. (2007). *Lesegeschwindigkeits- und –verständnistest für die Klassen 6-12 (LGVT 6-12).* Göttingen: Hogrefe.

Schilling, H. E. H., Rayner, K., & Chumbley, J. I. (1998). Comparing naming, lexical decision, and eye fixation times: Word frequency effects and individual differences. *Memory & Cognition,* 26(6), 1270-1281.

Schotter, E.R., Tran, R., & Rayner, K. (2014). Don't believe what you read (only once) comprehension is supported by regressions during reading. *Psychological science,* 25(6), 1218-1226.

Schuster, S., Hawelka, S., Hutzler, F., Kronbichler, M., & Richlan, F. (2016). Words in Context: The Effects of Length, Frequency, and Predictability on Brain Responses During Natural Reading. *Cerebral Cortex*, 26(10), 3889-3904.

Shapiro, K. A., Moo, L. R., & Caramazza, A. (2006). Cortical signatures of noun and verb production. *Proceedings of the National Academy of Sciences of the United States,* 103(5), 1644-1649.

Slattery, T. J. (2009). Word Misperception, the Neighbour Frequency Effect, and the Role of Sentence Context: Evidence from Eye Movements. *Journal of Experimental Psychology: Human, Perception and Performance,* 35(6), 1969-1975.

Slattery, T. J. & Rayner, K. (2009). The influence of text legibility on eye movements during reading. *Applied Cognitive Psychology*, 24(8), 1129-1148.

Tang, S., Reilly, R. G., & Vorstius, C. (2012). EyeMap: a software system for visualizing and analysing eye movement data in reading. *Behaviour Research Methods*, 44(2), 420-438.

Taylor, S. E. (1971). *The dynamic activity of reading: A model of the process*. Research Information Bulletin No 9. New York: Educational Developmental Laboratories.

Taylor S. E., Frackenpohl, H., & Petee, J. L. (1960). *Grade level norms for the components of the fundamental reading skill.* E.D.L. Research and Information Bulletin, 3, Huntington, NY: Educational Developmental labs, Inc.

Taylor, J. N. & Perfetti, C. A. (2016). Eye movements reveal readers´ lexical quality and reading experience. *Reading and Writing*, 29(6), 1069-1103.

Tyler, L. K., Russell, R., Fadili, J., & Moss, H.E. (2001). The neural representation of nouns and verbs: PET experiments. *Brain*, 124(8), 1619-34.

Underwood, N. R. & McConkie, G. W. (1985). Perceptual span for letter distinctions during reading. *Reading Research Quarterly*, 20(2), 153-162.

Vanyukov, P. M., Warren, T., Wheeler, M. E., & Reichle, E. D. (2012). The emergence of frequency effects in eye movements, *Cognition,* 123, 185-189.

Vitu, F. (1991). The existence of a center of gravity effect during reading. *Vision Research*, 31(7-8), 1289-1313.

Vitu, F. (2005). Visual extraction processes and regressive saccades in reading. In G. Underwood (Ed.), *Cognitive processes in eye guidance* (pp. 1–32). Oxford, England: Oxford University Press.

Vitu, F. & McConkie, G. W. (2000). Regressive saccades and word perception in adult reading. In A. Kennedy, R. Radach, D. Heller, & J. Pynte (Eds.), *Reading as a perceptual process* (pp. 301-326). Amsterdam, The Netherlands: North-Holland/Elsevier Science Publishers.

Vitu, F., McConkie, G. W., Kerr, P., & O'Regan, J. K. (2001). Fixation location effects on fixation durations during reading: an inverted optimal viewing position effect. *Vision Research*, 41, 3513-3533.

Vitu, E, McConkie, G. W., & Zola, D. (1998). About regressive saccades in reading and their relation to word identification. In G. Underwood (Ed.), *Eye guidance in reading and scene perception* (pp. 101-124). Oxford, England: Elsevier.

Von der Malsburg, T., Kliegl, R., & Vasishth, S. (2011). Determinants of Scanpath Regularity in Reading. *Cognitive Science*, 39, 1675-1703.

Von Karolyi, C. (2001). Visual-Spatial Strength in Dyslexia. *Journal of Learning Disabilities*, 34(4), 380-391.

Von Karolyi, C., Winner, E., Gray, W., & Sherman, G. F. (2003). Dyslexia linked to talent: Global visual-spatial ability. *Brain and Language*, 85, 427-431.

Vonk, W. & Cozijin, R. (2003). On the Treatment of Saccades and Regressions in Eye Movements Measures of Reading Time. In J. Hyönä, R. Radach & H. Deubel (eds.), *The Mind´s Eye: Cognitive and Applied Aspects of Eye Movement research* (pp. 291-311). Amsterdam: Elsevier Science BV.

Vonk, W., Radach, R., & van Rijn, H. (2002). Eye guidance and the saliency of wird beginnings in reading text. In A. Kennedy, D. Heller, & J. Pynte (Eds.), *Reading as a perceptual process* (pp. 269-299). Oxford, England: Elsevier.

Vorstius, C., Radach, R., & Lonigan, C. (2014). Eye movements in developing readers: A comparison of silent and oral sentence reading. *Visual Cognition*, 22(3-4), 458-485.

Vorstius, C., Radach, R., Mayer, M., & Lonigan, C. (2013). Monitoring local comprehension monitoring in sentence reading. *School Psychology Review*, 42, 191-206.

Warren, T., White, S. J., & Reichle, E. D. (2009). Investigating the causes of wrap-up effects: Evidence from eye movements and E-Z Reader. *Cognition,* 111, 132–137.

Weger, U. W. & Inhoff, A. W. (2007). Long-range regressions to previously read words are guided by spatial and verbal memory. *Memory & Cognition*, 35(6): 1293-1306.

Weiss, A. F., Kretzschmar, F., Schlesewsky, M., Bornkessel-Schlesewsky, I., & Staub, A. (2018). Comprehension demands modulate re-reading, but not first-pass reading behavior. *Quarterly Journal of Experimental Psychology*, 71(1), 198-210.

White, S. (2008). Eye Movement Control During Reading: Effects of Word Frequency and Orthographic Familarity. *Journal of Experimental Psychology: Human, Perception and Performance,* 34(1), 205-223.

White, S. J. & Liversedge, S. P. (2004). Orthographic familiarity influences initial eye fixation positions in reading. *European Journal of Cognitive Psychology*, 16, 52-78.

White, S. J., Rayner, K., & Liversedge, S. P. (2005). The influence of parafoveal word length and contextual constraints on fixation durations and word skipping in reading. *Psychonomic Bulletin & Review*, 12(3), 466-471.

White, S. J., Warren, T., & Reichle, E. D. (2011). Parafoveal preview during reading: effects of sentence position. *Journal of Experimental Psychology: Human Perception and Performance,* 37, 1221-1238.

Winner, E., von Karolyi, C., Malinsky, D., French, L., Seliger, C., Ross, E., & Weber, C. (2001). Dyslexia and Visual-Spatial Talents: Compensation vs Deficit Model. *Brain and Language*, 76, 81-110.

Wotschack, C. & Kliegl, R. (2013). Reading strategy modulates parafoveal-on-foveal effects in sentence reading. *The Quarterly Journal of Experimental Psychology*, 66, 548–562. *Contemporary Educational Psychology*, 18(4), 442-454.

https://sansforgetica.rmit, accessed 16 March 2019, 21:38.

8 Appendix

Reading materials experiment 1 and 2

Experimental sentences

word = target word

Die langen **Würmer** sieht man nach dem Regen aus tiefen **Löchern** kriechen.
Aus tiefen **Löchern** sieht man nach dem Regen die langen **Würmer** kriechen.

Die reifen **Tomaten** werden nach der Ernte in stabilen **Kisten** verpackt.
In stabilen **Kisten** werden nach der Ernte die reifen **Tomaten** verpackt.

Für strenge **Winter** wird von vielen Tieren ein großer **Vorrat** angelegt.
Ein großer **Vorrat** wird von vielen Tieren für strenge **Winter** angelegt.

Der gesuchte **Schatz** ist auf der Insel unter grünen **Palmen** versteckt.
Unter grünen **Palmen** auf der Insel ist der gesuchte **Schatz** versteckt.

Die kleinen **Fische** springen froh und munter in kalten **Bächen** herum.
In kalten **Bächen** springen froh und munter die kleinen **Fische** herum.

Die steile **Treppe** führt im alten Schloss zur versteckten **Kammer** hinauf.
Zur versteckten **Kammer** im alten Schloss führt die steile **Treppe** hinauf.

Viele schöne **Burgen** kann man aus festem Sand am leeren **Strand** bauen.
Am leeren **Strand** kann man aus festem Sand viele schöne **Burgen** bauen.

Der starke **Verkehr** ist in unserer Stadt am frühen **Morgen** gefährlich.
Am frühen **Morgen** ist in unserer Stadt der starke **Verkehr** gefährlich.

Auf vielen **Bergen** sind zum Ausruhen kleine gemütliche **Hütten** zu finden.
Kleine gemütliche **Hütten** sind zum Ausruhen auf vielen **Bergen** zu finden.

Die leuchtenden **Sterne** sieht man in klaren **Nächten** besonders deutlich.
In klaren **Nächten** sieht man die leuchtenden **Sterne** besonders deutlich.

Die saftigen **Beeren** schmecken auf frischem **Kuchen** besonders gut.
Auf frischem **Kuchen** schmecken die saftigen **Beeren** besonders gut.

Sehr viele **Pferde** leben auf dem Ponyhof in alten **Ställen** eng zusammen.
In alten **Ställen** auf dem Ponyhof leben sehr viele **Pferde** eng zusammen.

Das lustige **Gedicht** wird jetzt vom schlauen **Jungen** richtig aufgesagt.

Vom schlauen **Jungen** wird jetzt das lustige **Gedicht** richtig aufgesagt.

Die kaputten **Reifen** werden heute im staubigen **Keller** wieder repariert.
Im staubigen **Keller** werden heute die kaputten **Reifen** wieder repariert.

Die wichtigen **Briefe** werden hier im dicken **Ordner** sorgsam abgeheftet.
Im dicken **Ordner** werden hier die wichtigen **Briefe** sorgsam abgeheftet.

Der alberne **Scherz** wird zwischendurch vom netten **Lehrer** gerne erzählt.
Vom netten **Lehrer** wird zwischendurch der alberne **Scherz** gerne erzählt.

Das leere weiße **Papier** wird vorsichtig in den teuren **Drucker** gelegt.
In den teuren **Drucker** wird vorsichtig das leere weiße **Papier** gelegt.

Die leckere süße **Orange** wird in Spanien von sehr alten **Bäumen** geerntet.
Von sehr alten **Bäumen** wird in Spanien die leckere süße **Orange** geerntet.

Von den bösen **Räubern** wurde Schmuck aus dem sicheren **Tresor** gestohlen.
Aus dem sicheren **Tresor** wurde Schmuck von den bösen **Räubern** gestohlen.

Der neue weiße **Verband** wird fest um den verletzten **Finger** gebunden.
Um den verletzten **Finger** wird fest der neue weiße **Verband** gebunden.

Eine sehr schöne **Melodie** wird gerade auf dem alten **Klavier** vorgespielt.
Auf dem alten **Klavier** wird gerade eine sehr schöne **Melodie** vorgespielt.

Der dicke rote **Strumpf** wurde sicher in einem stabilen **Karton** verpackt.
In einem stabilen **Karton** wurde sicher der dicke rote **Strumpf** verpackt.

Der neue blaue **Füller** und der Stift werden in die teure **Tasche** gepackt.
In die teure **Tasche** werden der Stift und der neue blaue **Füller** gepackt.

Auf dem weißen **Teller** liegt zum Essen die sehr leckere **Waffel** bereit.
Die sehr leckere **Waffel** liegt zum Essen auf dem weißen **Teller** bereit.

Der streng geheime **Zettel** ist unten im kleinen **Schrank** gut versteckt.
Unten im kleinen **Schrank** ist der streng geheime **Zettel** gut versteckt.

Ein starker junger **Bursche** bekämpfte mit dem langen **Schwert** den Feind.
Mit dem langen **Schwert** bekämpfte ein starker junger **Bursche** den Feind.

Die schöne neue **Spange** sieht man in den blonden **Haaren** besonders gut.
In den blonden **Haaren** sieht man die schöne neue **Spange** besonders gut.

Die kleinen jungen **Hühner** sind mit dicken weichen **Federn** gut geschützt.
Mit dicken weichen **Federn** sind die kleinen jungen **Hühner** gut geschützt.

Das sehr teure **Gemälde** kann in einem großen **Museum** bewundert werden.
In einem großen **Museum** kann das sehr teure **Gemälde** bewundert werden.

Gelbe und reife **Birnen** werden mit einer langen **Leiter** sorgsam geerntet.
Mit einer langen **Leiter** werden gelbe und reife **Birnen** sorgsam geerntet.

Das dicke rosa **Schwein** wurde durch eine heiße **Dusche** wieder sauber.
Durch eine heiße **Dusche** wurde das dicke rosa **Schwein** wieder sauber.

Die neuen bunten **Stifte** werden auf dem kleinen **Hocker** schön sortiert.
Auf dem kleinen **Hocker** werden die neuen bunten **Stifte** schön sortiert.

Filler sentences

Auf der Straße warnt ein rotes **Schild** deutlich vor vielen Gefahren.
Vor vielen Gefahren warnt ein rotes **Schild** deutlich auf der Straße.

Beim Spielen wurde die leise **Klingel** von beiden Kindern nicht gehört.
Von beiden Kindern wurde die leise **Klingel** beim Spielen nicht gehört.

Kräftige Winde wehen auf den trockenen **Feldern** während des Winters.
Während des Winters wehen auf den trockenen **Feldern** kräftige Winde.

Alle größeren Schüler besuchen das spannende **Museum** mit ihrer Lehrerin.
Mit ihrer Lehrerin besuchen das spannende **Museum** alle größeren Schüler.

Im Park hat uns heute eine fremde **Familie** nach dem Weg zum Bus gefragt.
Nach dem Weg zum Bus hat uns eine fremde **Familie** heute im Park gefragt.

In Omas Keller liegt ein alter **Spiegel** versteckt unter einer Kiste.
Versteckt unter einer Kiste liegt ein alter **Spiegel** in Omas Keller.

Am liebsten möchte Lucas eine große **Giraffe** oder einen Löwen malen.
Einen Löwen oder eine große **Giraffe** möchte Lucas am liebsten malen.

Auf dem Spielplatz steht die steile **Rutsche** neben der blauen Wippe.
Neben der blauen Wippe steht die steile **Rutsche** auf dem Spielplatz.

Aus einem Brunnen kann **Wasser** mit einem stabilen Eimer geholt werden.
Mit einem stabilen Eimer kann **Wasser** aus einem Brunnen geholt werden.

Große Höhlen werden von Füchsen unter **Wiesen** und in Wäldern gegraben.
In Wäldern und unter **Wiesen** werden von Füchsen große Höhlen gegraben.

Ein Bericht und ein Foto des **Unfalls** sind in der Zeitung abgedruckt.
In der Zeitung sind ein Foto des **Unfalls** und ein Bericht abgedruckt.

Dunkle Wolken sind im warmen **Sommer** nur sehr selten am Himmel zu sehen.
Am Himmel sind im warmen **Sommer** nur sehr selten dunkle Wolken zu sehen.

Ein Schirm oder ein **Mantel** schützen bei schlechtem Wetter vor Nässe.
Bei schlechtem Wetter schützen ein **Mantel** oder ein Schirm vor Nässe.

Um zwölf Uhr müssen der kleine und der große **Zeiger** ganz oben stehen.
Der kleine und der große **Zeiger** müssen um zwölf Uhr ganz oben stehen.

Auf der Jagd nach Beute nutzen die klugen **Falken** ihre scharfen Krallen.
Ihre scharfen Krallen nutzen die klugen **Falken** auf der Jagd nach Beute.

Am Meer kann man **Drachen** nur mit einem starken Band steigen lassen.
Nur mit einem starken Band kann man **Drachen** am Meer steigen lassen.

Reading materials experiment 3

Experimental sentences

word = easy target word

word = difficult target word

Der schlaue **Mensch** versucht für seine Überlegungen sichere Evidenz zu finden.

Sichere Evidenz für seine Überlegungen versucht der schlaue **Mensch** zu finden.

In unsicheren **Zeiten** verschafft sich der erfahrene Ökonom den Durchblick.

Der erfahrene Ökonom verschafft sich in unsicheren **Zeiten** den Durchblick.

Seit unzähligen **Jahren** bildet der Vatikanstadt eine winzige Enklave in Rom.

Eine winzige Enklave bildet der Vatikanstadt seit unzähligen **Jahren** in Rom.

Zahlreiche **Fragen** wurden in der Prüfung zum historischen Olympia gestellt.

Zum historischen Olympia wurden in der Prüfung zahlreiche **Fragen** gestellt.

Die meisten **Kinder** werden noch immer von schmutzigen Pfützen angezogen.

Von schmutzigen Pfützen werden noch immer die meisten **Kinder** angezogen.

Durch fleißige **Arbeit** hat er sich zum angesehenen Lektor hochgearbeitet.
Zum angesehenen Lektor hat er sich durch fleißige **Arbeit** hochgearbeitet.

Durch farbenfrohe **Bilder** drückt der Künstler seine freudige Emotion aus.
Seine freudige Emotion drückt der Künstler durch farbenfrohe **Bilder** aus.

Der schwermütige **Gedanke** heute Nacht hat den begabten Lyriker inspiriert.
Den begabten Lyriker hat heute Nacht der schwermütige **Gedanke** inspiriert.

Eine zuverlässige **Person** wird sehr dringend als fähiger Zoologe gesucht.
Als fähiger Zoologe wird sehr dringend eine zuverlässige **Person** gesucht.

Ihr schönes **Gesicht** hat er immer wieder mit tausenden Juwelen verglichen.
Mit tausenden Juwelen hat er immer wieder ihr schönes **Gesicht** verglichen.

Besondere **Kräfte** wurden in alten Zeiten dem mythischen Phönix zugesprochen.
Dem mythischen Phönix wurden in alten Zeiten besondere **Kräfte** zugesprochen.

Triftige **Gründe** haben ihm diese Woche einen schlimmen Rüffel eingebracht.
Einen schlimmen Rüffel haben ihm diese Woche triftige **Gründe** eingebracht.

Keine trockene **Stelle** ist bei Seegang in der undichten Kajüte zu finden.
In der undichten Kajüte ist bei Seegang keine trockene **Stelle** zu finden.

Sein angeblicher **Freund** hat sich sehr schnell als totaler Egoist entpuppt.
Als totaler Egoist hat sich sehr schnell sein angeblicher **Freund** entpuppt.

Die steilen **Straßen** fährt noch jeden Tag das klapprige Vehikel hinauf.
Das klapprige Vehikel fährt noch jeden Tag die steilen **Straßen** hinauf.

Die gotische **Kirche** kann bei Konzerten mit besonderer Akustik glänzen.
Mit besonderer Akustik kann bei Konzerten die gotische **Kirche** glänzen.

Die liebende **Mutter** hat ihrer einzigen Tochter das kostbare Amulett vererbt.
Das kostbare Amulett hat ihrer einzigen Tochter die liebende **Mutter** vererbt.

Seine kräftige **Stimme** hält glücklicherweise die gesamte Auktion durch.
Die gesamte Auktion hält glücklicherweise seine kräftige **Stimme** durch.

Zu nächtlicher **Stunde** sollte die Gruppe geräuschvolle Exzesse unterbinden.
Geräuschvolle Exzesse sollte die Gruppe zu nächtlicher **Stunde** unterbinden.

Wenig frisches **Wasser** und Sonne reicht den meisten Kakteen zum Leben.
Den meisten Kakteen reicht Sonne und wenig frisches **Wasser** zum Leben.

In gewandter **Sprache** wird sie den Zuhörern einen historischen Exkurs geben.
Einen historischen Exkurs wird sie den Zuhörern in gewandter **Sprache** geben.

Wenig **Mittel** stehen den Studenten für benötigtes Zubehör zur Verfügung.
Für benötigtes Zubehör stehen den Studenten wenig **Mittel** zur Verfügung.

Bei der schwierigen **Aufgabe** hat sie leider die wichtige Vokabel vergessen.
Die wichtige Vokabel hat sie leider bei der schwierigen **Aufgabe** vergessen.

Die letzten **Monate** haben ihm und seiner Familie unerwartete Zäsuren gebracht.
Unerwartete Zäsuren haben ihm und seiner Familie die letzten **Monate** gebracht.

Beim kleinsten **Schritt** hat der regierende König sein goldenes Zepter dabei.
Sein goldenes Zepter hat der regierende König beim kleinsten **Schritt** dabei.

Die gefundenen **Sachen** lassen sich eindeutig dem eleganten Rokoko zuordnen.
Dem eleganten Rokoko lassen sich eindeutig die gefundenen **Sachen** zuordnen.

Die gedruckten **Briefe** werden sorgfältig in das passende Kuvert gesteckt.
In das passende Kuvert werden sorgfältig die gedruckten **Briefe** gesteckt.

Trockene **Gebiete** und viel Sonne kommen sehr häufig am heißen Äquator vor.
Am heißen Äquator kommen sehr häufig viel Sonne und trockene **Gebiete** vor.

Veruntreute **Gelder** haben bei mehreren Banken eine riesige Razzia ausgelöst.
Eine riesige Razzia haben bei mehreren Banken veruntreute **Gelder** ausgelöst.

Unzählige **Frauen** haben sich im alten China eine vornehme Blässe gewünscht.
Eine vornehme Blässe haben sich im alten China unzählige **Frauen** gewünscht.

Beneidende **Blicke** bekommt er häufig als eindeutiger Favorit zugeworfen.
Als eindeutiger Favorit bekommt er häufig beneidende **Blicke** zugeworfen.

In den letzten **Reihen** konnte das Publikum die akustische Gitarre kaum hören.
Die akustische Gitarre konnte das Publikum in den letzten **Reihen** kaum hören.

Filler sentences

Beim heutigen Treffen stimmen alle führenden **Staaten** dem aktuellen Gesetz zu.
Dem aktuellen Gesetz stimmen alle führenden **Staaten** beim heutigen Treffen zu.

Im Universum wurden noch keine anderen **Welten** von mutigen Astronauten entdeckt.
Von mutigen Astronauten wurden noch keine anderen **Welten** im Universum entdeckt.

Für Rudeltiere sind gefestigte **Gruppen** wichtig um ein sicheres Leben zu haben.
Um ein sicheres Leben zu haben sind gefestigte **Gruppen** für Rudeltiere wichtig.

Das verliebte Pärchen mochte die abgelegenen **Städte** auf der weiten Rundreise.
Auf der weiten Rundreise mochte die abgelegenen **Städte** das verliebte Pärchen.

In alternativen Schulen wird mit neuartigen **Formen** das schwere Rechnen geübt.
Das schwere Rechnen wird mit neuartigen **Formen** in alternativen Schulen geübt.

Neue Mitglieder braucht die verstaubte **Partei** um ihr Bestehen zu sichern.
Um ihr Bestehen zu sichern braucht die verstaubte **Partei** neue Mitglieder.

Im uralten Schloss sollen furchtbare **Geister** ihr gruseliges Unwesen treiben.
Ihr gruseliges Unwesen sollen furchtbare **Geister** im uralten Schloss treiben.

Dem glücklichen Gewinner konnte die **Million** am Wochenende ausgezahlt werden.
Am Wochenende konnte die **Million** dem glücklichen Gewinner ausgezahlt werden.

Bei wilden Pferden sind die weichen Nüstern eine empfindliche Stelle.
Eine empfindliche Stelle sind die weichen Nüstern bei wilden Pferden.

Der schnelle Zug fährt zur richtigen Uhrzeit in den belebten Bahnhof ein.
In den belebten Bahnhof fährt zur richtigen Uhrzeit der schnelle Zug ein.

In der dunklen Gruft schläft ein schwarzer Vampir in einem noblen Sarg.
In einem noblen Sarg schläft ein schwarzer Vampir in der dunklen Gruft.

Zu dem Kostüm fehlt noch ein langer Umhang damit sie als Hexe erkannt wird.
Damit sie als Hexe erkannt wird fehlt noch ein langer Umhang zu dem Kostüm.

In der Fabrik ist der schmackhafte Kümmel das am meisten verwendete Gewürz.
Das am meisten verwendete Gewürz ist der schmackhafte Kümmel in der Fabrik.

Der kleine Mann hat sich eine gemeine Abfuhr bei der hübschen Frau eingefangen.
Bei der hübschen Frau hat sich eine gemeine Abfuhr der kleine Mann eingefangen.

Das starke Pferd treibt sein gesundes Fohlen auf der grünen Wiese an.
Auf der grünen Wiese treibt sein gesundes Fohlen das starke Pferd an.

Die Hunde gefallen dem kritischen Züchter auf der bekannten Ausstellung gut.

Auf der bekannten Ausstellung gefallen dem kritischen Züchter die Hunde gut.

Reading materials experiment 4

Experimental sentences

word = target word

Lange **Würmer** mögen den warmen Regen. Sie kriechen aus tiefen **Löchern** hervor.

Lange **Würmer** mögen den warmen Regen und kriechen aus tiefen **Löchern** hervor.

Tiefe **Löcher** sind wichtige Lebensräume. Sie bieten langen **Würmern** Schutz.

Tiefe **Löcher** sind wichtige Lebensräume und bieten langen **Würmern** Schutz.

Die reifen **Tomaten** schmecken köstlich. Sie werden in stabilen **Kisten** verpackt.

Die reifen **Tomaten** schmecken köstlich und werden in stabilen **Kisten** verpackt.

Stabile **Kisten** wurden gekauft. Sie müssen nun mit reifen **Tomaten** bepackt werden.

Stabile **Kisten** wurden gekauft und müssen nun mit reifen **Tomaten** bepackt werden.

Der starke **Verkehr** nimmt weiter zu. Es ist am frühen **Morgen** besonders laut.

Der starke **Verkehr** nimmt weiter zu und ist am frühen **Morgen** besonders laut.

Am frühen **Morgen** ist es schon sehr laut. Der starke **Verkehr** nimmt weiter zu.

Am frühen **Morgen** ist es schon sehr laut und starker **Verkehr** nimmt weiter zu.

Kleine **Hütten** laden zur Rast ein. Sie werden auf vielen **Bergen** häufig genutzt.

Kleine **Hütten** laden zur Rast ein und werden auf vielen **Bergen** häufig genutzt.

Auf vielen **Bergen** kann man Pause machen. Man findet oft kleine **Hütten** dafür.

Auf vielen **Bergen** kann man Pause machen und findet oft kleine **Hütten** dafür.

Der dunkle **Himmel** strahlt. Er ist erleuchtet von hellen **Sternen** bei Nacht.

Der dunkle **Himmel** strahlt und ist erleuchtet von hellen **Sternen** bei Nacht.

Die hellen **Sterne** leuchten klar. Sie lassen den dunklen **Himmel** erstrahlen.

Die hellen **Sterne** leuchten klar und lassen den dunklen **Himmel** erstrahlen.

Der frische **Kuchen** backt im Ofen. Er wurde mit saftigen **Birnen** gefüllt.

Der frische **Kuchen** backt im Ofen und wurde mit saftigen **Birnen** gefüllt.

Saftige **Birnen** sind zum Backen geeignet. Sie geben frischem **Kuchen** Geschmack.

Saftige **Birnen** sind zum Backen geeignet und geben frischem **Kuchen** Geschmack.

Der kluge **Schüler** bewundert Goethe. Er lernte das schöne **Gedicht** auswendig.

Der kluge **Schüler** bewundert Goethe und lernte das schöne **Gedicht** auswendig.

Das schöne **Gedicht** handelt vom Frühling. Es wird vom klugen **Schüler** rezitiert.

Das schöne **Gedicht** handelt vom Frühling und wird vom klugen **Schüler** rezitiert.

Der strenge **Richter** liest das Urteil. Er verärgert den eifrigen **Anwalt** sehr.

Der strenge **Richter** liest das Urteil und verärgert den eifrigen **Anwalt** sehr.

Der eifrige **Anwalt** ist empört. Er möchte sich am strengen **Richter** rächen.

Der eifrige **Anwalt** ist empört und möchte sich am strengen **Richter** rächen.

Der junge **Lehrer** ist witzig. Er erzählt gerne lustige **Scherze** im Unterricht.

Der junge **Lehrer** ist witzig und erzählt gerne lustige **Scherze** im Unterricht.

Lustige **Scherze** sind beliebt. Sie werden vom jungen **Lehrer** gerne erzählt.

Lustige **Scherze** sind beliebt und werden vom jungen **Lehrer** gerne erzählt.

Das weiße **Papier** hat beste Qualität. Es wird in den modernen **Drucker** eingelegt.

Das weiße **Papier** hat beste Qualität und wird in den modernen **Drucker** eingelegt.

Der moderne **Drucker** ist neu. Er wird mit bestem weißen **Papier** befüllt.

Der moderne **Drucker** ist neu und wird mit bestem weißen **Papier** befüllt.

Der listige **Räuber** knackt den Tresor. Er stiehlt den teuren **Schmuck** der Frau.

Der listige **Räuber** knackt den Tresor und stiehlt den teuren **Schmuck** der Frau.

Der teure **Schmuck** war wertvoll. Er wurde von listigen **Räubern** gestohlen.

Der teure **Schmuck** war wertvoll und wurde von listigen **Räubern** gestohlen.

Der kleine **Finger** ist verletzt. Er wird mit einem festen **Verband** verbunden.
Der kleine **Finger** ist verletzt und wird mit einem festen **Verband** verbunden.
Der feste **Verband** schützt die Wunde. Er wurde um den kleinen **Finger** gebunden.
Der feste **Verband** schützt die Wunde und wurde um den kleinen **Finger** gebunden.

Das sonore **Klavier** ist neu gestimmt. Es bringt eine schöne **Melodie** hervor.
Das sonore **Klavier** ist neu gestimmt und bringt eine schöne **Melodie** hervor.
Die schöne **Melodie** verzaubert. Sie lässt das sonore **Klavier** erklingen.
Die schöne **Melodie** verzaubert und lässt das sonore **Klavier** erklingen.

Das scharfe **Schwert** ist sehr wertvoll. Nur ein tapferer **Bursche** bekommt es.
Das scharfe **Schwert** ist sehr wertvoll und ein tapferer **Bursche** bekommt es.
Der tapfere **Bursche** wird bewundert. Er hält sein scharfes **Schwert** mit Stolz.
Der tapfere **Bursche** wird bewundert und hält sein scharfes **Schwert** mit Stolz.

Die bunte **Klammer** setzt Akzente. Sie sieht in den blonden **Haaren** hübsch aus.
Die bunte **Klammer** setzt Akzente und sieht in den blonden **Haaren** hübsch aus.
Mit blonden **Haaren** fühlt sie sich wohl. Sie trägt oft eine bunte **Klammer** darin.
Mit blonden **Haaren** fühlt sie sich wohl und trägt oft eine bunte **Klammer** darin.

Das schöne **Gemälde** ist bekannt. Es kann im neuen **Museum** bewundert werden.

Das schöne **Gemälde** ist bekannt und kann im neuen **Museum** bewundert werden.

Das neue **Museum** ist sehr beliebt. Es stellt das schöne **Gemälde** aus.

Das neue **Museum** ist sehr beliebt und stellt das schöne **Gemälde** aus.

Die teure **Kamera** war ein Geschenk. Sie ging leider im letzten **Urlaub** verloren.

Die teure **Kamera** war ein Geschenk und ging leider im letzten **Urlaub** verloren.

Der letzte **Urlaub** war schön. Er wurde mit der teuren **Kamera** festgehalten.

Der letzte **Urlaub** war schön und wurde mit der teuren **Kamera** festgehalten.

Tosender **Beifall** füllt die Halle. Er gibt dem jungen **Musiker** ein gutes Gefühl.

Tosender **Beifall** füllt die Halle und gibt dem jungen **Musiker** ein gutes Gefühl.

Der junge **Musiker** spielt ein Konzert. Er genießt den tosenden **Beifall** sehr.

Der junge **Musiker** spielt ein Konzert und genießt den tosenden **Beifall** sehr.

Das laute **Telefon** klingelt. Es stört den akuten **Termin** des Professors.

Das laute **Telefon** klingelt und stört den akuten **Termin** des Professors.

Der akute **Termin** war leider erfolglos. Er wurde vom lauten **Telefon** gestört.

Der akute **Termin** war leider erfolglos und wurde vom lauten **Telefon** gestört.

Das leichte **Gepäck** ist handlich. Es wird von ihrem lieben **Partner** getragen.

Das leichte **Gepäck** ist handlich und wird von ihrem lieben **Partner** getragen.

Ihr lieber **Partner** ist hilfsbereit. Er trägt das leichte **Gepäck** nach Hause.

Ihr lieber **Partner** ist hilfsbereit und trägt das leichte **Gepäck** nach Hause.

Der barocke **Sessel** gefiel ihr. Er wurde sogleich im neuen **Katalog** bestellt.

Der barocke **Sessel** gefiel ihr und wurde sogleich im neuen **Katalog** bestellt.

Der neue **Katalog** bietet Antiquitäten. Er hat auch barocke **Sessel** im Angebot.

Der neue **Katalog** bietet Antiquitäten und hat auch barocke **Sessel** im Angebot.

Der hohe **Absatz** gefällt ihr sehr. Er macht die teuren **Schuhe** besonders.

Der hohe **Absatz** gefällt ihr sehr und macht die teuren **Schuhe** besonders.

Die teuren **Schuhe** sind reduziert. Sie stechen durch den hohen **Absatz** hervor.

Die teuren **Schuhe** sind reduziert und stechen durch den hohen **Absatz** hervor.

Das edle **Kissen** wird aus Seide genäht. Es ist mit einem feinen **Muster** bestickt.

Das edle **Kissen** wird aus Seide genäht und ist mit einem feinen **Muster** bestickt.

Das feine **Muster** stickt sie geduldig. Sie verschönert das edle **Kissen** damit.

Das feine **Muster** stickt sie geduldig und verschönert das edle **Kissen** damit.

Die jungen **Eltern** sind sehr stolz. Sie bringen ihre gesunde **Tochter** nachhause.
Die jungen **Eltern** sind sehr stolz und bringen ihre gesunde **Tochter** nachhause.
Die gesunde **Tochter** wird nachhause gebracht. Sie macht ihre jungen **Eltern** stolz.
Die gesunde **Tochter** wird nachhause gebracht und macht ihre jungen **Eltern** stolz.

Die ratlose **Polizei** sucht den Dieb. Sie findet in der leeren **Wohnung** Spuren.
Die ratlose **Polizei** sucht den Dieb und findet in der leeren **Wohnung** Spuren.
Die leere **Wohnung** zeigt keine Spuren. Sie hilft der ratlosen **Polizei** nicht.
Die leere **Wohnung** zeigt keine Spuren und hilft der ratlosen **Polizei** nicht.

Die starken **Bauern** arbeiten auf dem Feld. Sie lieben ihre grüne **Heimat** sehr.
Die starken **Bauern** arbeiten auf dem Feld und lieben ihre grüne **Heimat** sehr.
Die grüne **Heimat** wird sehr geliebt. Sie macht den starken **Bauern** viel Arbeit.
Die grüne **Heimat** wird sehr geliebt und macht den starken **Bauern** viel Arbeit.

Die bunten **Blätter** fallen von den Bäumen. Sie kündigen den kalten **Winter** an.
Die bunten **Blätter** fallen von den Bäumen und kündigen den kalten **Winter** an.
Der kalte **Winter** naht schnell. Er wird durch viele bunte **Blätter** angekündigt.
Der kalte **Winter** naht schnell und wird durch viele bunte **Blätter** angekündigt.

Der schwarze **Kaffee** ist heute stark. Er hilft bei der schweren **Prüfung** weiter.

Der schwarze **Kaffee** ist heute stark und hilft bei der schweren **Prüfung** weiter.

Die schwere **Prüfung** war früh. Sie wurde nur mit viel starkem **Kaffee** geschafft.

Die schwere **Prüfung** war früh und wurde nur mit viel starkem **Kaffee** geschafft.

Das große **Schiff** ist aufgelaufen. Es wurde vor dem schweren **Unglück** gerettet.

Das große **Schiff** ist aufgelaufen und wurde vor dem schweren **Unglück** gerettet.

Das schwere **Unglück** war unerwartet. Es macht dem großen **Schiff** Schwierigkeiten.

Das schwere **Unglück** war unerwartet und macht dem großen **Schiff** Schwierigkeiten.

Die edlen **Kleider** hängen im Laden. Sie werden im breiten **Spiegel** betrachtet.

Die edlen **Kleider** hängen im Laden und werden im breiten **Spiegel** betrachtet.

Der breite **Spiegel** steht im Laden. Er zeigt die edlen **Kleider** in ganzer Pracht.

Der breite **Spiegel** steht im Laden und zeigt die edlen **Kleider** in ganzer Pracht.

Der flinke **Kellner** ist geübt. Er balanciert die vollen **Teller** geschickt.

Der flinke **Kellner** ist geübt und balanciert die vollen **Teller** geschickt.

Die vollen **Teller** sind schwer. Sie werden von dem flinken **Kellner** serviert.

Die vollen **Teller** sind schwer und werden von dem flinken **Kellner** serviert.

Die jungen **Beamten** arbeiten fleißig. Sie setzen die geringe **Kritik** schnell um.

Die jungen **Beamten** arbeiten fleißig und setzen die geringe **Kritik** schnell um.

Die geringe **Kritik** sagt viel aus. Sie bringt den jungen **Beamten** einen Bonus.

Die geringe **Kritik** sagt viel aus und bringt den jungen **Beamten** einen Bonus.

Filler sentences

Die fleißige Biene bestäubt die **Blumen**. Sie genießt den warmen Sommer sehr.

Die fleißige Biene bestäubt die **Blumen** und genießt den warmen Sommer sehr.

Voller Angst erwacht er aus dem **Schlaf**. Er ist vom schlimmen Traum geschockt.

Voller Angst erwacht er aus dem **Schlaf** und ist vom schlimmen Traum geschockt.

Die dunklen Wolken bringen **Unheil**. Sie kündigen einen schlimmen Sturm an.

Die dunklen Wolken bringen **Unheil** und kündigen einen schlimmen Sturm an.

Das saftige Fleisch schneidet er mit dem **Messer**. Er genießt das deftige Essen.

Das saftige Fleisch schneidet er mit dem **Messer** und genießt das deftige Essen.

Den weiten Ozean liebt der **Kapitän**. Er steuert das große Schiff souverän.

Den weiten Ozean liebt der **Kapitän** und steuert sein großes Schiff souverän.

Die roten Blüten verschönern den **Garten.** Sie blühen den ganzen Sommer lang.

Die roten Blüten verschönern den **Garten** und blühen den ganzen Sommer lang.

Der üble Skandal steht in der **Presse.** Er bewegt die erhitzten Gemüter sehr.

Der üble Skandal steht in der **Presse** und bewegt die erhitzten Gemüter sehr.

Der junge Mann bearbeitet den **Antrag.** Er möchte heute früh Feierabend machen.

Der junge Mann bearbeitet den **Antrag** und möchte heute früh Feierabend machen.

Der strenge Prüfer erfragt die **Theorie.** Er lässt faule Studenten durchfallen.

Der strenge Prüfer erfragt die **Theorie** und lässt faule Studenten durchfallen.

In dunklen Stunden kommen ihr **Zweifel.** Sie stehlen ihr den benötigten Schlaf.

In dunklen Stunden kommen ihr **Zweifel** und stehlen ihr den benötigten Schlaf.

Die neue Firma versinkt in **Kosten.** Sie wird freiwillig Insolvenz anmelden.

Die neue Firma versinkt in **Kosten** und wird freiwillig Insolvenz anmelden.

Das spannende Buch hat viele **Kapitel.** Es wird von vielen Kindern gelesen.

Das spannende Buch hat viele **Kapitel** und wird von vielen Kindern gelesen.

Direkte Konkurrenz belebt den **Handel**. Sie lässt den täglichen Umsatz steigen.
Direkte Konkurrenz belebt den **Handel** und lässt den täglichen Umsatz steigen.

Die große Stadt unterstützt das **Theater**. Sie bietet so vielfältige Kultur an.
Die große Stadt unterstützt das **Theater** und bietet so vielfältige Kultur an.

Der fleißige Arzt erweitert die **Praxis**. Er kann mehr kranken Patienten helfen.
Der fleißige Arzt erweitert die **Praxis** und kann mehr kranken Patienten helfen.

Der mutige Soldat verweigert den **Befehl**. Er wird den Wehrdienst aufgeben.
Der mutige Soldat verweigert den **Befehl** und wird den Wehrdienst aufgeben.

Die junge Frau ist sehr tierlieb. **Katzen** lassen ihr Herz höher schlagen.
Die junge Frau ist sehr tierlieb und **Katzen** lassen ihr Herz höher schlagen.

Der Zoo hat Papageien bekommen. **Flügel** sollten bei ihnen gestutzt werden.
Der Zoo hat Papageien bekommen und **Flügel** sollten bei ihnen gestutzt werden.

Im Krankenhaus gibt es viele Verletzte. **Organe** werden sehr häufig gespendet.
Im Krankenhaus gibt es viele Verletzte und **Organe** werden sehr häufig gespendet.

Sie hat eine schlechte Nachricht erhalten. **Tränen** fließen über ihr Gesicht.
Sie hat eine schlechte Nachricht erhalten und **Tränen** fließen über ihr Gesicht.

Der neue Pokertisch wird eingeweiht. **Karten** werden an alle Spieler verteilt.
Der neue Pokertisch wird eingeweiht und **Karten** werden an alle Spieler verteilt.

Sie hat viel Geld geerbt. **Sorgen** muss sie sich nun keine mehr machen.
Sie hat viel Geld geerbt und **Sorgen** muss sie sich nun keine mehr machen.

Der alte Baum ist hoch gewachsen. **Wurzeln** ranken weit über den Boden.
Der alte Baum ist hoch gewachsen und **Wurzeln** ranken weit über den Boden.

Das Restaurant ist heute gut besucht. **Teller** fliegen nur so durch die Küche.
Das Restaurant ist heute gut besucht und **Teller** fliegen nur so durch die Küche.

Er freut sich auf das abendliche Vorlesen. **Märchen** hat er am liebsten.
Er freut sich auf das abendliche Vorlesen und **Märchen** hat er am liebsten.

Die schwere Klausur steht an. **Themen** werden zwei Tage zuvor bekannt gegeben.
Die schwere Klausur steht an und **Themen** werden zwei Tage zuvor bekannt gegeben.

Der Boden enthält viele Nährstoffe. **Gemüse** wächst darauf besonders gut.
Der Boden enthält viele Nährstoffe und **Gemüse** wächst darauf besonders gut.

Er hat einen großen Balkon. **Wäsche** trocknet hier ausgesprochen schnell.
Er hat einen großen Balkon und **Wäsche** trocknet hier ausgesprochen schnell.

Sie sortiert die Dokumente in Ordner. **Anlagen** kommen nach ganz hinten.
Sie sortiert die Dokumente in Ordner und **Anlagen** kommen nach ganz hinten.

Der Staat ist hoch verschuldet. **Banken** geben schon kein Geld mehr raus.
Der Staat ist hoch verschuldet und **Banken** geben schon kein Geld mehr raus.

Die Strömung der Nordsee trägt Sand ab. **Inseln** sind dadurch akut bedroht.
Die Strömung der Nordsee trägt Sand ab und **Inseln** sind dadurch akut bedroht.

Der Ball wird ein großes Spektakel. **Fürsten** aus ganz Europa sind geladen.
Der Ball wird ein großes Spektakel und **Fürsten** aus ganz Europa sind geladen.

Target word frequency and length

Experiment 1 and 2

Table 11. Word frequency and word length for all target words of experiment 1 and 2.

Experiment	Nr.	Target	Lemma	Word frequency	Word length
1, 2	1	Würmer	Wurm	9.48	6
1, 2	2	Löchern	Loch	28.60	7
1, 2	3	Tomaten	Tomate	6.06	7
1, 2	4	Kisten	Kiste	18.86	6
1, 2	5	Winter		51.23	6
1, 2	6	Vorräte	Vorrat	15.97	7
1, 2	7	Schatz	Schatz	23.44	6
1, 2	8	Palmen	Palme	5.87	6
1, 2	9	Fische	Fisch	42.66	6
1, 2	10	Bächen	Bach	21.63	6
1, 2	11	Treppe		39.50	6
1, 2	12	Kammer		37.06	6
1, 2	13	Burgen		15.01	6
1, 2	14	Strand		10.57	6
1, 2	15	Verkehr		60.41	7
1, 2	16	Morgen		92.96	6
1, 2	17	Bergen	Berg	58.21	6
1, 2	18	Hütten	Hütte	14.33	6
1, 2	19	Sterne	Stern	42.41	6
1, 2	20	Nächten*	Nacht	166.30	7
1, 2	21	Beeren*	Beere	3.43	6
1, 2	22	Kuchen		9.59	6
1, 2	23	Pferde	Pferd	76.31	6
1, 2	24	Ställe	Stall	16.44	6
1, 2	25	Gedicht		39.17	7
1, 2	26	Jungen	Junge	91.29	6
1, 2	27	Reifen		6.90	6
1, 2	28	Keller		24.93	6
1, 2	29	Briefe*	Brief	165.65	7
1, 2	30	Ordner*		2.19	6
1, 2	31	Scherz		11.27	6
1, 2	32	Lehrer		81.02	6
1, 2	33	Papier		44.42	6
1, 2	34	Drucker		6.89	7
1, 2	35	Orangen*	Orange	4.03	7
1, 2	36	Bäumen		74.56	6

1, 2	37	Räubern	Räuber	8.82	7
1, 2	38	Tresor*		0.76	7
1, 2	39	Verband		54.11	7
1, 2	40	Finger		66.47	6
1, 2	41	Melodie		17.80	7
1, 2	42	Klavier		12.32	7
1, 2	43	Strumpf		11.32	7
1, 2	44	Karton		5.23	6
1, 2	45	Füller*		0.34	6
1, 2	46	Tasche		38.08	6
1, 2	47	Teller		22.02	6
1, 2	48	Waffel*		0.43	6
1, 2	49	Zettel		15.42	6
1, 2	50	Schrank		15.98	7
1, 2	51	Bursche		22.59	7
1, 2	52	Schwert		14.89	7
1, 2	53	Spange*		0.81	7
1, 2	54	Haaren	Haar	79.30	6
1, 2	55	Hühner	Huhn	11.90	6
1, 2	56	Federn	Feder	22.56	6
1, 2	57	Gemälde		16.82	7
1, 2	58	Museum		27.84	6
1, 2	59	Birnen	Birne	5.40	6
1, 2	60	Leiter		69.21	6
1, 2	61	Schwein		18.36	7
1, 2	62	Dusche*		2.41	6
1, 2	63	Stifte	Stift	8.52	6
1, 2	64	Hocker*		2.02	6

*these targets were eventually excluded from the analysis.

Experiment 3

Table 12. Word frequency and word length for all target words of experiment 3.

Experiment	Nr.	Target	Lemma	Word frequency	Orthographic regularity	Word length
3	1	Mensch		714	33.82	6
3	2	Zeiten	Zeit	857.08	14.13	6
3	3	Jahren	Jahr	1085.85	10.7	6
3	4	Fragen	Frage	534.47	44.52	6
3	5	Kinder	Kind	493.52	20.98	6
3	6	Arbeit		391.06	10.7	6
3	7	Bilder	Bild	269.28	24.83	6
3	8	Gedanke		192.74	20.12	7
3	9	Person		188.04	62.93	6
3	10	Gesicht		239.15	69.35	7
3	11	Kräfte	Kraft	265.26	11.99	6
3	12	Gründe	Grund	295.84	14.55	6
3	13	Stelle		237.53	69.35	6
3	14	Freund		174.24	40.24	6
3	15	Straßen	Straße	167.03	117.72	7
3	16	Kirche		172.76	20.98	6
3	17	Mutter		264.38	14.55	6
3	18	Stimme		227.83	35.53	6
3	19	Stunde		229.47	41.52	6
3	20	Wasser		195.9	12.41	6
3	21	Sprache		167.66	35.53	7
3	22	Mittel		167.44	23.12	6
3	23	Aufgabe		227.33	30.82	7
3	24	Monate	Monat	152.18	47.09	6
3	25	Schritt		151.92	533.4	7
3	26	Sachen	Sache	224.45	19.26	6
3	27	Briefe	Brief	165.65	43.24	6
3	28	Gebiete	Gebiet	264.22	47.95	7
3	29	Gelder	Geld	173	33.4	6
3	30	Frauen	Frau	625.21	44.52	6
3	31	Blicke	Blick	172	12.84	6
3	32	Reihen	Reihe	152.02	38.1	6
3	33	Evidenz		2.9	2.14	7
3	34	Ökonom		2.71	0.86	6
3	35	Enklave		1.37	4.28	7
3	36	Olympia		2.36	3.85	7
3	37	Pfützen	Pfütze	2.51	2.14	7

3	38	Lektor		2.46	1.71	6
3	39	Emotion		2.89	0.43	7
3	40	Lyriker		2.4	3.42	7
3	41	Zoologe		1.13	4.28	7
3	42	Juwelen	Juwel	2.14	0.86	7
3	43	Phönix		1.83	0.56	6
3	44	Rüffel		1.44	1.71	6
3	45	Kajüte		2	3.85	6
3	46	Egoist		1.01	3.85	6
3	47	Vehikel		1.68	1.28	7
3	48	Akustik		1.24	1.28	7
3	49	Amulett		1.01	3	7
3	50	Auktion		2.13	1.71	7
3	51	Exzesse		2.67	2.14	7
3	52	Kakteen		1.05	3.85	7
3	53	Exkurs		2.38	1.71	6
3	54	Zubehör		2.93	2.14	7
3	55	Vokabel		2.22	3.85	7
3	56	Zäsuren	Zäsur	1.22	0.43	7
3	57	Zepter		1.35	2.14	6
3	58	Rokoko		2.95	4.28	6
3	59	Kuvert		2.31	0.43	6
3	60	Äquator		1.86	1.28	7
3	61	Razzia		1.33	4.71	6
3	62	Blässe		1.77	3.42	6
3	63	Favorit		1.28	4.71	7
3	64	Gitarre		2.6	4.28	7

Experiment 4

Table 13. Word frequency and word length for all target words of experiment 4.

Experiment	Nr.	Target	Lemma	Word frequency	Word length
4	1	Würmer	Wurm	9.48	6
4	2	Löchern	Loch	28.60	7
4	3	Tomaten	Tomate	6.06	7
4	4	Kisten	Kiste	18.86	6
4	5	Verkehr		60.41	7
4	6	Morgen		92.96	6
4	7	Hütten	Hütte	14.33	6
4	8	Bergen	Berg	58.21	6
4	9	Himmel		97.55	6
4	10	Sterne	Stern	42.41	6
4	11	Kuchen		9.59	6
4	12	Birnen	Birne	5.40	6
4	13	Schüler		81.39	7
4	14	Gedicht		39.17	7
4	15	Richter		69.22	7
4	16	Anwalt		15.00	6
4	17	Lehrer		81.02	6
4	18	Scherze	Scherz	11.27	7
4	19	Papier		44.42	6
4	20	Drucker		6.89	7
4	21	Räuber		8.82	6
4	22	Schmuck		13.77	7
4	23	Finger		66.47	6
4	24	Verband		54.11	7
4	25	Klavier		12.32	7
4	26	Melodie		17.80	7
4	27	Schwert		14.89	7
4	28	Bursche		22.59	7
4	29	Klammer		5.10	7
4	30	Haaren	Haar	79.30	6
4	31	Gemälde		16.82	7
4	32	Museum		27.84	6
4	33	Kamera		11.38	6
4	34	Urlaub		19.15	6
4	35	Beifall		31.82	7
4	36	Musiker		18.18	7
4	37	Telefon		16.10	7

4	38	Termin		16.50	6
4	39	Gepäck		9.60	6
4	40	Partner		30.52	7
4	41	Sessel		17.25	6
4	42	Katalog		10.81	7
4	43	Absatz		54.33	6
4	44	Schuhe		28.46	6
4	45	Kissen		11.87	6
4	46	Muster		25.52	6
4	47	Eltern		84.47	6
4	48	Tochter		78.31	7
4	49	Polizei		69.31	7
4	50	Wohnung		88.22	7
4	51	Bauern	Bauer	82.27	6
4	52	Heimat		53.26	6
4	53	Blätter	Blatt	93.25	7
4	54	Winter		51.23	6
4	55	Kaffee		39.04	6
4	56	Prüfung		42.48	7
4	57	Schiff		87.83	6
4	58	Unglück		28.47	7
4	59	Kleider	Kleid	47.85	7
4	60	Spiegel		41.43	7
4	61	Kellner		19.63	7
4	62	Teller		22.02	6
4	63	Beamten	Beamte	86.39	7
4	64	Kritik		72.87	6

Examples of experimental materials for experiments 1, 3, and 4

Hallo und herzlich Willkommen!

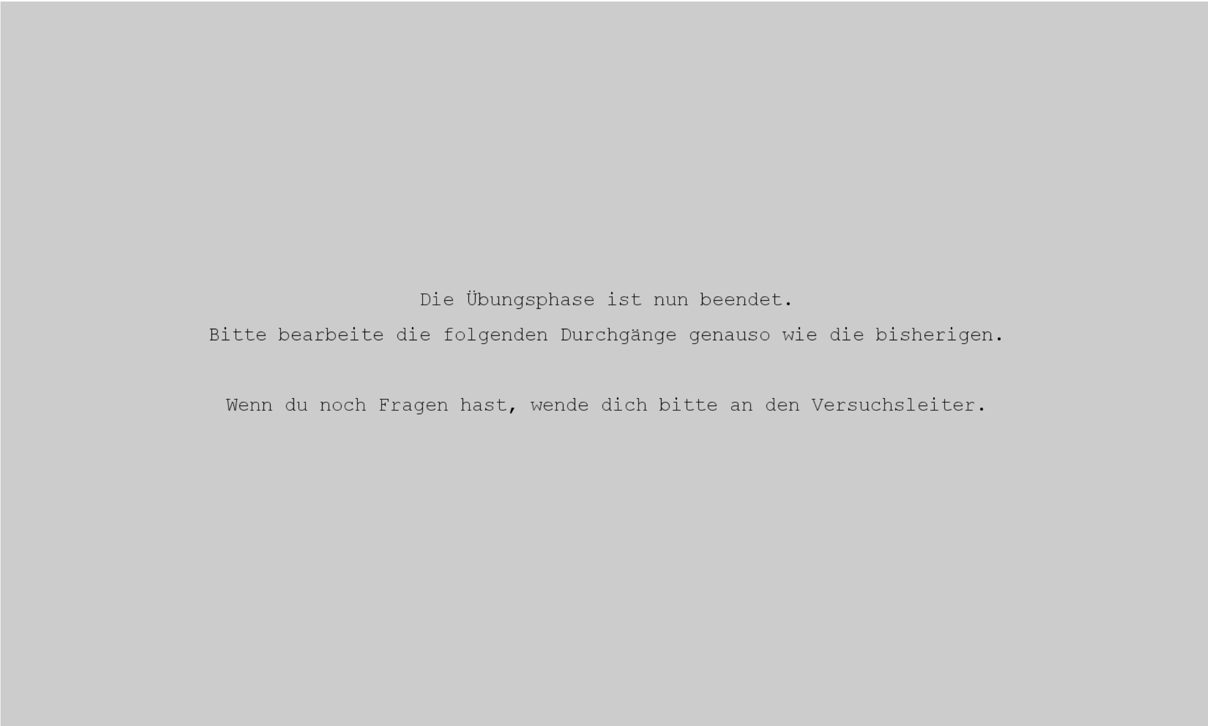
Im Folgenden werden dir einzelne Sätze präsentiert.
Bitte lies den Satz einmal und in deinem normalen Lesetempo leise durch.
Nachdem du den Satz gelesen hast, erscheint am Ende des Satzes ein Wort.
Schaue bitte **so schnell und genau wie möglich** auf dieses Wort in dem zuvor gelesenen Satz zurück und überprüfe, ob es einen Rechtschreibfehler enthält.

Ist ein **Fehler** enthalten, drücke bitte den **linken** Knopf.
Ist **kein Fehler** enthalten, drücke bitte den **rechten** Knopf.

Zwischen den Durchgängen werden dir manchmal Fragen zum letzten Satz gestellt, die du bitte beantwortest.

In diesem Experiment werden deine Augenbewegungen aufgezeichnet werden.
Um den Eyetracker zu kalibrieren werden dir zunächst nacheinander kleine Kreise präsentiert.
Bitte schau immer möglichst genau in das Zentrum dieser Kreise.
Im weiteren Verlauf der Untersuchung wird dieser Vorgang noch ein paar mal wiederholt.

Nun folgen einige Übungsdurchgänge.
Falls du noch Fragen hast, wende dich bitte jetzt an den Versuchsleiter.


Die Übungsphase ist nun beendet.
Bitte bearbeite die folgenden Durchgänge genauso wie die bisherigen.
Wenn du noch Fragen hast, wende dich bitte an den Versuchsleiter.

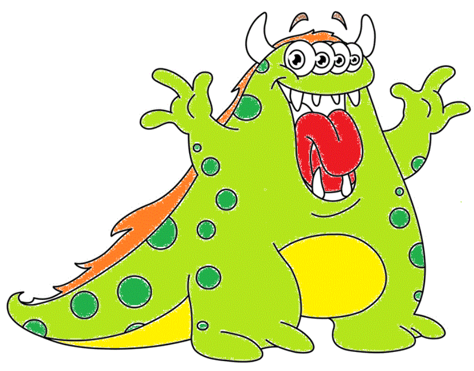

Du hast es geschafft!
Vielen Dank für deine Teilnahme!
Dieser Teil des Experimentes ist damit beendet.
Nun folgen noch ein paar Aufgaben ohne Eyetracker.

Experimental materials of experiment 2

Instructions for children were presented auditive and additionally explained by the experimenter, if needed.

Super, dass du heute hier bist und uns bei unserem Forschungsprojekt hilfst!
Find ich total klasse!
Ach ja, ich bin übrigens Flori – der Lesefuchs
(wie man -hier- hinter mir an den vielen Büchern erkennen kann!)
Willst du wissen, was wir forschen? Ich verrate es dir!
Wir versuchen herauszufinden, ob du ein Lesefuchs bist- so wie ich. Dafür zeige ich dir gleich ein paar Sätze, die du bitte **leise liest**. Lies die Sätze **so wie du sonst auch liest**. Lies bitte jeden Satz nur **ein Mal** und versuche ihn zu verstehen. Ich frag dich ab und zu, was du gelesen hast und du sollst dann mit „Ja" oder „Nein" antworten!

Heute ist auch mein Freund dabei, dass kleine Fehlermonster. Um mich ein bisschen zu ärgern, schummelt er immer wieder Fehler in meine Sätze.
Komm, hilf mir alle Fehler zu finden.



Wir schauen mal, wie das so klappt…!
Ach ja… Und damit du gut sehen kannst, musst du mit deinem Kinn auf der Kinnstütze bleiben und manchmal auf Kreise gucken… das üben wir jetzt erst einmal und die Frau Friede erklärt dir das gleich alles noch noch mal ganz genau.
Drücke einfach einen Knopf, dann geht´s los!

Nach den Übungssätzen:
Das machst du schon richtig gut!
Versuch noch ein bisschen **ruhiger** zu sitzen, damit du auch alles richtig lesen kannst.
Wenn du einen Satz zu Ende gelesen hast, dann denk daran **einen Knopf zu drücken**, damit wir weitermachen können.
Ähm…was hab ich gesagt, was war noch mal wichtig beim Lesen?...
Ach ja…so lesen wie du sonst auch liest und jeden Satz nur ein Mal lesen und versuchen ihn zu verstehen!
Nun viel Spaß mit meinen anderen Sätzen!
Konzentriere dich und versuch dein Bestes!
Hast du noch eine Frage? Wenn nicht, drücke einen Knopf und los geht's.

Pause:
Das machst du echt toll! Jetzt machen wir eine kurze Pause,
um uns ein bisschen locker zu machen. Und um zu schauen,
wie gut deine Augen sind. Vielleicht hast du ja so gute Augen wie ein Lesefuchs?!
Wir schauen einfach mal nach…

Nach Pause:

Nur noch ein paar Sätze…wenn du das schaffst, dann bist du so gut wie ein Lesefuchs!

Konzentriere dich und lies die Sätze nur ein Mal!

Ich drück dir die Daumen!

Es geht weiter, wenn du einen Knopf drückst.

Ende:

Ich bin echt begeistert von dir!!

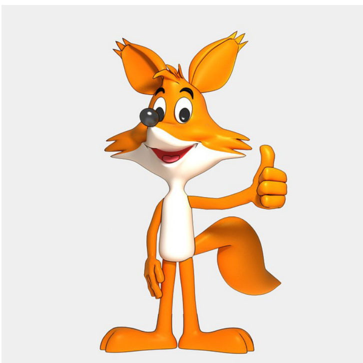
Mensch, du hast es geschafft alle Sätze zu lesen! Großartig!!!

Du hast uns Forschern bei unserem Projekt sehr geholfen! Vielen Dank dafür!

Und…ach ja…

Herzlichen Glückwunsch!

Du bist ein super Lesefuchs!!



Letters sent to primary schools

BERGISCHE
UNIVERSITÄT
WUPPERTAL

Anne Friede, M,Sc.

Bergische Universität Wuppertal, Anne Friede
Max-Horkheimer-Str. 20, 42119 Wuppertal

Adresse der Schule

Fachbereich G
Allgemeine und Biologische Psychologie
Max-Horkheimer-Str. 20, 42119 Wuppertal
RAUM Z.01.07
TELEFON +49 (0)202 439 2067
FAX +49 (0)202 4392926
MAIL afriede@uni-wuppertal.de

DATUM

Sehr geehrte Frau XXX (Direktorin der Schule),

vielen Dank für Ihre grundsätzliche Bereitschaft sich mit unserem Forschungsprojekt zum Thema **Leseerwerb** auseinanderzusetzen.

Wie bereits telefonisch besprochen, sende ich Ihnen weitere Informationen über Inhalt und Ablauf, als Grundlage für eine mögliche Zustimmung zur Studienteilnahme, zu.

Das Ziel unseres Projektes ist es, besser zu verstehen, wie Informationen beim Lesen aufgenommen und verarbeitet werden. Besonders interessiert uns dabei, welche Rolle das visuell-räumliche Gedächtnis bei der Entwicklung des Lesens spielt, denn immer wieder stellen Lehrer und Lehrerinnen sowie Therapeuten und Therapeutinnen fest, dass Kinder mit Leseschwäche Schwierigkeiten haben, sich in Texten zu orientieren, dadurch Zeilen und Wörter verlieren und dann viel langsamer lesen und mehr Fehler machen. Mit der Teilnahme Ihrer Schule können Sie uns helfen, den normalen Leseprozess besser zu verstehen, damit wir Ursachen für Leseschwächen finden und ein wissenschaftlich basiertes Therapieverfahren entwickeln können.

Dazu suchen wir zunächst **Kinder der Klassenstufe 4,** die an unserer Untersuchung (geplanter Zeitraum für alle Untersuchungen ist der 24.02.-04.04.14) teilnehmen möchten. Dabei sitzen die Kinder vor einem Computer, lesen uns Sätze vor und eine kleine Kamera, die sich unter dem Computerbildschirm befindet, filmt deren Blickbewegungen.

Zusätzlich möchten wir zwei Lesetests (ELFE 1-6, Lenhard & Schneider 2006; SLS 1-4, Mayringer & Wimmer 2003, beides Tests zur Untersuchung des Leseverstehens auf Wort- und Satzebene) sowie die Untertests zum visuell-räumlichen Arbeitsgedächtnis aus der Arbeitsgedächtnistestbatterie für Kinder von 5 bis 12 Jahren (AGTB 5-12, Hasselhorn et al. 2012) mit den Kindern durchführen. Insgesamt werden unsere Untersuchungen etwa eine Unterrichtsstunde gemeinsam mit allen Kindern, die teilnehmen, und noch einmal eine Unterrichtsstunde mit jedem Kind einzeln, dauern. Gern führen wir Untersuchungen auch im Nachmittagsbereich, mit den Kindern, die den offenen Ganztag besuchen, durch.

Sollten Sie sich für eine Teilnahme Ihrer Schule an unserer Studie entscheiden, bitten wir Sie, die Einverständniserklärung und die Beschreibung unserer Studie an die Kinder bzw. die Eltern zu verteilen. Zudem möchten wir Sie bitten, uns einen Raum zur Verfügung zu stellen, der abgedunkelt werden kann (durch Vorhänge/Sonnenrollo). Mehr Aufwand wird für ihre Schule nicht entstehen.

Die Ergebnisse der Untersuchung und der Lesediagnostik werden auf Wunsch den Eltern zur Verfügung gestellt.

Selbstverständlich stehen wir für Gespräche mit Lehrern und Eltern zur Verfügung.

Unser Team ist in der Forschung zu Lesen und Lesestörungen international ausgewiesen. Bei Fragen können Sie sich gerne über die oben genannten Kontaktdaten mit uns in Verbindung setzen.

Weitere Informationen über uns erhalten Sie auch im Internet unter:

http://www.allgemeinepsychologie.uni-wuppertal.de/team.html

BERGISCHE
UNIVERSITÄT
WUPPERTAL

Wir würden uns sehr freuen, wenn Sie sich entschließen, die Leseforschung mit Ihrer Schule zu unterstützen.

Mit freundlichen Grüßen

Prof. Dr. Ralph Radach
Lehrstuhl für Allgemeine
Prodekan Fachbereich G
„Bildungs- und Sozialwissenschaften“

M. Sc. Anne Friede
Lehr- und Forschungslogopädin

Parental information letter

BERGISCHE
UNIVERSITÄT
WUPPERTAL

Anne Friede, M.Sc.

Fachbereich G
Allgemeine und Biologische Psychologie
Max-Horkheimer-Str. 20, 42119 Wuppertal
RAUM Z.01.07
TELEFON +49 (0)202 439 2067
FAX +49 (0)202 4392926
MAIL afriede@uni-wuppertal.de

DATUM Februar 2014

Bergische Universität Wuppertal, Anne Friede
Max-Horkheimer-Str. 20, 42119 Wuppertal

An die Eltern der Klasse 4
der GGS XXX

Sehr geehrte Eltern,

in einem Forschungsprojekt zum Thema **Leseentwicklung** an der Bergischen Universität Wuppertal möchten wir das Lesen bei Grundschulkindern untersuchen. Besonders interessiert uns dabei, welche Rolle das räumliche Gedächtnis bei der Entwicklung des Lesens spielt, denn immer wieder stellen Lehrerinnen und Lehrer sowie Eltern fest, dass Kinder Schwierigkeiten haben, sich in Texten zu orientieren, dadurch Zeilen und Wörter verlieren und dann viel langsamer lesen und mehr Fehler machen. Mit der Teilnahme Ihres Kindes können Sie uns helfen den normalen Leseprozess besser zu verstehen, damit wir Ursachen für Leseschwächen finden können. Im Gegenzug werden wir Sie persönlich informieren, falls bei ihrem Kind Auffälligkeiten bestehen.

In der Untersuchung werden die Augenbewegungen Ihres Kindes beim Lesen von Sätzen am Computer aufgezeichnet. Das geschieht in einer für die Kinder vertrauten Vorgehensweise ähnlich einem Sehtest beim Augenarzt. Die Kinder sitzen vor einem Computer und lesen Sätze vor. Eine kleine Kamera, die sich unter dem Computerbildschirm befindet, filmt die Augenbewegungen. Gemeinsam mit den Daten aller anderen teilnehmenden Kinder können wir dann sehen, welche Lesestrategien Kinder anwenden und wie sich diese im Laufe der Zeit entwickeln. Zusätzlich möchten wir zwei weitere Lesetests zum Leseverstehen durchführen, alles zusammen wird etwa zwei Unterrichtsstunden dauern. Die Aufgaben werden von einer Mitarbeiterin unseres Teams in den Räumlichkeiten der Schule durchgeführt.

Die Daten Ihres Kindes werden selbstverständlich anonymisiert und nach den gesetzlichen Bestimmungen für Datenschutz behandelt. Haben Sie Interesse an den

Ergebnissen Ihres Kindes, erhalten Sie eine individuelle Übersicht sowie eine objektivierte Auswertung. Dazu kreuzen Sie bitte das entsprechende Kästchen in der Einverständniserklärung an, um einer Teilanonymisierung der Daten zuzustimmen. Infolgedessen können wir die erhobenen Daten einmalig Ihrem Kind zuordnen und Ihnen zugänglich machen, bevor sie anschließend wieder anonymisiert werden.
In der Vergangenheit hat sich gezeigt, dass Kinder unsere Untersuchungen interessant finden und sehr gerne daran teilnehmen.

Unser Team ist in der Forschung zum Lesen und zu Lesestörungen international ausgewiesen. Bei Fragen können Sie sich gerne über die oben genannten Kontaktdaten mit uns in Verbindung setzen.
Weitere Informationen über uns erhalten Sie im Internet unter:
http://www.allgemeinepsychologie.uni-wuppertal.de/team.html

Wir freuen uns, wenn Sie und Ihr Kind unsere Leseforschung unterstützen. Dazu geben Sie Ihrem Kind bitte die unterschriebene Einverständniserklärung mit in die Schule. Wir werden Sie dann frühzeitig darüber informieren, wann wir an die Schule Ihres Kindes kommen.

Mit freundlichen Grüßen

______________________________	______________________
Prof. Dr. Ralph Radach	Anne Friede, M.Sc.
Prodekan Fachbereich G	Lehr- und Forschungs-
Bildungs- und Sozialwissenschaften	logopädin

EINVERSTÄNDNISERKLÄRUNG

Hiermit erkläre ich mich damit einverstanden, dass mein Sohn/meine Tochter

____ ______

Vorname/ Nachname in Druckschrift Klasse

an der Studie „Räumliches Gedächtnis und Lesen" teilnimmt.

Duisburg,

____________________ ______________________________

Ort, Datum Unterschrift des/ der Sorgeberechtigten

☐ Ich würde gerne eine individuelle Übersicht und eine objektivierte Auswertung der Ergebnisse erhalten und stimme einer Teilanonymisierung der Daten zu.

☐ Ich bin damit einverstanden, dass die Ergebnisse meines Kindes an die Klassenlehrerin, Frau XXX, weitergegeben werden.

Feedback for classroom teacher

BERGISCHE UNIVERSITÄT WUPPERTAL

Anne Friede, M.Sc.

Bergische Universität Wuppertal, Anne Friede
Max-Horkheimer-Str. 20, 42119 Wuppertal

Fachbereich G
Allgemeine und Biologische Psychologie
Max-Horkheimer-Str. 20, 42119 Wuppertal
RAUM Z.01.07
TELEFON +49 (0)202 439 2067
FAX +49 (0)202 4392926
MAIL afriede@uni-wuppertal.de
DATUM

Adresse der Schule

Sehr geehrte Frau XXX (Klassenlehrerin),

herzlichen Dank für die Teilnahme Ihrer Klasse XX am Projekt „Räumliches Gedächtnis und Lesen“. Die Erfassung der Daten ist nun abgeschlossen und ich freue mich, Ihnen die Ergebnisse der teilnehmenden Schüler und Schülerinnen mitzuteilen.

Die Testergebnisse liefern eine orientierende Einschätzung, jedoch keine hinreichende Diagnostik des Leistungsniveaus der einzelnen Kinder. Die Testverfahren wurden nicht individuell auf die Kinder abgestimmt, sondern im Vorfeld für die Studie ausgewählt. Zur Erstellung individueller Leistungsprofile wären weitere Untersuchungen notwendig.

Einige Ergebnisse der Kinder werden im Folgenden in Prozenträngen berichtet. Der Normalbereich liegt innerhalb eines Prozentrangs von 15-85. Ab einem Prozentrang von 86 spricht man von einer überdurchschnittlichen Leistung, ab einem Prozentrang von 14 von einer unterdurchschnittlichen Leistung.

Dürchgeführt wurden:

1) Leseverständnistest (ELFE 1-6, Lenhard & Schneider, 2006)

Es wurden die Untertests zum Wortverständnis und zum Satzverständnis durchgeführt. Das Verständnis auf Wortebene wird durch Wort-Bild-Zuordnungsaufgaben überprüft. Dazu muss zu einem vorgegebenen Bild ein Wort aus einer Auswahlmenge von 4 Wörtern ausgewählt werden.

Beim Untertest Satzverstehen müssen Lückensätze mit einem Wort aus einer Auswahlmenge von 5 Wörtern vervollständigt werden. Für beide Aufgaben hat das Kind je 3 Minuten Zeit.

Das Diagramm in der folgenden Abbildung zeigt die Verteilung der Ergebnisse in beiden Klassen. XX% (= XX von XX Kindern) der Kinder erreichten durchschnittliche Leistungen, XX% (= XX von XX Kindern) erreichten überdurchschnittliche Ergebnisse. Kein Kind zeigte in diesem Test unterdurchschnittliche Leistungen.

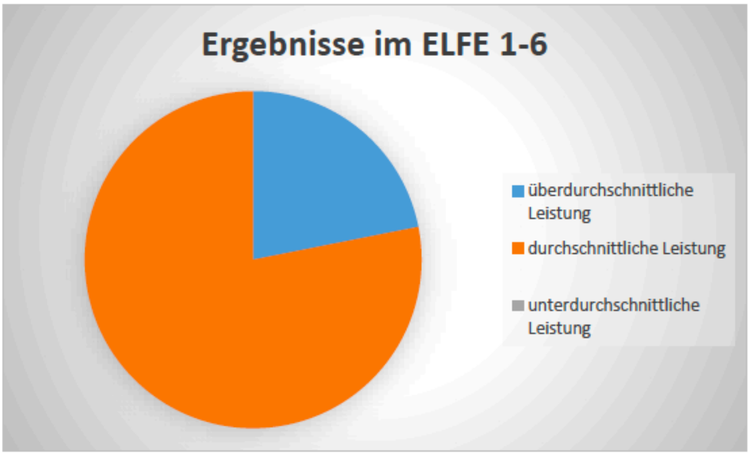


Abb. 1: Verteilung der Ergebnisse im ELFE 1-6 in der Klasse XX der GGS XX.

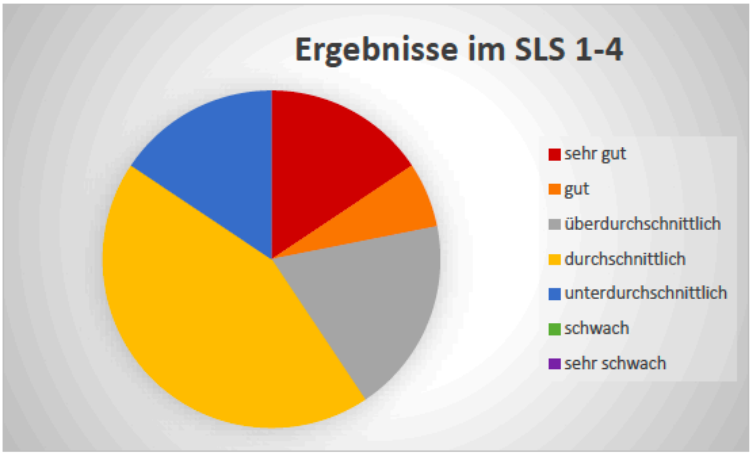
2) Salzburger Lesescreening (SLS 1-4, Mayringer & Wimmer, 2003)

Mit dem SLS 1-4 wird die Lesegeschwindigkeit gemessen. Den Kindern wird eine Liste richtiger und falscher Sätze vorgelegt (z.B. „Bananen sind blau“), die möglichst schnell gelesen und nach ihrer Richtigkeit beurteilt werden sollen. Gemessen wird, wie viele Sätze die Kinder in 3 Minuten korrekt bearbeiten können und daraus ein „Lesequotient“ berechnet.

Leseleistung (LQ)	Leistung
≥ 130	sehr gut
120 – 129	gut
110 – 119	überdurchschnittlich
90 – 109	durchschnittlich
80 – 89	unterdurchschnittlich
70 – 79	schwach
≤ 69	sehr schwach

Die Tabelle zeigt einen Überblick über die Zuordnung der erreichten Punkte zur entsprechenden Leseleistung.

XX% (= XX von XX Kinder) der Kinder der Klasse XX erreichten ein durchschnittliches Ergebnis im SLS 1-4. XX% (= XX von XX Kindern) konnten eine überdurchschnittliche Leistung erreichen, X% (= X von XX Kindern) eine gute und XX% (= X von XX Kindern) eine sehr gute Leistung. XX% (= XX von XX Kindern) der Kinder schnitten unterdurchschnittlich ab. Kein Kind erreichte eine schwache oder sehr schwache Leistung.



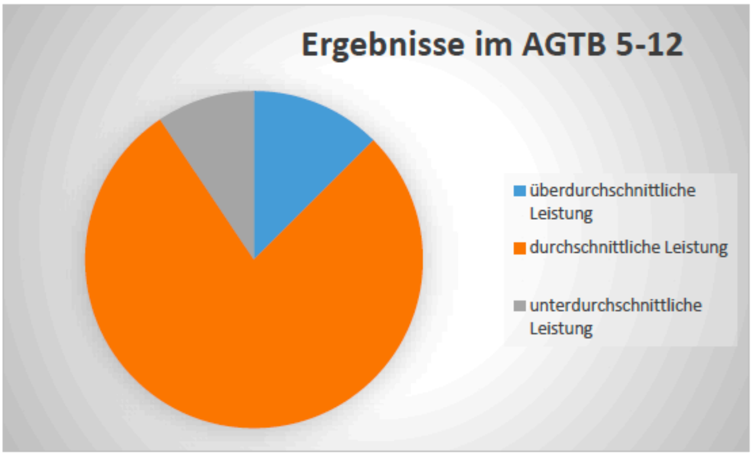
3) Visuell-räumliches Arbeitsgedächtnis

Mit der Arbeitsgedächtnisbatterie (AGTB 5-12, Hasselhorn et al., 2012) wurden die Leistungen des visuell-räumlichen Gedächtnisses durch Rekonstruktionsaufgaben überprüft.

Bei dem Test **„Matrix“** werden in einem Raster von 4 x 4 Feldern einige der Felder für kurze Zeit schwarz eingefärbt, so dass sich ein schwarz-weißes Muster ergibt. Danach soll durch Antippen der entsprechenden Felder auf einem Touchscreen das Muster rekonstruiert werden.

Beim Test **„Corsi-Block“** werden nacheinander Smileys in unterschiedlicher Reihenfolge in neun verschiedenen Feldern gezeigt. Direkt nach der Präsentation der Smileys soll das Kind in der richtigen Reihenfolge die Felder antippen, in denen es zuvor die Smileys gesehen hat.

Im Test des visuell-räumlichen Arbeitsgedächtnisses zeigten XX% (= XX von XX Kindern) der Kinder durchschnittliche Leistungen, XX% (= X von XX Kindern) überdurchschnittliche Leistungen und X% (= X von XX Kindern) unterdurchschnittliche Leistungen.



Bei Rückfragen stehe ich gern zur Verfügung.

Freundliche Grüße

Anne Friede, M.Sc.

Lehr- und Forschungslogopädin

Feedback for parents

BERGISCHE
UNIVERSITÄT
WUPPERTAL

Anne Friede, M.Sc.

Fachbereich G
Allgemeine und Biologische Psychologie
Max-Horkheimer-Str. 20, 42119 Wuppertal
RAUM Z.01.07
TELEFON +49 (0)202 439 2067
FAX +49 (0)202 4392926
MAIL afriede@uni-wuppertal.de

Bergische Universität Wuppertal, Anne Friede
Max-Horkheimer-Str. 20, 42119 Wuppertal

An Familie

«Familienname»

Duisburg

Sehr geehrte Familie XX,

herzlichen Dank für die Teilnahme Ihres Kindes XX am Projekt „Räumliches Gedächtnis und Lesen". Die Erfassung der Daten ist abgeschlossen und ich freue mich, Ihnen die Ergebnisse Ihres Kindes mitzuteilen.

Die Testergebnisse liefern eine orientierende Einschätzung, jedoch keine hinreichende Diagnostik des Leistungsniveaus. Die Testverfahren wurden nicht individuell auf ihr Kind abgestimmt, sondern im Vorfeld für die Studie ausgewählt. Zur Erstellung eines individuellen Leistungsprofils wären weitere Untersuchungen notwendig.

Einige Ergebnisse ihres Kindes werden im Folgenden in Prozenträngen berichtet. Der Normalbereich liegt innerhalb eines Prozentrangs von 15-85. Ab einem Prozentrang von 86 spricht man von einer überdurchschnittlichen Leistung, ab einem Prozentrang von 14 von einer unterdurchschnittlichen Leistung.

Übersicht der Ergebnisse

1) Leseverständnistest (ELFE 1-6, Lenhard & Schneider, 2006)

Es wurden die Untertests zum Wortverständnis und zum Satzverständnis durchgeführt. Das Verständnis auf Wortebene wird durch Wort-Bild-Zuordnungsaufgaben überprüft. Dazu muss zu einem vorgegebenen Bild ein Wort aus einer Auswahlmenge von 4 Wörtern ausgewählt werden.

XX hat im Wortverständnis einen Prozentrang[1] von XX erreicht. Das bedeutet, dass XX % der Kinder derselben Klassenstufe ein gleiches oder niedrigeres Verständnis beim Lesen von Wörtern haben.

Beim Untertest Satzverstehen müssen Lückensätze mit einem Wort aus einer Auswahlmenge von 5 Wörtern vervollständigt werden. Für beide Aufgaben hat das Kind je 3 Minuten Zeit.

Im Satzverständnis hat XX einen Prozentrang von XX erreicht. Das bedeutet, dass XX % der Kinder derselben Klassenstufe ein gleiches oder niedrigeres Verständnis beim Lesen von Sätzen haben.

2) Salzburger Lesescreening (SLS 1-4, Mayringer & Wimmer, 2003)

Mit dem SLS 1-4 wird die Lesegeschwindigkeit gemessen. Den Kindern wird eine Liste richtiger und falscher Sätze vorgelegt (z.B. „Bananen sind blau"), die möglichst schnell gelesen und nach ihrer Richtigkeit beurteilt werden sollen. Gemessen wird, wie viele Sätze die Kinder in 3 Minuten korrekt bearbeiten können.

Innerhalb von drei Minuten hat XX von insgesamt 70 Sätzen XX Sätze gelesen und korrekt beurteilt. Auf dieser Grundlage wurde ein „Lesequotient" von XX ermittelt. XX zeigt somit eine XX Leistung bei der Überprüfung der Lesefähigkeit (siehe Tabelle).

Leseleistung (LQ)	**Leistung**
≥ 130	sehr gut
120 – 129	gut
110 – 119	überdurchschnittlich
90 – 109	durchschnittlich
80 – 89	unterdurchschnittlich
70 – 79	schwach
≤ 69	sehr schwach

Die Tabelle zeigt einen Überblick über die Zuordnung der erreichten Punkte zur entsprechenden Leseleistung.

[1] Erreicht eine Person den Prozentrang der Höhe a, besagt dieser, dass a% der Testpersonen der Vergleichsstichprobe eine niedrigere (oder gleiche) Ausprägung des Testmerkmals aufweisen.

3) Visuell-räumliches Arbeitsgedächtnis

Mit der Arbeitsgedächtnisbatterie (AGTB 5-12, Hasselhorn et al., 2012) wurden die Leistungen des visuell-räumlichen Gedächtnisses durch Rekonstruktionsaufgaben überprüft.

Bei dem Test **„Matrix“** werden in einem Raster von 4 x 4 Feldern einige der Felder für kurze Zeit schwarz eingefärbt, so dass sich ein schwarz-weißes Muster ergibt. Danach soll durch Antippen der entsprechenden Felder auf einem Touchscreen das Muster rekonstruiert werden.

Beim Test **„Corsi-Block“** werden nacheinander Smileys in unterschiedlicher Reihenfolge in neun verschiedenen Feldern gezeigt. Direkt nach der Präsentation der Smileys soll das Kind in der richtigen Reihenfolge die Felder antippen, in denen es zuvor die Smileys gesehen hat.

XX hat in den Aufgaben zum visuell-räumlichen Gedächtnis einen Prozentrang von XX erreicht. Das bedeutet, dass XX % der Kinder derselben Klassenstufe ein gleiches oder niedrigeres Ergebnis erzielten.

4) Überprüfung des räumlichen Sehens

Das räumliche Sehen hat eine wichtige Funktion beim Lesen und wurde aus diesem Grund im Rahmen der Studie mit dem TNO-Test (Stereopsistest) überprüft.

XX hat einen Stereopsiswinkel von
XX Bogensekunden und somit ein
XX räumliches Sehen.

Stereopsiswinkel	Qualität des räumlichen Sehens
480 Sekunden 240 Sekunden	ausreichend bis befriedigend
120 Sekunden 60 Sekunden	gut
30 Sekunden 15 Sekunden	sehr gut

Bei Rückfragen stehe ich gern zur Verfügung.

Freundliche Grüße

____

Anne Friede, M.Sc.

Lehr- und Forschungslogopädin

Feedback for children

URKUNDE

verliehen an

____

für die erfolgreiche Teilnahme
an einem Experiment
im Forschungsbereich

Lesen und Leseentwicklung

Duisburg, den

Prof. Dr. Ralph Radach

Lehrstuhl für Allgemeine und
Biologische Psychologie

ANOVA Output experiment 1

Table 14. R-Output for Target Position X Shift X Repetition ANOVAs for length of initial regression of experiment 1.

Length of initial regression	df	Type III Sum of Squares	Mean Square	*F*-Value	*p*-Value
Target Position	1	0.545	0.5450	9.101	0.00274
Shift Condition	1	0.032	0.0316	0.528	0.46783
Repetition	1	0.064	0.0637	1.063	0.30313
Target Position X Shift Condition	1	0.145	0.1453	2.427	0.12014
Target Position X Repetition	1	0.068	0.0679	1.134	0.28775
Shift Condition X Repetition	1	0.010	0.0103	0.172	0.67821
Target Position X Shift Condition X Repetition	1	0.003	0.0025	0.042	0.83745
Residuals	352	21.077	0.0599		

Table 15. R-Output for Target Position X Shift X Repetition ANOVAs for number of regressions of experiment 1.

Number of regressions	df	Type III Sum of Squares	Mean Square	*F*-Value	*p*-Value
Target Position	1	31.282	31.282	69.910	0.0000
Shift Condition	1	1.38	1.378	3.079	0.0802
Repetition	1	0.20	0.204	0.455	0.5003
Target Position X Shift Condition	1	0.51	0.514	1.148	0.2846
Target Position X Repetition	1	0.00	0.002	0.005	0.9412
Shift Condition X Repetition	1	0.00	0.005	0.010	0.9194
Target Position X Shift Condition X Repetition	1	0.10	0.099	0.222	0.6378
Residuals	352	157.51	0.447		

Table 16. R-Output for Target Position X Shift X Repetition ANOVAs for regression time of experiment 1.

Regression time	df	Type III Sum of Squares	Mean Square	*F*-Value	*p*-Value
Target Position	1	1112579	1112579	54.184	0.0000
Shift Condition	1	62517	62517	3.045	0.0819
Repetition	1	1	1	0.000	0.9957
Target Position X Shift Condition	1	29387	29387	1.431	0.2324
Target Position X Repetition	1	2016	2016	0.098	0.7542
Shift Condition X Repetition	1	5521	5521	0.269	0.6044
Target Position X Shift Condition X Repetition	1	13251	13251	0.645	0.4223
Residuals	352	7227770	20533		

ANOVA Output experiment 2

Table 17. R-Output for Target Position X Shift X Repetition ANOVAs for regression error of experiment 1.

Regression error	df	Type III Sum of Squares	Mean Square	*F*-Value	*p*-Value
Target Position	1	3887	3887	91.525	0.000
Shift Condition	1	74	74	1.736	0.189
Repetition	1	59	59	1.379	0.241
Target Position X Shift Condition	1	52	52	1.224	0.269
Target Position X Repetition	1	39	39	0.927	0.336
Shift Condition X Repetition	1	2	2	0.039	0.844
Target Position X Shift Condition X Repetition	1	2	2	0.037	0.849
Residuals	352	14949	42		

Table 18. R-Output for Target Position X Shift X Repetition ANOVAs for length of initial regression of experiment 2.

Length of initial regression	df	Type III Sum of Squares	Mean Square	*F*-Value	*p*-Value
Target Position	1	1.49	1.4889	19.038	0.0000
Shift Condition	1	0.01	0.0074	0.094	0.7593
Repetition	1	0.40	0.4018	5.138	0.0238
Target Position X Shift Condition	1	0.00	0.0031	0.040	0.8420
Target Position X Repetition	1	0.04	0.0408	0.521	0.4706
Shift Condition X Repetition	1	0.00	0.0000	0.000	0.9859
Target Position X Shift Condition X Repetition	1	0.03	0.0271	0.347	0.5562
Residuals	525	41.06	0.0782		

Table 19. R-Output for Target Position X Shift X Repetition ANOVAs for number of regressions of experiment 2.

Number of regressions	df	Type III Sum of Squares	Mean Square	*F*-Value	*p*-Value
Target Position	1	99.4	99.36	77.184	0.0000
Shift Condition	1	0.7	0.71	0.553	0.4576
Repetition	1	0.4	0.38	0.298	0.5854
Target Position X Shift Condition	1	0.3	0.27	0.208	0.6489
Target Position X Repetition	1	5.5	5.53	4.293	0.0388
Shift Condition X Repetition	1	0.0	0.00	0.001	0.9811
Target Position X Shift Condition X Repetition	1	0.6	0.59	0.456	0.4999
Residuals	525	675.9	1.29		

Table 20. R-Output for Target Position X Shift X Repetition ANOVAs for regression time of experiment 2.

Regression time	df	Type III Sum of Squares	Mean Square	*F*-Value	*p*-Value
Target Position	1	6874023	6874023	68.062	0.0000
Shift Condition	1	81500	81500	0.807	0.3694
Repetition	1	20589	20589	0.204	0.6518
Target Position X Shift Condition	1	62906	62906	0.623	0.4303
Target Position X Repetition	1	380123	380123	3.764	0.0529
Shift Condition X Repetition	1	2794	2794	0.028	0.8680
Target Position X Shift Condition X Repetition	1	169552	169552	1.679	0.1957
Residuals	525	53022858	100996		

Table 21. R-Output for Target Position X Shift X Repetition ANOVAs for regression error of experiment 2.

Regression error	df	Type III Sum of Squares	Mean Square	*F*-Value	*p*-Value
Target Position	1	8346	8346	188.953	0.000
Shift Condition	1	33	33	0.756	0.38501
Repetition	1	9	9	0.212	0.64501
Target Position X Shift Condition	1	6	6	0.138	0.71010
Target Position X Repetition	1	392	392	8.883	0.00301
Shift Condition X Repetition	1	11	11	0.238	0.62564
Target Position X Shift Condition X Repetition	1	0	0	0.011	0.91532
Residuals	525	23189	44		

Table 22. R-Output for Target Position X Age ANOVAs for length of initial regression.

Length of initial regression	df	Type III Sum of Squares	Mean Square	*F*-Value	*p*-Value
Target Position	1	0.697	0.6975	13.849	0.000249
Age	1	0.169	0.1689	3.354	0.068323
Target Position X Age	1	0.000	0.0001	0.003	0.959336
Residuals	232	11.684	0.0504		

Table 23. R-Output for Target Position X Age ANOVAs for number of regressions.

Number of regressions	df	Type III Sum of Squares	Mean Square	*F*-Value	*p*-Value
Target Position	1	28.61	28.61	45.003	0.00000
Age	1	36.10	36.10	56.801	0.00000
Target Position X Age	1	5.62	5.62	8.835	0.00327
Residuals	232	147.47	0.64		

Table 24. R-Output for Target Position X Age ANOVAs for regression time.

Regression time	df	Type III Sum of Squares	Mean Square	*F*-Value	*p*-Value
Target Position	1	1639852	1639852	34.21	0.000000
Age	1	8316891	8316891	173.49	0.000000
Target Position X Age	1	543980	543980	11.35	0.000885
Residuals	232	11121860	47939		

Table 25. R-Output for Target Position X Age ANOVAs for regression error.

Regression error	df	Type III Sum of Squares	Mean Square	*F*-Value	*p*-Value
Target Position	1	3279	3279	112.337	0.0000
Age	1	148	148	5.057	0.0255
Target Position X Age	1	136	136	4.647	0.0321
Residuals	232	6772	29		

Landing position distribution of progressive and regressive saccades of experiment 1

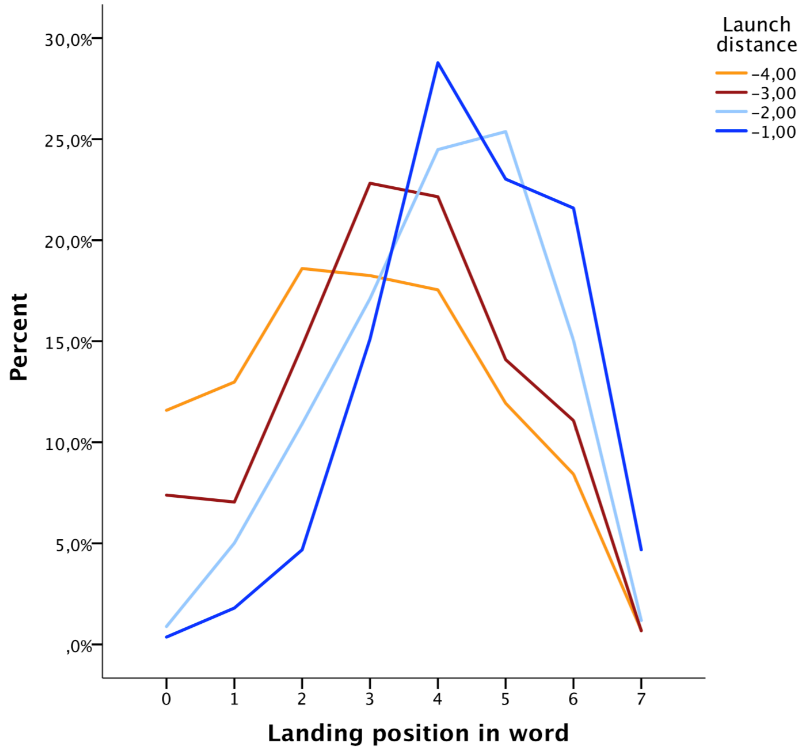


Figure 33. Landing position of the first fixation (first gaze) on the target word as a function of the launch distance in experiment 1 (adults).
Landing position 0 represents the blank space before the word.

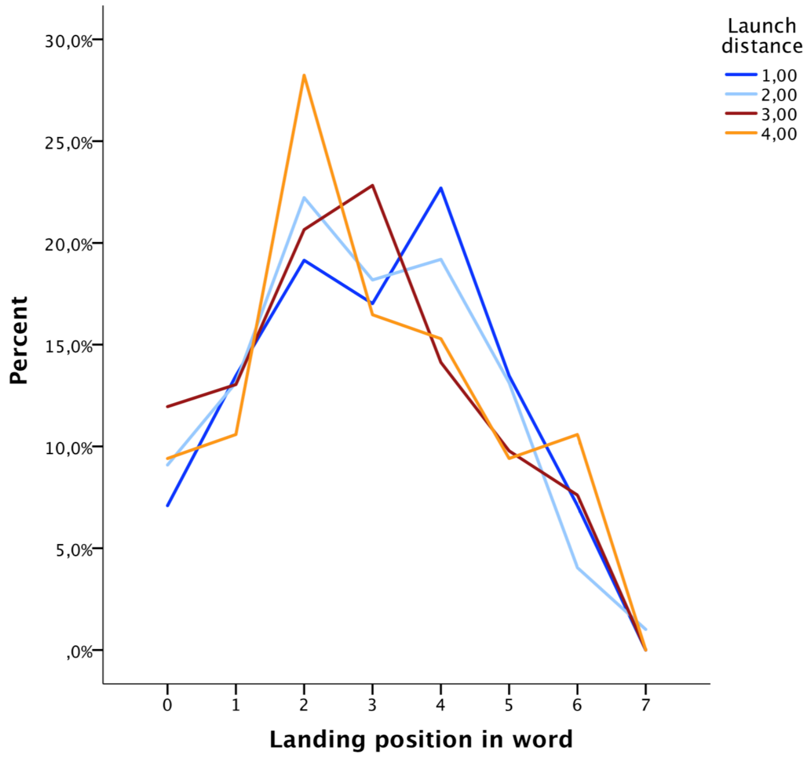


Figure 34. Landing position of regressions on the target word after crossing the boundary as a function of the launch distance in experiment 1 (adults).
Landing position 0 represents the blank space before the word.

Landing position distribution of progressive and regressive saccades of experiment 2

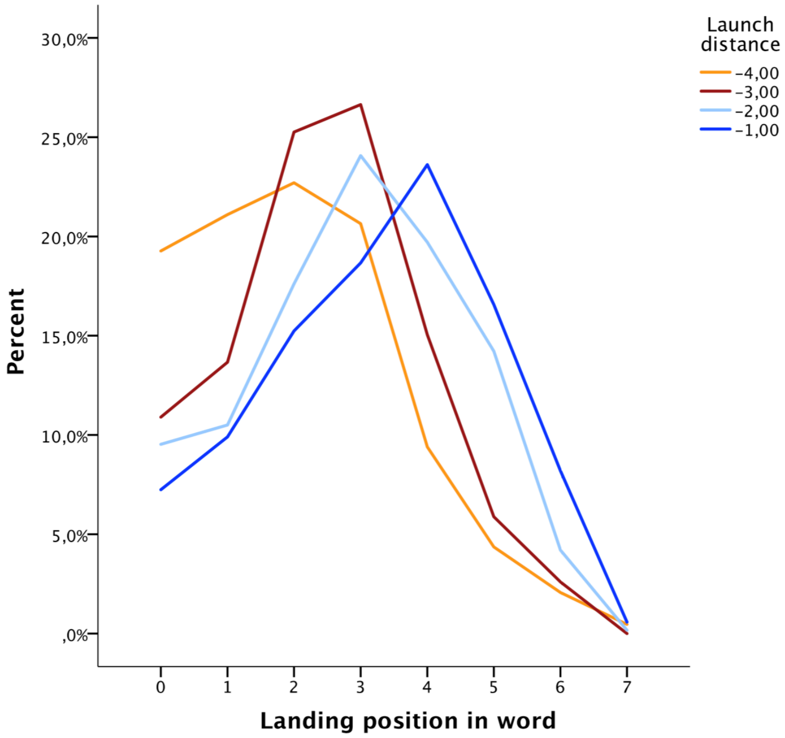


Figure 35. Landing position of the first fixation (first gaze) on the target word as a function of the launch distance in experiment 2 (children).
Landing position 0 represents the blank space before the word.

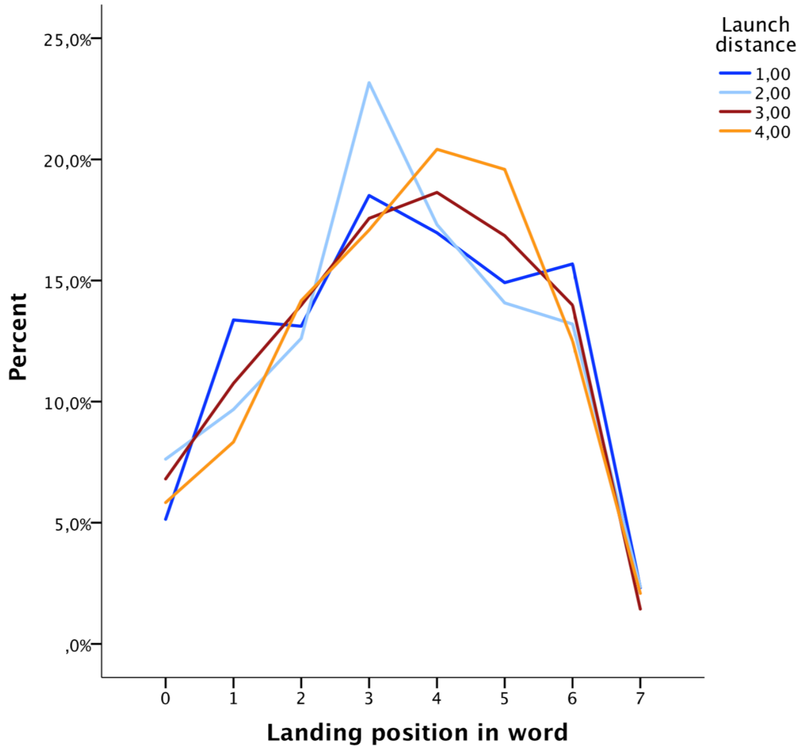


Figure 36. Landing position of regressions on the target word after crossing the boundary as a function of the launch distance in experiment 2 (children).
Landing position 0 represents the blank space before the word.

Output of LMMs for experiment 3

Length of initial regression

Minimal models

m1 <- lmer(log(lengthfirstfixback) ~ TargetPosition*Diff + (1|ID) + (1|ST), REML=F, data=R1)
print(summary(m1), cor=F)

m2 <- lmer(log(lengthfirstfixback) ~ Position + Diff + Position_Diff+ (1|ID) + (1|ST), REML=F, data=R1)
print(summary(m2), cor=F)

Maximal models

m3 <- lmer(log(lengthfirstfixback) ~ Position + Diff + Position_Diff +
(1+Position + Diff + Position_Diff|ID) +
(1+Position + Diff + Position_Diff|ST), REML=F, data=R1)

NOTE: Model failed to converge!

Compare the minimal and maximal models:

anova(m2,m3) *p* < .000***

The maximal model captures some important random variance that is not captured by the minimal model, however, the model did not converge.

m4 <- lmer(log(lengthfirstfixback) ~ Position + Diff + Position_Diff +
(1+Position + Diff + Position_Diff||ID) +
(1+Position + Diff + Position_Diff||ST), REML=FALSE, data=R1)

NOTE: Model failed to converge!

m5.1 <- lmer(log(lengthfirstfixback) ~ Position + Diff + Position_Diff+
(1+Position + Position_Diff|ID) +
(1+ Position + Position_Diff|ST), REML=FALSE, data=R1)

NOTE: this model did converge

I now re-run the model without random factor correlations

m5.2 <- lmer(log(lengthfirstfixback) ~ Position + Diff + Position_Diff+
(1+Position + Position_Diff||ID) +
(1+ Position + Position_Diff||ST), REML=FALSE, data=R1)

anova(m5.1, m5.2) *p* = .5731

m5.3 <- lmer(log(lengthfirstfixback) ~ Position + Diff + Position_Diff+
(1+Position|ID) +
(1+Position|ST), REML=FALSE, data=R1)

anova(m5.2, m5.3) *p* < .000***
anova(m2, m5.2) *p* < .000***

The position random effect is needed, as model m5.2 is better than the minimal model m2. m5.2 is the final and used model.

Table 26. R-Output of different models following LMM Analyses for length of initial regression (experiment 3).

Model	Factor	Estimate	Standard Error (SE)	/t/-Value
m2	Target Position	0.29091	0.01283	22.669
	Difficulty	-0.02833	0.01311	-2.161
	Target Position X Difficulty	0.34738	0.05047	6.883
m3	Target Position	0.28900	0.0290	9.96
	Difficulty	-0.02500	0.02392	-1.05
	Target Position X Difficulty	0.33940	0.04874	6.96
m4	Target Position	0.28951	0.02887	10.028
	Difficulty	-0.02643	0.02377	-1.112
	Target Position X Difficulty	0.34052	0.04828	7.053
m5.1	Target Position	0.28915	0.02884	10.026
	Difficulty	-0.02517	0.02237	-1.125
	Target Position X Difficulty	0.33908	0.04862	6.974
m5.2	Target Position	0.28969	0.02878	10.066
	Difficulty	-0.02618	0.02234	-1.172
	Target Position X Difficulty	0.34018	0.04815	7.065
m5.3	Target Position	0.28971	0.02881	10.056
	Difficulty	-0.02487	0.02238	-1.112
	Target Position X Difficulty	0.33996	0.04827	7.043

Number of regressions

Minimal models

m1 <- lmer(log(FixNumTargDetect) ~ TargetPosition*Diff + (1|ID) + (1|ST), REML=F, data=R1)

m2 <- lmer(log(FixNumTargDetect) ~ Position + Diff + Position_Diff+ (1|ID) + (1|ST), REML=F, data=R1)

Maximal models

m3 <- lmer(log(FixNumTargDetect) ~ Position + Diff + Position_Diff +
(1+Position + Diff + Position_Diff|ID) +
(1+Position + Diff + Position_Diff|ST), REML=F, data=R1)

NOTE: Model failed to converge!

anova(m2,m3) *p* < .000***

The maximal model captures some important random variance that is not captured by the minimal model, however, the model did not converge.

m4 <- lmer(log(FixNumTargDetect) ~ Position + Diff + Position_Diff +
(1+Position + Diff + Position_Diff||ID) +
(1+Position + Diff + Position_Diff||ST), REML=FALSE, data=R1)

NOTE: Model failed to converge!

m5.1 <- lmer(log(FixNumTargDetect) ~ Position + Diff + Position_Diff+
(1+Position + Position_Diff|ID) +
(1+ Position + Position_Diff|ST), REML=FALSE, data=R1)

NOTE: Model failed to converge!

m5.2 <- lmer(log(FixNumTargDetect) ~ Position + Diff + Position_Diff+
(1+Position + Position_Diff||ID) +

NOTE: Model failed to converge!

m5.3 <- lmer(log(FixNumTargDetect) ~ Position + Diff + Position_Diff +
(1+Position|ID) +
(1+Position|ST), REML=FALSE, data=R1)

NOTE: this model did converge

anova(m2, m5.3) *p* < .000***

The position random effect is needed, as model m5.3 is better than the minimal model m2. m5.3 is the final and used model.

Table 27. R-Output of different models following LMM Analyses for number of regressions (experiment 3).

Model	Factor	Estimate	Standard Error (SE)	/t/-Value
m2	Target Position	0.50251	0.01711	29.363
	Difficulty	-0.19612	0.01744	-11.248
	Target Position X Difficulty	0.19632	0.05300	3.704
m3	Target Position	0.49873	0.04706	10.597
	Difficulty	-0.20012	0.02739	-7.306
	Target Position X Difficulty	0.19451	0.05871	3.313
m4	Target Position	0.50088	0.04695	10.668
	Difficulty	-0.19983	0.02582	-7.740
	Target Position X Difficulty	0.19588	0.05819	3.366
m5.1	Target Position	0.49932	0.04652	10.73
	Difficulty	-0.19978	0.02562	-7.80
	Target Position X Difficulty	0.19450	0.05754	3.38
m5.2	Target Position	0.50122	0.04676	10.720
	Difficulty	-0.19919	0.02515	-7.922
	Target Position X Difficulty	0.19672	0.05742	3.426
m5.3	Target Position	0.50061	0.04659	10.744
	Difficulty	-0.19905	0.02513	-7.921
	Target Position X Difficulty	0.19696	0.05424	3.631

Regression time

Minimal models

m1 <- lmer(log(RegTime) ~ TargetPosition*Diff + (1|ID) + (1|ST), REML=F, data=R1)

m2 <- lmer(log(RegTime) ~ Position + Diff + Position_Diff+ (1|ID) + (1|ST), REML=F, data=R1)

Maximal models

m3 <- lmer(log(RegTime) ~ Position + Diff + Position_Diff +
(1+Position + Diff + Position_Diff|ID) +
(1+Position + Diff + Position_Diff|ST), REML=F, data=R1)

NOTE: Model failed to converge!

anova(m2,m3) $p < .000$***

The maximal model captures some important random variance that is not captured by the minimal model, however, the model did not converge.

m4 <- lmer(log(RegTime) ~ Position + Diff + Position_Diff +
(1+Position + Diff + Position_Diff||ID) +
(1+Position + Diff + Position_Diff||ST), REML=FALSE, data=R1)

NOTE: Model failed to converge!

m5.1 <- lmer(log(RegTime) ~ Position + Diff + Position_Diff+
(1+Position + Position_Diff|ID) +
(1+ Position + Position_Diff|ST), REML=FALSE, data=R1)

NOTE: Model failed to converge!

m5.2 <- lmer(log(RegTime) ~ Position + Diff + Position_Diff+
(1+Position + Position_Diff||ID) +
(1+ Position + Position_Diff||ST), REML=FALSE, data=R1)

NOTE: Model failed to converge!

m5.3 <- lmer(log(RegTime) ~ Position + Diff + Position_Diff+
(1+Position|ID) +
(1+Position|ST), REML=FALSE, data=R1)

NOTE: this model did converge

anova(m2, m5.3) $p < .000$***

The position random effect is needed, as model m5.3 is better than the minimal model m2. m5.3 is the final and used model.

Table 28. R-Output of different models following LMM Analyses for regression time (experiment 3).

Model	Factor	Estimate	Standard Error (SE)	/t/-Value
m2	Target Position	0.39776	0.01548	25.700
	Difficulty	-0.17447	0.01576	-11.070
	Target Position X Difficulty	0.18292	0.04498	4.067
m3	Target Position	0.39409	0.03840	10.3
	Difficulty	-0.17700	0.02382	-7.4
	Target Position X Difficulty	0.18402	0.04908	3.7
m4	Target Position	0.39566	0.03831	10.3
	Difficulty	-0.17657	0.02261	-7.8
	Target Position X Difficulty	0.18271	0.04830	3.8
m5.1	Target Position	0.39487	0.03808	10.4
	Difficulty	-0.17703	0.02244	-7.9
	Target Position X Difficulty	0.18487	0.04803	3.8
m5.2	Target Position	0.39558	0.03817	10.365
	Difficulty	-0.17530	0.02210	-7.931
	Target Position X Difficulty	0.18153	0.04789	3.791
m5.3	Target Position	0.39555	0.03808	10.387
	Difficulty	-0.17516	0.02210	-7.926
	Target Position X Difficulty	0.18281	0.04611	3.965

Output of LMMs of experiment 4

Length of initial regression

Minimal model

m1 <- lmer(log(lengthfirstfixback) ~ TargetPosition*Punctuation + (1|ID) + (1|ST), REML=F, data=R1)

m2 <- lmer(log(lengthfirstfixback) ~ Position + Punctuation + Position_Punctuation + (1|ID) + (1|ST), REML=F, data=R1)

Maximal models

m3 <- lmer(log(lengthfirstfixback) ~ Position + Punctuation + Position_Punctuation +
(1+Position + Punctuation + Position_Punctuation|ID) +
(1+ Position + Punctuation + Position_Punctuation|ST), REML=F, data=R1)

anova(m2,m3) *p* < .000***

The maximal model captures some important random variance that is not captured by the minimal model.

m4 <- lmer(log(lengthfirstfixback) ~ Position + Punctuation + Position_Punctuation +
(1+Position + Punctuation + Position_Punctuation||ID) +
(1+ Position + Punctuation + Position_Punctuation||ST), REML=F, data=R1)

NOTE: Model failed to converge!

m5.1 <- lmer(log(lengthfirstfixback) ~ Position + Punctuation + Position_Punctuation +
(1+ Position |ID) +
(1+ Position |ST), REML=FALSE, data=R1)

m5.2 <- lmer(log(lengthfirstfixback) ~ Position + Punctuation + + Position_Punctuation +
(1+Position ||ID) +
(1+ Position ||ST), REML=FALSE, data=R1)

anova(m5.1, m5.2) *p* = .5052

The model without random factor correlations is not significantly different from the model with random factor correlation.

anova(m2, m5.2) *p* < .000***

The position random effect is needed, as model m5.2 is better than the minimal model m2. m5.2 is the final and used model.

As can be seen, the effect of punctuation are negligible. I drop this factor and its interactions from model m5.2

```
m6 <- lmer(log(lengthfirstfixback) ~ Position +
            (1+Position |ID) +
            (1+ Position |ST), REML=FALSE, data=R1)
```

anova(m5.1, m6) *p* = .5318

Inclusion of punctuation and its interactions does not improve the model. This means that the model can be simplified to one main effect.

Table 29. R-Output of different models following LMM Analyses for length of initial regressions (experiment 4).

Model	Factor	Estimate	Standard Error (SE)	/t/-Value
m2	Target Position	3.234e-01	1.430e-02	22.612
	Punctuation	3.343e-04	1.431e-02	0.023
	Target Position X Punctuation	-2.423e-02	2.908e-02	-0.833
m3	Target Position	0.325202	0.034951	9.304
	Punctuation	0.003059	0.016597	0.184
	Target Position X Punctuation	-0.032596	0.028267	-1.153
m4	Target Position	3.247e-01	3.496e-02	9.287
	Punctuation	2.332e-03	1.651e-02	0.141
	Target Position X Punctuation	-3.024e-02	2.751e-02	-1.099
m5.1	Target Position	3.250e-01	3.512e-02	9.255
	Punctuation	2.601e-03	1.373e-02	0.189
	Target Position X Punctuation	-3.062e-02	2.761e-02	-1.109
m5.2	Target Position	3.250e-01	3.512e-02	9.254
	Punctuation	2.475e-03	1.373e-02	0.180
	Target Position X Punctuation	-3.058e-02	2.762e-02	-1.107
m6	Target Position	0.32503	0.03509	9.263

Number of regressions

Minimal models

m1 <- lmer(log(FixNumTargDetect) ~ TargetPosition*Punctuation + (1|ID) + (1|ST), REML=F, data=R1)

m2 <- lmer(log(FixNumTargDetect) ~ Position + Punctuation + Position_Punctuation + (1|ID) + (1|ST), REML=F, data=R1)

Maximal models

m3 <- lmer(log(FixNumTargDetect) ~ Position + Punctuation + Position_Punctuation +
(1+Position + Punctuation + Position_Punctuation |ID) +
(1+ Position + Punctuation + Position_Punctuation |ST), REML=F, data=R1)

NOTE: Model failed to converge!

anova(m2,m3) *p* < .000***

(the maximal model captures some important random variance that is not captured by the minimal model, however, the model did not converge)

m4 <- lmer(log(FixNumTargDetect) ~ Position + Punctuation + Position_Punctuation +
(1+Position + Punctuation + Position_Punctuation||ID) +
(1+ Position + Punctuation + Position_Punctuation||ST), REML=F, data=R1)

m5.1 <- lmer(log(FixNumTargDetect) ~ Position + Punctuation + Position_Punctuation +
(1+ Position |ID) +
(1+ Position |ST), REML=FALSE, data=R1)

m5.2 <- lmer(log(FixNumTargDetect) ~ Position + Punctuation + Position_Punctuation +
(1+Position ||ID) +
(1+ Position ||ST), REML=FALSE, data=R1)

anova(m5.1, m5.2) *p* < .05*

The model without random factor correlations is significantly different from the model with random factor correlation. Hence, I continue with m5.1

anova(m2, m5.1) *p* < .001***

The Position random effect is not needed as the model m5.1 is better than the minimal model m2.

As can be seen, the effects of punctuation are negligible. I drop this factor and its interactions from model m5.1

m6 <- lmer(log(FixNumTargDetect) ~ Position +
(1+Position |ID) +
(1+Position |ST), REML=FALSE, data=R1)

anova(m5.1, m6) *p* = .1497

Inclusion of punctuation and its interactions does not improve the model. This means that the model can be simplified to one main effect.

Table 30. R-Output of different models following LMM Analyses for number of regressions (experiment 4).

Model	Factor	Estimate	Standard Error (SE)	/t/-Value
m2	Target Position	0.53721	0.01741	30.861
	Punctuation	0.03311	0.01742	1.900
	Target Position X Punctuation	-0.01328	0.03536	-0.376
m3	Target Position	0.53602	0.04085	13.121
	Punctuation	0.03213	0.01890	1.700
	Target Position X Punctuation	-0.01318	0.04030	-0.327
m4	Target Position	0.53561	0.04083	13.118
	Punctuation	0.03186	0.01709	1.864
	Target Position X Punctuation	-0.01180	0.03629	-0.325
m5.1	Target Position	0.53585	0.04081	13.130
	Punctuation	0.03242	0.01698	1.909
	Target Position X Punctuation	-0.01341	0.03398	-0.394
m5.2	Target Position	0.53559	0.04082	13.122
	Punctuation	0.03231	0.01699	1.902
	Target Position X Punctuation	-0.01208	0.03399	-0.355
m6	Target Position	0.53548	0.04084	13.11

Regression time

Minimal models

m1 <- lmer(log(RegTime) ~ TargetPosition*Punctuation + (1|ID) + (1|ST), REML=F, data=R1) m2 <- lmer(log(RegTime) ~ Position + Punctuation + Position_Punctuation + (1|ID) + (1|ST), REML=F, data=R1)

Maximal models

m3 <- lmer(log(RegTime) ~ Position + Punctuation + Position_Punctuation +
(1+Position + Punctuation + Position_Punctuation|ID) +
(1+ Position + Punctuation + Position_Punctuation|ST), REML=F, data=R1)

anova(m2,m3) *p* < .000***

(The maximal model captures some important random variance that is not captured by the minimal model.)

m4 <- lmer(log(RegTime) ~ Position + Punctuation + Position_Punctuation +
(1+Position + Punctuation + Position_Punctuation||ID) +
(1+ Position + Punctuation + Position_Punctuation||ST), REML=F, data=R1)

m5.1 <- lmer(log(RegTime) ~ Position + Punctuation + Position_Punctuation +
(1+ Position |ID) +
(1+ Position |ST), REML=FALSE, data=R1)

m5.2 <- lmer(log(RegTime) ~ Position + Punctuation + Position_Punctuation +
(1+Position ||ID) +
(1+ Position ||ST), REML=FALSE, data=R1)

anova(m5.1, m5.2) *p* < .000***

The model without random factor correlations is significantly different from the model with random factor correlation. Hence, I continue with m5.1

anova(m2, m5.1) *p* < .001***

The Position random effect is not needed as the model m5.1 is better than the minimal model m2.

As can be seen, the effects of punctuation are negligible. I drop this factor and its interactions from model m5.1

m6 <- lmer(log(RegTime) ~ Position +
(1+Position |ID) +
(1+ Position |ST), REML=FALSE, data=R1)

anova(m5.1, m6) *p* = .3259

Inclusion of punctuation and its interactions does not improve the model. This means that the model can be simplified to one main effect.

Table 31. R-Output of different models following LMM Analyses for regression time (experiment 4).

Model	Factor	Estimate	Standard Error (SE)	/t/-Value
m2	Target Position	0.46630	0.02388	19.524
	Punctuation	0.03570	0.02383	1.498
	Target Position X Punctuation	-0.02540	0.03399	-0.747
m3	Target Position	0.46625	0.04355	10.707
	Punctuation	0.03360	0.02843	1.182
	Target Position X Punctuation	-0.02602	0.03891	-0.669
m4	Target Position	0.46565	0.03429	13.580
	Punctuation	0.03381	0.02338	1.446
	Target Position X Punctuation	-0.02448	0.03319	-0.737
m5.1	Target Position	0.46624	0.03887	11.996
	Punctuation	0.03389	0.02327	1.457
	Target Position X Punctuation	-0.02580	0.03296	-0.783
m5.2	Target Position	0.46588	0.03430	13.583
	Punctuation	0.03389	0.02335	1.451
	Target Position X Punctuation	-0.02449	0.03319	-0.738
m6	Target Position	0.45322	0.03520	12.88

Erklärung

Hiermit erkläre ich, dass ich gemäß § 10 der Promotionsordnung vom 21.05.2015, die eingereichte Arbeit selbstständig verfasst, nur die in der Dissertation angegeben Hilfsmittel benutzt und alle wörtlich oder inhaltlich übernommenen Stellen als solche gekennzeichnet habe. Die Dissertation wurde in der gegenwärtigen oder einer anderen Fassung keinem anderen Fachbereich einer wissenschaftlichen Hochschule vorgelegt.

Wuppertal, im November 2019

____

Anne Friede