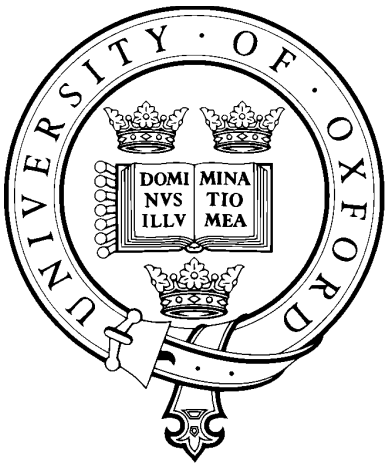
Predict or Bridge? Investigating inference making strategies in skilled readers

Maria Friedmannova

St. John's College



Thesis submitted for the degree of Doctor of Philosophy

Department of Experimental Psychology

University of Oxford

December, 2014

Table of contents

Acknowledgements …… iii

Short Abstract ...…..…… .iv

Long Abstract …… . v

Chapter 1
General introduction and overview of the thesis ……. 1

Chapter 2
The effect of supporting context on the construction of bridging inferences during word-to-text integration (Experiment 1)…….…… 27

Chapter 3
Rapid word-to-text integration: a result of lexical repetition or supporting context? (Experiment 2) ……..……. 57

Chapter 4
Addressing methodological issues of previous work and exploring the effect of supporting context on instantiation of inferred information (Experiment 3)……..… 94

Chapter 5
Predict or Bridge? Investigating skilled readers' use of predictive inferences as a default strategy for word integration (Experiment 4) ……...…… 137

Chapter 6
General Discussion ……..……. 183

References …… 209

Appendices…… 221

Acknowledgements

First of all, I wish to thank my supervisor, Professor Kate Nation. Kate, my greatest thanks to you for your guidance and support, and for the endless hours you have invested in helping me produce this thesis! I know it has been challenging, and so I want to thank you for sticking with me! I equally wish to thank my co-supervisor, Dr Holly Joseph. Holly, thank you for introducing me to the world of eye-tracking and inference-making! I really appreciate your help and eye-tracking expertise, which many times saved both my work and our eye-tracker! I deeply enjoyed our morning coffees and conversations about inference-making in the Jericho Café, and hope to manage to get in a few more!

To the whole LCD group, past and present – Robin, Gurpreet, Niina, Nick, Fiona, Bram, Paul, Hei-Won and Titus – thank you for your camaraderie, support and help in matters far beyond statistics and research design. I also wish to thank St. John's College MCR and Oxford University's Czech and Slovak Society for providing a wonderful student community and lasting friendships that I will treasure. For financial support, I want to thank the Economic and Social Research Council.

Thanks are also especially due to Janice, Natasha, Barbara, Kristina, William, Matt, Kubo, Matej, Joel, Stefan and Jakub for their unyielding support, generosity, wisdom, love and friendship! Thank you for keeping me going and for being such wonderful people! I also wish to thank Jackie, Max and Miles for inspiring conversations and help on bad research days. Last, but certainly not least, I want to thank my parents and siblings for their endless love and support, I look up to every one of you!

Short Abstract

Predict or Bridge? Investigating inference making strategies in skilled readers

Maria Friedmannova

St. John's College, University of Oxford

Thesis submitted for the degree of Doctor of Philosophy

Inference-making is one of the crucial skills necessary to achieve good reading comprehension. This thesis explored which inference-making strategies are employed by skilled readers when they integrate new information in their mental text representation. Specifically, I examined how context affected readers' choice of inference-making strategy and its time-course in the word-to-text integration process. Four eye-tracking experiments were carried out to observe inference-making as reading happens. By manipulating the amount of supporting context, the first three experiments studied how different properties of target word referents presented in the previous context influenced generation of bridging inferences. In particular, the effect of having lexically explicit, direct referent as opposed to a contextually-related but indirect referent was examined. The results of the first three experiments conflicted suggesting that readers either used predictive inferences instead of bridging to integrate the target words, or that context offered no advantage for the drawing of bridging inferences. To address this question, Experiment 4 examined whether participants integrated the target words by making predictive inferences even in weakly constraining contexts. Participants were presented with texts where context provided either strong or only weak indication of future text development, which was then confirmed or disconfirmed by the target word. The results suggested that even if context provided only limited indication of future development, skilled readers would generate predictive inferences to integrate new concepts. Only when predictive inferences fail to integrate new words, bridging inferences would be employed.

Long Abstract

Predict or Bridge? Investigating inference making strategies in skilled readers

Maria Friedmannova

St. John's College, University of Oxford

Thesis submitted for the degree of Doctor of Philosophy

Previous research has shown that ability to make inferences is crucial for achieving good reading comprehension (Cain & Oakhill, 1999; Cain, Oakhill, Barnes, & Bryant, 2001; Yuill & Oakhill, 1988). Through inference-making, people are able to create a rich and coherent representation of the text they are reading, a process which is key to good understanding. Inferences require people to actively process newly read information. As such, inferences are the main vehicle for the integration of new words into a mental text representation. When reading a new word, readers need to establish how the new concept relates to what they already know in order to integrate it with the rest of text. To do so, people draw two specific types of inferences - bridging (McKoon & Ratcliff, 1992; Yang, Perfetti & Schmalhoffer, 2007) and predictive (Keefe & McDaniel, 1993; Singer, 1993). Bridging inferences help readers establish coherence between the just read word and previous content by establishing a bridge between the two. As such, bridging inferences are drawn after the word has been read. Predictive inferences on the other hand are generated prior to reading of the new word on the basis of previous context. If previous context is constraining enough, readers will generate predictions about future text development allowing rapid word integration of new concepts which fit with the predictions.

Previous work has established that these inferences are drawn, but since the literature predominantly used measures applied after the text had been read (offline measures) it was not possible to say when exactly these inferences are generated during reading. This thesis thus aims to answer the question: which inference making strategy

– bridging or predictive – readers use to integrate new words in their text representation as reading happens. Specifically, it examines how context affects the choice and the subsequent time-course of the chosen inference-making strategy.

To do so, four eye-tracking experiments have been conducted. Experiment 1 reported in Chapter 2 explored how readers utilized available contextual information when drawing inferences to integrate the target word in their text representation. Participants were presented with two-sentence paragraphs containing the same target word in three different conditions, varying the availability of a target word referent. As the target word referent availability decreased, context became less supportive. Consequently, it became more difficult to construct the inference necessary for the target word integration. I hypothesized that as inferential demands rose and supporting context became less available, skilled readers would spend more time constructing bridging inferences.

The results of Experiment 1 confirmed that readers used bridging inferences as their main integration strategy, taking more time to complete the inferences in the less contextually supporting condition. When the context explicitly invoked the inference, rapid target word integration was observed during early text processing. However, prior context in the explicit condition (the baseline condition) presented an identical word as both target word referent and the target word itself. This may have resulted in unintended lexical priming. Since lexical priming rather than supporting context might have produced the early integration effect seen in the first experiment, the second experiment attempted to distinguish between the two causes. To compensate for the inadequate baseline, a new condition which substituted the explicit, lexically identical target word referent with a synonym was included in the design of Experiment 2, reported in Chapter 3. However, Experiment 2 failed to replicate the effect of the main

manipulation altogether and suggested that supporting context provided no benefit for the drawing of bridging inferences. Chapter 3 concluded by outlining methodological problems of Experiment 1 and 2 which might have affected the results of previous work.

Chapter 4 addressed these methodological problems in Experiment 3. Experiment 3 sought to clarify the effect of contextually supporting information on the target word integration. This study also explored whether readers were equally likely to instantiate information from the conditions which gave less contextual support as from those which gave strong contextual support. Since the third experiment confirmed no effect of highly supporting context on the time course of target word integration, a fourth experiment set out to distinguish whether this is due to readers using predictive rather than bridging inferences, even in conditions which presented the least amount of contextually supporting information.

In Experiment 4 reported in Chapter 5, participants were presented with paragraphs where prior context suggested a certain direction of what might happen next in the text. The prior context either provided a strong indication of future text development or only a weak one. The target word then either confirmed the text direction set out by the previous context or suggested different development. I hypothesized that if participants used predictive inferences, conditions where the target words confirmed text direction outlined by the prior context would be processed faster, relative to those where the target words did not confirmed it. Furthermore, if predictive inferences were used to integrate the target word across all conditions in previous experiments, conditions providing strong and weak indications of future development should be processed equally fast as long as the target word confirmed this development. Results of the last experiment confirmed my predictions.

Thus, I argued on the basis of evidence presented in the experimental chapters that skilled readers used predictive inferences rather than bridging as their primary strategy for word integration in my materials, even in less constraining contexts, provided the texts they read were easy. When texts became more demanding, for example by developing in an unexpected direction, predictive inferences became insufficient to integrate the target word and skilled readers generated bridging inferences as well to complete the word integration. My findings were addressed in the General discussion, Chapter 6, with regard to text processing strategies and an altered view of two-stage text processing account was proposed. I argued my findings indicated that when reading easy text, skilled readers adopt a fast, heuristics-driven yet error-prone approach. Once the texts became more demanding, they would switch to slower strategic and goal-oriented text processing.

Chapter 1: General introduction and overview of the thesis

This chapter provides a broad overview of the existing literature relating to research presented in this thesis. It begins by reviewing the importance of our ability to draw inferences in order to achieve good reading comprehension. Specific focus is directed at two inference-making strategies crucial for successful integration of new words in people's mental text representation; bridging and predictive inferences. Evidence for the use of each strategy is reviewed and the effect of context on generation of inferences is discussed. The need to study inference-making as reading happens is emphasized and the chapter concludes by presenting research questions addressed by experiments in this thesis and its methodological approach.

We love good stories – those that suck you in to a point you forget the world around you. We praise the authors for their gripping talent, but without readers, no good stories would ever be appreciated. Perhaps then, readers are almost as important for the existence of good stories as their authors. A reader will only be able to immerse herself in the magical world of author's imagination if her reading comprehension ability is good. Without being able to understand what is happening on the page, even the best story will fail to interest. Apart from losing out on the wonderful world of literature, reading comprehension problems impact one's life in more profound ways by hindering academic progress and consequently, leading to worse employment opportunities. Understanding how readers turn words on the page into meaning is an important key to helping struggling readers overcome their reading difficulties.

This thesis focuses on the process by which people integrate new concepts into their mental text representation through inference making. The main question I am trying to answer is how and when inferences are generated and what are the strategies that skilled readers employ when they use inferences to integrate new concepts in their mental representation of a text. Additionally, I placed emphasis on the role of context in inference generation during reading; I am specifically trying to elucidate how different contexts influence the type of inferences that readers use to integrate new concepts, and when during the reading process they do so. In this introduction, I will first explain why inference making is a crucial skill for successful reading comprehension. Then, I will describe the types of inferences that seem the most important for good comprehension; predictive and bridging inferences. I will characterize each type and review the evidence for when and how readers generate them. After that, I will briefly discuss how context influences inference making and conclude by introducing eye tracking, the prime methodology used in this thesis.

Inference making as a crucial skill in reading comprehension

To learn to read words, people need a plethora of different skills; phonological awareness (Goswami & Bryant, 1990; Hulme & Snowling, 1992), vocabulary (Perfetti, 1985, Perfetti & Hart, 2001) and the ability to decode print (Hoover and Gough, 1990; Perfetti & Hogaboam, 1975; Stanovich, 1982). But mastering the mechanics of learning to read is not enough to make someone a good reader. Although word decoding problems are associated with poor comprehension (Perfetti, 1985), good decoding does not guarantee good comprehension. In fact, between 10 - 15% of children have impaired reading comprehension despite having age-appropriate decoding skills (Cain & Oakhill, 2006; Cain, Oakhill, & Bryant, 2000; Nation & Snowling, 1997). Such children have been termed poor comprehenders, and they show difficulties with text-level processing. Mainly, these skills include ability to engage in constructive processing (Oakhill, 1982), using context efficiently (Nation & Snowling, 1998) and importantly, the ability to make inferences (Cain & Oakhill, 1999; Cain, Oakhill, Barnes, & Bryant, 2001; Oakhill, 1984; Yuill & Oakhill, 1988). This is the main focus of this thesis.

The reason why these skills are important for comprehension is that they help readers to create a rich and coherent mental model of the text they are reading (Zwaan & Radvansky, 1998; Kintsch & van Dijk, 1978). To build such model, readers need to compute the semantic and syntactic relationships among the successive words, phrases and sentences and join them in a meaningful representation. Such a model is thus much more than just a recollection of what readers had read or an understanding of individual words. It contains information about the different actors and objects in the text, how they relate to and act upon each other, how the content fits with what the reader already knows from previous context, and how it fits with their prior topic knowledge and their general world knowledge.

To build such model and create a database of relations connecting the textual elements, readers need to establish links with information and knowledge not directly mentioned in the text itself. For example, to understand the sentences in (1), the following links need to be made: water extinguished the fire, putting out the campfire is a sign that that the party is over, when the party finished people went to sleep.

(1)Laura poured a bucket of water on the campfire. After she did so, everybody went to sleep.

The process by which readers generate the links and integrate phrases that seem unrelated to the previous content, such as 'went to sleep' in (1) is called inference making.

The role of inference-making in comprehension

Inference making is one of the most crucial processes people need to carry out in order to understand what they read. A large body of research suggests that ability to draw inferences is central to reading comprehension (Garnham, Oakhill, & Johnson-Laird, 1982; Magliano & Mills, 2003; Oakhill, 1984; Oakhill, Yuill & Parkin, 1986). There are three main reasons for this. Firstly, inferences form a way of combining information from the text with knowledge already known. Without combining the two, it is difficult to generate a text representation that would be rich, coherent and easy to remember. Linking read text with information that a reader already has in their memory helps them to remember the text better and subsequently form a better text representation.

Secondly, inferences engage readers with the text. They force readers to be active when treating the information in a text, rather than passive. Active reading can

lead to better comprehension. For example, Einstein, McDaniel, Owen and Cote (1990) presented participants with texts where sentence order was scrambled and asked them to put the sentences in the correct order. Participants who read sentences in a scrambled order and had to re-order them showed better and more detailed free recall of the texts than those in the control group who read the same texts in the correct form. Similarly, when good readers are presented with texts that are too cohesive, making otherwise inferred information explicit (via a high degree of argument overlap, for example), they show more shallow comprehension than when texts are less cohesive (McNamara & Kintsch, 1996; McNamara, Kintsch, Songer & Kintsch, 1996). This is because less cohesive texts 'force' readers to compute more inferences in order to keep their text representation accurate and understand the text.

Thirdly, inferences help readers to monitor their comprehension. When a reader is asked to reconcile information from the text with his general knowledge, or join bits of information from different parts of the text, they do so by making inferences. If during this process readers find they cannot integrate the information together because it does not make sense, they are provided with an important signal that comprehension has broken down (Paris & Myers, 1981). The ability to monitor comprehension is crucial for good reading: if people do not notice they have lost understanding, they have no opportunity to check whether their mental model fits with the text they are reading. Thus, they lose a chance to correct arising misanalyses and their comprehension suffers.

Not all types of inference matter equally

It is not just drawing of any inferences that matters. The critical criterion is how much the reader needs to engage the text with his previous knowledge, be it the knowledge of the prior context or world knowledge. The more readers engage with the

text and draw inferences that go beyond what is explicitly mentioned in the text, the better their understanding. For example, after reading (2), only a text-based inference is needed to answer the question 'What letter did Annie receive?'

(2) Annie received her university application results and she let out a deep sigh.

However, to answer the question 'Was Annie accepted to the university?, a reader must go beyond the explicit content in the text and engage with their world knowledge. Is letting out a deep sigh a mark of disappointment rather than happiness? If so, her disappointment seems to be connected with the letter she received, and this must mean she was probably rejected by the university. Indeed, a large body of research had indicated that less skilled readers make fewer inferences that go beyond the explicit meaning of the text and they tend to ignore the conceptual gaps, or fail to repair them. Skilled readers, on the other hand, are more likely to generate inferences which are necessary to fill the conceptual gaps between sentences or paragraphs (Cain & Oakhill, 1999; Cain, Oakhill, Barnes & Bryant, 2001; Long, Oppy & Seely, 1994).

Two particular types of inference making processes seem central for deeper engagement with the text: bridging inferences and predictive inferences. These are of particular interest because they act as vehicles of word integration during reading. Whenever a new word or a phrase is read, it needs to be readily integrated with the previous content in order to keep the mental model accurate and up-to-date. Without this process, the newly acquired concept does not fit with reader's representation of the text they are reading and comprehension breaks down. Bridging and predictive inferences are not limited to word integration though; they are also important for reconciling information from different parts of the text and from the text and reader's knowledge base (knowledge base contains everything reader knows about the topic;

from previous context as well as from general knowledge). Let us have a closer look at them and start with bridging inferences, describing what they are, how they work and importantly, when in the comprehension process readers generate them.

Bridging inferences

Bridging inferences act as a bridge between a word and the previous context in order to keep the text coherent. They are generated frequently during reading (Bloom, Fletcher, van der Broek, Reitz & Schapiro, 1990; Klin & Myers, 1993; McKoon & Ratcliff, 1992; Noordman, Vonk & Kempff, 1992). They are typically drawn to establish the connection between a just-read clause and a previously read clause; they also serve maintain global coherence, especially if sentence-level coherence is lost (McKoon & Ratcliffe, 1992). Generally, bridging inferences are needed when readers cannot retrieve a referent from their memory for the information they are currently reading about. Thus, they need to make a 'bridge' to the previous sentence by making the connection themselves through constructing an inference (Singer, Andrusiak, Reisdorf & Black, 1992).

Singer et al. (1992) presented readers with sentences such as (3) and argued that since the first sentence offered no referent for the word 'smell', readers had to construct a 'bridge' to the previous content.

(3) Carol found that the milk was three weeks old. The smell turned her stomach.

To do so, they had to activate world knowledge that three week old milk turns bad and bad milk smells. To show that participants made bridging inferences, Singer et al. compared people's response times to questions about explicit text information, such as

'Did Carol find the milk?' to those requiring a bridging inference, such as 'Did the milk go bad?'. Readers took the same time to answer both questions, presumably because they effortlessly made the bridge to integrate the seemingly incoherent sentence about smell. Singer et al. thus interpreted their results as evidence that both pieces of information were equally activated in the text model because people drew bridging inference during reading.

Similarly, Keenan and Kintsch (1974) provided evidence that bridging connections are well integrated into the memory representation of a text. For example, when reading sentences such as (4), response times for the statement 'Bob Birch is the man who is hiding' were equally fast for the text that required a bridging inference as for another version of the text that made the inference explicit.

> *(4) Police are hunting a man in hiding. The wife of Bob Birch disclosed illegal business practices in an interview on Saturday.*

Importantly, while these two experiments demonstrate that readers generate bridging inferences, they did not show when this happens during the comprehension process. Keenan and Kintsch (1974) presented the readers with the test statement 15 minutes after they read the text; it was thus impossible to say whether readers drew the bridging inference during reading or just during the question time. Similarly, Singer at al. (1992) measured reaction times on answering comprehension questions after they finished reading the texts. Again, it is not clear when the inference was drawn.

Haviland and Clark (1974) measured reading times for the target sentence like 'This Christmas, he got very drunk again' in sentence pairs such as (5a) and (5b).

(5a) Last Christmas, Eugene went to a lot of parties. This Christmas, he got very drunk again.

(5b) Last Christmas, Eugene became absolutely smashed. This Christmas, he got very drunk again.

Participants took longer to read (5a) than (5b) which Haviland and Clark suggested was because participants drew bridging inferences while reading (5a), but not while reading (5b). However, sentence reading time is not a good measure of online processing since it offers only a crude measure of whether or not an inference had been drawn. It does not offer any insight at which point in the sentences readers drew the inference – was it on the word 'drunk' or perhaps only after they had completed the whole sentence? Thus, we need a better measure of when readers draw bridging inferences.

It is important to know when readers drew bridging inferences in order to understand how they are constructed during text comprehension. The evidence described above is consistent with the idea that inferences are drawn immediately when people read sentences that do not obviously follow from the previous clause. This would indicate that the inferences were automatic and formed an inherent part of readers' text representation. It is also possible, however, that readers only draw an inference when it is important for the reading goal – such as when they are asked to answer a comprehension question.

To tap whether inferences are drawn during reading, some studies have employed tasks such as lexical decision or single word reading, with the rationale that if an inference had been drawn immediately when required, readers will show faster responses to the inference word. A good example comes from Potts, Keenan and Golding (1988) who explored whether people draw predictive or bridging inferences

during reading. They reported evidence to suggest that bridging inferences are made during the course of reading. They presented participants with short paragraphs, some of which required a bridging inference and some did not. Both the inference paragraph (6a) and control paragraph (6b) presented participants with the same inference inducing words such as *angry, husband, threw, delicate, vase, against, the wall*; importantly however, only (6a) was reasoned to induce the inferred concept '*break'*.

> *(6a) No longer able to control his anger, the husband threw a delicate porcelain vase against the wall. It cost him well over one hundred dollars to replace the vase.*
>
> *(6b) In one final attempt to win the delicate porcelain vase, the angry husband threw the ball at the bowling pins that stood against the wall. He had never won anything was determined not to miss this time.*

After reading the paragraph, participants were presented with a single word probe that tapped the inference participants were meant to draw, in this case the word *break.* They were then asked to either name the probe, or make a lexical decision on it. Potts et al. found that readers were faster to process inference words in both tasks, relative to the unrelated control words, suggesting that they activated the inference during reading. Even though lexical decision and word naming tasks come close to studying the reading process as it happens, they still do not tap it directly because they only test performance on tasks associated with the texts that were read. Furthermore, both tasks are administered only after the texts had been read and therefore do not provide insight into reading process as it happens.

Studies using event-related potentials (ERP) and self-paced reading provide more direct insight into online reading process and are thus better equipped to answer

whether people make bridging inferences immediately during reading (St George, Mannes & Hoffman, 1997; Yang, Perfetti & Schmalhoffer, 2007). In one such study, St George et al. (1997) found that readers showed diminished N400 amplitude when they read the final sentence in a paragraph which required a bridging inference, compared to paragraphs which did not require one to be made. The negative 400 amplitude is sensitive to processing difficulties and increases, for example, when people read anomalous words in the text. The reduction in N400 thus indicates easier word processing, in this case presumably because participants effortlessly drew a necessary inference to integrate the target word. St George et al. also found that readers showed reduced N400 for the bridging condition regardless of their working memory span, but differed in generation of other types of inferences. Their results provided further evidence that bridging inferences are a basic component for comprehension and are computed as text is read.

However, it still remains unclear whether people generate inferences immediately when they come across words that require them, or whether they wait until they have information from the entire sentence. The studies described so far cannot provide an answer to this question because they only looked at inference activation once the entire text had been read. Their results would be the same regardless of when the inferences were drawn. Even if the presented evidence shows that bridging inferences are generated during reading, the exact conditions of this process are yet to be clarified. This thesis hopes to elucidate when exactly during reading people generate inferences and aims to specifically examine the role of context in this process. Before we examine the importance of context for inference generation in general, let us consider the second type of inferences that are important for successful comprehension – predictive inferences.

Predictive inferences

While bridging inferences are drawn after the reader had come across information which requires them to be made, predictive inferences are drawn before the reader arrives to that specific concept. Unlike bridging inferences, predictive inferences are not primarily drawn to keep the text cohesive, but are drawn in anticipation of what might come next in the text. Evidence for their generation is mixed and they have a more controversial place in the literature than the well-established bridging inferences. However, many studies which focused on the generation of predictive inferences have methodological problems that either prevent detection of predictive inferences or cannot reliably distinguish bridging inferences from predictive. By describing these problems, I aim to show that evidence against the use of predictive inferences is often unreliable and I propose that predictive inferences might play a more prominent role in word-to-text integration during reading than commonly expressed in the literature. The reason why predictive inferences are important is that their generation might be one of the critical features for good comprehension. In this section, I will review the evidence concerning the generation of predictive inferences, and will describe how and when people generate them during reading, as well as discuss their importance for skilled comprehension.

Predictive inferences are inferences about what might come up next in the text. They are activated through cues in the current context; for example, when reading about heavy dark clouds, the concept of rain or storm might become activated. This activation causes a generation of predictive inference that rain or storm might occur in the following text. The activation happens before the text actually requires that readers make an inference to integrate a new concept. In this respect, predictive inferences differ from bridging inferences because bridging inferences are drawn when a new

concept is presented. Because the predictive inference is 'pre-activated', it is available immediately when the concept which requires it appears in the text, allowing very fast integration. Based on this assumption, we would expect texts with strong cues for further development to be processed faster than texts without such cues. Highly causally-related sentences are an example of such texts, meaning that people should process them more quickly than causally-unrelated texts. Indeed, when people read highly causally related sentences such as (7), they read the phrase "went out" faster than in (8), where the sentences were not as strongly related (Singer, 1993)

(7) Dorothy poured the bucket of water on the fire. The fire went out.

(8) Dorothy placed the bucket of water by the fire. The fire went out.

Correspondingly, research has further supported the conclusion that highly causally related sentences are easier to process and recall than sentences not causally related (Kuperberg, Paczynski, & Ditman, 2011; Myers, Shinjo & Duffy, 1987; Trabasso & van den Broek, 1985). The important characteristic of causally related sentences with respect to predictive inferences is that they are highly indicative of potential text development. But, predictive inferences can be generated for sentences other than those causally related as long as the context sufficiently constrains future text direction. Keefe and McDaniel (1993) presented participants with contextually constraining sentences such as (9) and measured naming times for the subsequent inference word 'sat'. The control condition provided the same lexical cues as the constraining sentence, but the overall context was not predictive, as in (10).

(9) After standing through a three-hour debate, the tired speaker walked over to his chair.

(10) The tired speaker moved the chair that was in his way and walked to the podium to continue his three-hour debate.

They found that participants pronounced the word 'sat' significantly faster in constraining context than in the control condition. Moreover, participants pronounced the word 'sat' equally fast as when they read sentences which made the inference explicit. Many other studies report similar results (Cook, Limber & O'Brien, 2001; McKoon & Ratcliff, 1986; Schmalhofer, McDaniel & Keefe, 2002 for a review).

Predictive inferences are thus generated rapidly when the context allows it but they also appear to decay rapidly if they are not confirmed by the following content. This may be a reason why many studies which looked for predictive inferences failed to find them. For example, an already mentioned study by Potts, Keenan and Golding (1988) presented participants with four scenarios describing the breaking of a vase. In the 'explicit' scenario, the inference probe word 'break' was made explicit in the text. In the 'predicted' scenario, the probe word was to be automatically inferred. For the 'coherent' scenario, authors avoided backward inferences to the probe word by providing ensuing context which held no association to it. Finally, the 'control' scenario controlled for transient activation of concepts related to the words presented in the text, as shown earlier in (6b). They found that participants showed faster naming times for the inference target word 'broke' only in the explicit and coherent conditions, not in the predicted condition. Let us though have a closer look at the scenarios Potts et al. used:

(11) Predicted: *No longer able to control his anger, the husband threw the delicate porcelain vase against the wall. He had been feeling angry for weeks, but had refused to seek help.*

(12) Explicit: *No longer able to control his anger, the husband threw the delicate porcelain vase against the wall and broke it. He had been feeling angry for weeks, but had refused to seek help.*

(13) Coherent: *No longer able to control his anger, the husband threw the delicate porcelain vase against the wall. It cost him well over one hundred dollars to replace the vase.*

Potts et al. measured responses to the inferential target word after participants had read the second sentence. Importantly however, the second sentences in the 'predicted' (11) and 'coherent' (13) condition serve to develop a reader's text representation in directions which may not keep the concept of 'break' activated. Predictive inferences might be generated right at the time when context invites them, but they will degrade quickly if the following context does not confirm them or elaborate them further. Thus, if Potts et al. had measured the activation of the predictive inference after the first sentence, they might have found an effect. This is what Keefe and McDaniel did (1993) and they indeed found that readers named inferential target words faster.

Thus, it appears that readers generate predictive inferences when context leads them to do so, but they degrade quickly if they are not confirmed. Contextual constraints determine how strongly the predictive inference becomes activated. The stronger the contextual constraints, the more active and specific the inferences are. Readers make predictions about a number of associated concepts as they are reading. Lassonde and O'Brien (2009) found that when reading about Jimmy throwing a rock at a car, both 'dent' and 'damage' showed faster naming times compared to the control condition in which Jimmy threw a soft ball. This is what we would expect good readers to do since they integrate current text content with their previous knowledge to create a

rich and connected text representation. In doing so, numerous concepts might be associated with a particular situation leading to an activation of different predictive inferences.

Predictive inferences are thus a critical component of good comprehension since readers can only generate them if they actively engage what they read with their knowledge of the topic. Furthermore, when explicitly instructed to read and think about likely consequences of events, actions and states described in the text, people generate more predictive inferences (Calvo, Castillo & Schmalhoffer, 2006) and their comprehension improves (Cain & Oakhill, 1999; Duke & Pearson, 2002). It is yet to be determined though which circumstances decide whether people use predictive or bridging inferences to integrate new words in their text representation, and work presented in this thesis aims to do so. Before I describe experiments undertaken in this thesis, I will examine the role of context in the inference making process since context plays a crucial role in the generation of both inference types.

Context and inference generation

Generally, the easier it is for people to generate inferences, the better their comprehension. This happens because inferences create links between different parts of the text and integrate them together. Context is one factor that determines how easy it is to make inferences and thus plays a very important role in the inference generation process. I will now review evidence showing how context influences inference-making in the word integration process. If the context is highly constraining, it would propel generation of predictive inferences to integrate new words. Contextual support also facilitates drawing of bridging inferences in texts where sentences are not linked as closely as in highly constraining texts. Furthermore, context can over-ride the effects of

otherwise anomalous texts and ensure smooth integration of concepts which would normally be disruptive.

As already discussed, contextual constraints are important for the generation of predictive inferences. Constrained texts, such as highly causally-related sentences, are easier to process as they lead to a strong activation of associated concepts. It is thus easy to integrate incoming information so long as it is congruent with the predictions. In terms of bridging inferences, congruent contexts which present readers with likely sequences of events make bridging inferences easier. For instance, the second sentence in (14) needs to be integrated by bridging inference because the first sentence does not give enough cues to actively predict the second, yet the integration is achieved easily because the events are congruent.

(14) Jane looked out of the window. A small squirrel played in the garden.

Yang, Perfetti and Schmalhoffer et al. (2007) explored how previous context influences integration of the first word in the following sentence. As this study will be further described in the Introduction to Chapter 2, I will mention it only briefly here. Yang et al. presented readers with contexts that gave varying levels of supporting information necessary for the integration of a target word, to be presented in a subsequent sentence. By measuring event-related potentials, they found that contexts which provided more supportive information resulted in the target word being integrated more rapidly, regardless of whether the supportive information was explicit or only semantically related. Participants thus integrated the target word ‘hospital’ (in bold) in (15a) and (15b) equally quickly, regardless of whether they were explicitly led to expect it (15a) or not (15b). However, integration was slower in (15c) which

provided the least amount of supportive information, even though the text was equally congruent (Yang, Perfetti, & Schmalhoffer, 2007; p. 73).

(15a) Explicit condition: *Allen's baby was very ill so he put the baby in the car and rushed off to the hospital. The **hospital** had a long waiting line.*

(15b) Semantically related condition: *Allen's baby was very ill so he put the baby in the car and rushed off to the emergency room. The **hospital** had a long waiting line.*

(15c) Inference condition: *Allen's baby was very ill so he put the baby in the car and rushed off. The **hospital** had a long waiting line.*

Context is thus a strong determinant of how easily readers are able to generate inferences, leading to the prediction that texts which give a high amount of supporting context should be easier to understand. In fact, context can be strong enough to override the effect of what is normally an anomalous sentence. We know that readers show longer reading times on words that are anomalous for a described situation such as "inflated carrots" in sentence (16) (Rayner, Warren, Juhasz & Liversedge, 2004).

(16) John used a pump to inflate the large carrots for dinner.

The longer reading times reflected problematic integration of the target word with the previous context. Since carrots violated the theme set up by the verb 'to inflate', people struggled to make a meaningful inference to place the new concept in their text representation. However, when an anomalous sentence like (16) is presented in an appropriate context, it can be processed without disruptions. Such a context may for example be that of a fairy tale world. Filik (2008) presented participants with sentences such as (17) in two different contexts; either preceded by a sentence indicating a fictional set up or a sentence indicating a real-world set up.

(17) The mouse took the dynamite.

When the real-world set up sentence preceded (17), readers read the target word 'dynamite' for longer. However, when the set-up sentence referred to the popular 'Tom and Jerry' cartoon, participants no longer showed difficulties reading the very same sentence (Filik, 2008). In this way, even slight contextual hints can help people to adjust their processing and integrate otherwise anomalous concepts easily. Thus, context has very strong influence on how readers integrate new concepts in their text representation and how they generate inferences in doing so. This thesis will further explore how specific types of contextual information influence the time course and type of inferences that readers make. To do this, I needed a methodology that allowed monitoring of reading behaviour online, as reading happens. Eye-tracking provides an excellent tool for the study of online reading behaviour. I will now describe reasons why I chose it as my primary methodology for the work presented in this thesis.

Eye-tracking as methodology to investigate reading processes

As described earlier, investigations of inference making have used methodologies that do not allow the monitoring of inference making as reading happens. Measurements such as reaction times on lexical decision tasks, word probe naming, sentence reading times and accuracy on comprehension questions only tap processing that happens once the reader finished reading. These methods are certainly very informative and useful, but a method that allows observation of processing as reading happens is necessary to answer the question of when skilled readers generate inferences during reading, and how they use context to do so.

Thus, I chose eye-tracking as a primary methodology for work presented in this thesis. Monitoring of eye-movements provides an excellent online behavioural measure of cognitive processes that underlie reading. There is a large body of research that uses eye-movements to study reading (Rayner, 1998; Staub & Rayner, 2007 for a detailed review). Multiple aspects of lexical processing such as frequency and length (Rayner & Duffy, 1986; Reichle, Pollatsek & Rayner, 2006), syntactic processing (Traxler and Pickering, 1996; Traxler, Morris & Seely, 2002) and semantic processing (Morris, 1994; Morris & Folk, 1998) have all been investigated with eye-movements to elucidate how these processes work, as reading happens. Before characterizing how eye-movements can be useful for psycholinguistic research, I will give a basic overview of their features in reading.

A normal eye-movement record consists of a series of fixations and saccades. During points of fixation, eyes are still as participants extract visual information. After they have done so they move to the next point of fixation by making a saccade. Saccades are rapid, jerky movements of the eyes during which people still carry on processing of the earlier fixated information (Liversedge & Findley, 2000; Rayner, 1998). When people read sentences, their fixations are typically between 60 and 500ms long, being on average 250 ms. Fixations usually land between the beginning and middle of the word, but there is some evidence that readers fixate informative portions of words (O'Regan, 1992). People do not fixate every word in the sentence and tend to skip short function words and highly frequent words such as 'the', 'has', 'from' and so on (Rayner, 1998). People do not only process text during fixation, they also pre-process it in the direction they are reading before actually fixating it (Rayner & McConkie, 1976). When reading English, readers make eye-movements from left to right approximately 85% of the time and from right to the left about 15% of the time

(Egeth & Yantis, 1997). These regressive saccades are an important feature for good comprehension since readers make them to correct their understanding when they misanalysed sentences they were reading (Just & Carpenter, 1980).

When monitoring eye movements, there are a multitude of eye-movement measures that can be taken and these are considered to indicate different linguistic processes. Generally, these measures can be categorized in to those reflecting early processing and tapping 'lower level' linguistic processes, and those reflecting later processing and 'higher level' linguistic factors. The earliest processing measure is the duration of the first fixation that readers make on a word. The duration of a fixation on a word is affected by lexical properties such as frequency, with low frequency words being fixated for longer than high frequency words. When words are longer they are more likely to be fixated more than once, producing longer gaze durations (also called first pass times; Rayner & McKonkie, 1976). First pass times reflect all fixations readers make on a word when they read it for the first time including intra word re-reading before moving past the word to the right. Longer first pass times can reflect both lower level linguistic factors, such as word disambiguation when readers read ambiguous words (e.g., bank), or higher level comprehension processes such as detecting sentence anomalies (Rayner et al., 2004).

The eye-movement measures used for studying higher-order or discourse level phenomena such as semantic processing usually centre around regressive eye movements and re-reading time (Liversedge & Findlay, 2000). When readers struggle to make sense of the text they may make regressive saccades to re-read and reconfigure earlier material, as proposed by the Selective Reanalysis hypothesis (Frazier & Rayner, 1982). Regressive eye-movements are thus one of the strongest indicators of additional

discourse processing although there are certain problems with over-reliance on this measure.

The first problem is that regressive eye-movements may not be entirely governed by linguistic factors, as proposed by Frazier and Rayner (1982) in their Selective Reanalysis hypothesis. There is evidence that regressive saccade landing sites can be influenced by low-level spatial properties such as how close they are to the point from which the regression originated (Mitchell, Shen, Green, & Hodgson, 2008). Mitchell et al. (2008) manipulated the formatting of linguistically identical garden-path sentences and showed that the physical lay-out of the text reliably influenced landing site of the first regression from the critical word. Moreover, the proportion of regressive eye-movements usually reaches between 20-30% only in texts which are difficult to process (Staub & Rayner, 2007). This is still a small number and suggests that in most cases, a regression from the critical text or back into it will not be made, yet people would still process the text successfully.

Despite these shortcomings, regressive eye-movements and additional measures of later processing such as total time readers spent fixating the word (i.e. adding together all fixations) still provide a very good account of how comprehension is constructed during reading. Experiments exploring syntactic processing when reading garden path sentences (Christianson, Hollingworth, Halliwell, & Ferreira, 2001), the effects of word typicality on anaphor resolution (Duffy & Rayner, 1990) and the influence of context and world knowledge on processing of otherwise anomalous information (Cook & Myers, 2004; Filik 2008) have all showed that total word reading times, regressive eye-movements and, to an extent, gaze durations provide measures that help us understand discourse processing, as it happens while people read.

A growing body of evidence from eye-movement monitoring indicates that semantic processing does not necessarily happen strictly at the location of critical information, but it may spill over into the region immediately downstream from the critical word (Traxler & Pickering, 1996) or manifest at the final sentence regions as a 'wrap up' process (Rayner, Kambe & Duffy, 2000). Thus, eye-movement monitoring provides extremely informative account of how readers process different aspects of text immediately during reading because it allows us to study processing in all parts of written text, and not just at the critical words as is the case in lexical decision or word naming tasks.

Eye-tracking offers a major advantage over another methodology commonly used to study online reading behaviour, the event-related potentials (ERPs), because it allows naturalistic reading during experiments. Unlike eye-tracking, ERPs present text to the readers in an unnatural way. Since ERPs are sensitive to eye movements, these have to be minimized during testing. Thus, readers are presented with only one word at a time during the ERP experiments and each word is presented for a fixed amount of time. Given the nature of eye-movements during reading I described in this section, such text presentation may give an imprecise account of online reading as readers cannot process the texts as they would naturally – by making saccades and adjusting fixation locations and fixation times as they need. Furthermore, it may increase processing load by placing higher demands on reader's memory as only one word is available at the time. Given advantages that eye-tracking offers, research presented in this thesis uses eye-movements to explore how skilled readers generate inferences to integrate new concepts in their mental representation. In doing so, I hope to extend findings from work using off-line measures and event-related potentials to study inferences and verify that the same processes take place during naturalistic reading.

Overview of this thesis

The aim of this thesis is to elucidate how people use inferences to integrate new concepts in their mental text representation; when during reading the inferences are generated, which inference-making strategies are used, and how skilled readers use previous context to help them with inference generation. By looking at populations of skilled readers, I want to further understand the strategies which allow readers to achieve good text comprehension. Previous work has studied the online generation of inferences in word-to-text integration by using event-related potentials; this does not allow participants to read texts naturally (Yang, Perfetti & Schmalhoffer, 2006, 2007). To allow natural text presentation, I thus chose eye-tracking as the main method for my work.

Traditionally, bridging inferences have been considered the prominent strategy for integration of new concepts into reader's text representation. The first aim of this thesis was thus to explore whether skilled readers used previous context to facilitate the drawing of bridging inferences during online word-to-text integration. To do so, I examined eye-movements for target words which required participants to draw a bridging inference in order to integrate them in mental text representation. In Experiment 1, I varied the amount of supporting context participants were given to make an inference needed for the integration of a target word. This was achieved through manipulating the availability of a direct referent for the target word in the previous context because new concepts are easily integrated when direct referents are available. As skilled readers are known to efficiently use context, I predicted they

would compute bridging inferences faster in conditions which provided a direct referent for the target word.

Experiment 1 indicated that only referents sharing the lexical and surface form reliably facilitated target word processing. It was therefore unclear whether participants' ability to quickly draw bridging inferences benefited from lexical or semantic contextual properties. To find out, an extra condition was included in the design of Experiment 2 which separated lexical and semantic aspects of the target word referents. However, the results of Experiment 2 conflicted with the results of Experiment 1 and indicated that participants may not have used bridging inferences to integrate the target word or that context provided no benefit for drawing of bridging inferences. Before a conclusion could be drawn, a number of methodological problems was uncovered and needed addressing.

This was the purpose of Experiment 3. A new set of stimuli was designed for Experiment 3 to minimize methodological problems associated with the previous two experiments. Additional manipulation targeted at increasing processing difficulty of the new stimulus set was included since previous experiments suggested problems with a potential ceiling effect. To provide further confirmation that participants truly integrated the target word and drew necessary inferences, an offline memory task was included in Experiment 3. The memory task measured how likely participants were to falsely recall seeing the inferred information explicitly presented in the text. As an offline task (that is, administered after the paragraphs had been read), the memory test provided evidence that the necessary inferences had been drawn; participants were more likely to falsely recall seeing inferred information in the text if the context was more supportive. However, eye-movement measures suggested that bridging inferences had not been drawn during online reading. A question thus arose whether participants used predictive

rather than bridging inferences to integrate the target word as even the least contextually supporting conditions might have provided participants with sufficient cues to anticipate the target word, thereby eliminating differences in the eye-movement record.

Thus, the second aim of this thesis was to establish whether readers used predictive inferences rather than bridging for word-to-text integration in texts similar to those used in Experiment 1, 2 and 3. To do so, Experiment 4 compared target word processing in conditions where the target word either confirmed the development set out by the preceding context or did not confirm it. Furthermore, the strength of contextual cues indicating future text development was manipulated to see whether skilled readers would integrate new concepts through predictive inferences even in less constraining contexts. Together, Experiments 1 to 4 provide an investigation into which inference-making strategies are employed by skilled readers during word-to-text integration in online reading, and how preceding context influences this choice.

Chapter 2: The effect of supporting context on the construction of bridging inferences during word-to-text integration (Experiment 1)

Experiment 1 explored whether availability of supporting context influenced the construction of bridging inferences as skilled readers integrated new words into their text representation. Participants read 27 two-sentence paragraphs containing the same target word in three different conditions while their eye-movements were recorded. The control condition provided explicit information for the inference together with a direct referent for the target word. The paraphrase condition provided a contextually related referent for the target word integration, and the inference condition provided less contextual support with no direct referent for the target word. I hypothesized that skilled readers would be sensitive to the availability of supporting context and use it to facilitate generation of inferences. As supporting context becomes less available, skilled readers were expected to spend longer integrating the target word. The results confirmed my prediction as participants were faster to construct the necessary inference and integrate the target word in the control condition relative to the other two conditions and this effect showed early on during target word processing. However, target word reading was only marginally shorter in the paraphrase condition compared to the inference condition. This finding suggested that contextual support did not reliably extend beyond providing explicit information for the inference construction.

Introduction

Successful reading comprehension is a result of efficient and accurate text processing which allows readers to construct a rich and coherent representation of the text they have read (Zwaan & Radvansky, 1998). Such a representation would need to hold information about text agents and their actions and mutual relationships. The information would need to be incorporated with readers' world knowledge, organized chronologically and revised according to the main reading goal. When new information renders past concepts irrelevant, the representation would need to be up-dated and irrelevant concepts discarded. Thus, good readers need to master not only lower level language skills such as decoding and lexical retrieval, but skills which allow them to actively work with information they have read. Three such skills indispensable for building a coherent and integrated text representation are inference making and information integration, comprehension monitoring and knowledge of text structure (Cain, Oakhill& Bryant, 2004). Inference making and information integration is of particular interest in this work.

To build a coherent and integrated text representation which is dynamically updated, readers need to incorporate explicitly stated new information into their mental model together with its implications and connotations. Consider the following sentence (1):

> *(1) Ann looked at the dark grey clouds above her head. She should have taken her umbrella with her.*

To create a coherent text representation and understand this text correctly, the reader needs to draw on their world knowledge that dark grey clouds are a source of rain. They also need to activate their knowledge that people do not like to get rained on

and umbrellas are used to protect people from the rain. Without activating this extra information and incorporating it into the text model, it would be difficult to smoothly integrate the word 'umbrella' with the previous text and understand it fully. The process of filling in implications of explicitly stated information is referred to as inference making. The ability to draw correct inferences is crucial for comprehension (Cain & Oakhill, 1999; Oakhill, 1984).

During reading, readers routinely fill in information in order to make sense of the text. Sometimes, this is achieved via conscious, problem-solving inferences, such as combining the clues in a murder mystery novel to find the murderer. The present work however focuses on more automatic inferences that are instrumental in keeping text coherent at both local (working with those propositions of text that are currently active in the working memory) and global levels (connecting explicit pieces of information into an overall causal chain). The role of global level inferences is to keep the overall text representation coherent and come to aid when local level inferences do not compute smoothly (McKoon & Ratcliffe, 1992). In terms of immediate word-to-text integration, I am more interested in local level inferences as they are critical for integrating adjacent information. Such inferences often appear in the form of bridging inferences.

Bridging inferences are important for the construction of a coherent text representation and are specifically computed to establish coherence between two sentences in a text. For example, in order to understand (2), a reader must make an inference across the two sentences that Jane is a member of the band. Without this inference, information from the second sentence would make little sense in the context of the first sentence.

(2) It was the day of the band's biggest performance. Jane was so nervous she couldn't even finish her breakfast.

When reading the word 'nervous', readers need to make a bridge back to the previous sentence by inferring that Jane is a member of the band. Otherwise, they will find it hard to meaningfully integrate the word 'nervous' into their text representation. By creating a bridge with preceding text, bridging inferences are crucial for successful word-to-text integration and subsequent text comprehension. Indeed, both children (Cain & Oakhill 1999; Cain et al, 2004; Yuill, Oakhill, & Parkin, 1989) and adults (Landi & Perfetti, 2007) with reading comprehension impairments are poor at making such inferences.

While it is clear that the ability to draw relevant inferences enables readers to form an active and well-integrated text representation as they read (Kintsch, 1994), readily integrating new information (Whitney, Ritchie & Clark, 1991) and drawing on world knowledge (Cain, Oakhill, Barnes & Bryant, 2001), it is less clear when and how these inferences are drawn. The question I am specifically interested in is whether bridging inferences are drawn immediately when new information is read, or whether this process happens later. It is possible that the time course of drawing inferences changes as a function of how easy it is to draw them. In turn, the ease of drawing inferences often strongly depends on the context. Thus, it is important to consider the effects of context on the ease and timing of drawing inferences.

For situations where context implies a causal relationship, bridging inferences tend be drawn on-line immediately as text is processed. For example Kuperberg, Paczynski and Ditman (2011) presented readers with the following texts while they recorded their ERPs:

(3) Jill had very fair skin. She forgot to put the sunscreen on. She had sunburn on Monday.

(4) Jill had very fair skin. She usually remembered to put the sunscreen on. She had sunburn on Monday.

The word 'sunburn' was processed sooner in (3) compared to (4), as evidenced by a reduction in N400 amplitudes. This suggests that when the bridging inference is causal – like in the case of getting sunburn because of forgetting to put the sunscreen on – readers draw it automatically compared to the situations when it is not causal.

However, the picture is less clear for situations when bridging inferences are not causal. When readers are presented with highly constraining context as in (5) they display shorter gaze durations for the word 'knife' than in the less constraining sentence (6) (O'Brien, Shank, Myers & Rayner, 1992):

(5) When the woman screamed, the mugger stabbed her with his weapon. He threw the knife away.

(6) When the woman screamed, the mugger assaulted her with his weapon. He threw the knife away.

This suggest that if context is constraining enough, readers draw bridging inferences immediately while reading the target word, but they are delayed in doing so if the strong context is absent, presumably because it is more difficult to make a bridge in such cases.

In order to find out whether readers draw bridging inferences immediately when integrating a word into the previous context we need to look at inference making as it happens. A large body of research in inference making uses mostly off-line measures such as probe naming tasks, lexical decision, accuracy of answering comprehension

questions and free recall (McKoon & Ratcliffe 1992; Haviland and Clark, 1974; Cain et al 2001; Cain et al, 2004). A problem with these methodologies is that their findings are difficult to interpret because they might stem from post-reading factors, rather than being a reflection of immediate processing while readers read inference evoking words.

To tap immediate processing, Yang, Perfetti and Schmalhoffer (2005, 2007) measured word-to-text integration in relation to preceding context in a series of ERP studies comparing skilled and less skilled readers. They aimed to discover whether readers incrementally mapped each newly read word into their semantically interpreted text representation for comprehension or whether they simply encoded words into their memory, integrating their meaning only later when more information became available. In order to map words into their text representation, readers would need to link them with the previous content by finding appropriate referents. To construct the links successfully, readers would need to draw inferences. Thus, the ease of making a link would be proportional to the ease of generating an appropriate inference. Yang et al. manipulated the ease of constructing links by varying the accessibility of referents for the target words in preceding text.

In the easiest condition, the explicit condition, they used an explicit referent which matched the target word in both lexical and surface form. Thus, it was very easy to map the target word onto its previous referent and readers only had to generate a minimal inference to do so. In the following example, integration of the target word 'explosion' would be very fast and easy due to the availability of its referent 'exploded'. The only inference that readers need to draw is that the explosion comes from the bomb mentioned in the first sentence:

(7) Explicit condition: *The bomb hit the ground and exploded. The explosion was later reported to the commander.*

For the second condition, the paraphrase condition, Yang et al. substituted the explicit referent with a phrase which provided a semantic link with the target word, but did not share its lexical and surface form. The phrase used was not synonymous with the explicit antecedent but it paraphrased its meaning. Thus, integration of the target word can only be achieved through establishing a more demanding inferential link than in the explicit condition. However, since the referent in the paraphrase condition was semantically related to the target word, it provided a strong mediating link between the two, as in the example (8) below:

(8) Paraphrase condition: *The bomb hit the ground and blew up. The explosion was later reported to the commander*.

In the most demanding condition, the inference condition, Yang et al. further manipulated the difficulty of target word integration by not providing a clear referent for the target word at all. Instead, readers had to lean on their world knowledge to compute an inference that would allow them to integrate the critical word in their text representation, as in (9):

(9) Inference condition: *The bomb hit the ground. The explosion was later reported to the commander.*

In order to establish that the target word was truly integrated, Yang et al. created a 'baseline' condition which provided minimal semantic or world knowledge antecedent for the target word integration as in the example (10). Integration of the target word

'explosion' in this case reflected baseline processing of the target word in the absence of earlier supporting context.

(10) Baseline condition: *The bomb was stored safely on the ground. The explosion was later reported to the commander.*

Such integration should be more demanding than in the previous conditions where context set up the scene for the target word. Yang et al. hypothesized that if context provided processing advantage for the target words and these were immediately integrated into textual representation as people read, the other three conditions should display reduced N400 amplitudes relative to the baseline. The N400 effect is the negative voltage shift between 300 and 500 ms after the onset of the word and it is argued to reflect semantic processing in ERP studies (via violation of semantic expectancy).

As hypothesized, they found that the baseline condition (which provided no antecedent for the target word) showed evidence of integration difficulties by displaying high N400 amplitude compared to the explicit and paraphrase conditions. Yang et al. found no difference between the explicit and paraphrase condition suggesting that readers integrated words equally easily if an antecedent was directly available, regardless of its nature. The inference condition though only showed marginally significant reduction in N400 compared to the baseline condition. Yang et al. concluded that readers used bridging inferences to integrate the target word rather than predictive inferences, reasoning that had readers pre-activated the inference prior to reading the target word (as would be the case, if predictive inferences were being made), integration should have been similarly difficult across the three contextually supporting conditions.

It was interesting that Yang et al. found no advantage for the explicit condition over the paraphrase condition when it came to the ease of integration is rather interesting. However, Yang et al. used word-by-word presentation paradigm with one word displayed at a time. Such reading differs from natural reading, where people routinely make regressive eye-movements during reading to re-read earlier portions of text, and process incoming words parafoveally prior to foveal fixation (Rayner, 1998; Schotter, Tran & Rayner, 2013). None of this is possible during word-by-word reading. It is thus possible that the advantage of having the explicit antecedent near the target word disappeared during word-by-word presentation as readers did not have the full text at their disposal. Therefore, it is important to establish whether the same pattern of word-to-text integration via drawing bridging inferences from available context happens in more natural reading situations.

A number of other issues arise from Yang et al.'s methodology. All of the target words were positioned as the first word in the sentence following the context-giving sentence. This position was of interest to them because the only context that the first word of a new sentence can be linked to is that from the preceding sentence. However, even words positioned in the middle of a sentence often need to be integrated with the content from the preceding sentences. In some sense, such integration may even be more demanding because the preceding sentence is not immediately available for the words in middle positions of the subsequent sentence. There is also a problem with Yang et al.'s baseline condition in that it presented readers with anomalous texts. It is thus not clear whether high N400 amplitudes for this condition resulted from the lack of previous referents, or from the text being anomalous.

In order to establish the time course of word integration via inference-making in natural reading, I adapted Yang et al.'s experimental paradigm for use in an eye-

tracking experiment. I varied the ease of mapping a target word to its referent in the previous context and used equivalents of explicit, paraphrase and inference condition. This way, I manipulated the difficulty of the inference readers needed to draw for successful word-to-text integration. I dropped the baseline condition due to its anomalous nature and used the explicit condition as the control condition. Since integration in the explicit condition requires only a minimal inference (because the antecedent for the target word is readily available), it served as the baseline for target word integration. The target words appeared closer to the middle of sentences since word integration from these positions is more frequent during naturalistic reading (for every sentence, only one word is integrated from the initial position).

I predicted that readers will draw bridging inferences immediately when reading the target word. Since evidence from the anaphor processing literature shows that sentences are read faster when the anaphor and antecedent share the same lexical and surface form during natural reading (Forster, 1981; Haviland & Clark, 1974), I expected the target word in the control condition to be read the fastest. As the mapping between the target word and its referent is less apparent in the paraphrase condition, reading times for the target word were predicted to increase, relative to reading the same word in the control condition. When readers were presented with predictive contexts such as in (11), they showed reduction in second-pass reading times and regressions back to the target word 'studied' compared to the control context which was not predictive (Calvo, Messeguer & Carreiras, 2001).

(11) Three days before the exam, the student went to the library and opened his notebook. He studied for an hour.

Because the paraphrase condition provided more predictive context than the inference condition, I expected shorter reading time on the target word in the paraphrase condition relative to the inference, especially with regard to the time used for re-analysis. The inference condition provided no referent for the target word other than the situational set up. The construction of the bridging inference was therefore the most demanding in this condition as participants had to rely on their world knowledge to integrate the target word. I thus predicted them to spend the longest time reading the target word in the inference condition.

An advantage of using eye-tracking measures is their ability to provide insight into the timecourse of word processing as reading happens. I anticipated that the rising demands on word integration and inference making would be reflected in the total word reading times and time spent on the reanalysis of the target word. These measures (also referred to as measures of *later processing*) are thought to reflect higher order comprehension processes such as word-to-text integration, build-up of a coherent text representation and up-dating of a situation model (Liversedge & Findlay, 2000; Rayner, 1998). The effect might also be seen in the first pass times, measures of *earlier processing*, although this may be less likely since first pass times tend to show processing effects on anomalous words for the given context, such as the phrase 'carrots for dinner' in (12) (Rayner, Warren, Juhasz & Liversedge, 2004).

(12) John used a pump to inflate the large carrots for dinner

Processing differences between conditions were not expected to arise in the first fixation times since such early measures are mostly thought to reflect fast and automatic initial processing at the word-level, such as lexical access, as reflected by the effects of word length and frequency (Rayner, 1998).

Method

Participants:

Thirty-nine participants were recruited from the population of Oxford University undergraduate and graduate students (18 males and 21 females). All participants were monolingual native English speakers between 20 and 27 years old with normal or corrected-to-normal vision and no history of language or reading difficulties.

Apparatus:

An SR Research EyeLink 1000 desktop eye-tracker with chin and head support to minimize the head movements was used to monitor the participants' eye movements. Participants read text presented on a Dell LCD monitor with a screen resolution of 1024 x 768 pixels at a distance of 60 cm. Viewing was binocular though the tracker measured gaze location from the right eye only by monitoring the pupil and corneal reflection at a sampling rate of 1000 Hz. The tracker was calibrated to allow no larger eye position error than 0.4˚ with 3 characters subtending to 1˚ of visual angle to allow precise tracking for the words only three characters long. The sentences were presented in three to four line paragraphs aligned to the left of the screen with a white Courier New 16-point font on a black background.

Materials and Procedures:

Twenty-seven two sentence paragraphs containing the same target word were constructed. To integrate the target word into their text representation, readers needed to make an inference drawing on the context of the previous sentence, or their world knowledge. Each paragraph was on average 34 words long and presented on three to four lines of text. The three conditions manipulated the difficulty of making a bridging

inference by varying the amount of contextual information necessary for easy mapping of the target word to its referent. The *control condition* provided an explicit referent for the target word, matching in lexical and surface form. The *paraphrase condition* provided a referent equivalent in meaning to the explicit one for the given context. The *inference condition* provided a limited contextual referent for the target word meaning that readers had to draw on their world knowledge and detailed text representation to complete the inference (see *Example 1* for illustration and *Appendix B* for the full stimulus set).

Three versions of each experimental item were constructed (control, paraphrase, inference). Paragraphs were presented in a counterbalanced design so that each participant received nine paragraphs from each of the three conditions without repetition. This way, each participant read each experimental paragraph in one condition only and all versions of 27 paragraphs were presented equally often across the participants. The paragraphs were presented in the random order.

Each paragraph was followed by a comprehension question. Participants were presented with two possible answers where the incorrect answer always contained referents to the paragraph to ensure readers had integrated the correct antecedent and referent information. Participants indicated their choice by pressing the corresponding button on a button box. The paragraphs remained present on the screen as participants read the question, allowing them to look back at the text if they wished.

Example 1 *shows the three different conditions of the same paragraph in Experiment 1. The target word is in italics. Each paragraph was followed by a comprehension question with two answers to choose from. The position of the correct answer (left vs. right) varied across the experiment.*

Control

Cathy had a beautiful blue lamp on her table when suddenly an earthquake hit the city and the lamp broke. Tiny pieces of *broken* blue glass lay on the floor.

Paraphrase

Cathy had a beautiful blue lamp on her table when suddenly an earthquake hit the city and the lamp fell down. Tiny pieces of *broken* blue glass lay on the floor.

Inference

Cathy had a beautiful blue lamp on her table when suddenly an earthquake hit the city where she lived. Tiny pieces of *broken* blue glass lay on the floor.

What happened to Cathy's lamp?

It fell and broke It lay on the floor

Procedure:

Participants were told that they would be reading short stories on the computer screen and their task was to answer the comprehension question afterwards. There was a 130 seconds limit on completing each trial.

Before commencing the reading session, a calibration procedure to ensure precise tracking was carried out. A white fixation dot appeared in the middle of the screen and subsequently moved in a random order between each screen corner and the middle screen position. Participants were instructed to follow the dot with their gaze until satisfactory calibration was achieved.

Participants were instructed to always read the story first and then read the comprehension question at the end. Before the start of each trial, a white fixation point appeared at the centre of the screen. The participants were instructed to fixate on this point. Afterwards, a white square appeared at the left edge of the screen at the position which marked the location of the beginning of the first sentence in the narrative. The participants were instructed to fixate on the white square which triggered the presentation of the text. The participants finished the reading of the paragraph by answering the comprehension question at the end, using either the right or left trigger on a hand-held controller. Pressing a trigger finished the trial and erased the paragraph from the screen.

Results

Comprehension accuracy

Readers displayed about the same level of accuracy for all three conditions; the control condition showed on average 96% accuracy (SD= 9.7%), the paraphrase condition scored 97% of accurate answers (SD= 5.8%) and the inference condition scored 95% correct answers (SD=7.5%). There was no significant difference between conditions in terms of comprehension question accuracy, ($F1(1,33)<1$, $F2(1,25)<1$). Based on these data, it is clear that readers successfully integrated the target word and drew the necessary inferences equally well in all three conditions.

Eye-tracking data preparation

Following standard procedures, fixations shorter than 80ms were combined with a nearby fixation. If there was no fixation within three letter space, the fixations were deleted. Fixations longer than 1200 ms were also deleted.

Out of the 39 participants tested, five were excluded from the analysis due to the deficient quality of the eye-tracker's recordings. From the items side, item 24 had to be excluded due to script error. The total number of subjects available for analysis was 34 and the total number of items was 26. Two trials out of 918 were excluded due to tracker loss (i.e. no fixations) in the regions of interest.

To facilitate the analysis, paragraphs were divided into two regions (see Example 2). The first region of interest is *the target word region,* comprising the target word (shown in italics) and the space before the target word. The second region of interest is the *post-target region* which consisted of the space and the word following the target word. If the word following the target word was three letters long, next word was included in the post-target region as three-letter words have a high likelihood of being skipped during reading.

Example 2 *shows an experimental paragraph with the target and post target word regions, the target word region is in italics.*

Cathy had a beautiful blue lamp on her table when suddenly an earthquake hit the city and the lamp fell down. Tiny pieces of */ broken/* blue*/* glass lay on the floor.

Eye-tracking analysis of the target word region

To assess the effect of increased demands on inferential integration in early and later processing (Liversedge & Findlay, 2000; Rayner, 1998; Rayner, Chace, Slattery & Ashby, 2006), the following measures were computed for the target word region: first fixations, first pass times (gaze duration), second pass times, total time, go-past time, the probability of regressing into the region and the probability of regressing out of the region *(see Appendix A for the list of all eye-movements measured and analysed in the current experiment together with cognitive processes the eye-movements are thought to reflect)* . A one-way repeated measures ANOVA with condition type as a within subject factor was carried out to explore the effect of increased inference-making demands on reading time, as tapped by the eye-movement measures. When applicable, Bonferroni corrected post-hoc pairwise comparisons were carried out to explore the direction of the effect. The analysis was performed across both subjects and items data. For the full summary of results see Table 2.1.

Table 2.1

Main effect of condition for eye-movements of early and later processing in the target word region in Experiment 1. Mean reading times (ms) and standard deviations (in parentheses) for all three conditions.

	N=34	M / SD			
	Measure (ms)	**Control**	**Paraphrase**	**Inference**	**F**
Early processing measures	First fixations	190_1	210_2	207_2	$F_1(2,66)=9.64$, $p<.01$
		(29)	(35)	(39)	$F_2(2,50)=4.94$, $p<.05$
	First pass time	237_1	255_2	262_2	$F_1(2,66)=3.23$, $p<.05$
		(55)	(46)	(64)	$F_2(2, 50)= 2.36$, $p=.8$
Later processing measures	Total time	304_1	357_2	400_2	$F_1(2,66) = 3.7$, $p<.01$
		(83)	(84)	(137)	$F_2(2,50)= 5.54$, $p<.05$
	Second pass time[1]	323	345	447	$F_1(2,16) = 1.2$, $p=.3$
		(140)	(119)	(299)	$F_2(2,16)= 1.12$, $p=.3$
	Go past time	301_1	337_2	358_2	$F_1(2,66)=2.5$, $p=.09$
		(94)	(121)	(125)	$F_2(2,50)=3.57$, $p<.05$
	Measure (probability of regressing)				
	Regressions in	0.15	0.18	0.21	$F_1(2,66)=2.7$, $p=.07$
		(0.14)	(0.15)	(0.14)	$F_2(2,50) = 2.02$, $p=.1$
	Regressions out	0.14	0.18	0.22	$F_1(2,66)=2.04$, $p=.1$
		(0.16)	(0.18)	(0.21)	$F_2(2,50) = 2.63$, $p=.08$

Note: *measures with same subscript number do not differ from each other significantly*

There was a main effect of condition type in both measures of early processing; first fixations and first pass times. For first fixations, both the paraphrase and inference condition elicited longer fixations to the same word than in the control condition.

In the same manner, first pass times on the target word were significantly longer in the paraphrase and inference condition than in the control condition. However, the difference was only significant in the by subjects analysis with a trend towards

[1] Degrees of freedom may differ for different eye-tracking measures, as the comparison between the three conditions was carried out only when data were available for each of the conditions. This will produced reduced degrees of freedom especially for second pass times.

significance in the by item analysis. The difference between paraphrase and inference conditions for the first fixations and first pass times was not significant for either of the early processing measures (see Figure 1a,b).

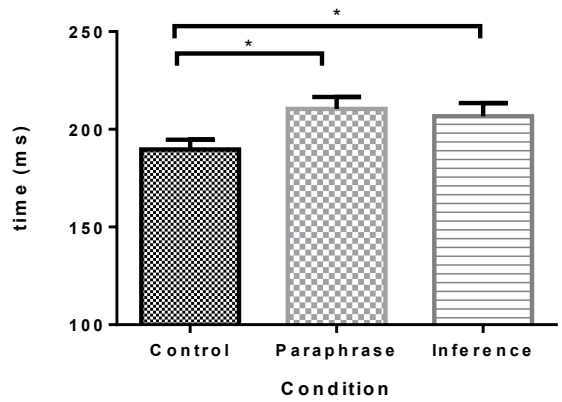


Figure 1a *Mean first fixation durations (ms) on the target word in all three conditions of Experiment 1. Error bars indicate SEM. *p< 0.05*

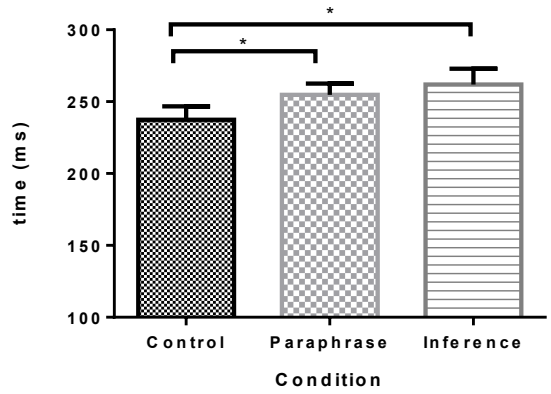


Figure 1b *Mean first pass reading times (ms) on the target word in all three conditions of Experiment 1. Error bars indicate SEM. *p< 0.05*

Consistent with my predictions, there was a main effect of condition type in the measures of later processing; total looking times and go-past times. Just like with the early measures, readers showed the shortest total looking times in the control condition compared to the paraphrase and inference conditions. However, the difference in total

reading time between paraphrase and inference condition was only marginal, p=.06 (see Figure 2a).

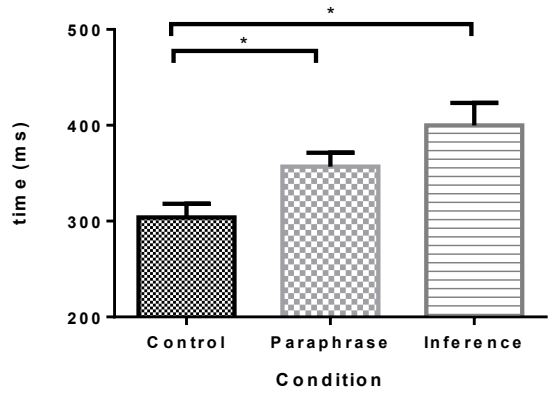


Figure 2a *Mean total reading times (ms) on the target word in all three conditions of Experiment 1. Error bars indicate SEM. *p< 0.05*

The go-past times revealed similar pattern as total times although the main effect of condition was significant only in the by items analysis, with a trend towards significance in the by subject analysis. Participants spent more time reading the target word before they moved past it in the inference condition when compared to the control condition (see Figure 2b).

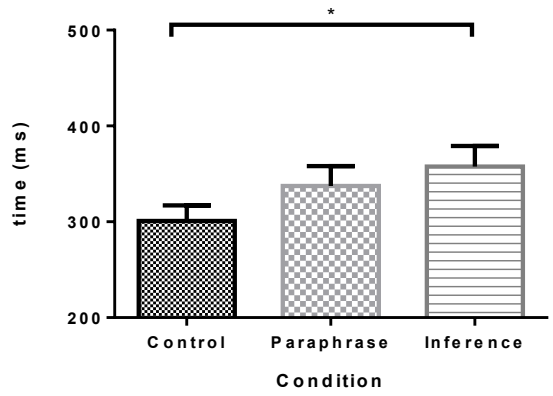


Figure 2b *Mean go-past reading times (ms) on the target word in all three conditions of Experiment 1. Error bars indicate SEM. *p< 0.05*

No other measures of late processing showed a main effect of condition, although the general trend was always in a similar direction: participants spent less time

reading the target word in the control condition compared to the other two conditions. Despite not reaching significance, the probability of regressing out of the target word to re-read previous portions of the text and the probability of regressing back to the target word showed the same trend. Readers regressed less in the control condition than in the other two conditions implying that readers needed to re-process the target word more as preceding context provided fewer direct referents for the word's integration.

Eye-tracking analysis of the post-target word region

The post-target region comprised the word or word phrase immediately following the target word. If integration of a word into text is demanding, processing of the target word may be delayed and spill over into processing of the following word (Calvo, 2010; Rayner, 1998). However, it seems that the target word processing was rather straightforward this experiment and there was no evidence for effects on delayed processing. Eye-movement measures for the post-target region showed no effect of condition type ($Fs<1$) (see Table 2.2 for the full summary of the data).

Table 2.2

Main effect of condition for eye-movements of early and later processing in the post-target word region in Experiment 1. Mean reading times (ms) and standard deviations (in parentheses) for all three conditions.

	N=34	M / SD			
	Measure (ms)	**Control**	**Paraphrase**	**Inference**	**F**
Early processing measures	First fixations	200	200	205	$F_1(2,66) <1$
		(41)	*(32)*	*(33)*	$F_2(2,50)<1$
	First pass time	211	222	221	$F_1(2,66)=1.2,p=.8$
		(58)	*(68)*	*(62)*	$F_2(2, 50)= 1.3, p=.8$
Later processing measures	Total time	299	331	325	$F_1(2,66) <1$
		(89)	*(92)*	*(123)*	$F_2(2,50) <1$
	Go past time	250	300	287	$F_1(2,66) <1$
		(60)	*(138)*	*(113)*	$F_2(2,50) <1$
	Second pass time	82	98	89	$F_1(2,16) <1$
		(82)	*(81)*	*(96)*	$F_2(2,16)< 1$
	Measure (probability of regressing)				
	Regressions in	0.17	0.14	0.15	$F_1(2,66)=1.4, p=.7$
		(0.14)	*(0.13)*	*(0.14)*	$F_2(2,50) = 1.1, p=.3$
	Regressions out	0.14	0.14	0.13	$F_1(2,66) <1$
		(0.19)	*(0.14)*	*(0.14)*	$F_2(2,50) < 1$

Thus, the integration of the target word appeared rather straightforward even in contextually less supporting conditions than the control, as no spill-over processing was recorded on the post-target word. The lack of main effect in the post-target region also provides further evidence that readers drew bridging inference as they read the target word, rather than completing the inference once the target word had been read.

Discussion

The aim of this experiment was to shed further light on the time course of word-to-text integration via inference making during natural reading. By manipulating the level of contextual information and availability of the antecedent for the target word, I varied the difficulty of the inferences readers had to draw in order to integrate the target word into their text representation. Thus, I wanted to establish whether readers use contextual information efficiently and immediately when making inferences during reading.

I found that readers spent proportionally longer integrating the target word as the amount of contextual information necessary to establish its antecedent decreased. Integration was fastest in the control condition which provided an explicit antecedent for the target word. Reading times were longest in the inference condition which did not provide a direct referent for the target word in the context-giving sentence. By observing this pattern, my data are consistent with the conclusions of Yang et al. (2007), namely that readers use bridging inferences to integrate target words in the paraphrase and inference condition. However, I did not replicate Yang et al.'s finding of facilitated processing in the paraphrase condition, relative to the inference condition. In this discussion, I will focus on the time course of word integration during natural reading via inference making and consider differences in this process in relation to reading speed. Afterwards, I will discuss limitations of the current study and propose alterations for future work.

I predicted that readers would try to incorporate the target word into their mental representation immediately upon reading it. To do that, they would compute inferences drawing on antecedents available in the preceding text, making active use of the

previously available context. Since our target words were not presented as the first words of the new sentence, readers had to choose the relevant antecedent from information gathered both in the previous sentence and the one containing the target word. As antecedents became less readily available (from control to paraphrase to inference), readers should display a stepwise increase in their target word reading times, reflecting increased word integration processing costs while immediately reading it. These processing costs would mainly affect eye-movements measures reflecting later stages of text processing since these are associated with integration of incoming words into reader's text representation and subsequent comprehension. If readers did not show the increased processing costs across conditions, it would suggest that word integration either did not happen on-line immediately while reading, or that it happened via predictive inferences rather than bridging.

In line with my predictions, the results confirmed that skilled readers integrated incoming words in their text representation immediately while reading them and that the readers were sensitive to the difficulty of this process. Since reading times for the same target word increased as inference making demands rose, the data confirmed that skilled readers integrated the target words by making bridging inferences. That the integration was successful and correct inferences were drawn was further confirmed by the high accuracy with which comprehension questions were answered, since these required the necessary inference to be made. Thus, my data agree with the ERP results of Yang et al. (2007) regarding the time course of word-to-text integration and its means; readers integrate words immediately upon reading them by constructing bridging inferences. The actual pattern of our results, however, is different to Yang et al's. The control condition in my experiment displayed faster processing of the target word relative to both the paraphrase and inference condition. In contrast, target word processing for the

equivalents of the control and the paraphrase condition did not differ in Yang et al.'s work. Moreover, I did not find any evidence for the paraphrase effect reported in their ERP study. In fact, target word integration in the paraphrase condition seemed almost equally difficult as in the inference condition, since the difference between the two conditions was only marginally significant.

In terms of early processing, I observed shorter looking times for the control condition compared to the paraphrase and inference conditions. The first pass reading times are thought to reflect early comprehension processing, such as rapid integration in the situation model while the text contents are active in memory (Cook & Myers, 2004; Duffy & Keir, 2004; Traxler, Foss, Seeley, Kaup & Morris 2000). This finding thus shows that participants benefited from having an explicit direct referent by immediately integrating the target word through a rapidly generated bridging inference. Because initial reading times for the paraphrase and inference conditions were longer relative to the control, it appears that readers were sensitive to the increased word integration demands already during the initial processing. The fact that the paraphrase and the inference condition only started to show differences in later processing sits well with the dual-stage processing view of reading comprehension. This proposes that comprehension happens in two stages: 'bonding' – automatic integration of information with the text content active in memory reflected in the early processing measures (first pass times); and 'resolution' – later evaluation of information in the discourse context reflected in the later processing measures (Cook & Myers, 2004; Garrod & Terras, 2000; Pollatsek, Reichle, Rayner, 2006).

However, it is quite surprising that I saw the same pattern in the first fixation durations for to target word as well. This measure is considered to reflect aspects of word-level rather than discourse-level processing; for example, effects of lexical

ambiguity (Duffy, Morris & Rayner, 1988), word length (Rayner & McConkie, 1976) and frequency (Raney & Rayner, 1995) are seen in early measures. My results suggest that skilled readers are sensitive to the increased difficulty of word integration right when they encounter it. Otherwise, there would be no difference in first fixation durations between the control condition and the other two conditions. Shorter first fixations in the control condition relative to the other two seem to suggest that rapid comprehension processes can also take place when the word is fixated for the first time, apart from just word-level processing. Such comprehension processes seem to target integration of the word with the text contents immediately active in memory. As the explicit referent in the control condition is immediately active in reader's memory, word-to-text integration for the target word is very fast.

This explanation is based on assumption that first fixations on the target word are longer in the more demanding conditions (i.e., inference and paraphrase) than in the control condition. However, an alternative and more likely explanation for the pattern seen in early processing is that the fixations in the control condition are shorter than what they should be. This is due to the overlap in surface and lexical form between the target word and its referent. Since the target word and its referent share the surface and lexical form, having read the referent in the previous sentence might have induced a priming-like effect, leading to shorter first fixations in the control condition (Van Petten, Kutas, Kluender, Mitchiner, & McIsaac, 1991). The results would then reflect faster lexical access rather than actual comprehension processes. However, it is also possible that shorter first fixations reflected only purely perceptual feature overlap between the explicit referent and the target word in the control condition (DeSchepper & Treisman, 1996; Rugg & Nagy, 1987). I will return to this issue in the discussion of

future work where I propose ways to distinguish between the potential causes of this early effect.

Similar to the early processing measures, later processing measures showed almost the same pattern across conditions. The control condition showed shorter total reading times and go past times than both paraphrase and inference conditions. Readers also regressed back to the target word and out of it more often in the paraphrase and inference conditions relative to the control, although the difference was not significant. The findings together with the trend in regressions provide evidence that people took advantage of having supporting context when constructing bridging inferences to integrate the target word. The major advantage was provided by having an explicit, direct antecedent for the target word and evidence for contextual facilitation was weaker in the paraphrase condition.

It is important to note that the difference between the paraphrase and the inference condition was only marginal and apparent in the total reading times only. This suggests that, contrary to my initial predictions and Yang et al.'s results, people found the paraphrase and the inference conditions equally demanding. It seems that the additional contextual information presented in the paraphrase condition was not sufficient to make the target word integration easier and the generation of bridging inferences more immediate than in the inference condition. This is surprising given that the paraphrase condition provided a similar level of contextual information as the control condition; the main difference between them was the ease of mapping the target word to its antecedent in the first sentence.

Furthermore, the anaphor resolution literature suggested that when context is constraining enough, sentences such as the target word sentence in my experiment are

read equally fast in paraphrased context as in an explicit context (Garnham, 2001). Unfortunately, sentential reading times in the current experiment could not be examined since our data reflected additional sentence re-reading induced by the comprehension question. However, should this have been the case for the present study too, our results would suggest that context facilitation does not take place during the immediate processing of the target word. Thus, it would be interesting to compare sentence-level and word-level context facilitation in the future work.

To return back to the problem of the paraphrase condition not showing facilitated processing over the inference condition, it is possible that the paraphrase condition might have given rise to potential referent competition during sentence processing. The paraphrase condition might have created an additional processing load for the reader by asking for an extra inference in linking up two different referents for the same concept. To illustrate, take example (13) where a reader needs to make an additional inference that 'riding off' refers to the same action as 'pedalling off' in the previous sentence in order to comprehend the text successfully.

(13) After playing with his pony, Ben picked up his bike and pedalled off. He rode off at a very high speed to get home before his dad returned from work.

Linking up two different referents to the same concept is a costly process, as findings from the anaphoric resolution literature show (Duffy & Rayner, 1990; Garrod & Sanford, 1980). When an anaphor is preceded by an antecedent other than the typical one for its category, processing of the anaphor takes longer. For example, the word 'bird' is processed for longer if it is preceded by an antecedent such as 'goose' compared to when it is preceded by a more typical category example, such as 'robin'(Duffy & Rayner, 1990). Since the inferences readers needed to make in the

paraphrase condition were similar to these needed for anaphor resolution, it is possible that I have not used antecedents which were typical enough and thus created an extra processing load.

There is also a more simple explanation for the lack of a significant difference between the paraphrase and inference conditions and it is to do with methodological problems of this experiment. It is possible that the paraphrase condition provided only marginally stronger contextual support than the inference condition, rendering the two conditions too similar. If this was the case, it would explain why the general trend of shorter target word reading times in the paraphrase condition relative to the inference failed to reach significance for most later processing measures. Only total target word reading times were significantly shorter in the paraphrase condition compared to the inference, although the significance was marginal. The direction of the trend in total reading times suggested that people were sensitive to the amount of supporting information provided in the previous context during word integration. Perhaps with stronger set up in the paraphrase condition, the difference between the paraphrase and inference condition target word reading times would reach significance.

Upon further inspection, it became apparent that a target word synonym instead of a contextually-related-but-not-synonymous referent was used in 7 (26%) items in the paraphrase condition. The paraphrase condition explicitly aimed to use contextual referents which were not normally synonymous with the target word, yet clearly referred to it in the given context. The use of synonyms violated the paraphrase condition set-up and increased the stimuli variance in doing so. Moreover, synonymous referents are more vulnerable to the referent competition problem discussed earlier because they often present a less typical term for the target word concept. Results of this experiment suggested that readers showed little benefit from having contextual support

other than an explicit target word referent. However, skilled readers should actively use information they gather in text, and thus it is surprising that they would not take advantage from having a contextual referent for the target word. It is therefore important to address problems of paraphrase condition in future experiments in order to describe how people use different types of supporting context for drawing of bridging inferences.

To extend the findings of current experiment and confirm their validity, future work should thus focus on two aspects. First, it is important to distinguish whether the very fast processing of the target word in the control condition resulted from the ease of finding an antecedent for the target word, or lexical priming effects arising from word repetition. In order to shed light on this question, future work should focus on testing out these two different possibilities, perhaps by including an extra condition which would provide an equally rich context and easily available antecedent for the target word without lexical repetition.

Second, to decide whether paraphrase condition indeed brought no benefit over the inference condition pertaining to the construction of bridging inferences, paraphrase items need to be constructed more carefully, avoiding the use of synonymous referents for the target word. I proposed that referent competition in the paraphrase condition might have caused the lack of clear difference between the paraphrase and the inference condition that I saw in my data. Thus, it would be beneficial to present readers with a paraphrase condition where the contextual antecedent for the target word would be apparent, yet would not create any direct competition for the concept referred to by the target word. This was the main aim of Experiment 2.

Chapter 3: Rapid word-to-text integration: a result of lexical repetition or supporting context? (Experiment 2)

Chapter 2 demonstrated that participants benefitted from supporting context when constructing bridging inferences to integrate the target word, but only to a limited extent. When context provided an explicit and direct referent for the target word in the control condition, participants integrated it rapidly, as evidenced by the early processing measures. Since the referent used in this condition was lexically identical to the target word, it was not possible to tell whether early integration was a result of priming due to lexical repetition, rather than support from context. Furthermore, problems with stimuli construction were identified in the paraphrase condition of Experiment 1. Experiment 2 included an extra condition to the design of Experiment 1 to distinguish the effect of lexical repetition from the effect of supporting context on the target word integration. It also addressed stimuli problems in paraphrase condition. In contrast to Experiment 1, the results demonstrated no effect of supporting context or lexical repetition on target word integration.

Introduction

Experiment 1 explored skilled readers use of context to draw inferences during word-to-text integration. There was potential evidence that readers integrated the target words by drawing a bridging inference and using context, as evidenced by faster target word reading times in the control condition compared to the paraphrase and inference condition. However, participants used context to facilitate target word integration to a lesser degree than predicted since target word reading times in the paraphrase condition were only marginally different from the inference condition. An alternative explanation to my findings might be that the fast processing of target words in the control condition, potentially due to lexical repetition, was an inadequate baseline against which to compare performance in the other two conditions. Thus, I decided to follow up the results and explore reasons for early integration in the control condition and the lack of a clear difference in context facilitation in the paraphrase condition when compared to the inference condition.

In this introduction, I will first recap the basic principles behind the design of Experiment 1 and the results obtained. I will then discuss potential explanations for the early integration seen in the control condition; namely the lexical repetition effect and the effect of explicit context. Then, a new experimental manipulation will be proposed to distinguish which might be driving the early integration. Finally, predictions for both scenarios, lexical repetition and explicit context, will be spelled out.

A rich and coherent mental representation of text is crucial for good comprehension, and multiple studies have demonstrated that people form such representations when reading (Cain, Oakhill, & Bryant, 2004; Long, Oppy, & Seely, 1997; Rayner, Chace, & Slattery, 2006). Crucial for the creation of a rich and

connected text representation and subsequent comprehension is the ability to draw inferences (Cain, Oakhill, Barnes, & Bryant, 2001; Oakhill, 1982). By drawing inferences, readers not only incorporate explicitly stated information into their mental model, but also its implications and relation to world knowledge. For example, to create an accurate representation of (1) readers need to infer that the stain is on the new dress, it came from the spilled wine and the dress is now ruined, despite these features not being mentioned explicitly.

> *(1) Jessica had just put on a new dress when her sister accidentally spilled red wine on her. The stain was too big to do anything about.*

The question I wanted to answer is how readers use context to integrate words such as 'stain' when reading similar sentences. In particular, do people use context immediately to draw appropriate inferences when reading words which invoke them, such as 'stain', or only later after the words have been read?

Most relevant studies in the literature have used offline measures, such as the speed and accuracy of answering comprehension questions. Because these measures are taken when readers have finished reading, it is impossible to separate processing which happened immediately during reading from processing which happened after the word had been read. It is important to establish how and when readers integrate words into their text representation as reading happens in order to shed light on mechanisms of successful and efficient text processing.

Yang, Perfetti and Schmalhoffer (2006, 2007) carried out a series of ERP studies where they addressed a similar question. The advantage of ERPs over behavioural measures is that the temporal qualities of ERPs provide an online observation of reading

processes. Yang et al. presented readers with paragraphs which varied the ease of mapping a target word to its previous referent. They introduced four conditions:

a.) ***explicit***, where the referent and the target word matched in lexical and surface form *(e.g. The bomb exploded. The explosion was later reported*.)

b.) ***paraphrase***, where the referent was semantically related to the target word in a given context (e.g. *The bomb blew up. The explosion was later reported*.),

c.) ***inference***, where the referent is only implied by the context and needs to be constructed by the reader (e.g. *The bomb hit the ground. The explosion was later reported*.), and

d.) ***baseline***, where there is no previous referent for the target word (e.g. *The bomb was stored safely. The explosion was later reported.*).

They proposed that readers would try to integrate the word into their text representation immediately upon reading it by drawing a bridging inference back to the previously read sentence. When the previous sentence offered an easily available referent, the bridging inference should be computed rapidly. In contrast, when the previous referent is more difficult to obtain, computing the bridging inference will be costly, as reflected in larger N400 amplitude. Since the baseline condition provided no referent, it should produce the largest N400. The results largely followed the predictions, although the paraphrase condition showed equal N400 reduction as the explicit condition; the N400 reduction between the baseline condition and the inference condition failed to reach significance.

However, Yang et al., like many ERP experiments, used a single word-by-word presentation paradigm, so as to minimize signal disruption from ocular movements. This is unlike natural reading. For example, parafoveal processing is not feasible under

single-word presentation and potentially, reading word-by-word imposes a higher working memory load. To confirm that the generation of bridging inferences during word-to-text integration is the same in natural reading, Experiment 1 aimed to replicate Yang et al.'s results using eye-tracking methodology.

As in Yang et al., I manipulated the ease of mapping the target word to its referent in previous text. To do so, I varied the amount of contextual information given in the sentence preceding the sentence containing the target word. In the control condition, an explicit referent and the highest amount of contextual information was provided. In the paraphrase condition, semantically related expression was used as a referent and less contextual information was provided. In the inference condition, the least amount of contextual information was provided and there was no direct referent for the target word. In line with our predictions, readers spent the shortest total time reading the target word in the control condition, followed by the paraphrase and the inference condition. Somewhat contrary to the results reported by Yang et al., the difference between the paraphrase and the inference condition was small. There was a trend for readers to regress back into and out of the target word in the inference condition compared to the other two conditions, but this did not reach significance.

Target words in the control condition were fixated for the shortest amount of time, suggesting rapid integration with the previous text, compared to the other two conditions. Importantly however, reading times for the target word as tapped by the early processing measures were significantly shorter in the control condition. Early measures, especially first fixation times, are considered to reflect lower level aspects of reading such as word recognition and lexical access, rather than comprehension processes (Rayner, 1998). Thus, an interesting question arising from this result is what drives the apparent rapid integration of the target word in the control condition

compared to the other conditions. Two explanations could account for our results: first, the lexical repetition effect and second, the effect of having very explicit context.

Lexical repetition versus contextual facilitation

Let us consider the effect of lexical repetition first. When presented with a written word encountered previously, people tend to display faster processing of that word, so called repetition priming (Forster & Davis, 1984; Jacoby & Dallas, 1981; Scarborough, Cortese & Scarborough, 1977). This has been observed in probe-recognition, naming and lexical decision. Since many of the experiments which explored the effect of lexical repetition examined processing of individual words presented in lists, there is room for doubt whether the effects of lexical repetition extend to meaningful text. This is especially so given evidence that message level-factors during discourse processing can override lexical level factors. For example, words show semantic priming effect in discourse processing only if context supports the semantic relations, but not in the absence of the contextual support (Morris, 1994).

Evidence for lexical repetition effect in discourse is somewhat contrary. On the one hand, there are studies which show that lexical repetition produces faster reading times for words embedded in context (Andreson & Holcomb, 2005; Raney, Therriault, & Minkoff, 2000; Traxler, Foss, Seely, Kaup, & Morris, 2000). When Traxler et al. (2000) used primes that were identical to the target word during discourse processing, they saw the effect of lexical repetition manifested in early processing measures, namely the first fixations and first pass times. The effect only lasted while they used plausible contexts and vanished once the context was implausible, but such findings fit well with my findings from Experiment 1 since I used plausible contexts and also found an effect of lexical repetition in early measures.

On the other hand though, there is a problem with lexical repetition being the explanation for the fast target word processing in the control condition of Experiment 1. This concerns issues with co-reference. When lexical repetition involves coreferential expressions, processing costs in the form of the so called 'repeated name penalty' can manifest in word processing. Two linguistic expressions are said to be coreferential if they refer to the same semantic entity, just like an antecedent and its anaphor in anaphoric resolution. Thus, in (2), the words 'John' and 'he' are coreferential.

(2) John was very tired. He slept barely two hours last night.

In my control condition, the explicit referent and the target word are also coreferential . Coreference becomes problematic when the repeating referents are prominent in the discourse; for example, they are the subject of a sentence (Gordon & Hendrick, 1998).

To give an example of what is meant here: If we substituted the pronoun 'he' in (2) with 'John' , as in (3), reading time on the second John would grow longer (Gordon, Grosz, and Gilliom, 1993).

(3) John was very tired. John slept barely two hours last night.

If the prominent referent's name appears in the discourse shortly after being presented, readers' discourse model assumes that the second name establishes a new agent. If it does not, it creates a conflict in the model and resolution of such a conflict is costly. Based on this view, lexical repetition should not have produced faster initial target word processing in the control condition because many of our target words were prominent. Thus, it is possible that an alternative (or additional) process to lexical repetition contributed to my results.

Aside from lexical repetition, another potential explanation driving the apparent rapid integration of the target word in the control condition compared to the paraphrase and inference conditions is the availability of explicit context. Because the control condition made it extremely easy to map the target word onto its referent in the previous sentence (as they shared the same lexical form), integration of the target word into the text representation may have been completed very rapidly. This could partly be due to the fact that readers might have generated predictive inferences in the control condition since the context was highly constraining. There is good evidence that if the context is constraining enough, predictive inferences are generated (Graesser, Singer & Trabasso, 1994; Mak & Sanders, 2013). Thus, it is impossible for us to say whether fast processing of the target word in the control condition of Experiment 1 reflected rapid comprehension based on the generation of a predictive inference, or faster lexical processing due to repetition priming.

In order to distinguish which of the two accounts is responsible for the fast initial reading of the target word seen in the control condition, Experiment 2 includes a 'synonym' condition as an addition to Experiment 1. This condition is similar to the control condition, but the context-giving sentence includes a synonym as a referent for the target word, instead of the lexically identical referent. For example, instead of having 'accident' as both an explicit referent and the target word as in (4), the word 'collision' is the referent in the context-giving sentence in (5).

(4) Anne saw the car ahead of her but could not stop the accident from happening. The accident was later reported.

(5) Anne saw the car ahead of her but could not stop the collision from happening. The accident was later reported.

If lexical repetition drove the rapid processing of the target word in the control condition in Experiment 1, I should see shorter reading times in early measures in the control condition compared to the synonym condition in Experiment 2. However, if shorter reading times in the control condition arose as a consequence of constraining context that allowed readers to generate a predictive inference, reading times should be equivalent in the synonym condition and control condition. This is because the context is equally constraining in both conditions. Furthermore, the synonym condition also aimed to clarify the effect of the paraphrase condition. In Experiment 1, 26% of items in the paraphrase condition involved a synonym, making the data from this condition difficult to interpret. It was thus not possible to confidently state the true effect of having contextually related referents for the target word integration. By separating synonym and paraphrase item sets, Experiment 2 aimed to investigate this matter further.

Method

Participants:

Forty participants were recruited from the population of Oxford University undergraduate and graduate students (10 males and 30 females, mean age 21.3 years). All were native English speakers who first learned to read in English. Six participants were bilingual speakers with English being their dominant language. Participants had normal or corrected-to-normal vision and no history of language or reading difficulties.

Apparatus:

As in Experiment 1, an EyeLink 1000 remote eye-tracker with chin and head support was used to monitor the participants' eye movements as they read text presented on the computer screen with the resolution of 1024 x 768 pixels at a distance of 60 cm. The tracker measured gaze location from the right eye by monitoring the pupil and corneal reflection at a sampling rate of 1000 Hz. The calibration was set to a maximal deviation of 0.48˚ with 3.9 characters subtending to 1˚ of visual angle. The sentences were presented in three to four line paragraphs aligned to the left of the screen with a white Courier New 16-point font on a black background.

Materials:

Paragraphs used in Experiment 1 were adapted altered according to the new experimental paradigm to add a new 'synonym' condition and to ensure the paraphrase condition did not contain any synonyms. Thus, 14 paragraphs in the paraphrase condition were changed while the control and the inference conditions were identical to those conditions in Experiment 1. To increase power, 10 new paragraphs (based on stimuli from Yang, et al., 2007) were added to the stimulus set to make a total of 36

two-sentence paragraphs. Each participant saw nine paragraphs per condition. As in Experiment 1, the target word was crucial for the bridging inference to be made. Each paragraph was on average 34 words long and presented over three lines of text. See *Appendix C* for the full stimulus set.

I varied the amount of contextual information available for mapping of the target word to its referent, thus manipulating the difficulty of making a bridging inference (see Example 1). The control condition provided an explicit referent for the target word, matching its lexical and surface form; the paraphrase condition provided a referent equivalent in meaning to the explicit one for the given context; the inference condition did not provide a direct referent for the target word meaning that readers had to draw on their world knowledge and rich representation from the previous text to complete the inference. The new synonym condition was included to answer the question of whether rapid processing seen in the control condition in Experiment 1 arose from lexical repetition or contextual benefits. The synonym condition gave the same amount of contextual information as the control condition, but it replaced the explicit referent with an appropriate synonym. The synonym carried the same contextual information, yet was distinct from the target word in its lexical and surface form.

The paragraphs were presented in a counterbalanced design so that each participant read each experimental paragraph in one condition only and all versions of 36 paragraphs were presented equally across the participants. The paragraphs were presented in a random order.

In order to see whether the word-to-text integration happened successfully, each paragraph was followed by a forced-choice comprehension question. Participants were presented with two answers where the correct answer contained the correct target word

referent and the incorrect answer contained an incorrect referent from the context-giving sentence. This was to ensure that readers integrated the target word with the correct referent. Participants indicated their choice by pressing the corresponding button on the button controller they were holding in their hands. The paragraphs remained on the screen during answering the question.

Example 1 *shows the four different conditions of the same paragraph. The target word is in italics followed by the post-target region. Each paragraph was followed by a comprehension question with two answers to choose from. The position of the correct answer was varied across the experiment.*

Control condition

Cathy had a beautiful blue lamp on her table when suddenly an earthquake hit the city and the lamp broke. Tiny pieces of/ *broken*/ blue/ glass lay on the floor.

Paraphrase condition

Cathy had a beautiful blue lamp on her table when suddenly an earthquake hit the city and the lamp fell down. Tiny pieces of/ *broken*/ blue/ glass lay on the floor.

Inference condition

Cathy had a beautiful blue lamp on her table when suddenly an earthquake hit the city where she lived. Tiny pieces of/ *broken*/ blue/ glass lay on the floor.

Synonym condition

Cathy had a beautiful blue lamp on her table when suddenly an earthquake hit the city and the lamp smashed. Tiny pieces of/ *broken*/ blue/ glass lay on the floor.

Question: What happened to Cathy's lamp?

It fell and broke It lay on the floor

Procedure:

Participants sat 60 cm away from a computer monitor with their chin and forehead rested on the apparatus to constrain head movements. Participants were told that they would be reading short stories on the computer screen and their task was to answer the comprehension question afterwards. There was a 130 seconds limit on completing the reading and comprehension question on each paragraph. Participants were instructed to always read the story first and then read the comprehension question at the end. The eye-tracker was then calibrated to a value of 0.48 maximum in order to allow precise tracking of fixation.

Before the start of each trial, a white fixation point appeared at the centre of the screen. The participants were instructed to fixate on this point. Afterwards, a white square appeared at the left edge of the screen at the position which marked the location of the beginning of the first sentence in the narrative. The participants were instructed to fixate on the white square which triggered the text to appear on the screen. After the participants finished reading of the paragraph, they moved to the next trial by answering the comprehension question at the end using either the right or left trigger on a hand-held controller. Pressing a trigger finished the trial and erased the paragraph from the screen.

Results

Comprehension questions accuracy

The mean level of accuracy in answering the comprehension question was 95% (SD= 9%) and there were no differences across the conditions (*Fs* < 1). The accuracy rates for individual condition were: 95% (SD=7%) for the control condition, 94% (SD=10%) for the synonym condition, 93% (SD=8%) for the paraphrase condition and 96% (SD=10%) for the inference condition.

Eye-tracking data preparation

Following the standard procedure in reading time studies, fixations shorter than 80ms were combined with a nearby fixation. If there was no fixation within three letter space, the fixations were deleted. Fixations longer than 800 ms were also deleted. One participant was excluded from the analysis due to tracker loss. One experimental item out of 36 was excluded due to a mistake in the presentation script. Two regions of interest were marked out for the eye-tracking analysis; the target word region and the post-target word. The target region comprised the space before the target word and the target word itself. The post-target word region comprised of the word (or a word phrase where appropriate) following the target word.

Eye-tracking analysis of the target word region

Descriptive data are shown in Table 3.1. Overall, the reading times for the target word were very similar across conditions. Two one-way repeated measures ANOVA with four levels for condition were carried out to explore its effect on the eye-movement measures, one by-subjects and one by-items. There was no main effect of condition in

the measures of early processing, first fixations and first pass times. Unlike in Experiment 1, target words in the control condition were not processed any faster in comparison with other conditions. First fixation durations and first pass reading times for the control condition were marginally shorter than in the other three conditions, but the differences were so negligible they do not reflect any trend.

Table 3.1 *Main effect of condition for eye-movements of early and later processing in Experiment 2. Mean reading times (ms) on the target word and standard deviations (in parentheses) for all four conditions.*

	N=39	*M* / *SD*				
	Measure (ms)	**Control**	**Paraphrase**	**Inference**	**Synonym**	**F**
Early processing measures	First fixations	192	199	197	201	$F_1(3,117)= 1$
		(30)	*(32)*	*(34)*	*(34)*	$F_2(3,102)= 1,$
	First pass	236	241	241	240	$F_1(3,117) <1$
		(52)	*(44)*	*(57)*	*(49)*	$F_2(3,102) <1$
Later processing measures	Total time	296	328	328	315	$F_1(3,117)=1.5$, $p=0.2$
		(72)	*(120)*	*(92)*	*(75)*	$F_2(3,102)= 1.2$ $p=0.3$
	Second pass	287	261	279	263	$F_1(3,18) <1$
		(79)	*(87)*	*(89)*	*(79)*	$F_2(3,39)<1$
	Go-past time	294	330	317	313	$F_1(3,102)= 1.3$, $p>0.2$
		(81)	*(135)*	*(104)*	*(95)*	$F_2(3,102)= 1.4$, $p>0.2$
	Measure (probability of regressing)					
	Regressions in	0.18	0.19	0.19	0.17	$F_1(3,117) <1$, $p=0.9$
		(0.14)	*(0.15)*	*(0.16)*	*(0.13)*	$F_2(3,102)<1$, $p=0.8$
	Regressions out	0.19	0.20	0.21	0.19	$F_1(3,117) <1$
		(0.15)	*(0.17)*	*(0.19)*	*(0.16)*	$F_2(3,102) <1$

Table 3.1 also shows there was no main effect of condition for any of the late processing measures: total looking time, go-past time and second pass time. Equally, the number of regressions into and out of the target word region showed no main effect of condition type.

These analyses show that reading time on the target word was not influenced by the amount of supporting contextual information in the context-giving sentence, or the availability of a direct referent. It appears that participants either drew an inference to integrate the target word with equal ease in all conditions or that they did not draw an inference while they read the target word. To explore whether participants drew the inferences in a delayed manner, reading times on the post-target region were examined.

Eye-tracking analysis of the post-target word region

The integration of the target word may in certain situations, for example when presented in an improbable context, result in a spill-over of processing costs into the following word. Reading times in the post-target region were therefore analysed to explore whether participants processed the target word in a delayed manner. A series of one-way repeated measures ANOVAs with condition as a within subject factor was carried out by-subjects and by-items. Table 3.2 presents the summary of main effects for each dependent measure together with means and standard deviations.

Table 3.2 *Main effect of condition for eye-movements of early and later processing in the post-target region in Experiment 2. Mean reading times (ms) on the post-target word region and standard deviations (in parentheses) for all four conditions.*

	N=39	*M* / *SD*				
	Measure (ms)	**Control**	**Paraphrase**	**Inference**	**Synonym**	**F**
Early processing measures	First fixations	196	189	186	198	$F_1(3,114) = 2.1$, p=.1
		(33)	*(35)*	*(31)*	*(32)*	$F_2(3,102)<1$
	First pass	239	223	237	253	$F_1(3,114)= 2.5$, p=.06
		(61)	*(50)*	*(62)*	*(63)*	$F_2(3,102)= 1.2$, p=.3
Later processing measures	Total time	314	274	309	314	$F_1(3,114) = 3.7$, p< .05
		(90)	*(63)*	*(94)*	*(84)*	$F_2(3,102) = 1.8$, p= .15
	Second pass	318	238	246	279	$F_1(3,18) = 2.6$,p=.06
		(155)	*(104)*	*(91)*	*(151)*	$F_2(3,39)<1$
	Go-past time	329	302	359	317	$F_1(3,102) = 1.9$, p= .1
		(153)	*(133)*	*(177)*	*(107)*	$F_2(3,102) = 2.3$, p= .08
	Measure (probability of regressing)					
	Regressions in	0.17	0.16	0.14	0.19	$F_1(3,117) <1$,
		(0.13)	*(0.12)*	*(0.15)*	*(0.15)*	$F_2(3,102)<1$,
	Regressions out	0.17	0.17	0.20	0.13	$F_1(3,117) =1.5$, p=.2
		(0.17)	*(0.2)*	*(0.17)*	*(0.15)*	$F_2(3,102) <1$

In terms of the early processing measures, first fixations duration did not differ across conditions. First pass reading times on the post-target word region indicated a trend towards significance for the main effect of condition in the by subject analysis, but this did not generalise across items. As Figure 1 shows, this trend was driven by shorter target word reading times in the synonym condition. Since the by items analysis was not significant, it is hard to consider the trend reliable though.

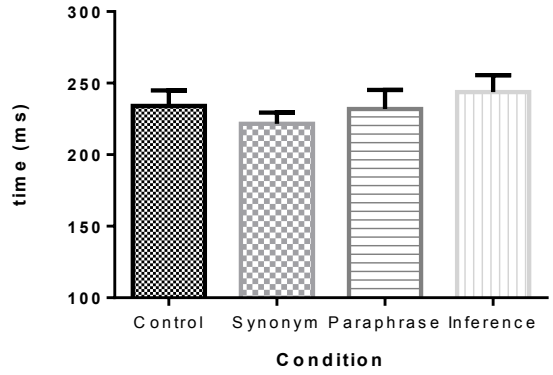


Figure 1 *Mean first pass reading times (ms) on the post-target word in all four conditions in Experiment 2. Error bars indicate SEM.*

The later processing measured displayed a similar picture. With the exception of total reading times, there were no main effects. Total reading times on the post-target region, shown in Figure 2, revealed a main effect of condition by subjects, and a trend by items. Bonferroni corrected post-hoc comparisons for the by subject analysis indicated that reading times on the post-target region in the synonym condition were shorter than in the other three conditions, which did not differ from each other. This finding is in line with the trend seen in the first pass times, but it does not follow the results from the target word region. It is thus difficult to explain what this finding means and it is probably spurious.

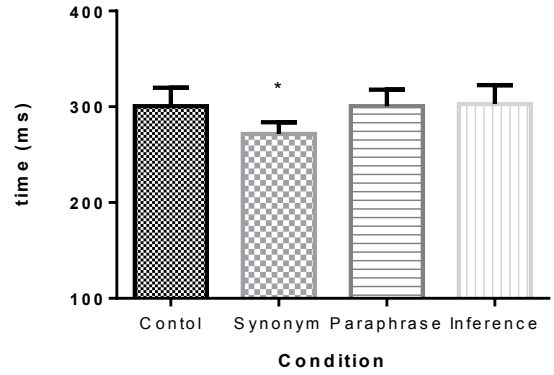


Figure 1 *Mean total reading times (ms) on the post-target word in all four conditions in Experiment 2. Synonym condition displays significantly shorter reading time than the other three conditions. Error bars indicate SEM. * p< .05*

To summarize, the amount of contextually supporting information available did not affect reading times on the target word, and the effect on the post-target word region was questionable. The values were similar across all four conditions in both regions, with the exception of synonym condition in the post-target region. Here, there was a trend for faster reading times but it reflected in one measure only and it was not reliable across items. Unlike Experiment 1, Experiment 2 presented no evidence for an effect akin to repetition priming in the control condition. Similarly, the synonym condition did not provide any benefit over the paraphrase condition with respect to the target word reading times. Overall, the findings suggest that target words were integrated with equivalent ease across all four conditions.

Discussion

One aim of Experiment 2 was to explore the influence of lexical repetition on the fast processing of target words in the control condition in Experiment 1. Furthermore, as the paraphrase condition in Experiment 1 often contained a referent that was synonymous with the target word rather than contextually related, Experiment 2 set to clarify the difference between these two types of referents. Thus, an extra condition was included: the synonym condition in which a synonym of the target word served as its referent in the context-giving sentence.

As in Experiment 1, I predicted that target word reading times would increase when less contextual information was provided to support the drawing of a bridging inference. Thus, I expected longest target word reading times in the inference condition, followed by the paraphrase condition and finally the synonym and the control conditions. I predicted that if lexical repetition caused the rapid early processing seen in the control condition in Experiment 1, reading times on the target word in the control condition in Experiment 2 would be shorter compared to the other conditions. If, on the other hand, the rapid integration arose due to the highly constraining context and a subsequent generation of a predictive inference, I expected the control condition and synonym condition to provide equally fast reading times. Lastly, I predicted that the synonym condition would elicit shorter looking times than the paraphrase condition since it provided more explicit contextual support for the bridging inference.

Clearly, my predictions for Experiment 2 were not supported since there was no main effect of condition on the target word reading times. Quite the contrary, I found no difference between the conditions. This finding suggests that the availability of more contextual information brought no advantage to participants as they read the target

word. Consequently, it is not possible to answer the question of what caused the fast processing of the target word in the control condition in Experiment 1. The control condition and the synonym condition displayed very similar target word reading times which would normally suggest that the lexical repetition did not cause the rapid early processing of the target word seen in Experiment 1. Since no effect of condition on target word reading times was found though, such a conclusion cannot be made.

Why should the results of Experiment 2 differ so radically from those seen in Experiment 1? There are four possible explanations for the lack of a difference between conditions in Experiment 2. First, it is possible that our participants did not draw inferences immediately when they read the target word or the post-target word region and instead made them off-line, once they had finished reading. Second, participants drew the inferences to integrate the target word when they read it, but did not use previous context to facilitate the generation of bridging inferences. Third, although participants use context to aid generation of bridging inferences during reading, my stimuli were too easy to show contextual benefits. Thus, I observed an equivalent of a ceiling effect. Fourth, participants did not draw bridging inferences to integrate the target word in their representation but instead drew predictive inferences.

Before I discuss these four potential accounts further, it is important to consider the striking differences in the pattern of findings between Experiment 1 and 2. Given their similarity, it is surprising that the main findings did not replicate across the two experiments. It is therefore important to carry out a detailed comparison between Experiment 1 and Experiment 2 to identify aspects in which the experiments differed from each other. It then needs to be examined whether the differences might have accounted for the different pattern of findings across the two experiments. Apart from comparing Experiment 2 to Experiment 1, it is important to consider whether specific

participants or a particular combination of items could have produced the current pattern of findings by chance. Thus, a thorough post-hoc investigation is needed, before I can conclude that the findings of Experiment 2 are reliable and worthy of further discussion. The next section presents additional analyses and comparisons between Experiments 1 and 2 in an attempt to establish the reliability of their findings.

Results of explorative post-hoc analyses of Experiment 2 and comparisons with Experiment 1

Exploration of experiment specific differences between Experiment 1 and 2

I compared target word reading times between Experiment 1 and Experiment 2 to find out whether there are any experiment specific differences in reading pattern. To do so, I compared the performance on control, paraphrase and inference conditions across both experiments. I selected the 25 items which were used in both experiments. It must be noted though that the paraphrase condition was changed considerably in Experiment 2, containing only 15 of the original 25 stimuli. However, the control and the inference conditions were identical and so performance on these conditions should not differ across the two experiments. Descriptive data are summarised in Table 3.3.

A two-way mixed analysis of variance with experiment (Experiment 1 and Experiment 2) as a between subjects factor and condition as a within subject factor was carried out.

Table 3.3 *Mean reading times (ms) on the target word and standard deviations (in parentheses) for data from both Experiment 1 and Experiment 2, including only those items used in both experiments.*

	Experiment 1 (N=34)	**M *(SD)***			**Experiment 2 (N=40)**	**M *(SD)***	
	Measure (ms)	**Control**	**Paraphrase**	**Inference**	**Control**	**Paraphrase**	**Inference**
Early processing measures	First fixations	190	210	207	192	199	198
		(29)	*(35)*	*(39)*	*(26)*	*(29)*	*(21)*
	First pass	237	255	262	234	240	239
		(55)	*(46)*	*(64)*	*(56)*	*(39)*	*(54)*
Later processing measures	Total time	304	357	400	293	320	321
		(83)	*(84)*	*(137)*	*(101)*	*(129)*	*(102)*
	Second pass	323	345	447	278	256	283
		(140)	*(119)*	*(299)*	*(109)*	*(113)*	*(99)*
	Go past time	301	337	358	288	327	315
		(94)	*(121)*	*(125)*	*(73)*	*(114)*	*(98)*
	Measure (probability of regressing)						
	Regressions in	0.15	0.18	0.21	0.18	0.19	0.19
		(0.14)	*(0.15)*	*(0.14)*	*(0.18)*	*(0.18)*	*(0.18)*
	Regressions	0.14	0.18	0.22	0.18	0.20	0.21
		(0.16)	*(0.18)*	*(0.21)*	*(0.12)*	*(0.16)*	*(0.17)*

In terms of early target word processing, there was no main effect of experiment for either first fixation times ($F(1, 73) <1$) or first pass target word reading times ($F(1, 73) =2.25$, $p=0.1$). Similarly, there was no interaction between experiment and condition in the first fixation durations and the first pass times (all $Fs < 1$). In terms of initial processing, it appears that participants in Experiment 1 and Experiment 2 were no different in their reading performance. With respect to the main effect of condition, early processing measures reflected the same pattern seen in the individual analysis of Experiment 1: first fixation durations ($F(2, 146)=5.1$, $p<0.05$, $\eta^2= 0.07$) with the control condition eliciting significantly shorter reading times on the target word than paraphrase and inference condition. Equally, first pass reading times also demonstrated significantly shorter looking times ($F(2,144)=3.7$, $p<0.05$, $\eta^2= 0.04$) in the control condition than paraphrase and inference conditions.

Interestingly, the measures of later processing showed experiment-specific differences. There was a significant main effect of experiment (F (1, 73)= 5.32, p<0.05, η^2=0.07) qualified by a significant interaction between experiment and condition (F (2,144)=4, p<0.05, η^2=0.05) in total target word reading times. The participants in Experiment 1 took more time to read the target words (M=353, SD= 101) than those in Experiment 2 (M=311, SD=110) (Figure 3). The interaction arose as reading times in the inference condition differed significantly between Experiment 1 and 2. Post hoc comparisons revealed that participants in Experiment 1 (M=400, SD=137) took significantly longer time to read the target word in this condition than those in the identical condition in Experiment 2 (M=321, SD=102).

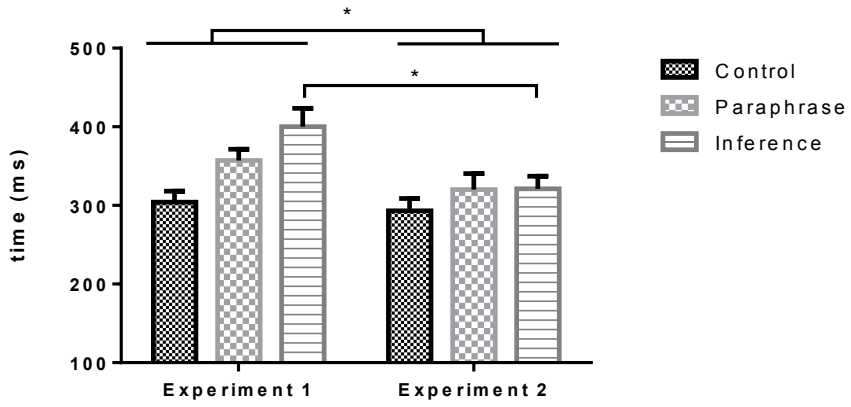


Figure 3 Mean total reading times (ms) on the target word in Experiment 1 and Experiment 2 for items used in both experiments. Participants in Experiment 2 took less time to read the target words in the inference condition than participants in Experiment 1. Error bars indicate SEM.

A number of later processing measures also showed a significant main effect of condition: total target word reading time (F (2,146)=11, p<0.01, η^2=0.13), go-past reading time(F (2, 146)=3.9, p<0.05, η^2=0.53) and second pass reading times (F (2,106)=3.1, p<0.05, η^2=0.05) all showed a main effect of condition. Paired comparisons for total reading time on the target word demonstrated that the control

condition (M=297, SE=18) elicited shorter overall reading times than both paraphrase (M=345, SE=25) and inference (M=362, SE=26) conditions. There was no strong trend towards significance for the difference between the paraphrase and the inference conditions, unlike in the Experiment 1. Paired comparisons for go-past reading times demonstrated that participants read the target word and preceding text for longer before going past it in the paraphrase (M=308, SE=17) and inference (M=313, SE= 16) conditions compared to the control condition (M =268, SE=13).

Investigation of participant differences in Experiment 2

The significant interaction between experiment and condition revealed that specific experiment factors might be the cause of discrepant results across the two experiments. The comparison of the two experiments indicated that there may have been important differences between participant samples used in each experiment. Participants in Experiment 2 processed the target words faster in terms of both early and later processing and displayed significantly faster processing of the target word in the inference condition than participants in Experiment 1. Thus, I decided to scrutinize participant sample used in Experiment 2 further.

First, I removed all outlier values above 2.5 standard deviations from the mean to make sure that results in Experiment 2 are not driven by anomalous participants or items. However, only six trials were excluded out of 1384 based on the 2.5 standard deviation difference criterion and their removal did not alter the main findings at all.

I next asked whether the two experiments differed in terms of population sample: it is possible that some participants in Experiment 2 displayed an unusual reading pattern which distorted the overall trend in the data. The comparison of participant samples revealed that there were no bilingual participants in Experiment 1,

yet six out of 40 participants in Experiment 2 were bilingual. It is possible that these participants have processed target words in an altered manner due to interference from their second language. I thus removed the bilingual participants from the Experiment 2 sample and analysed the data set again. However, the pattern of results remained unaltered.

Investigation of stimuli differences in Experiment 2

Recall that the item set for Experiment 2 was extended by 10 new items and two items were altered. This may have added variance to the data due to their potentially different nature. To see whether data from Experiment 2 would replicate the findings from Experiment 1 when only items used in both experiments were analysed, the 10 new items and two altered items were excluded from the next analysis on Experiment 2 data.

The reading times for target word in the original three conditions (control, paraphrase and inference) were analysed (the synonym condition was excluded). One-way repeated measures ANOVA with condition as a within subject factor was carried out on the remaining 25 items used in both experiments. See Table 3.4 for the summary of main effects, means and standard deviations.

Table 3.4 *Mean reading times (ms) on the target word and standard deviations (in parentheses) for data in Experiment 2, including only those items used in both Experiment 1 and Experiment 2.*

	N=40	M / SD			
	Measure (ms)	**Control**	**Paraphrase**	**Inference**	**F**
Early processing measures	First fixations	195	203	198	$F_1(2,78)<1$,
		(33)	*(35)*	*(38)*	$F_2(2,50)<1$
	First pass	235	249	248	$F_1(2,78)<1$
		(55)	*(51)*	*(64)*	$F_2(2,50)=1.1$, $p=.3$
Later processing measures	Total time	298	349	334	$F_1(2,78)=3.2$, $p=.05$
		(79)	*(141)*	*(102)*	$F_2(2,50)=2.9$, $p=.06$
	Second pass time	280	329	306	$F_1(2,48)<1$
		(94)	*(213)*	*(133)*	$F_2(2,26)<1$
	Go past time	290	341	335	$F_1(2,78)=2$, $p=.1$
		(95)	*(178)*	*(126)*	$F_2(2,50)=2.7$, $p=.07$
	Measure (probability of regressing)				
	Regressions in	0.19	0.23	0.18	$F_1(2,78)=1.2$, $p=.3$
		(0.17)	*(0.18)*	*(0.15)*	$F_2(2,50)<1$
	Regressions out	0.17	0.19	0.22	$F_1(2,78)=1.1$, $p=.3$
		(0.19)	*(0.22)*	*(0.21)*	$F_2(2,50)=1.1$, $p=.3$

The analysis of the original items yielded mixed results and did not fully replicate the pattern seen in the Experiment 1. The early processing measures showed no main effect of condition, unlike Experiment 1 which found an effect in both first fixations and first pass times. The later processing measures revealed a marginally significant effect of condition in total reading times (Figure 4) and a trend towards significance for go-past reading times. This, again, was different to Experiment 1 where later processing measures revealed a significant difference between the control condition and the other two conditions, and a trend towards significance between the paraphrase and inference conditions. Paired comparisons for total times in Experiment 2 at least showed a trend for the control condition being read faster than the paraphrase

and inference condition, but there was no hint of difference between the paraphrase and inference conditions.

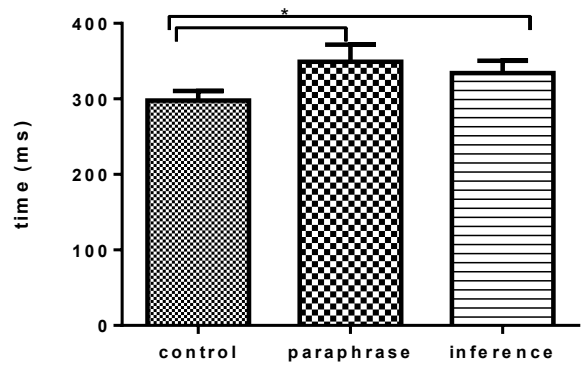


Figure 4 *Mean total reading times (ms) on the target word in Experiment 2 for items used in both Experiment 1 and 2. Error bars indicate the standard error of the mean (SEM). *p< .05*

Apart from re-analysing the Experiment 2 stimuli quantitatively, they were also examined qualitatively. Three independent raters with English as their first language were asked to screen the experimental paragraphs in all conditions for any anomalies. The results indicated that five of the synonyms used as target word referents were perceived as unnatural word choice for the given sentence. This finding though does not explain the different pattern of results in Experiment 2.

To conclude, the additional analyses unfortunately did not reveal a clear explanation for the failure of Experiment 2 to replicate the main effect of condition seen in Experiment 1. People were faster overall in Experiment 2, and there was a significant interaction between Experiment and condition in overall target word reading times, caused by slowed reading times in the inference condition of Experiment 1 relative to Experiment 2. The implications of the analyses described in this section together with main findings from Experiment 2 will now be discussed.

General Discussion

After observing faster early processing of the target word in the control condition compared to the paraphrase and inference condition in Experiment 1, I decided to explore this effect further. To distinguish whether faster first fixation durations and first pass times resulted from repetition priming or from having strongly supporting context, I included a synonym condition in the design of Experiment 1. This condition used a synonym of the target word as its direct referent in the context-giving sentence, instead of the lexically identical referent used in the control condition. This way, the synonym condition provided an equally strong context as the control condition but without the repetition priming problems associated with the explicit referent. Moreover, twelve items in Experiment 1 used a synonym referent instead of a contextually related referent in the paraphrase condition. Thus, the synonym condition aimed to clarify the effect of having a synonymous referent versus contextually related referent on the target word integration. However, the results of Experiment 2 failed to replicate the main effect of supporting context on the ease of drawing a bridging inference when reading the target word. Thus, the two questions Experiment 2 sought to address remained unanswered.

To identify potential causes for the failure to replicate results of Experiment 1 in Experiment 2, a set of exploratory post hoc analyses was carried out. The analyses directly compared results from Experiment 1 and 2, and further explored the particularities of participant sample and stimuli set used in Experiment 2. Even though the post-hoc analyses did not identify clear reasons for the lack of difference between conditions in Experiment 2, they indicated that findings from at least one of the experiments were probably spurious. This is because performance on identical items differed significantly between Experiment 1 and 2, a surprising finding given the

participant samples were drawn from the same population of Oxford University students. The results of post-hoc analyses will now be discussed in further detail followed by the discussion of null findings from Experiment 2 and their implications for how skilled readers use context to draw bridging inferences. This section will finish with the discussion of methodological problems that need to be addressed in future work.

Direct comparison with Experiment 1

To investigate whether there were any experiment-specific differences in reading pattern, data from Experiment 1 and 2 were directly compared, with experiment as a between subject factor. Only items and conditions shared between Experiment 1 and 2 were included in the analysis. This reduced the number of items in the paraphrase condition since most paraphrase condition items were changed in Experiment 2, but the control condition and the inference conditions remained unchanged. This comparative analysis yielded two interesting findings which indicated that experiment-specific factors influenced my findings. It is possible that findings from one of the two experiments were spurious and did not reflect the true effect of supporting context on the drawing of bridging inferences during processing on-line.

First, participants in Experiment 2 showed significantly faster later target word processing compared to the participants in Experiment 1. It thus appears that participants in Experiment 2 found target word integration easier than participants in Experiment 1 since they spent less time on any additional target word processing. From this point of view, it seems that participants in Experiment 2 might have been in some respects more skilled readers than participants in Experiment 1.

Second, the difference between participants in Experiment 2 and 1 was especially pronounced for the inference condition. A significant interaction between experiment and condition factors revealed that participants in Experiment 1 took significantly longer to integrate the target word in the inference condition than participants in Experiment 2, despite the two groups reading identical sentences. There were no obvious differences between the two groups in terms of outliers but one possible factor might be that Experiment 1 and 2 differed in terms of the proportion of bilingual speakers included. There were no bilingual speakers among participants of Experiment 1, but six out of forty participants in Experiment 2 were bilingual. People who are bilingual display slightly altered language processing because the languages interact and influence each other (Marian, Spivey, & Hirsch, 2003; Meuter & Allport, 1999). Since second language might have influenced drawing of bridging inferences in bilingual participants, they were excluded and the data were re-analysed. However, the results remained the same and did not reveal a main effect of condition or even any trend towards one. Thus, it appeared that participant differences were probably not at the root of replication problems in Experiment 2.

Since participant differences did not explain why Experiment 2 failed to replicate the results of Experiment 1, the stimuli used in Experiment 2 were scrutinized to find out whether a particular alteration to the stimuli might have changed the pattern of results. First, it was possible that the ten new items added to the previous stimuli set were somehow anomalous and responsible for creating a different pattern of results than observed in Experiment 1. To test this out, the 10 new items were excluded and the 25 items used in both Experiment 1 and 2 were re-analysed on their own, excluding the synonym condition and looking only at the three original conditions. However, the pattern of results did not change and no main effect of condition emerged in the data.

Thus, the addition of the 10 new items and inclusion of an extra condition was probably not responsible for differences seen between Experiment 1 and 2.

Implications of the lack of main effect in Experiment 2

Given that post-hoc exploratory analyses did not offer a clear explanation for the different pattern of findings in Experiment 2, and since Experiment 2 was more carefully designed than Experiment 1, let us assume that participants integrated the target word equally easily in all four conditions. That is, the results of Experiment 1 are spurious and the null effects of Experiment 2 stand. One of the aims of Experiment 2 was to show whether the rapid target word integration seen in the control condition of Experiment 1 resulted from repetition priming or from the ease of mapping the target word to the referent in the context-giving sentence. Since no differences in target word reading times were found between conditions in Experiment 2, it was impossible to answer this question. However, the lack of difference between target word reading times opened a related yet bigger question than the one Experiment 2 sought to address originally – did readers draw inferences to integrate new concepts immediately during reading in the first place?

The findings of Experiment 2 or rather the lack of thereof, offer four possible interpretations. First, it is possible that our participants did not draw inferences immediately when they read the target words or in the immediate the post-target region. Since participants answered the comprehension questions with accuracy rate of over 95% in Experiment 2, I know that the inferences were drawn at some point. It is questionable though whether this happened during reading of the target word or later, perhaps after the comprehension question had been read. Based on the substantial body of evidence suggesting that readers try to integrate new words in their mental text

representation immediately as they read them (Just & Carpenter, 1980; Rayner, Warren, Juhasz, & Liversedge, 2004; Robin, 1994; Yang et al., 2007), it seems unlikely that inferences were not drawn at least partially when the target word was read.

If we accept that participants probably drew the necessary inferences while reading, the lack of any difference between the conditions in Experiment 2 indicates two possible explanations. These are as follows: the nature of the previous context did not serve to make inference generation easier or harder, or the stimuli used in the experiment were too easy to reveal processing differences between the four conditions. Since a body of work suggests that readers use context during word-to-text integration as noted in the previous paragraph, it is not unreasonable to assume that there ought to be differences between the conditions and that my stimuli were too easy to reveal an effect of context on the generation of bridging inferences. It is possible that the current stimuli might have displayed a version of a ceiling effect where inferences were equally easy to compute in all conditions. This argument brings us to another potential interpretation of data from Experiment 2: if the computation of inferences was equally easy in all conditions, readers might have drawn predictive inferences rather than bridging inferences.

Is it possible that the materials in all conditions were easy enough for the participants to make a sufficiently strong expectation of what would happen next in the text, thus generating a predictive inference? Many studies have demonstrated that if a context supports the activation of predictive inferences, such inferences would become available (McKoon & Ratcliffe, 1992; Morris, 1994; Peracchi & O'Brien, 2004). For example, Cook, Limber and O'Brien (2001) found that naming times for the word 'dead' were significantly faster if the probe word was preceded by highly predictive context such as in (6), compared to the control context in (7).

(6) The building was fifteen stories high. The scaffolding he was using was a little wobbly and Mark lost his balance and fell.

(7) The building was only one story high. The scaffolding he was using was a little wobbly and Mark lost his balance and fell.

Since all our conditions were designed to provide some kind of context for target word integration, it is possible that even the less-contextually supporting conditions (paraphrase and inference) provided enough context for the readers to generate a predictive inference. On this view, even the inference condition gave a rather clear contextual set-up for the target word, possibly facilitating target word integration to a degree, comparable to the more explicit conditions that contained either a synonym or lexical repetition. Unfortunately, this speculation cannot be verified until reading times for the target words in neutral contexts are obtained and compared with reading times in the four conditions. If the target word reading times in the inference condition were faster than those in a neutral context condition, this would provide further support that readers generated predictive inferences in Experiment 2.

The stimuli I used in Experiment 1 and 2 thus cannot provide a definite answer whether participants drew inferences during target word integration and whether they used previous context in doing so. It appears the stimuli have been too easy to process, producing an equivalent of a ceiling effect. This would mean that even the less explicit context conditions might have provided sufficient information for equally fast target word integration as the more contextually explicit conditions. Thus, it is important to increase stimulus difficulty in the next experiment. Besides that, qualitative comparison of Experiment 1 and 2 revealed important methodological problems in the stimuli set which I will now review. Thus, to tell which experiment reflected the true effect of

supporting context on inference generation during word-to-text integration, it is important to first address these methodological problems.

First, there was considerable variation in the number of intervening words between the explicit referent and the target word in the control condition. This is problematic since any effect of repetition priming would be stronger for items where the two words occur next to each other and weaker in the conditions where there are words separating them (Branigan, Pickering, & Cleland, 1999; Ratcliff, Hockley, & McKoon, 1985). In the current stimulus set, there were between zero and eight words separating the explicit referent and the target word, introducing extra variance in the data and making it more difficult to determine the strength of the lexical repetition effect.

Second, 12 out of the 36 items still contained a synonym in the paraphrase condition, making the two conditions less distinctive and thus causing a potential confound. Third, five of the synonyms used in Experiment 2 were marked an 'unnatural word choice' as indicated by three independent raters during the post-hoc examination of Experiment 2 stimulus set. Thus, reading times for the synonym condition might have reflected the strangeness of the target word for the context, rather than the effect of manipulation.

Third, the comprehension questions presented were also problematic since they varied in nature: some asked about the inference that had to be computed for the target word to be integrated whereas others did not. Thus, if readers regressed back to the text after reading the question to review relevant parts again, it might have changed the pattern of reading data observed. The same questions were used across conditions but as different information was provided in the context-giving sentence, the comprehension questions might have provoked different re-reading pattern across conditions.

Furthermore, the experiment lacked filler items and it is thus possible that participants noticed the main manipulation, potentially giving rise to strategic effects. The next important step is therefore to address these problems in an experiment.

Conclusion

The goal of creating the synonym condition in Experiment 2 was to determine whether the rapid processing in the control condition seen in the Experiment 1 was due to lexical repetition or a consequence of it having provided explicit contextual support. Additionally, the synonym condition aimed to further explore the effect of having a contextually-related referent in the paraphrase condition rather than a referent synonymous with the target word. Stimuli used in Experiment 1 (as well as those used by Yang et al., 2006, 2007) suffered from the problem of using synonymous referents in the paraphrase condition. By establishing a separate synonym condition, I wanted to clarify the effect of having a referent which is not strongly lexically related to the target word as a synonym is, but it instead provides a contextual link with the target word.

However, the results failed to confirm the predictions or even replicate the main effect of condition seen in Experiment 1. Despite thoroughly comparing the stimuli used in both experiments and scrutinizing the participant samples for any relevant differences, it is not clear why the two experiments generated a different pattern of results. However, important methodological issues were identified in both experiments. The stimuli varied considerably in how well they adhered to the rules for each condition; for example how many words should be intervening between the explicit referent and the target word in control condition; it also seems likely that my stimuli were easy for skilled readers to process, thus masking potential between-condition

differences. Both experiments suffered from lacking statistical power. Because of this, their results were vulnerable to chance fluctuations influencing what is and what is not statistically significant. Thus more stimuli and more participants should be used in the next experiment, and the stimuli need to be carefully designed to minimize variance across individual conditions.

Experiment 3 aims to address the methodological problems identified in Experiment 1 and 2 by increasing stimulus difficulty and tightening the rules applied to the design of each experimental condition. Problems regarding the nature of the comprehension question presentation and lack of filler items will also be addressed, and the number of items and participants will be increased to improve statistical power. Thus, Experiment 3 will hopefully provide a clear answer to the question of whether supporting context influences inference generation and word-to-text integration.

Chapter 4: Addressing methodological issues of previous work and exploring the effect of supporting context on instantiation of inferred information (Experiment 3)

Experiment 2 did not replicate the results of Experiment 1 and showed no effect of supporting context on construction of bridging inferences during target word integration. Chapter 3 concluded by comparing performance on identical items used in Experiment 1 and Experiment 2 and found significant differences in target word processing between experiments, suggesting that the results of one experiment were spurious. A number of methodological problems were highlighted though none identified as the source of discrepant results. Experiment 3 addressed these methodological problems to clarify the effect of supporting context on the construction of bridging inferences. A distance manipulation to separate the context-giving sentence from the target word sentence was included to increase the processing difficulty of the stimuli. Furthermore, an off-line memory task was added to explore whether readers were more likely to instantiate inferred information if presented with supporting context. The on-line eye-movement record in Experiment 3 again revealed no benefit of supporting context on drawing of bridging inferences, despite methodological problems having been resolved. Contrastingly, results of the off-line instantiation task indicated that participants were more likely to instantiate inferred information in conditions which provided more contextual support relative to those with less contextual-support.

Introduction

Experiment 1 and 2 explored how highly skilled readers drew inferences on-line and used context to support this process. In this introduction, I will briefly discuss the theoretical background for the current study followed by the findings from Experiment 1 and Experiment 2. I will then address the methodological shortcomings that the two studies suffered from and their implications for the design of Experiment 3. Inference generation in relation to the instantiation paradigm will also be considered and a new task will be introduced to examine their relationship. The introduction will conclude with the predictions for the new eye-tracking study as well as the instantiation task.

An ability to draw inferences efficiently and rapidly is crucial for comprehension as it allows readers to form a rich and connected text representation that incorporates information gained from the text with world knowledge (Cain, Oakhill, Barnes & Bryant, 2001; Cain & Oakhill, 1999; Graesser, Singer & Trabasso, 1994). It is also essential for word-to-text integration: in order to integrate the word with the rest of the text, readers need to decide how a word is related to the information they gained previously and this decision often demands an inference to be made. For example, to integrate the word 'rain' with the previous text in the following sentence (1)

(1) *Heavy grey clouds dominated the sky for days and people prayed for the rain to stop as the rivers rose*

readers need to draw an inference from their world knowledge that clouds are the source of rain, and that the heavy grey clouds mentioned earlier in the text are the source of the word 'rain' they just read.

Thus, it is known that inference-making is an essential process for proficient reading. What still needs to be explored is how exactly this process happens and

whether readers draw inferences to integrate individual words into their text representation immediately while reading them or after they had read the words. In a series of experiments, Yang, Perfetti and Schmalhoffer (2006, 2007, 2008) demonstrated that skilled readers integrated words into their textual representation immediately as they read, actively using preceding context to facilitate the integration process. They recorded ERPs on critical words in sentences where the critical word (for example 'explosion') matched the referent in the preceding sentence either explicitly, as in (2,) contextually, as in (3), or when the direct referent was missing but the critical word directly followed from the previous sentence as in (4) compared against a baseline control as in (5).

(2) Explicit condition: *The bomb hit the ground and exploded. The explosion was quickly reported.*

(3) Paraphrase condition: *The bomb hit the ground and blew up. The explosion was quickly reported.*

(4) Inference condition: *The bomb hit the ground. The explosion was quickly reported.*

(5) Baseline condition: *The bomb was stored safely on the ground. The explosion was quickly reported.*

Yang et al. found that integration of the critical word, was more difficult in the absence of a direct referent in the preceding context. In the sentences such as (2) and (3), critical words were integrated with ease, as evidenced by reduced N400 amplitude. This was especially the case for sentences like (2) where an explicit direct referent was available. However, integration became more difficult with the absence of a direct referent. Even though critical words produced smaller N400 amplitude in sentences such as (4) where the critical word followed the scenario set up by the preceding context

than in (5), where it did not follow, the difference was not significant. Yang et al. interpreted their results as evidence that readers completed their word-to-text integration through drawing of bridging inferences. They reasoned that had readers used predictive inferences to integrate the critical words, similar N400 reductions should have been observed in all three conditions where the target word followed from the preceding context because the target concept would have been pre-activated. Differential reduction in the N400 amplitude thus suggested that readers made bridging inferences instead to integrate the critical words.

It is important to note that Yang et al.'s experiments used single word-by-word presentation; arguably, this might place different demands on word-text integration processes than occur under natural reading. I thus set off to replicate their results using a more naturalistic methodology: measuring readers' eye movements as they read passages of text. Like Yang et al., I hypothesized that readers would draw bridging inferences when they read critical words and that they would actively use previous context to facilitate their inference making. As referents for the target words became less available in the preceding context, I expected a stepwise increase in readers' target word reading times. I omitted the baseline condition where the target word did not follow the previous context at all as in the example (5) because the resulting texts were rather anomalous and thus it was problematic to determine whether integration difficulties observed in this condition resulted from the lack of available referent, or simply from the processing costs imposed by anomalous contexts (Rayner, Warren, Juhasz & Liversedge, 2004).

Thus, I created three conditions analogous to the explicit, paraphrase and inference conditions form Yang et al.'s studies to investigate whether skilled readers take context into account immediately while integrating a word during natural reading

in my Experiment 1. I hypothesized that readers would spend increasingly longer time processing the target word as the amount of contextual information decreases. The prolonged reading times would reflect the increasing difficulty of drawing a bridging inference when they were integrating the target word into their text representation. Since word-to-text integration draws on context and utilizes higher level reading skills, I expected processing effects to become more evident in later processing, as tapped by additional re-reading in the eye movement record (Rayner, 1998). In line with my predictions, Experiment 1 showed a stepwise increase in reading time as the inferential demands rose, although the difference between the paraphrase and the inference condition was marginal. Somewhat unexpectedly, I also found that the explicit condition elicited an early processing effect seen in the first fixations and the first pass reading times already. It was not clear whether this effect arose due to quicker integration in the explicit condition or slowed down integration in the paraphrase and inference conditions.

It was possible that shorter first fixation durations and shorter first pass times in the explicit condition resulted from lexical priming effects brought about by the shared lexical and surface form between the explicit referent and the target word, such as ‘exploded’ and ‘explosion’ in example (2). Furthermore, methodological problems were noted which might have influenced the pattern of the findings. The paraphrase condition used synonyms in 26% of the cases and thus could produce facilitation due to using lexically and semantically closely associated terms rather than via providing rich contextual relationship between the referent and the target word. To further explore whether rapid processing seen in the control condition arose as a result of lexical priming and to establish the effect of contextual relationship in the paraphrase condition, Experiment 2 included a ‘synonym’ condition alongside the original

conditions. Additionally, ten extra items were added to the stimulus set to increase statistical power and two faulty items which did not require readers to draw inferences when integrating the target word were replaced with new items.

The synonym condition replaced the explicit referent used in the control condition with a synonym. I predicted that if the rapid processing seen in the control condition of Experiment 1 was driven by easier word-to-text integration, there would be no difference between the control and the synonym condition because they provided the same content. However, if short initial reading times in the control condition reflected lexical priming, the control condition would elicit faster reading times than the synonym condition in Experiment 2. The synonym condition also aimed to further explore whether integration in the paraphrase condition benefitted from lexical-semantic associations or contextual relationships. If people are sensitive to the lexical-semantic associations during message-level processing, the synonym condition should have elicited faster reading times than the paraphrase condition. If however integration relies on a clear contextual relationship between the referent and the target word (like in (3) where 'blow up' is contextually-related to 'explosion'), reading times for the two conditions should have been similar.

However, Experiment 2 failed to replicate effects previously seen in the Experiment 1 and yielded no main effect of condition, making it impossible to verify my predictions. This was rather puzzling since Experiment 2 presented readers with better controlled stimuli than Experiment 1. Moreover, when performance on identical items was compared across the two experiments, reading times differed significantly for the inference condition implicating that results of one group did not reflect the true population mean reading times. Since Experiment 2 was methodologically more rigorous than Experiment 1, there are grounds to consider the results of Experiment 1

spurious. If correct, this challenges Yang et al.'s conclusions that readers draw bridging inferences while reading to integrate target words into their text representation, using previous context to actively support this process. Experiment 3 was designed to explore this issue further and to address methodological problems uncovered by the detailed examination of Experiment 1 and 2.

The comparison of Experiment 1 and 2 revealed a number of methodological problems with the stimulus construction. First, the number of words between the explicit referent and the target word in the control condition varied from zero to eight. This is a problem since processing of the target word becomes more difficult the further away it is from its referent and this lack of uniformity thus introduced extra variation to the data. The lexical priming effect would be especially vulnerable to this variation because items with zero intervening words would show it strongly, but items with eight intervening words might not show it anymore (Ratcliff, Hockley & McKoon, 1985). Second, the paraphrase condition still presented synonyms on 12 occasions despite efforts to avoid this when constructing Experiment 2. This is problematic since it creates a confound between the paraphrase and the synonym condition, making it impossible to test out the additive effect of lexical-semantic association between the target word and its synonymous referent against the effect of contextual relationship in the paraphrase condition.

Third, five of the synonyms used in the Experiment 2 were marked as unnatural word choices by three independent raters in the follow up examinations of Experiment 2. Since unnatural words produce higher reading times (Rayner et al, 2004) it is possible that reading times on the synonym condition in the Experiment 2 might have been inflated. Fourth, an inference could be drawn from the context-giving sentence alone in nine items, thus violating the basic principle behind the inference condition. Finally,

although comprehension questions occurred on every trial, they did not always ask about the inference needed to integrate the target word. This is a problem because the questions were presented on the same screen as the stimuli and we know that readers read text differently based on their reading goal (Greasser, Singer & Trabasso, 1994; Van Der Schoot, Vasbinder, Horsley& Van Lieshout, 2008). Thus, if the comprehension question asked readers to make a different inference than the one needed for target word integration, any re-reading induced by the question might have affected the reading patterns on the target word. However, this additional processing would not necessarily reflect integration of the target word as such.

Besides problems with the construction of the stimuli, it is possible that the sentences were too easy, resulting in the equivalent of a ceiling effect masking true differences between conditions. Readers might have been sensitive to the amount of contextual information given when integrating a word, yet because the materials were easy, inferences were generated equally well even in the inference condition which did not provide a direct referent for the target word. In Experiment 3, I thus made alterations to increase the difficulty of the stimuli. Texts which require readers to store more information in their working memory are more difficult to process and I therefore increased the working memory load my stimuli imposed during reading (Daneman & Carpenter, 1980; King & Just, 1991). This was achieved by inserting additional text between the context-giving sentence and the target word sentence, as in (6). The target word 'birth' was presented in the final sentence, but the contextual referents necessary for the target word's integration were presented in the first sentence. This presentation was more demanding than in previous experiments where the target word sentence directly followed the context-giving sentence, as in (7).

(6) Mary was having severe stomach cramps and so she went to the hospital to deliver. Many friends and relatives came to see her. They brought lots of flowers and cards. Mary's whole family celebrated the birth of her first baby.

(7) Mary was having severe stomach cramps and so she went to the hospital to deliver. Mary's whole family celebrated the birth of her first baby.

Such a manipulation required readers to hold information from the context-giving sentence in their working memory for longer, making it less readily available when they arrived at the target word as a result. Because the referents for the target word were not available immediately as in cases where the target word sentence followed the context-giving sentence, the process of drawing a bridging inference should be more difficult. Moreover, readers also needed to decide that the information gained from the intervening texts was not relevant for drawing of the bridging inference, thus adding further processing load because they had to suppress it. In Experiment 3, half of the stimuli were presented in the long distance condition with two sentences separating the context-giving sentence and the target word as in the example (6), and the other half in the short distance condition without intervening sentences, as in Experiments 1 and 2.

Apart from inducing a higher memory load, the insertion of additional text offered potential to explore the following questions:

a.) do readers integrate words equally efficiently if the referents are further apart compared to when they are close?

b.) does integration across intervening text happen immediately or later?

We know from previous research that readers integrate incoming concepts with information from previous content because they form inferences based on them even when these pieces of information are separated by several sentences, or even paragraphs (Cook, Lassonde, Splinter, Guéraud, Stiegler-Balfour, & O'Brien, 2014). However, it has been shown that integration slows as the distance between the critical pieces of information increases. For example, during anaphoric processing, additional text between the anaphor and its antecedent results in longer reading times on the anaphor, as long as the anaphor and antecedent are semantically related (Duffy & Rayner, 1990). That is, target word reading times increased for anaphor-antecedent pairings like 'sword – weapon' which are semantically linked, but not for pairings such as 'sword – object'. In Experiment 3, the target word referents provided in the context-giving sentence were always semantically-related to the target words. Thus, linking of the target words to their referents should be sensitive to separation by extra text, producing longer reading times on the target word reading times in the long distance condition.

The results of Experiment 1 and Experiment 2 were also unable to address whether readers utilized the supporting contextual information immediately during reading of the target word, and whether this ability to quickly use supporting context led to the rapid integration seen in the explicit condition of Experiment 1. To finally answer these questions, I decided to use the same paradigm in the Experiment 3 again and address all methodological issues from Experiment 1 and 2. I created entirely new stimulus set of forty experimental items that strongly adhered to the rules created for each condition based on what I learnt from examination of the previous experiments. The rules are further described in the Materials section later in this chapter. In addition, comprehension questions tapped the inferences required by the target words and were presented on a separate screen to prevent potential alterations of reading pattern. To

control for any potential strategic effects, I included twenty filler items to disguise the main manipulation.

Apart from increasing stimulus difficulty, I added an extra task to the design. Previous work has shown that skilled readers tend to infer particular meanings of words according to the context in which they appeared via a process of instantiation (Anderson, Pichert, Goetz, Schallert, Stevens, & Trollip, 1976; Oakhill, 1983). This is when information strongly associated with the context of a word is stored in the mental representation of a text, even if not explicitly stated in the text. Since words can have more specific meanings in context than in isolation, people often instantiate a particular meaning representation for the word in their mental model (Anderson et al., 1976) For example, readers might think of a shark when they are reading about a fish in (8) but not in (9).

(8) The fish attacked the swimmer.

(9) The fish avoided the swimmer.

Indeed, when Anderson et al. (1976) presented participants with a recall test after they had read sentences such as (8) and (9), they were more likely to falsely recall reading about sharks when asked about (8), but not (9). Thus, participants assigned a contextually appropriate meaning to the word 'fish' in (8), even though it was not mentioned in the text. The purpose of the extra task in Experiment 3 was to explore whether skilled readers instantiated information they had to infer in the paraphrase and inference conditions. Since skilled readers should form rich text representations by incorporating inferences in their mental models, they should be more likely to instantiate inferred information in their memory, despite not reading it directly in the text (Cook, Lassonde, Splinter, Guéraud, Stiegler-Balfour & O'Brien, 2014; Oakhill,

1983). Previous work showed that after reading sentences such as (10), skilled readers were more likely to falsely recall having seen the word 'igloo' in the text than their less skilled counterparts (Oakhill, 1983).

(10) A man built a house made of ice.

Since my participants are proficient readers, I wanted to establish that they drew the inferences and instantiated them in their memory. Thus, I created a recall test that was administered after the paragraph reading task. This presented participants with information that should have been inferred in all but the explicit condition, and asked them to indicate whether the information had appeared in the earlier reading task or not. I predicted that if participants actively used contextually supporting information, conditions which were more contextually constraining would show higher false recall rate; the synonym and the paraphrase condition should therefore display higher false recall rate than the inference condition.

To summarise the predictions for Experiment 3, I expect to see the same reading pattern as in Experiment 1 if readers integrated target words immediately upon reading them by generating bridging inferences. Reading times on the target word should be longer as the amount of contextual information decreases and mapping of the target word with the previous referent becomes more difficult. This would produce a stepwise pattern in reading times with target words in the control condition being read the fastest and those in the inference condition taking the longest. Furthermore, if early processing of the target word in the control condition resulted from lexical priming, target word reading times in control condition will be shorter than in the synonym condition. Because processing of intervening text creates extra working memory demands, target words separated from their context-giving sentences would show longer reading times

and less facilitation from contextually supporting information. In terms of performance on the instantiation memory task, I predicted that participants will be more likely to falsely recall inferred information in conditions which provided more contextual support, because their context was more constraining. By addressing methodological issues, introducing a distance manipulation and an instantiation memory test, Experiment 3 aimed to clarify the confusing picture presented by Experiments 1 and 2 where basic effects did not replicate across the two experiments.

Method

Participants:

Forty-eight participants were recruited from the population of Oxford undergraduate and graduate students (16 males and 32 females, mean age 20 years). All participants were native speakers of English speakers who also first learned to read in English. Three participants were bilingual speakers with English being their dominant language. Participants had normal or corrected-to-normal vision and no reported history of language or reading difficulties.

Apparatus:

As in the previous experiments, an EyeLink 1000 remote eye-tracker with chin and head support was used to monitor the participants' eye movements as they read text presented on the computer screen with the resolution of 1024 x 768 pixels at a distance of 60 cm. The tracker measured gaze location from the right eye by monitoring the pupil and corneal reflection at a sampling rate of 1000 Hz. The calibration was set to a

maximal deviation of 0.4˚ with 3.9 characters subtending to 1˚ of visual angle. The sentences were presented in three-to-four line paragraphs aligned to the left of the screen with a white Courier New 16-point font on a black background.

Materials:

(i) Reading task

Participants read 60 four-sentence paragraphs; 40 experimental items and 20 filler items. The filler items were included to address potential strategic effects and disguise the main manipulation. To maintain participants' attention and focus, twenty comprehension questions were presented randomly throughout the experiment.

Bridging inference difficulty was manipulated by varying the amount of contextual information provided for an inference to be made. The experiment comprised the same four conditions as used in Experiment 2 (control, paraphrase, inference and synonym). The control condition provided an explicit referent which matched the target word in its lexical and surface form. The synonym condition presented identical content as the control condition, but used a synonym as the target word referent instead of the explicit referent. The paraphrase condition provided a contextually related expression as a target word referent. The inference condition did not provide any direct referent for the target word. The amount of supporting context was thus modified across conditions to vary the ease of target word integration into readers' text representation.

Special care was taken to keep the materials within each condition as consistent as possible and following alterations were made to specifically address methodological problems identified in Experiments 1 and 2. A distance of at least three words

separating the explicit referent and the target word was introduced in the control condition. The synonym condition only used synonyms which were natural in the given sentence settings. The paraphrase condition never used synonyms. The inference condition was created so that both sentences were always necessary for the inference to be made.

To check that the target words were equally predictable across conditions, cloze probabilities for each condition were acquired for the stimuli. Forty participants (who did not participate in the main experiment) were presented with the experimental stimuli in a counterbalanced order so that each participant saw each paragraph in one condition only. The target word was replaced with a blank and participants were asked to fill in a fitting word. The probability of participants filling in a target word is shown in Table 4.1. This was compared across conditions using a one way repeated measures ANOVA which revealed that target word cloze probabilities were similar across all conditions, ($F(3,117)=1.8$, $p=.14$, $\eta^2=.04$)

Table 4.1 *Mean target word cloze probabilities for each of the four conditions, standard deviations in parentheses*

		M		
N=40		*(SD)*		
Measure	**Control**	**Synonym**	**Paraphrase**	**Inference**
Cloze probability	0.35 *(0.23)*	0.45 *(0.31)*	0.42 *(0.31)*	0.41 *(0.32)*

To explore whether increased memory demands influenced processing, a distance manipulation was introduced as a between-subjects factor. A sample trial is shown in Example 1. In the short distance condition, the target word sentence immediately followed the context-giving sentence as in Experiments 1 and 2. Two filler

sentences, together an average of 18 words long, followed the target sentence and contained no information relevant for the inference. In the long distance condition, the two filler sentences were positioned between the context-giving sentence and the target word sentence. Thus, participants in this condition experienced a higher processing load as they needed to hold the information from the first sentence in memory for longer and process irrelevant intervening text before they needed to draw the inference to integrate the target word.

Example 1 *shows the distance manipulation used to create higher processing load. Filler sentences which were used to create the distance manipulation are in italics, the target word is displayed in bold.*

Short distance condition:

Mary was having severe stomach cramps and so she went to the hospital in a hurry. Mary's whole family celebrated the **birth** of her first baby. *Many friends and relatives came to see her. They brought her a lot of flowers and cards.*

Long distance condition:

Mary was having severe stomach cramps and so she went to the hospital in a hurry. *Many friends and relatives came to see her. They brought her a lot of flowers and cards.* Mary's whole family celebrated the **birth** of her first baby.

(ii) Memory task

The memory task was administered after the eye-tracking procedure had been completed to assess the likelihood of falsely recalling information which was not explicitly stated in the eye-tracking paragraphs, but readers needed to infer it and

include it in their text representation. The memory task aimed to measure how strongly readers instantiated inferred information in their text model. Participants were presented with 60 sentences and were asked after completing each sentence whether or not the sentence appeared in the texts they had read during the paragraph reading phase of the experiment. From the total of 60 sentences, 40 sentences related to the 40 experimental items in the eye-tracking part, 10 sentences related to the filler items from the eye-tracking experiment and 10 sentences were unrelated to paragraphs presented during the eye-tracking task. The presentation of the sentences was randomized.

The forty sentences related to the experimental paragraphs presented in the memory task were the context-giving sentences from the experimental paragraphs in the control condition. These sentences were chosen because they explicitly stated the information which was inferred in the other three conditions. Each participant only saw 10 paragraphs in the control condition which meant that only 10 of the 40 sentences were related to the experimental items that had actually appeared. The other thirty sentences in the memory task were slightly different because participants had seen their synonym, paraphrase or inference condition version during the paragraph reading phase of the experiment.

Procedure

Forty-eight participants were assigned to either a short or a long distance condition, making a total of 24 participants in each condition.

(i) Reading task

The experimental paragraphs were presented in a counterbalanced design so that each participant only read one version of each experimental paragraph. Ten paragraphs

were presented in each of the four conditions. The paragraphs were presented in the random order. Twenty comprehension questions were presented randomly throughout the experiment to maintain participants' attention. The questions were presented on separate screens and participants responded by pressing a button on a hand-held controller.

Participants sat 60 cm away from a computer monitor with their chin and forehead rested on a chinrest to constrain head movements. They were told that they would be reading short stories on the computer screen and answering comprehension questions about some of the stories. After they finished the stories, they would complete a memory test about what they had read.

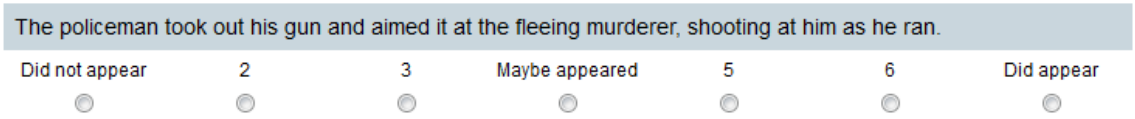
Participants were instructed to read the stories as they normally do and refrain from skim reading. They were asked to restrict blinking if possible and to try reading with their eyes wide open to allow for better eye-movement recording. The eye-tracker was then calibrated to a value of 0.4 maximum in order to allow precise tracking even for words only three characters long.

Before the start of each trial, a white fixation point appeared at the centre of the screen. The participants were instructed to fixate on this point. After fixating the point, a white square appeared at the left edge of the screen at the position which marked the location of the beginning of the first sentence in the narrative. The participants were instructed to fixate on the white square until the text appeared on the screen and press a trigger button on a hand-held controller after they finished reading to move on to the next paragraph.

(ii) Memory task

Following the eye-tracking session, participants were presented with 60 sentences and asked to rate whether each sentence with the exact wording appeared in the stories they had read earlier, and how certain they were about their knowledge. They responded with clicking a corresponding value on a 5-point Likert scale, as shown in the Example 2. The memory test was presented on a laptop screen and created with the on-line survey research platform Qualtrics. All sentences were displayed on a single screen and participants scrolled down through them.

Example 2 *A sample sentence from the memory task administered after participants finished the eye-tracking part. Participants were asked to indicate whether the sentence appeared in the eye-tracking part of the experiment.*



Results

Comprehension accuracy

On average, participants achieved 94.5% (SD= 8%) accuracy of answering the comprehension questions. Individual conditions displayed similar levels of accuracy: the control condition had an average 96% (SD = 8%) accuracy, the synonym condition scored 93% (SD= 6%) correct answers, the paraphrase condition 97% (SD= 9%) and the inference condition showed 92% (SD = 8%) of correct answers.

Eye-tracking data preparation

Following the standard procedure in reading eye-tracking studies, fixations shorter than 80ms were combined with a nearby fixation. If there was no fixation within three letter space, the fixations were deleted. Fixations longer than 800 ms were also deleted. Data from one participant from the 'short' condition were excluded due to tracker loss. No other participants or trials were excluded.

As in the previous experiments, the regions of interest are the target word with the preceding space shown in italics and the post-target word following immediately after the target word (see Example 2). The post-target word comprised a word or a word phrase if the following word was less than four characters long, and the space before the word.

Example 2 *shows an experimental paragraph in the 'Short distance' condition marking out the regions of interest, the target word region is in italics, the post-target word region follows immediately afterwards.*

Short distance condition:

Mary was having severe stomach cramps and so she went to the hospital in a hurry. Mary's whole family celebrated the/ *birth*/ of her/ first baby. Many friends and relatives came to see her. They brought her a lot of flowers and cards.

Eye-tracking analysis

Eye-tracking analysis of the target word region

Data were analysed both by participants and by items using a mixed analysis of variance to explore the effect of condition type and distance on reading. The analysis treated condition (four levels) as a repeated measure and distance (two levels) as a

between-subjects factor. For the full summary of means and standard deviations in the four conditions, see Table 4.2 for the Short distance group and Table 4.3 for the Long distance group. Figures 1a-b show mean initial reading measures for the target word and figures 2a-c depict mean values for later reading measures on the target word. Table 4.4 presents an overview of main effects for the repeated factor (condition) and the between-subject factor (distance) in the target word region.

Table 4.2 *Mean reading times (ms) on the target word and standard deviations (in parentheses) in all four conditions for participants in the Short distance group*

		M			
	N=24		*(SD)*		
	Measure (ms)	**Control**	**Synonym**	**Paraphrase**	**Inference**
Early processing measures	First fixations	182	201	189	203
		(27)	*(42)*	*(33)*	*(38)*
	First pass	218	235	216	233
		(57)	*(58)*	*(39)*	*(55)*
Later processing measures	Total time	273	284	267	282
		(84)	*(73)*	*(76)*	*(71)*
	Second pass	253	249	234	246
		(137)	*(121)*	*(128)*	*(101)*
	Go past time	246	244	314	258
		(81)	*(86)*	*(255)*	*(101)*
	Measure (probability of regressing)				
	Regressions in	0.17	0.14	0.11	0.11
		(0.16)	*(0.12)*	*(0.09)*	*(0.12)*
	Regressions	0.24	0.19	0.19	0.18
		(0.17)	*(0.14)*	*(0.15)*	*(0.16)*

Table 4.3 *Mean reading times (ms) on the target word and standard deviations (in parentheses) in all four conditions for participants in the Long distance group*

	N=24	*M* (*SD*)			
	Measure (ms)	**Control**	**Synonym**	**Paraphrase**	**Inference**
Early processing measures	First fixations	199	201	198	198
		(34)	*(38)*	*(40)*	*(31)*
	First pass	223	231	222	223
		(47)	*(41)*	*(46)*	*(46)*
Later processing measures	Total time	306	276	276	314
		(64)	*(55)*	*(47)*	*(78)*
	Second pass	311	237	260	261
		(151)	*(96)*	*(126)*	*(131)*
	Go past time	310	305	305	365
		(112)	*(119)*	*(90)*	*(211)*
	Measure (probability of regressing)				
	Regressions in	0.17	0.16	0.18	0.22
		(0.14)	*(0.12)*	*(0.14)*	*(0.16)*
	Regressions	0.17	0.17	0.22	0.24
		(0.12)	*(0.14)*	*(0.21)*	*(0.15)*

Table 4.4 *Main effects for the repeated factor (condition type), the between-subjects factor (distance) and their interaction in the target word region by subjects (F1) and by items (F2).*

	N=48			
	Measure	**Condition**	**Distance*Condition**	**Distance**
Early processing measures	First fixations	$F_1(3,138)=2.38$, p=.07	$F_1(3,138)=2.02$, p=.1	$F_1(1,46)<1$
		$F_2(3,231)=1.21$, p=.3	$F_2(3,231)=1.5$, p=.2	$F_2(1,77)<1$
	First pass	$F_1(3,138)=1.55$, p=.2	$F_1(3,138)<1$	$F_1(1,46)<1$
		$F_2(3,231)<1$	$F_2(3,231)=1.92$, p=.1	$F_2(1,77)<1$
Later processing measures	Total time	$F_1(3,138)=2.1$, p=.1	$F_1(3,138)=1.59$, p=.2	$F_1(1,46)=1.25$, p=.2
		$F_2(3,231)=1.01$, p=.4	$F_2(3,231)=1.58$, p=.3	$F_2(1,77)<1$
	Go past time	$F_1(3,138)<1$	$F_1(3,138)=1.46$, p=.2	$F_1(1,46)=5.02$, p<.05
		$F_2(3,231)<1$	$F_2(3,231)=1.57$, p=.2	$F_2(1,77)=4.46$, p<.05
	Second pass	$F_1(3,93)<1$	$F_1(3,93)=1.38$, p=.3	$F_1(1,31)<1$
		$F_2(3,75)<1$	$F_2(3,75)=3.01$, p=.05	$F_2(1,25)<1$
	Measure (probability of regressing)			
	Regressions in	$F_1(3,138)<1$	$F_1(3,138)=2.26$, p=.08	$F_1(1,46)=3.87$, p=.05
		$F_2(3,231)<1$	$F_2(3,231)=1.94$, p=.1	$F_2(1,77)=3.05$, p=.08
	Regressions out	$F_1(3,138)<1$	$F_1(3,138)=2.53$, p=.06	$F_1(1,46)<1$
		$F_2(3,231)<1$	$F_2(3,231)=1.5$, p=.2	$F_2(1,77)<1$

To briefly summarise the results, there was no significant main effect of condition type on looking times. A general lack of interaction effects and an inconsistent pattern for the main effect of distance suggests that introducing extra text between the context-giving sentence and the target word sentence did not alter processing difficulty in the current population.

In terms of early processing, there was no main effect of condition for either measure of early processing; first fixations or first pass times. Equally, there was no main effect of distance and no interaction between distance and condition. It thus appeared that initial word integration did not benefit from having an explicit referent in the context-giving sentence. Intervening text introduced in the long distance condition seemed to have no effect on target reading time as reflected by the early processing measures.

Measures of later processing showed the same pattern as first fixation durations and first pass times. There was no main effect of condition type for any of the later processing measures. There was also no reliable pattern of interaction between the condition type and distance in our data, although regressions into the target word, regressions out of the target word and second pass times did show a trend towards a significant interaction. The trends followed no clear pattern though and manifested only by subjects or by item analysis, but never in both (see Table 4).

There was no main effect of distance suggesting that introducing text to separate the context-giving sentence from the target word sentence did not affect target word reading times. The only exceptions were the go-past times and regressions back to the target word. Participants in the Long distance group spent more time reading the regions preceding the target word and the target word itself than participants in the Short

distance group. However, the fact that this difference did not project into target word total reading times which were similar for both groups makes this finding less reliable and meaningful. Similarly, the analysis of regressive eye-movements back to the target word showed that participants in the Long distance group re-read the word more often (M=0.18, SD=0.14) than those in the Short distance group (M=0.13, SD=0.12). Just like with the go-past times though, additional re-reading of the target word did not affect the overall processing time and it is thus difficult to interpret. Results from the go-past reading times and additional re-reading suggests that it was more difficult to process the target word when it was separated from the context-giving sentence by distance, but it appears that participants quickly overcame these difficulties and their overall word integration remained unaffected.

In summary, findings from analyses of the target word region suggest that readers did not draw inferences during reading of the target word. Alternatively, the current manipulation may not have been suitable to study bridging inferences in word-to-text integration through eye-movement records. Before either conclusion can be drawn, reading times in the post-target word region should be examined. Target word processing costs may spill over into processing of the immediately following word if the word integration becomes delayed. I therefore analysed reading times for the post-target word region.

Eye-tracking analysis of the post-target word region

Just as with the target word region, data were analysed by subjects and by items with a mixed analysis of variance. Condition type was treated as a within subject factor (four levels) and distance as a between subjects factor (two levels). For the full summary of means and standard deviations in the four conditions see Table 4.5 for the Short distance group and Table 4.6 for the Long distance group. Table 4.7 presents a summary of main effects by subjects and by items for the within subject factor condition, the between subject factor distance and their interaction. Figure 3 shows general trends initial reading of the post-target region and Figure 4 depicts mean overall reading times for this region.

Table 4.5 *Mean reading times (ms) on the post-target word and standard deviations (in parentheses) in all four conditions for participants in the Short distance group in Experiment 3.*

		M			
	N=24		*(SD)*		
	Measure (ms)	**Control**	**Synonym**	**Paraphrase**	**Inference**
Early processing measures	First fixations	191	184	194	200
		(25)	*(34)*	*(35)*	*(31)*
	First pass	223	218	232	213
		(46)	*(79)*	*(59)*	*(39)*
Later processing measures	Total time	251	248	276	241
		(62)	*(103)*	*(95)*	*(55)*
	Go past time	193	198	204	163
		(146)	*(103)*	*(115)*	*(55)*
	Second pass	179	223	229	192
		(45)	*(67)*	*(83)*	*(60)*
	Measure (probability of regressing)				
	Regressions in	0.20	0.11	0.19	0.22
		(0.19)	*(0.14)*	*(0.12)*	*(0.16)*
	Regressions	0.15	0.14	0.11	0.12
		(0.17)	*(0.13)*	*(0.11)*	*(0.14)*

Table 4.6 *Mean reading times (ms) on the post-target word and standard deviations (in parentheses) in all four conditions for participants in the Long distance group in Experiment 3.*

	N=24	*M* (*SD*)			
	Measure (ms)	**Control**	**Synonym**	**Paraphrase**	**Inference**
Early processing measures	First fixations	207	197	197	201
		(40)	*(36)*	*(44)*	*(50)*
	First pass	235	230	237	245
		(64)	*(59)*	*(61)*	*(78)*
Later processing measures	Total time	293	301	295	326
		(82)	*(96)*	*(79)*	*(91)*
	Go past time	367	376	310	436
		(269)	*(234)*	*(139)*	*(223)*
	Second pass	278	281	301	274
		(142)	*(134)*	*(126)*	*(96)*
	Measure (probability of regressing)				
	Regressions in	0.25	0.28	0.24	0.23
		(0.15)	*(0.1)*	*(0.2)*	*(0.2)*
	Regressions	0.13	0.15	0.17	0.25
		(0.16)	*(0.18)*	*(0.18)*	*(0.18)*

Table 4.7 *Main effects for the repeated factor (condition type), the between subject factor (distance) and their interaction in the post-target word region in Experiment 3.*

	N=48			
	Measure	**Condition**	**Distance*Condition**	**Distance**
Early processing measures	First fixations	$F_1(3,138)=1.37$, p=0.25	$F_1(3,138)<1$	$F_1(1,46)<1$
		$F_2(3,207)<1$	$F_2(3,207)<1$	$F_2(1,69)=2.36$, p=0.1
	First pass	$F_1(3,138)<1$	$F_1(3,138)<1$	$F_1(1,46)=1.2$, p=0.27
		$F_2(3,207)<1$	$F_2(3,207)<1$	$F_2(1,69)=3.19$, p=0.07
Later processing measures	Total time	$F_1(3,138)<1$	$F_1(3,138)=2.38$, p=0.07	$F_1(1,46)=6.7$, p<0.05
		$F_2(3,228)=1.47$, p=0.3	$F_2(3,228)<1$	$F_2(1,76)=9.1$, p<0.01
	Go past time	$F_1(3,138)<1$	$F_1(3,138)=1.98$, p=0.1	$F_1(1,46)=38.8$, p<0.01
		$F_2(3,222)<1$	$F_2(3,222)=1.31$, p=0.2	$F_2(1,74)=15.2$,p <0.01
	Second pass time	$F_1(3,87)<1$	$F_1(3,87)<1$	$F_1(1,29)=11.3$, p<0.01
		$F_2(3,60)=2.57$, p=0.06	$F_2(3,60)<1$	$F_2(1,20)=1.04$, p=0.3
	Measure (probability of regressing)			
	Regressions in	$F_1(3,138)<1$	$F_1(3,138)=2.87$, p<0.05	$F_1(1,46)=5$, p<0.05
		$F_2(3,228)=1.09$, p=0.3	$F_2(3,228)<1$	$F_2(1,76)=3.17$, p=0.08
	Regressions out	$F_1(3,138)<1$	$F_1(3,138)=1.95$, p=0.1	$F_1(1,46)=3.27$, p=0.07
		$F_2(3,207)<1$	$F_2(3,207)<1$	$F_2(1,76)<1$

Just as for the target word region, there was no main effect of condition type in any of the eye-movement measures in the post-target region. However, there was a main effect of distance for total reading time (Figure 1a) and go-past times (Figure 1b) such that participants in the Long distance group spent significantly longer reading and re-reading the post-target word region than those in the Short distance group. The other measures of later processing; second pass times and regressions into the post-target word showed a strong trend towards significance in the same direction of total times and go-past times. It thus appears that separating the target word from its reference sentence did make the integration process more difficult, but not immediately. The costs associated with having to process additional text were reflected in delayed processing of the target word, suggesting that word integration continued after the target word had been read.

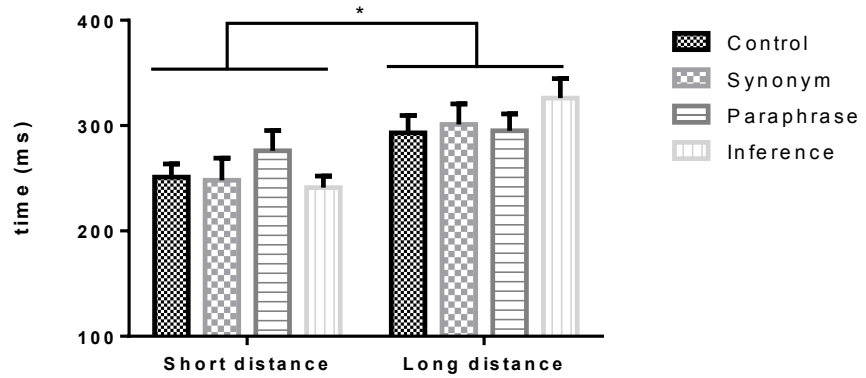


Figure 1a *Mean total reading times (ms) on the post-target word for participants in the Short and Long distance group. Participants in the Long distance group spent more time reading the post-target word than those in the Short distance group. Error bars indicate SEM. * p< .05*

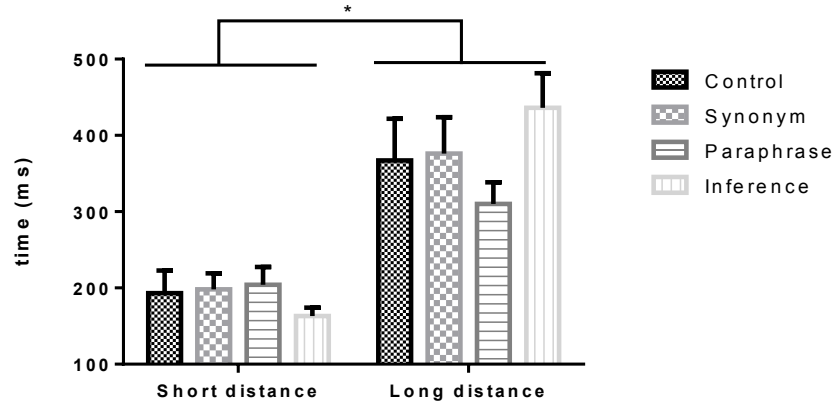


Figure 1b *Mean go-past reading times (ms) on the post-target word for participants in the Short and Long distance group. Participants in the Long distance group spent more time reading the previous text and the post-target word than those in the Short distance group. Error bars indicate SEM. * p< .05*

Analysis of the performance on memory task

To explore whether participants instantiated inferred information, scores from the recall task for each condition were compared. Participants indicated on a Likert scale how certain they were that a presented sentence appeared or not in the stories they had read during eye-tracking. The sentences explicitly mentioned information that participants needed to infer in all but the control condition.

Scores higher than 4 indicated growing certainty that participants had seen the sentence during eye-tracking whereas scores lower than 4 indicated growing certainty that they had not seen the sentence during the eye-tracking part. The score of 4 meant that participants did not know whether the sentence was presented during the eye-tracking part or not. Thus, scores higher than 4 for the synonym, paraphrase and inference conditions reflected an increased rate of false recall and instantiation of inferred information. High scores for the control condition indicated that participants correctly recalled reading the presented sentences during the eye-tracking part.

Figure 3 shows mean scores for each condition by distance group. The scores were also analysed with a mixed analysis of variance with condition type as within subject factor and distance as a between subject factor.

There was a significant main effect of condition type ($F_1(3,132)=47.14$, $p<0.01$, $\eta^2=0.52$; $F_2(3,231)=26.2$, $p<0.05$, $\eta^2=0.34$). Memory scores decreased across conditions in a stepwise manner, indicating good recall of presented information in case of the control condition and decreasing rate of false recall from the synonym to the inference condition. Post-hoc tests revealed that the control condition (M=5.9, SE=1.26) received a significantly higher score than the other three conditions. The synonym condition (M=5.5, SE=1.76) received a significantly lower score than the control condition, but higher than both the paraphrase (M=5, SE=2.12) and the inference (M=4.2, SE=2.38) condition, and the paraphrase condition received significantly higher scores than the inference condition yet lower than the other two conditions (see Fig 5).

This suggests that readers were sensitive to the amount of contextual information provided in the previously read paragraphs and were more likely to falsely recall reading about the inferred information proportionately to the level of support provided. The main effect of distance and the interaction between condition type and distance were not significant ($Fs < 1$).

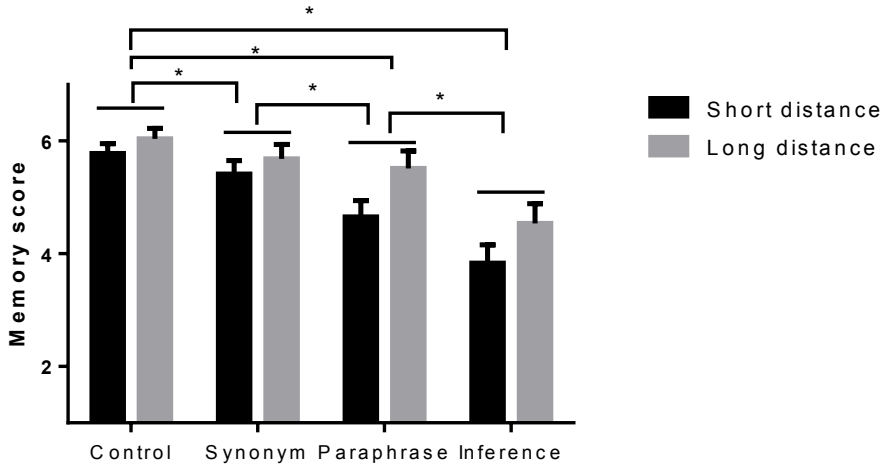


Figure 3 *Mean scores of participants' certainty that inferred information explicitly appeared in the eye-tracked paragraphs in Experiment 3. Scores higher than 4 indicated participants were more certain that the inferred information explicitly appeared, scores lower than 4 indicated participants were more certain the information did not explicitly appear. Score of 4 indicated participants were uncertain whether the information was presented. Error bars indicate the SEM.*p< .05*

General Discussion

Experiment 3 aimed to finally establish whether skilled readers actively used previous context to draw bridging inferences as they integrate new concepts into their text representation. Since Experiment 1 and Experiment 2 yielded inconclusive results, Experiment 3 addressed methodological issues of the first two experiments and aimed to answer the question once again. Despite addressing these methodological concerns and introducing a distance manipulation the predictions for Experiment 3 were not supported. Furthermore, Experiment 3 also aimed to explore whether skilled readers instantiated inferred information into their text representation proportionally to the amount of contextually supporting information they were given. In this discussion, I will review the overall findings of Experiment 3 and lay out a potential explanation for discrepancies between my results and those of Yang et al (2006, 2007),who used a similar manipulation. I will then discuss my findings in relation to the "reverse cohesion effect" described by McNamara and Kintsch (1996) and propose that the readers in my experiments might have used predictive inferences to integrate the target words in their text representation, rather than making a bridging inference. Since the current manipulation is not suitable to verify whether readers drew predictive inferences instead of bridging, I will conclude by suggesting directions to differentiate them in future work.

Review of the experimental manipulation and predictions

Experiment 3 varied the difficulty of bridging inferences that participants needed to draw to integrate the target word into their text representation. This was achieved by manipulating contextual information which provided referents for the integration of the target word. The availability of direct referent for the target word in the context-giving sentences was manipulated, as in Experiments 1 and 2, with the

lowest inferential demands in the control condition and the highest in the inference condition. The control condition provided an explicit direct referent for the target word which shared the same lexical and surface form. Since an early integration effect was observed in the control condition of Experiment 1, a synonym condition was added to explore whether early integration resulted from lexical priming rather than contextual and semantic factors. The synonym condition provided identical content to the control condition, but replaced the explicit referent with a synonym to control for lexical priming. The paraphrase condition provided a less direct but nevertheless contextually related referent for the target word while the inference condition provided no direct referent for the target word. Thus, bridging inferences were the most difficult to draw in the inference condition. This design was inspired by Yang et al. (2006, 2007) who used a similar manipulation to explore on-line word integration with the use of event-related potentials (ERPs). I also included an extra manipulation in Experiment 3 to increase its difficulty, since the results of Experiments 1 and 2 suggested that the stimuli might have been too easy. Thus, I included a distance manipulation which separated the context-giving sentence from the target word by two sentences irrelevant for the target word integration.

Since previous work showed that skilled readers formed an active text representation and their word-to-text integration was immediately sensitive to the contextual information available to them (Shroeder, 2011; Stine-Morrow, Milinder, Pullara, & Herman, 2001; Yang, Perfetti and Schmalhoffer, 2006, 2007), I expected participants to immediately utilize available context in their target word integration. Thus, I predicted that as the level of contextual information decreased and inference making demands rose, skilled readers would take longer to read the target word. Furthermore, I expected participants in the 'Long' distance group to show longer target

word reading times and less facilitation from the contextual information because they had to process additional text separating the target word from its context-giving sentence. With respect to the memory task, I predicted that participants would display further contextual sensitivity and show a higher rate of false recall in conditions which provided more contextual support for instantiation of inferred information.

Review of the main findings

My predictions that readers would actively use supporting contextual information during target word integration were not supported. There were no differences in initial or later reading times on the target word across conditions. This replicates the null findings reported in Experiment 2. Thus, it appears that target word integration was equally easy across the four conditions. Furthermore, target word reading times were largely unaffected by having to process additional irrelevant text in the Long distance condition, although overall word integration did prove more difficult. Participants spent more time reading the target word in the Long distance condition, as indicated by increases in go-past reading times and more regressions back to re-read the target word. However, this additional processing did not translate into overall target word reading times, suggesting the finding is not robust.

Interestingly however, reading times in the post-target word region were all affected by having to process additional text. Normally, spillover processing from the target word affects initial reading times on the post-target word (Rayner, 1998). In my data though, all measures of post-target word later processing showed spillover processing costs: go-past reading times, second pass reading times and total word reading times were all longer in the Long distance group. Regressions back to the post-target region were more frequent and so were the regressions out of it, although the

latter were only marginally significant. Thus, processing of additional irrelevant text created indirect difficulties with target word integration, which did not manifest immediately while participants read the critical word. Such a processing pattern provided evidence that some portion of word-to-text integration carries on after the critical word itself had been read. Thus, integration of a new concept into readers' text representation continues as readers gain following sentential information. It is important to note though that the distance manipulation did not differentially affect processing in individual conditions, further suggesting equal target word integration across conditions.

With respect to the results of the memory task, my predictions were supported. Participants were more likely to falsely recall inferred information in conditions which provided more contextual support. Skilled readers thus instantiated information proportionally to how constraining the context proved to be. This finding supports the idea that skilled readers encoded inferred information into their text representation by making use of context and world knowledge to restrict the meaning of words (e.g., Andreson & Ortony, 1975; Anderson et al., 1976; Oakhil, 1983). It is surprising that our participants showed contextual sensitivity in the offline memory task, yet displayed no hint of it during reading itself, as the target words were being integrated. Let me now discuss the lack of effect in the target word reading times more closely, starting with a comparison of Yang et al.'s (2006, 2007) findings.

Comparison with Yang et al.'s (2006, 2007) findings

In the event-related potential study, Yang et al. (2006, 2007) used a very similar manipulation and their findings demonstrated that skilled readers used previous context immediately to draw bridging inferences when integrating the target word. They found

that the N400 amplitude was significantly reduced in conditions which provided either an explicit referent or a strong contextual referent for the target word, compared to the baseline and inference conditions. Their findings further showed that skilled readers readily responded to even slightly increasing inferential demands, but less skilled readers did not (Yang, Perfetti & Schmalhoffer, 2006). However, there is a considerable difference between presentation of text in reading studies that use eye-tracking and those that measure reading processes using ERPs. In studies using ERPs, text is typically displayed on the screen one word at a time. Readers fixate the centre of the screen while words are presented at a fixed pace. This prevents saccades from occurring and it is different to the natural reading in multiple ways. In normal reading, readers determine how long a word is fixated for and which word is fixated next, with many words being skipped or refixated multiple times through regressive eye-movements. Also, during normal reading, upcoming words are pre-processed parafoveally (Rayner, 1998). These patterns of eye-movements in reading are considered markers of the cognitive processes involved in reading (Rayner, 1988).

Another important difference between natural reading and the word-by-word presentation used in ERP studies is the speed of visual input. In Yang et al.'s (2006, 2007) experiments, words were presented with the stimulus-onset asynchrony of 600 ms whereas average reading fixation lasts only 200-250ms. It is therefore difficult to straightforwardly map findings from the ERP studies to eye-tracking studies. Even though both ERPs and eye-tracking studies are sensitive to word predictability and discourse congruence, direct mapping between ERPs and eye-movement effects is limited (Camblin, Gordon, & Swaab, 2007; Dimigen, Sommer, Hohlfeld, Jacobs, & Kliegl, 2011). There is consensus that the N400 is more likely to reflect later processing since the effect only starts to arise at around 250ms, although Dimigen et al.'s (2011)

model of mapping the EEG effects to eye-movements related the N400 effect to first pass reading times. Dimigen et al. (2011) concluded that when they locked the N400 onset in their EEG signal temporally to the first fixations and first pass times, it mapped well to the eye-movement measures. Moreover, the N400 onsets shared a large degree of variance with the first fixations and first pass times.

Thus, it appears that ERPs are better equipped to capture effects corresponding to early processing reflected in the eye-movement record, but I was more interested in later processing measures such as total time that reflect additional re-reading. Following Dimigen et al.'s analysis, it is questionable whether ERP results are a reliable indicator of later processing effects in the eye-movement record. Additionally, Dimigen et al. used a natural reading display to obtain their EEG data instead of single word presentation display used by Yang et al. (2006, 2007). It is therefore a reasonable question to ask whether the effects observed by Yang et al. (2006, 2007) related to the same processes as those observed in my experiments. Aside from this, it is now important to address potential reasons why my experimental work failed to show an effect of supporting context on drawing of bridging inferences during word-to-text integration.

Explanation for the lack of effect in my experiments:

Inferences possibly not drawn on the target word but later

One possible interpretation of my data is that inferences were not drawn on-line while reading the critical word. However, they were certainly drawn at some point as evidenced by 94.5% success on the comprehension questions, and the patterns obtained on the memory task. Given the current design, the possibility that inferences were not drawn on-line cannot be ruled out since I do not know the target word reading times in a

neutral setting when there is no need to integrate them. This is because even the control condition in the current design required target words to be integrated into participants' text representation. Thus, if any integration processing happened on-line, it would be reflected in target word reading times potentially making them shorter, or longer, than when the word is read on its own. Perhaps word integration happened during sentential wrap-up processing at the end of a clause when readers had gained all necessary information (Just, Carpenter, & Woolley, 1982; Rayner, Kambe, & Duffy, 2000). Such a finding would fit with literature showing that sentence reading times stay equal when readers read sentences such as our control condition, paraphrase and inference condition (Garrod & Sanford, 1977). However, this explanation seems unlikely, given the large amount of eye-tracking literature which finds evidence of inferences being drawn online.

Another possible interpretation of my results is that the presence of additional contextual information did not confer any advantage for drawing inferences to integrate the target word and instead made text processing more difficult. Such explanation might sound counterintuitive, but research exploring effects of text cohesiveness and anaphor resolution suggests that overly explicit context might be detrimental for comprehension. The work I am referring to described two interesting phenomena that occurred when 'too much' context was provided and text processing slowed down, namely the repeated name penalty (Gordon, Grosz, & Gilliom, 1993) and the reverse cohesion effect (McNamara & Kintsch, 1996; McNamara, Kintsch, Butler, Songer & Kintsch, 1996). I briefly discussed the repeated name penalty in the Introduction to Experiment 2 and discussed how it might explain the lack of facilitation seen in the control condition. The reverse cohesion effect offers a good general explanation as to why adding extra context

might not help skilled readers to generate bridging inferences more quickly. I'll now expand on both of these effects further and relate them to the findings of Experiment 3.

The problems of overly explicit contexts - repeated name penalty and reverse cohesion effect

The data showed no evidence of facilitated processing in the control condition when an explicit referent was provided, relative to the other conditions. In contrast, results of Experiment 1 showed very fast target word integration in the control condition. At first sight, it is puzzling that providing readers with an explicit referent for the target word brought no processing benefit. However, early experiments on anaphoric processing (e.g., Garrod & Sanford, 1977; Garrod & Sanford, 1982; Anderson, Garrod & Sanford, 1983) demonstrated that reading times for sentences where an antecedent was explicitly mentioned such as (11) showed no difference in reading time compared to the sentences where the antecedent was only implied, such as (12).

(11) Kent took his car to London. The car had been overhauled.

(12) Kent drove to London. The car had been overhauled.

When the same word constitutes an anaphor and an anaphoric reference at the same time such as 'Kent' in example (13), it may in some cases slow down the anaphor resolution.

(13) Kent drove to London. Kent had been pulled over for speeding on the motorway.

As already mentioned in the Introduction to Experiment 2, this happens when the anaphoric referent is prominent in the discourse and stands for example as the subject of a sentence (Gordon & Hendrick, 1998). If the 'name' of such referent appears in the text again shortly after it had been presented, people assume that it establishes a new agent with the same name. For example, people might assume that second 'Kent' in example (13) introduces a new agent also called 'Kent', since otherwise a pronoun would be used as an anaphor. In such cases, sentence reading time increases compared to when a pronoun is used (Gordon, Grosz & Gilliom, 1993) indicating that readers' anaphoric expectations are violated. Since skilled readers engage deeply with the text and hold information active in their working memory (Zwaan & Radvansky, 1998), the repeated name penalty in their case might be quite severe.

The repeated name penalty arises because readers are repeatedly presented with information they already know and thus do not expect it to be explicitly stated in the text again. This principle also offers a good explanation for why the synonym and paraphrase conditions might have failed to offer contextual facilitation for drawing of bridging inferences during target word integration. It has been found that highly cohesive texts which provide too much supporting information may at times be more difficult to process than less cohesive ones, leading to a so called 'reverse cohesion effect' (McNamara & Kintsch, 1996). Cohesiveness provides people with the majority of information they need to maintain text coherence by connecting adjacent sentences with many connectives and increasing the amount of argument overlap. Thus, the need for readers to engage with prior knowledge and actively generate inferences during reading is minimized. Consequently, it has been found that under certain circumstances, less cohesive texts led to better reading comprehension than the same texts in highly

cohesive format (McNamara & Kintsch, 1996; McNamara, et al. 1996; O'Reilly & McNamarra, 2007).

McNamara et al. (1996) found that increasing text cohesiveness did not lead to better comprehension unless participants' knowledge of a particular topic covered in the text was low. Even then, increasing text cohesiveness only helped low knowledge readers to improve their text-based comprehension, but not their ability to draw bridging inferences (O'Reilly & McNamarra, 2007; Ozuru, Dempsey, McNamara, 2009). When readers with good knowledge of a given topic were presented with highly cohesive texts, their comprehension did not improve and even decreased compared to their scores on low cohesion texts (McNamara & Kintsch, 1996; McNamara, et al. 1996; O'Reilly & McNamarra, 2007). Good prior knowledge grants readers quick and relatively effortless access to relevant information which they activate by using cues from the information stated explicitly in the text base. When this information is less cohesive and contains small gaps, readers are encouraged to construct bridging inferences in order to integrate content they had just read with their background knowledge.

On the contrary, when readers with good topic knowledge read highly cohesive texts, they encounter a lot of information they already know. This might induce them to process the text in a shallow manner because they are not required to engage actively by generating inferences to fill in potential gaps (McNamara et al., 1996). Moreover, unnecessary details in highly cohesive texts might appear unexpected to the readers since there is no need for them. Similar to the repeated name penalty, the unnecessary details might interfere with the knowledge that is already activated and impact text processing because they violate the reader's expectation of what should be mentioned (Kalyuga, Ayers, Chandler & Sweller, 2003). Consequently, highly cohesive texts might create an overload of unnecessary information resulting in either shallow text

processing or costs associated with violation of readers' expectations whereas less cohesive texts might encourage readers to process the written content more actively. Thus, less cohesive texts would produce better comprehension.

It needs to be pointed out that McNamara's research used expository texts whereas my experiment used narratives. However, the scenarios used in my experiments were usual everyday situations and thus it is reasonable to assume that our readers would be considered high-knowledge. By increasing the level of context available, text cohesiveness increased and since the participants are high-knowledge, they did not benefit from this intervention. However, we still need to consider more deeply which processes drive the effects of reverse cohesion and repeated name penalty and whether an alternative explanation could better account for the lack of effect in my data.

Predictive inferences as a mechanism of target word integration

Both the repeated name penalty and the reverse cohesion effect depend on a reader's ability to easily access knowledge concepts associated with what they read and then to form subsequent expectations about what might ensue. The process of forming such expectation is closely linked to the formation of predictive inferences. When drawing predictive inferences, people use cues from the text they had just read to activate associated concepts which are likely to appear in the following text. A body of research reviewed in Chapter 1 documents that readers make predictive inferences during reading, especially if the context is sufficiently constraining, as in causally related texts for example (Calvo & Castillo, 1996; Graesser & Kreuz, 1993; Garrod, O'Brien, Morris & Rayner, 1990; Murray, Klin & Myers, 1993). Whether a context is

deemed sufficiently constraining will strongly depend on people's knowledge of the topic covered in text (Chiesi, Spillich, & Voss, 1979).

As already mentioned in the previous section about reverse cohesion effect, high knowledge readers form links between information contained in the text and their background knowledge faster and with less effort than low knowledge individuals. Because they have rich and detailed knowledge of the topic, they can activate and engage their background knowledge with the text content even in the absence of explicit cues. Thus, they could form equally rich representations from contextually-less supporting texts as from the more constraining ones. It is therefore possible that participants in my studies did not show any facilitation of having extra context simply because all conditions provided them with cues that were strong enough to activate the target word concept in advance, that is, all conditions allowed predictive inferences to be made.

The differences between the conditions in my experiment were rather small. All introduced the same scenario and only differed in the amount of extra context they provided. Since my stimuli captured common every-day situations and my participants were highly skilled readers, it is possible that they formed predictive inferences based on their world knowledge in all conditions. Thus, even the context-giving sentence in the inference condition might have given sufficient contextual cues to anticipate the target word, despite not providing any obvious referent. Consider this following example (14) of an inference condition sentence set from my experiment.

(14) A suspicious man stuck a hanger in the window of Tom's car. In the morning, Tom found out somebody stole his car.

It is quite probable that even indirect indicators such as a 'suspicious man' and 'sticking a hanger in the car window' were strong enough to activate a concept of stealing in our participants, especially given their good world knowledge. Had this been the case, I would see no differences between individual conditions as participants would be able to activate the target word concept in the control condition or the inference condition.

Thus, future work needs to address whether the participants integrated the target word via predictive inferences rather than bridging inferences in the past three experiments presented in this thesis. If the target words were not integrated through predictive inferences, then it is important to further explore how increasing text cohesiveness and repeating prominent information affects bridging inferences during word integration. Furthermore, knowledge of the text topic and reading skill in general will also play a crucial role in determining whether readers draw predictive or bridging inferences to integrate new words in their text representation. Future work needs to address these issues to understand why supporting context failed to facilitate drawing of bridging inference in skilled readers, as suggested by the results work presented in this thesis.

Chapter 5: Predict or Bridge? Investigating skilled readers' use of predictive inferences as a default strategy for word integration (Experiment 4)

Experiment 3 showed that even after addressing methodological mistakes identified in Experiment 1 and 2, supporting context had no measurable effect on inference construction in word-to-text integration during online reading. Thus, skilled readers either did not draw inferences as they read, or used predictive rather than bridging inferences to complete word integration. To answer this question, a final experiment explored whether participants drew predictive inferences, even in weakly constraining contexts, such as the inference condition of previous experiments. Participants were presented with paragraphs where the target word either confirmed the text development suggested by the prior context or not. The prior context provided either a strong indication of future development (like the control condition in previous experiments), or weak (like the inference condition). If participants used predictive inferences, target words should be processed faster when they confirm the development suggested by the prior context than when they do not, regardless of indication strength. If participants used bridging inferences, target word reading times should increase in a stepwise manner, reflecting gradually increasing difficulty of constructing the bridge. The results were consistent with the view that predictive inferences were used to complete target word integration.

Introduction

The previous experiments in this thesis studied how varying the amount of available contextual support influences readers' ability to generate inferences. In this introduction, I will briefly describe the rationale and results of previous experiments. Then, I will lay out a possible explanation for the null findings found in Experiment 2 and 3 and propose that predictive inference making might account for the similar reading times seen across all experimental conditions. Based on this explanation, I will propose a new experiment to test whether readers generate predictive inferences in contexts like those used in my previous experiments. The new study will also aim to explore whether readers use bridging or predictive inferences as their default integration strategy. I will conclude by spelling out predictions for target word reading patterns depending on which inference strategy readers have used.

In accordance with the literature that previous context immediately affects word processing (Cook & Myers, 2004; Yang et al. 2007; Rayner, Warren, Juhasz & Liversedge, 2004; Staub & Rayner, 2007), Experiment 1 revealed that readers readily used relevant contextual information to facilitate integration of a new concept introduced by the target word. However, Experiment 2, which aimed to clarify whether the fast integration seen in the first experiment arose as a result of lexical repetition or availability of supporting semantic information, failed to replicate Experiment 1, with no evidence of any effect of previous context.

Since both Experiment 1 and Experiment 2 suffered from methodological problems which might have contributed to their discrepant results, Experiment 3 was designed to address these problems and finally provide an answer as to how previous context affects inference making, and its time course during word-to-text integration.

Despite addressing the methodological problems of Experiment 1 and Experiment 2, Experiment 3 also failed to demonstrate any evidence that readers used supporting contextual information to facilitate the inference-making needed to integrate the target word.

There are two possible explanations for the lack of effect seen in Experiments 2 and 3. First, it is possible that readers do not draw inferences on-line and do so only after they have finished reading the target word, without regressing back to it. On this view, comprehension processes might not happen immediately but may be put on hold until readers acquire further information. Thus, rather than processing discourse incrementally with every new piece of information, readers might wait until they acquire more information and process it 'in chunks', perhaps when they reach a comma or a full stop. Such an account would fit with research showing that readers spend longer reading sentence- or clause-final words during 'wrap-up'; the processing associated with integration of overall information (Just & Carpenter, 1980; Liversedge & Findlay, 2000; Rayner, Kambe & Duffy, 2000). However, this fits uncomfortably with the vast body of evidence suggesting that readers integrate information from incoming words into their text representation immediately, and even delayed processing is manifested by re-reading and frequent regressions back to the problematic words (Staub & Rayner, 2007). Thus, it seems unlikely that entire integration would take place during wrap-up processing and almost none at the target word itself.

A second potential explanation for the lack of effect seen in earlier experiments is that the primary integration strategy used by participants was to make predictive inferences, rather than bridging inferences. My rationale in the previous experiments was that readers used bridging inferences to complete their word-to-text integration, following similar claims in the literature (McKoon & Ratcliffe, 1992; Yang et al. 2006,

2007). I reasoned that if readers are skilled, they would use previous context efficiently and this would translate directly into their immediate word processing. Thus, when they draw bridging inferences to integrate new words in their text representation as seen in Yang et al.'s work (2006, 2007), integration should be easier when the previous context serves to provide supporting information. I predicted this to be the case especially for the control condition in Experiment 1, 2 and 3. This is because the control condition strongly invoked the target word concept by using a lexically-identical referent in the context-giving sentence. Thus, making of the bridging inference on the target word such as 'accident' in the second sentence in (1) should have been extremely easy. However, Experiment 3 failed to show a processing advantage for this condition.

> *(1) Anne looked up and saw the bus was heading in her direction, but it was too late to stop the accident from happening. The accident was later reported on the news.*

If I agree that it seems unlikely that readers did not integrate the target word on-line and only did so after they finished reading, it follows that readers must have integrated the target word during reading. One possibility is that integration happened through predictive inference making rather than bridging. The explicit condition provided readers with such constraining context that it was easy to expect the target word in later content. For example, when reading about illness and medical help, the target word 'hospital' is likely to get activated. It might be that the less supportive conditions also provided enough cues for target word activation, and this would explain why I did not see any differences in target word reading times across conditions: readers could integrate the target word equally easily because it had been pre-activated by the context-giving sentence in all conditions.

It has been argued that readers engage in predictive inference making when they read contexts that are highly constraining (Garrod & Terras, 2000; Schmalhofer, McDaniel & Keefe, 2002 for a review). For example, when reading a sentence such as (2) the phrase 'went out' is read faster than when it is preceded by a less constraining sentence as in (3).

(2) Mary poured a bucket of water on the fire. The fire went out.

(3) Mary placed a bucket of water by the fire. The fire went out.

This is because readers know from their world knowledge that water extinguishes fire and since pouring of water causes direct contact with the fire while placing the bucket by the fire does not, the concept of fire going out becomes more available in the first instance (Singer, 1993). Similarly, when readers read about professions which are stereotypically male or female, they immediately activate the corresponding stereotype and expect the gender before it is mentioned directly (Duffy & Keir 2004, Sturt 2003). Thus, if a pronoun violates this gender stereotype, for example if 'herself' refers to a 'pilot', reading times on the pronoun are inflated.

Instances of predictive inference are not necessarily tied to causal events or strong stereotypes; they also arise in situations that do not have such strongly established schemas. For example, Cook et al. (2001) found that when people read about a little boy throwing stones at a car, subsequent naming times for the words 'dent' and 'damage' were faster than when they read about a little boy throwing foam balls at a car. Faster naming times indicated that 'dent' and 'damage' concepts have been activated in the first scenario, but not in the second. Thus, critical for the generation of predictive inferences is whether previous context provides cues that might invoke them. However, whether a cue is strong enough to activate a prediction is not just a function

of how explicit is the text itself. It also depends on world knowledge and readers' ability to integrate it with their text representation. Readers' world knowledge interacts with information from the text as a whole and immediately preceding content. If strong enough signals come from the text content to world knowledge, they lead to the activation of a specific concept. Thus, text about rocks being thrown at a car could activate associated concepts from the world knowledge such as car getting damaged.

The generation of predictive inferences has been argued to be automatic meaning that once activated, a predictive inference is generated even if the previous context does not support it (Cook, Lassonde, Splinter, Guéraud, Stiegler-Balfour & O'Brien, 2014). Cook et al. (2014) presented readers with narratives that described the main protagonist, waitress Carol, as someone who used to be short-tempered but changed and is now calm. However, when readers read the sentence (4), they named the word 'dump' faster than in the control condition.

(4)Carol lifted a plate of spaghetti over [the customer's] head.

Despite previous context telling them that Carol is now a calm individual, the readers still generated a prediction that Carol dumped the spaghetti on the customer's head. The signals from initial information about Carol's past short-temper and her action of lifting the spaghetti over the customer's head were enough to activate the general world knowledge concept of short-tempered people acting rashly.

Thus, readers do not need only strongly constraining and explicit contexts to generate predictive inferences. If world knowledge is engaged within their text representation, a small cue might be sufficient to activate a concept and trigger a prediction. With respect to Experiments 2 and 3, it is therefore possible that even those conditions with less contextual support (i.e. the inference and paraphrase conditions)

provided cues that were strong enough to activate the target word concept in advance. For example, the 'inference condition' from the previous experiments provided the following context for the target word 'hospital':

(5) Allen's baby was very ill so he put the baby in the car and rushed off.

If readers are skilled and engage their general knowledge with their text model, the above sentence will resonate with their real-world experience that 'ill people are usually taken to get help'. Thus, the concept of a 'hospital', 'medical help' or 'doctor' will become activated before it is mentioned in the following sentence.

It is important to establish whether participants in the previous experiments activated the target word concept equally in all conditions and used predictive inferences. If this is not the case, previous evidence that readers' draw inferences on-line will need to be re-considered. Previous experimental manipulations were not sufficient to distinguish between bridging and predictive inferences. To establish whether readers might have used predictive inferences as a default strategy in the previous experiments, they need to be presented with texts which can distinguish between these two types of inferences.

If the default strategy to integrate new concepts is via a bridging inference, reading time on the target word should reflect the difficulty of making the bridge. In contexts which provide ample prior referents for the target word, it should be easy to integrate it and processing time for bridging inference would be short. When the integration proves more demanding, for example when the text direction takes an unexpected turn, computation of the bridging inference will be more difficult and hence take more time.

In contrast, if the defaults strategy to integrate new concepts is via predictive inferences rather than bridging, no differences in reading times between conditions will be seen as long as previous context provided enough cues to reach the activation threshold. This is because readers activated the target word concept and thus generated a predictive inference prior to reading the target word. When previous context does not provide cues for advance activation of the target word concept, people will not anticipate the upcoming target word concept in the ensuing text. When they encounter it, they will struggle to integrate the target word even more than in the bridging inference scenario because they would need to alter predictions they had already generated. Predictive inferences are therefore more vulnerable to unexpected directions than bridging inferences.

It is impossible to know whether bridging or predictive inferences were used in previous experiments before setting up the way to distinguish them. A good way to distinguish them would be to present participants with contexts that either give indications pointing to the target word or not. However, manipulating only the target word expectancy would not be sufficient to reveal whether readers use mainly predictive or bridging inferences to integrate new concepts. A demanding bridging inference and a violated predictive inference are both difficult to process and their reading patterns might look rather similar as both might elicit longer reading times and a higher rate of re-reading. It is therefore important to add another manipulation where bridging and predictive inferences differ and that is the strength of previous context, as used in the Experiment 2 and 3. Experiment 4 attempts to do this.

Rational and design for Experiment 4

Experiment 4 will thus manipulate both the strength of previous context and the expectancy of the target word. If the main difference in target word reading times is caused by expectancy, then readers use predictive inferences as a default integration mechanism. This is because target word reading times will not be sensitive to having extra supporting information, once contextual cues activated the target word concept. However, if the target word expectancy and strength of the previous context interact in their effect on target word reading times, it suggests that readers predominantly use bridging inferences. By providing clear referents for the target word in prior context, it becomes easier to build a bridge to the previous text and reading times on the target word should thus be shorter. I will now describe the design in further detail.

Two factors will be manipulated: contextual strength and target word expectancy. The contextual strength manipulation provides readers with either strong contexts or weak contexts, similarly to the manipulation in the previous experiments. Just as in the control condition of Experiment 3, 'strong' context provides readers with both a direct and an indirect indication for the target word concept. The 'weak' context only provides the indirect indication, as in the 'Inference' condition of Experiment 3. Example (6) illustrates the 'strong context' condition which gives both a direct indication ('the infirmary) and an indirect ('ill') indication for the target word '*hospital'*:

(6) Allen's baby became very ill so he put the baby in the car and rushed off to the nearby infirmary. The hospital had a long queue.

The 'weak context' condition contains only the indirect indication, 'ill' as in (7):

(7) Allen's baby became very ill so he put the baby in the car and rushed off. The hospital had a long queue.

Target word expectancy was manipulated on the basis of the previous context by comparing confirmed and unconfirmed outcomes. In the confirmed condition, the strong and weak indications pointed to the target word concept whereas in the unconfirmed condition they did not. Examples (6) and (7) show the case of the confirmed expectations evoked by reading about an ill baby and infirmary. In the expectation unconfirmed condition, the weak indication 'ill' gets substituted for 'unhappy', and strong indication 'infirmary' for 'work', as in (8):

(8) Allen's baby became very unhappy so he put the baby in the car and rushed off to his work that was nearby. The hospital had a long queue.

In order to integrate the target word '*hospital'* in (8), readers need to make an inference that it is Allen's workplace, but the context-giving sentence does not provide prior ground for this inference. The weak context version of the unconfirmed condition is a little easier to process as readers only need to infer that the hospital is where Allen rushed to, as in (9).

(9) Allen's baby became very unhappy so he put the baby in the car and rushed off. The hospital had a long queue.

By crossing the 'context' and 'expectation' factors, four conditions were created; the strong confirmed, the weak confirmed, the strong unconfirmed and the weak unconfirmed. The strong confirmed and the weak confirmed conditions directly map to the 'Synonym' and 'Inference' conditions respectively from the previous experiments. I chose the 'Synonym' condition rather than the 'Control' condition

because it provides supporting contexts that are semantically equal but avoids problems associated with lexical repetition. Experiment 4 thus sets to establish whether limited contexts in the 'Inference' condition in earlier experiments were in fact sufficient for readers to generate predictive inferences.

The predictions for results of Experiment 4 are as follows. If participants in the previous experiments integrated the target words by predictive inferences in all conditions, reading times for both the 'strong' confirmed and 'weak' confirmed conditions in Experiment 4 should not differ. This would suggest that even weak indications on their own are sufficient to activate the target word concept and trigger a predictive inference. However, both confirmed conditions should have significantly shorter target word reading times than the unconfirmed conditions because prior context in the unconfirmed conditions does not allow readers to pre-activate the target word concept. If readers generate specific and well-formulated predictions during reading, the differences in reading pattern should be visible in early measures such as first pass reading times already. Because the unconfirmed conditions challenge initial predictions that readers may have formed after reading the context-giving sentence, I expect a higher rate of target word re-reading in these conditions. Since words which are difficult to process often affect processing of the next word, I expect longer reading times on the post-target word region in the unconfirmed conditions.

On the other hand, if readers are drawing bridging inferences, I should see a gradual increase in reading times as follows: The 'strong confirmed' condition should show the shortest reading time since it provides multiple referents for the target word and a bridging inference can thus be completed very quickly. The 'weak confirmed' condition should show significantly longer reading times since the weak indication of the target word does not provide an obvious referent for the target word. The 'weak

unconfirmed' condition should also show longer processing than the 'strong confirmed' condition, but it is quite similar to the 'weak confirmed' condition in terms of difficulty of making the bridge to the previous context. Finally, the 'strong unconfirmed' condition should be the hardest to integrate because the previous context does not provide any referents for the target word and requires the reader to make a make a large conceptual jump to compute the bridge. Since bridging inferences can be quite costly compared to the predictive inferences that can arise through automatic associations, I expect the reading differences to manifest in later processing through longer and more frequent re-reading.

Furthermore, to uncover potential differences in the use of predictive and bridging inferences with regard to participants' comprehension ability, participants were split into two groups based on their performance on comprehension questions in the experiment. Participants scoring above the median comprehension score were included in the higher comprehension group, participants scoring at the median or below were included in the lower comprehension group. Because skilled comprehenders are known to have an active text representation well connected with their world knowledge, I expect the higher comprehension group to show a more pronounced use of predictive inferences than the lower comprehension group. Given the high skill level of my participants, however, it is quite likely that no differences in target word integration will emerge between the groups since comprehension differences are likely to be marginal. The examination of processing differences based on comprehension performance is thus primarily of explorative value.

Method

Participants:

Forty-eight participants were recruited from the population of Oxford University undergraduate and graduate students (20 males and 28 females, mean age 21 years and 4 months). All participants were native English speakers who first learned to read in English. Four participants were bilingual speakers with English being their dominant language. Participants had normal or corrected-to-normal vision and no history of language or reading difficulties.

Apparatus:

As in Experiments 1-3, an EyeLink 1000 remote eye-tracker with chin and head support was used to monitor the participants' eye movements as they read text presented on the computer screen with the resolution of 1024 x 768 pixels at a distance of 60 cm. The tracker measured gaze location from the right eye by monitoring the pupil and corneal reflection at a sampling rate of 1000 Hz. The calibration was set to a maximal deviation of 0.4˚ with 3.9 characters subtending to 1˚ of visual angle. The sentences were presented in three to four line paragraphs aligned to the left of the screen. The sentences were written in a white Courier New 16-point font on a black background.

Materials:

Sixty two-sentence paragraphs comprised the stimuli set; there were 40 experimental paragraphs and 20 filler paragraphs which mimicked the experimental manipulation (see Appendix E). Additionally, 20 comprehension questions were presented to the participants during reading to keep participants' attention and assess their comprehension both on experimental and filler paragraphs.

Two within subject factors were manipulated in the experiment: contextual constraint (strong and weak) and prediction confirmation (confirmed and unconfirmed). Contextual constraint denotes how strongly the context constrains towards making a prediction about what might come next in the text. The context manipulation either gave participants a strong or weak contextual indication for the inference necessary to integrate the target word in the next sentence. In the 'strong context' condition, participants were given both the weak and the strong indication for the target word inference. In the 'weak context' condition, only the weak indication was given. See the example below:

Context manipulation

The weak indication is underlined and the strong indication is in italics. The target word is 'accident':

Strong:

Anne looked up and saw the bus was heading in her direction but it was too late to stop the *collision*. The accident was later reported on the news.

Weak:

Anne looked up and saw the bus was heading in her direction while honking at her. The accident was later reported on the news.

In the above example, the inference necessary to integrate the word 'accident' is that the bus and Anne collided. The weak indication of this event is that the bus was heading in Anne's direction, the strong one is the direct mention of there being a collision.

The second factor manipulated is prediction confirmation. In the 'Confirmed' condition, participants are given a strong and a weak indication to activate the predictive inference

for the target word, just as words 'heading' and 'collision' activate the target word concept 'accident' in the example above. In the 'Unconfirmed' condition, the phrases giving weak and strong indication were changed to not lead the reader to activate the inferential prediction while the passage still retained overall plausibility. The target word sentence remained the same in both conditions. See the example below:

Prediction manipulation
The weak indication is underlined and strong indication is in italics. The target word is 'accident'. In the Unconfirmed condition, the indication phrases are changed to be less constraining to the target word inference than in the Confirmed condition.

Confirmed

Anne looked up and saw the bus was heading in her direction but it was too late to stop the *collision*. The accident was later reported on the news.

Unconfirmed

Anne looked up and saw the bus was passing in her direction but it was too late to stop *it and get on*. The accident was later reported on the news.

The two factors 'contextual constraint' and 'prediction confirmation' were crossed to yield four conditions. The 'Strong confirmed' and the 'weak confirmed' conditions were identical in their structure to the 'Synonym' and the 'Inference' condition from the previous experiments in order to allow for a direct comparison. See all four conditions in Example 1.

Example 1
Four experimental conditions created by crossing the 'contextual constraint' manipulation and 'prediction confirmation' manipulation. The weak indications are in italics, strong indications are underlined and the target word 'accident' is in bold.

Strong Confirmed

Anne looked up and saw the bus was *heading* in her direction but it was too late to stop the collision. The accident was later reported on the news.

Weak Confirmed

Anne looked up and saw the bus was *heading* in her direction while honking at her. The **accident** was later reported on the news.

Strong Unconfirmed

Anne looked up and saw the bus was *passing* in her direction but it was too late to stop it and get on. The **accident** was later reported on the news.

Weak Unconfirmed

Anne looked up and saw the bus was *passing* in her direction while honking at her. The **accident** was later reported on the news.

Procedure:

The experimental paragraphs were presented to the participants in a counterbalanced design so that each saw 10 items in each of the four conditions and never saw any of the items more than once. During the experimental session, participants read 60 paragraphs of which 40 were experimental and 20 were filler paragraphs. The paragraphs were presented in a random order.

Twenty filler items were included in order to address potential strategic effects and disguise the main manipulation. To maintain the participants' attention and assess their comprehension, 20 comprehension questions about the experimental and filler items were presented randomly throughout the experiment. The paragraphs were not presented on the screen during reading and answering the question.

Participants sat 60 cm away from a computer monitor with their chin and forehead rested on a chinrest to constrain head movements. They were told that they would be reading short stories on the computer screen and answering comprehension questions about some of the stories.

Participants were instructed to read the stories as they normally read and refrain from skim reading. They are asked to restrict blinking and to open their eyes wide as they were reading to allow for better eye-movement recording. The eye-tracker was then calibrated to a value of 0.4 maximum in order to allow precise tracking of fixation, even for the words that were only three characters long.

Before the start of each trial, a white fixation point appeared at the centre of the screen. The participants were instructed to fixate on this point. After fixating the point, a white square appeared at the left edge of the screen at the position which marked the location of the beginning of the first sentence in the narrative. The participants were instructed to fixate on the white square until the text appeared on the screen and press a trigger button on a hand-held controller after they finished reading to move on to the next paragraph. The controller was also used to answer the comprehension questions.

Results

Eye-tracking data preparation

Readers answered the comprehension questions with average accuracy of 0.86 (SD= 0.06). As in the previous experiments, fixations shorter than 80ms were combined with a nearby fixation within three letter spaces or deleted. Fixations longer than 800ms were deleted.

As in the previous experiments, the regions of interest are the target word with the preceding space shown in italics (“ accident”) and the post-target word (“was later”) immediately following the target word (see Example 2). The post target word comprised the space before a word or a word phrase where the immediate word was less than four letters long that followed the target word. Both the target word and the post-target word region are located in the second sentence. Additionally, the weak indication region in the context-giving sentence preceding the target word sentence was analysed.

Example 2 *shows an experimental paragraph in the 'strong confirmed' and 'weak confirmed' condition. The regions of interest are marked in brackets. For the target word sentence, the target word region is in italics and the post-target word region follows immediately afterwards. For the context giving sentence, the weak indication region is marked out.*

Strong Confirmed

Anne looked up and saw the bus was/ heading/ in her direction but it was too late to stop the collision. The/ *accident*/ was later/ reported on the news.

Weak Confirmed

Anne looked up and saw the bus was/ heading/ in her direction while honking at her. The/ *accident*/ was later/ reported on the news.

Eye-tracking analysis

Eye-tracking analysis of the target word region

Data were analysed by subjects and by participants using a 2x2 repeated measures ANOVA with two within-subject factors; prediction (confirmed and unconfirmed) and context (strong and weak) to test whether readers generated predictive or bridging inferences when integrating target words. For the full summary of means and standard deviations for all four conditions see Table 5.1.

Table 5.1 *Mean reading times (ms) on the target word and standard deviations (in parentheses) in all four conditions for all participants in Experiment 4*

		M			
	N=48		*(SD)*		
	Measure (ms)	**Strong/ Confirmed**	**Weak/ Confirmed**	**Strong/ Unconfirmed**	**Weak/ Unconfirmed**
Early processing measures	First fixations	213	207	208	215
		(36)	*(29)*	*(35)*	*(36)*
	First pass	250	237	242	242
		(47)	*(51)*	*(59)*	*(51)*
Later processing measures	Total time	335	333	368	369
		(82)	*(103)*	*(99)*	*(125)*
	Second pass	299	276	298	311
		(115)	*(109)*	*(106)*	*(130)*
	Go past time	345	321	318	351
		(145)	*(137)*	*(108)*	*(174)*
	Measure (probability of regressing)				
	Regressions in	0.15	0.19	0.24	0.24
		(0.14)	*(0.15)*	*(0.16)*	*(0.19)*
	Regressions out	0.23	0.16	0.19	0.16
		(0.19)	*(0.15)*	*(0.19)*	*(0.17)*

Early processing measures showed no difference between the conditions. In terms of initial reading of the target word, there was no effect of context ($Fs<1$) or prediction confirmation ($Fs<1$) in the first pass times and first fixation durations. Equally, the interaction between prediction confirmation and context was not

significant. The similarity of target word reading times in the early processing measures across all conditions thus suggested that any predictions about future text direction readers might have generated during reading were rather underspecified.

In contrast, measures of later processing showed a different picture. In terms of total reading time that readers spent on the target word, there was a strong main effect of prediction type ($F_1(1,47)=17.5$, $p<0.01$, $\eta^2=0.27$; $F_2(1,39)= 8.4$, $p<0.01$, $\eta^2=0.18$) such that participants spent significantly less time reading the target word in the confirmed prediction conditions than in the unconfirmed conditions (see Figure 1a and Table 1). The lack of main effect of context ($Fs< 1$) suggested that participants did not receive processing benefits from having extra contextually supporting information. Thus, even weak indication was enough for readers to activate associated concepts and generate a prediction about what might ensue in the text in the confirmed conditions. The interaction between the context and prediction type was not significant ($Fs <1$).

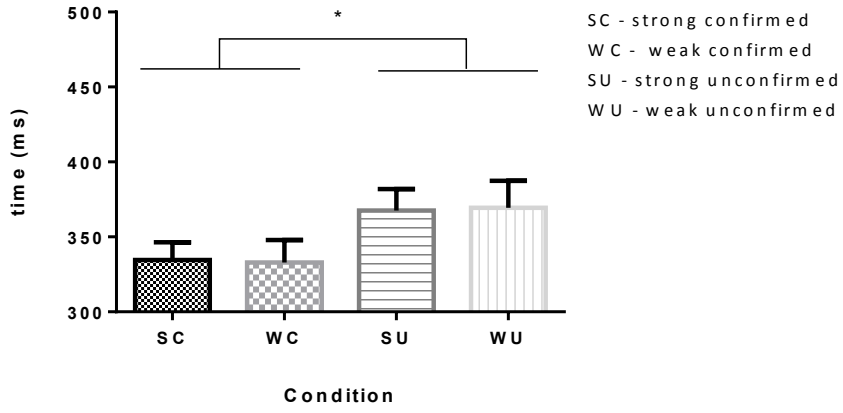


Figure 1a *Mean total target word reading times (ms) for participants in Experiment 4. Confirmed conditions were read faster the unconfirmed. Error bars indicate the SEM. *p< .05*

Since there was no difference between the conditions in the early processing measures, the difference seen in the total reading time must come from additional re-reading of the target word. In line with this logic, readers made significantly more regressions back to the target word in the unconfirmed conditions than in the confirmed ($F_1(1,47)=15.1$, $p<.01$, $\eta^2= 0.24$, $F_2(1,39)=9.9$, $p<.01$, $\eta^2=0.2$) (Figure 1b). Contextual constraints had no effect ($F_1(1,47)=1.19$, $p>.05$; $F_2(1,39)<1$) on the likelihood of regressing back into the target word and there was no interaction between context and prediction either ($Fs<1$).

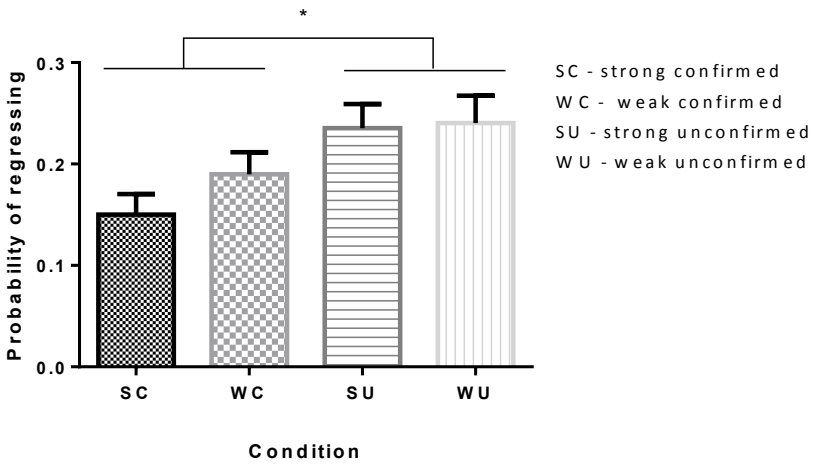


Figure 1b *Mean regressions back to the target word for participants in Experiment 4. Fewer regressions made in the confirmed conditions than the unconfirmed. Error bars indicate the SEM. *p< .05*

Interestingly and somewhat counter-intuitively, there was a main effect of context for regressions out of the target word ($F_1(1,47)=6.15$, $p<.05$, $\eta^2=0.11$; $F_2(1,39)=4.91$, $p<.05$, $\eta^2=.11$) (Figure 1c) such that readers regressed out of the target word significantly more often in the strong context condition than following weak context. This finding is rather puzzling since it would indicate that strong context condition is harder to process than the weak context condition. This might be understandable for the Strong/Unconfirmed condition where the context-giving

sentence does not provide readers with the cues to anticipate the target word, it is hard to understand such finding for the Strong/Confirmed condition. The lack of interaction effect though suggests that strong context seems to be more difficult to process in general.

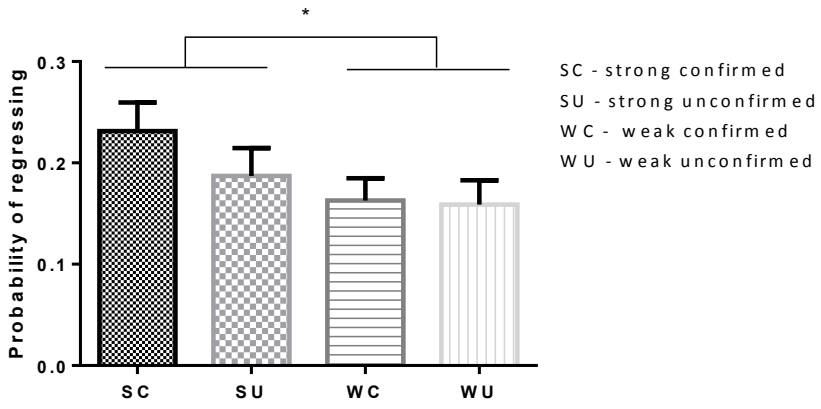


Figure 1c *Mean regressions out of the target word for participants in Experiment 4. Participants regressed more often in strong context conditions relative to weak context conditions. Error bars indicate the SEM. *p< .05*

The remaining measures of later processing; go-past times (*prediction*: *Fs*<1; *context*: *Fs*<1) and second-pass times (*prediction*: $F_1(1,35)= 1.82, p>.05$; $F_2<1$; *context:* Fs<1) , showed no significant main effects or interactions. This implies that once readers decided to go back to the target word, they took a similar amount of time to re-read across all conditions, despite making more frequent regressions back to the target word in the unconfirmed conditions.

Relating this to findings from earlier experiments, the current finding gives support to the view that readers generated predictive inferences in all conditions in Experiment 2 and 3. The weak /confirmed condition in Experiment 4 is the same as the inference condition in the previous experiments and the strong/confirmed condition the

same as the synonym condition in Experiments 2 and 3. The fact that target word reading times in the weak/confirmed and strong/confirmed conditions did not differ from each other yet were shorter than in the unconfirmed conditions thus suggests that readers drew predictive inferences in the synonym as well as inference conditions in Experiment 2 and 3.

Eye-tracking analysis of the post- target word region

As noted in the results section of Experiments 2 and 3, processing of the target word may spill over to the immediately following region, especially if processing is more demanding. Thus, I also looked at the processing of the post-target region using the same procedure as for the target word region; a two-way repeated measures ANOVA. For the full summary of means and standard deviations see Table 5.2. Figure 2a-c depicts overall post-target word reading times, go-past reading times and regressions back into the post-target word.

Table 5.2 *Mean reading times (ms) on the post-target word and standard deviations (in parentheses) in all four conditions for all participants in Experiment 4*

	N=48	*M* / *SD*			
	Measure (ms)	**Strong/ Confirmed**	**Weak/ Confirmed**	**Strong/ Unconfirmed**	**Weak/ Unconfirmed**
Early processing measures	First fixations	199	214	198	200
		(34)	*(44)*	*(40)*	*(37)*
	First pass	247	262	252	249
		(69)	*(62)*	*(65)*	*(66)*
Later processing measures	Total time	339	357	376	369
		(97)	*(126)*	*(115)*	*(126)*
	Second pass	274	269	278	279
		(106)	*(108)*	*(111)*	*(110)*
	Go past time	339	403	408	355
		(169)	*(222)*	*(204)*	*(134)*
	Measure (probability of regressing)				
	Regressions in	0.25	0.21	0.28	0.23
		(0.17)	*(0.15)*	*(0.18)*	*(0.17)*
	Regressions out	0.13	0.17	0.17	0.17
		(0.15)	*(0.17)*	*(0.13)*	*(0.18)*

In terms of early processing, there was a trend towards significance in the first fixation times for both main effects; context ($F_1(1,47)= 5.25$, $p<.05$, $\eta 2=.12$; $F_2(1,39)<1$) and prediction ($F_1(1,47)=3.9$, $p=.05$, $\eta^2= .1$;$F_2(1,39)=2.87$, $p=.09$, $\eta^2=.07$). This trend was probably caused by slow times for the 'weak confirmed' condition yet it seems rather spurious since no other eye-movement measures showed similar values for the 'weak confirmed' condition. Thus, the first fixations pattern is probably just due to chance. The first pass reading times showed no main effects or interaction effect ($Fs < 1$).

Following the pattern of findings on the target word, the total reading times (Figure 2a) for the post-target word region were significantly longer in the unconfirmed than in the confirmed conditions ($F_1(1, 47)=4.44$, $p<.05$, $\eta^2=.09$; $F_2(1,39)=4.14$, $p<.05$,

η^2=.09). This suggests that processing of the target words and its integration in the unconfirmed condition was more demanding than in the confirmed condition, thus spilling into the processing of the following words. Since there is no difference between the conditions in the early processing measures, the difference should stem from additional re-reading of the post-target region. The main effect of context and interaction between prediction confirmation and context were not significant for the post-target word region (all *Fs* <1).

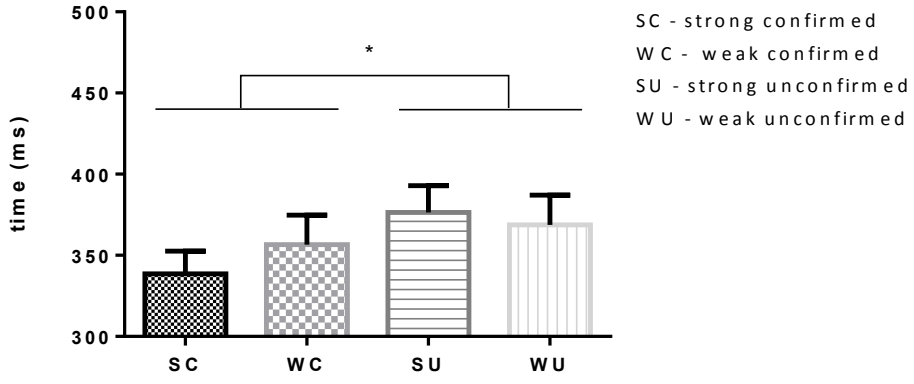


Figure 2a *Mean total reading times for the post-target word in all four conditions in Experiment 4. Post-target word read for longer in the unconfirmed conditions. Error bars indicate SEM. *p< .05*

Correspondingly, go-past times (Figure 2b) revealed a significant interaction between the prediction confirmation and context (F_1(1,47)=6.29, p<.05, η^2=0.1; F_2(1,39)=4.1, p=.05, η^2=0.09) such that the 'weak confirmed' condition showed significantly longer go-past times than the 'strong confirmed' condition. This suggests that 'weak confirmed' condition might be more demanding to process than the 'strong confirmed' condition, but the processing costs only manifest themselves in the post-target region and not during the processing of the target word itself. The 'strong confirmed' condition also showed significantly shorter go-past times than the 'strong

unconfirmed' condition indicating that readers found it more difficult to integrate the target word in the latter condition and these difficulties affected processing of the post-target words.

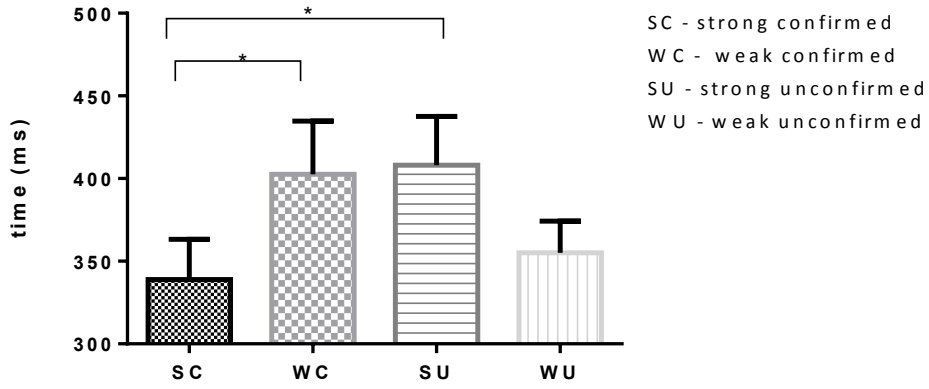


Figure 2b *Mean go-past times for the post-target word in all four conditions in Experiment 4. Post-target word read for longer in the weak confirmed and strong unconfirmed conditions relative to the other two. Error bars indicate SEM. *p< .05*

Regressions back into the target word (Figure 2c) showed a surprising trend toward a main effect of context ($F_1(1, 47)=5.38$, $p<.05$, $\eta 2=.09$; $F_2(1,39) <1$) such that 'strong context' elicited more regressions back into the target word region than weak context. This is slightly puzzling since a high number of regressions back into the post-target word region in the 'strong unconfirmed' condition be expected yet the 'strong confirmed' condition also elicited relatively many regressions back to the post-target word region.

Second pass times and regressions out of the target word did not show any effect of prediction confirmation or context manipulations and no interaction effects between them (all *Fs* <1).

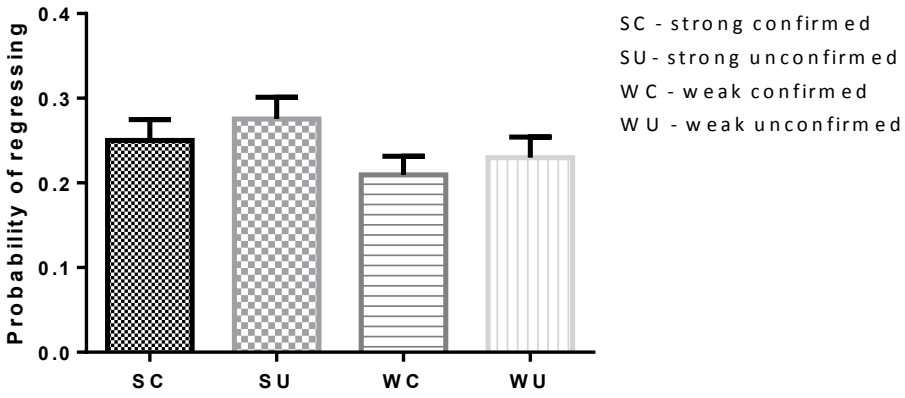


Figure 2c *Probability of regressing back into the post-target word in all four conditions in Experiment 4 showing a trend towards more frequent post-target word re-reading in the strong context conditions. Error bars indicate SEM.*

Recall from the Introduction, I predicted that it will be more difficult to integrate the target word with the text representation in the unconfirmed conditions because the target word does not confirm the text direction set out in the context-giving sentence. Thus, I expected a potential spill-over processing in the post-target word region in the unconfirmed conditions. The results largely followed this prediction, especially in terms of later processing measures. The go-past reading times for the post-target region indicated that the weak/confirmed condition was, in the end, more demanding to process than the strong/confirmed condition, but this difficulty only manifested once the target word had been read. Somewhat unexpectedly, the analysis also revealed an increased rate of post-target word re-reading in the strong context conditions.

Eye-tracking analysis of weak indication region

In addition to the target word and post-target word regions, this experiment afforded an opportunity to analyse the region that provided a weak indication for the potential prediction. This region was identical in both 'strong' and 'weak' context

conditions. I thus reasoned that this region will experience more re-reading in the 'weak confirmed' condition compared to the 'strong confirmed' condition since it is the only indication that readers have for generation of the predictive inference in the 'weak confirmed' condition. To test this reasoning, I compared the total reading times, regressions back to the region and second pass times of the weak indication region between the 'strong confirmed' and 'weak confirmed' conditions using a one way repeated measures ANOVA with 'context' as two-level within subject factor. See Table 5.3 for mean reading measures and standard deviations. However, there was no significant difference in overall reading time, second pass reading time or number of regressive eye-movements across the two conditions (*Fs* <1) despite the higher values in the weak/confirmed condition.

Table 5.3 *Mean measures of additional processing and standard deviations (in parentheses) for the weak indication region in the confirmed conditions in Experiment 4.*

N=48	*M* (*SD*)	
Measure (ms)	**Strong/ Confirmed**	**Weak/ Confirmed**
Total time	407	415
	(123)	*(157)*
Second pass	301	307
	(122)	*(156)*
Measure (probability of regressing)		
Regressions in	0.18	0.21
	(0.14)	*(0.16)*

Analysing readers' individual differences: Comprehension performance median split for the target word region

To explore whether there are any strategic differences between readers who performed highly on the comprehension questions in the experiment and those who performed less well, I split the participants around the median value for their comprehension accuracy (median accuracy = 0.85) into a Higher (M=0.92, SD= 0.03) and Lower (M=0.82, SD= 0.04) group and analysed the target word data again, adding Comprehension group as a between subject variable. Only by subject analysis was performed, since this analysis was primarily of explorative nature. Due to careful construction of the stimuli in Experiment 4, I assumed that by items variation would be small enough to consider by-subject analysis reasonably indicative. Table 5.4a and 5.4b provide means and standard deviations for each comprehension group.

Table 5.4a *Mean reading times (ms) on the target word and standard deviations (in parentheses) in all four conditions for participants in the Lower comprehension group*

	N=28	*M* / *SD*			
	Measure (ms)	**Strong/ Confirmed**	**Weak/ Confirmed**	**Strong/ Unconfirmed**	**Weak/ Unconfirmed**
Early processing measures	First fixations	208	204	206	209
		(36)	*(27)*	*(38)*	*(36)*
	First pass	236	229	241	236
		(38)	*(42)*	*(64)*	*(53)*
Later processing measures	Total time	330	323	384	368
		(93)	*(98)*	*(116)*	*(139)*
	Second pass	295	265	313	318
		(123)	*(100)*	*(104)*	*(128)*
	Go past time	344	304	325	341
		(168)	*(95)*	*(121)*	*(183)*
	Measure (probability of regressing)				
	Regressions in	0.18	0.2	0.27	0.26
		(0.14)	*(0.16)*	*(0.19)*	*(0.2)*
	Regressions out	0.27	0.18	0.21	0.15
		(0.21)	*(0.17)*	*(0.18)*	*(0.17)*

Table 5.4b *Mean reading times (ms) on the target word and standard deviations (in parentheses) in all four conditions for participants in the Higher comprehension group*

	N=20	*M* / *SD*			
	Measure (ms)	**Strong/ Confirmed**	**Weak/ Confirmed**	**Strong/ Unconfirmed**	**Weak/ Unconfirmed**
Early processing measures	First fixations	220	212	210	225
		(35)	*(31)*	*(31)*	*(34)*
	First pass	269	248	243	250
		(53)	*(61)*	*(53)*	*(49)*
Later processing measures	Total time	340	347	345	372
		(67)	*(110)*	*(66)*	*(106)*
	Second pass	308	291	278	303
		(101)	*(120)*	*(108)*	*(135)*
	Go past time	347	346	309	365
		(109)	*(181)*	*(88)*	*(163)*
	Measure (probability of regressing)				
	Regressions in	0.11	0.17	0.18	0.21
		(0.13)	*(0.13)*	*(0.11)*	*(0.17)*
	Regressions out	0.17	0.14	0.16	0.17
		(0.15)	*(0.12)*	*(0.2)*	*(0.17)*

In terms of early processing, there were no significant main effects for the within-subject factors (prediction, context) or the between subject factor (comprehension group) in first fixation durations (all *Fs* <1). The interactions between the factors, too, were not significant (*Fs* <1). First pass reading times showed a marginally significant interaction between Comprehension group and prediction type ($F(2, 46)= 3.7$, $p=.06$, $\eta^2= 0.07$): in the confirmed conditions, the Lower comprehension group's first pass times for the target word were shorter than those of the Higher comprehension group (Figure 3). While this difference is interesting, it is only marginally significant and thus needs to be treated cautiously. Other main effects and interactions were not significant for first pass reading times (*Fs* <1).

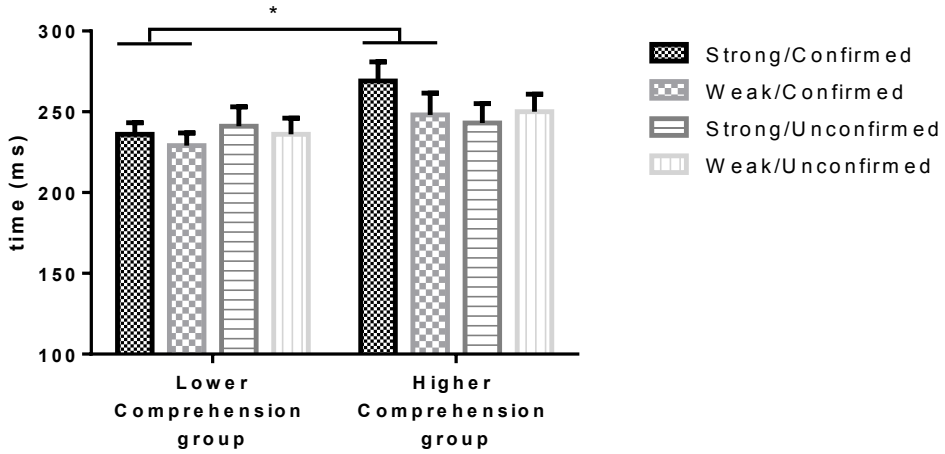


Figure 3 *Mean first pass reading times (ms) for participants in Experiment 4 by comprehension group. Participants in Higher comprehension group took longer to read the target word in the confirmed conditions than those in the Lower comprehension group. *p= .06*

Similarly, total reading times (Figure 4a) revealed a significant interaction between prediction and Comprehension group (F (2, 46)= 4.4, p<.05, η^2= 0.08) showing that the Lower comprehension group spent significantly less time reading the target word in the confirmed conditions than in the unconfirmed conditions. However, the Higher comprehension group did not show a significant difference between the time they took to read the target word in the confirmed and unconfirmed conditions (p > .05). The Higher comprehension group thus displayed a more uniform pattern of target word processing across the four conditions than the Lower comprehension group; spending more time reading the confirmed conditions and less time reading the unconfirmed conditions than the other group.

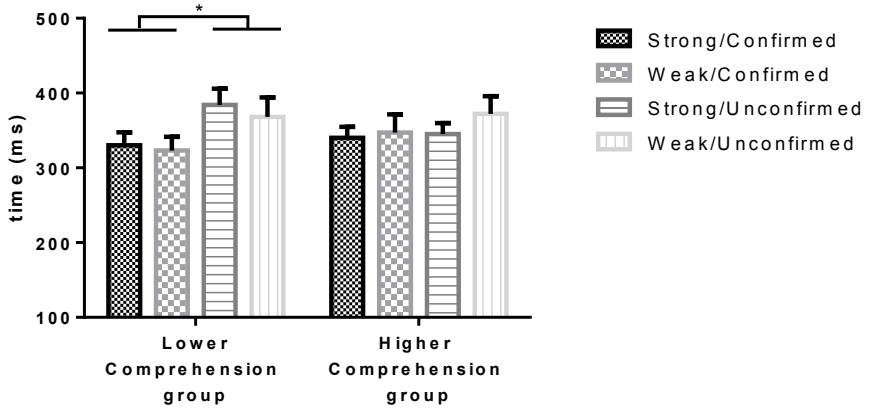


Figure 4a *Total reading times (ms) on the target word for participants in Experiment 4 by comprehension group. Participants in Lower comprehension group took longer to read the target word in the unconfirmed conditions than in the confirmed. Error bars indicate SEM. *p< .05*

For second pass times, go-past times and regressions out of the target word there was no main effect of Comprehension group and no interactions between this and the other factors . However, regression back into the target word showed an interesting trend (Figure 4b). There was a trend towards a main effect of group ($F(1, 47)=3.3$, $p=.07$, $\eta^2=0.06$), suggesting that the Lower comprehension group made more regressions back into the target word than the Higher comprehension group. Interestingly, it appears that Higher comprehension group was slower when processing the 'confirmed' conditions than the Lower comprehension group yet showed fewer processing costs associated with processing in the 'unconfirmed' conditions.

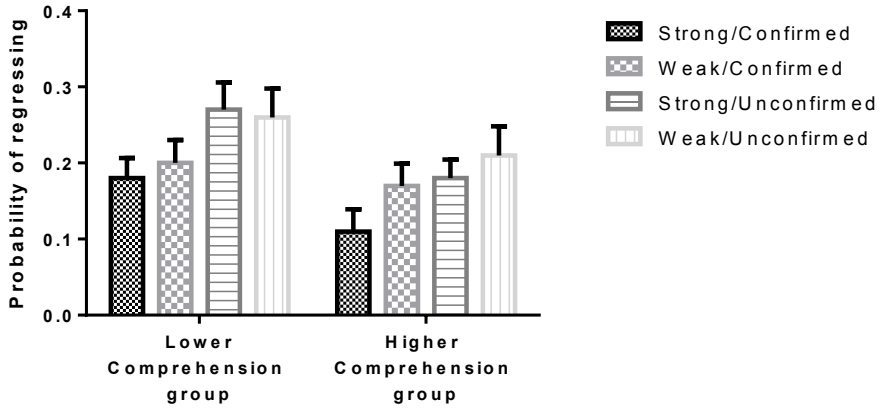


Figure 4b *Regressions back into the target word for participants in Experiment 4 by comprehension group. Participants in Lower comprehension group showed a trend towards regressing to the target word more frequently than the other group. Error bars indicate SEM.*

Although speculative, taken together these findings thus suggest that the Higher comprehension group potentially engaged in deeper overall processing than Lower comprehension group and implies that the groups might use slightly different strategies to integrate new information during reading.

Overall, the results of Experiment 4 thus largely supported the view that skilled readers generate predictive inferences as they integrate new information into their textual model and use bridging inferences when their initial predictions are not met. The computation of bridging inferences was reflected in increased rate of re-reading after the initial prediction had not been met.

Discussion

Experiment 4 sought to establish whether people used predictive rather than bridging inferences when integrating new concepts into their text representation in contexts similar to those used in the previous experiments in this thesis. Furthermore, it aimed to elucidate which textual characteristics allow readers to generate predictive inferences and whether even indirect indications of future text development suffice to encourage generation of predictive inferences. Experiment 4 demonstrated that readers routinely generated predictive inferences when integrating new information into their text representation during reading. The results revealed that people engaged in bridging inferences only after they failed to integrate the new information through predictions.

I will discuss the results in relation to a broader question of whether readers integrate new information in written texts predominantly by predictive or bridging inferences. First, I will re-state the predictions I spelled in the Introduction and evaluate how the results from the target and post-target word fit with them. I will then specifically compare the results for the 'strong confirmed' and 'weak confirmed' condition because they further explore the manipulation used in previous experiments in this thesis. Afterwards, I will interpret the pattern of current results with reference to a 'good enough' processing account. This will be followed by discussing the rather surprising processing displayed in the conditions with strong context. Finally, reading patterns of readers with higher vs. lower comprehension scores will be contrasted arguing that readers with good comprehension are less likely to show pattern similar to 'good enough' processing.

In the Introduction, I hypothesized that if readers integrate new information during reading primarily through the use of predictive inferences, the target word in the 'prediction confirmed' conditions will be read faster than in the 'prediction

unconfirmed' conditions. This is because in the unconfirmed condition, readers will have generated predictions based on the previous context yet the target word would not confirm them. To integrate the target word with their current text representation, readers would be required to make a bridge to the previous content in order to maintain coherence. This would result in longer reading times for the target word compared to the confirmed conditions, where coherence is pre-established by the prediction. I also hypothesized on the basis of the results presented in the previous chapters that readers generate predictive inferences even from the weak indications given in the previous contexts. If this is the case there should be no difference between the 'strong confirmed' and 'weak confirmed' conditions. Furthermore, if the predictive inferences generated by readers were well-formulated and specific, I expected to see the effects in early processing measures.

The results largely followed these initial predictions. I found that readers took longer to read the target word in the 'unconfirmed' than in the 'confirmed' conditions. Interestingly, there were no differences between the conditions in the early processing measures: readers spent the same amount of time reading the target word across all four conditions on first encounter. The differences observed in the target word total reading times thus came from the additional processing after the target word had been read for the first time. Total reading times for the target words were longer in the 'unconfirmed' conditions because the word was re-read more often in these conditions than in the 'confirmed' conditions. This was especially evident in the higher number of regressions made from other parts of the text back to the target word in the 'unconfirmed' conditions. There are two implications of these findings. First, in sentences like these, readers use predictive inferences as a default mechanism for word integration, turning to bridging inferences when their predictions fail. Second, these predictions are rather

general and underspecified, indicating shallow-but-fast text processing strategies akin to so called 'good enough' processing which I will discuss later.

Let us discuss the first implication that predictive inferences act as a default mechanism for word integration with bridging inferences only coming into play when predictions fail. If readers used bridging inferences as a default for integrating new information into their text representation, their target word reading times would be a function of how easy it is to establish a connection between the previous context and the new word. Thus, reading times for the target word would show a stepwise pattern with 'strong confirmed' condition being read the fastest , followed by the 'weak confirmed', 'weak unconfirmed' and 'strong unconfirmed' conditions respectively.

Similarly, there would be little reason to expect a marked difference between the 'confirmed' and 'unconfirmed' conditions because the differences between individual conditions would be gradual. However, if readers activated associated concepts that may come up next, I would expect them to show a difference between the 'confirmed' and 'unconfirmed' conditions. This is because the target word in the unconfirmed condition is not the concept they are expecting, thus violating their expectations. To retain cohesiveness of their mental text representation, readers would need to compute bridging inferences in order to integrate the new concept, resulting in extra processing time. This is exactly the pattern seen in the data.

Secondly, the results imply that predictions generated during reading are rather general and underspecified. This is based on the fact that that the four conditions did not differ in their initial reading times; only in the additional processing measures were differences seen in the 'unconfirmed' conditions. This suggests that the predictions formed initially were probably rather weak and underspecified; containing a number of

potential fitting candidate words. For example, a specific prediction for a sentence such as (10) would be the word 'doctor', yet an underspecified prediction would include a number of candidate words, such as 'his mother, an ambulance, a doctor..'.

(10) The boy was ill and so the teacher decided to call…

If the predictions participants generated were more concrete, I would have seen the differences in the early processing as I know specific predictions are formed in highly constraining contexts (Ashby, Rayner & Clifton, 2005; Duffy & Keir, 2004; Koornneef & Van Berkum, 2006; Oakhill, Garnham & Reynolds, 2005). A good example of this is the expectation of a character's gender based on their stereotypical role (e.g. doctor, nurse, electrician, etc). Duffy and Keir (2004) found that when reading texts that presented unexpected gender pronoun for the role such as in 'The electrician taught herself…', first pass times for the word 'herself' were elevated compared to the conditions when the previous context explicitly stated that the electrician is female.

Similarly, when Ashby et al. (2005) looked at differences in reading times for readers reading highly contextually constraining sentences as in (11) versus low contextually restraining sentences as (12), they found shorter first pass times for the target word 'spoon' in the highly contextually restraining condition (11).

(11) John stirred the soup with a broken spoon.

(12) John scraped food from his dinner spoon.

Thus, the predictions our readers formed must have been rather general as otherwise the first pass times would be shorter in our 'confirmed' conditions too.

A corresponding pattern was found for the post-target word region too, showing a spillover of processing costs associated with the unconfirmed conditions. Usually, the

spillover effect manifests through longer first fixations or first pass times for the post-target word immediately following the target word (Koorneef & Van Berkum, 2006; Rayner, 1998; Staub & Rayner, 2007). Interestingly, in this case the spillover effect was seen in later processing measures only; the total times for the 'unconfirmed' conditions and the go-past times for the 'weak confirmed' and 'strong confirmed' conditions. This suggests that readers did not register integration difficulties immediately after they read the target word, otherwise early measures on the post-target word would show it. Instead, they only realized the problems after they acquired later information in the sentence. At that point, they returned to the target word and re-read it, along with the immediately following content. In fact, the average proportion of regressions back into the post-target word was higher than into the target word.

The post-target region was thus re-read more frequently than the target word itself, regressing back to it from later parts of the text. This is not too surprising: only the sentential content *after* the target word provided information necessary for target word integration. What this suggests is that target word integration difficulties only became apparent after readers gained more information. It seems that their strategy for cases when they cannot immediately integrate a new concept because it does not fit with their immediately available predictions is to 'wait' till they acquire more information. Only when this happens, do readers fully 'realize' the integration difficulties and show additional processing, in order to complete the integration of the new concept. Thus, unless the new concept is in obvious conflict with expectations, readers put the integration process 'on hold' until they acquire extra information and then return to complete it.

Another interesting finding was the lack of difference between the 'strong confirmed' and 'weak confirmed' conditions. The 'strong confirmed' condition gave a

highly constraining context for the inference necessary to integrate the target word. Readers were thus easily able to generate predictions which contained the target word concept with regard to what will happen next. For example, upon reading that Allen's baby was ill and taken to the nearby infirmary, the concept of the target word 'hospital' becomes readily activated and is integrated easily, as evidenced by little additional processing. The fact that equivalent processing was seen in the 'weak confirmed' condition (which only provided a weak indication for the concept of hospital) suggests that as skilled readers read, they activate all potential concepts associated with the context they just read, even if the context points towards them rather weakly. Thus, it is no longer surprising that Experiments 1-3 failed to find an effect of increasing inference-making demands, since even the most demanding condition provided readers with enough information to comfortably generate a predictive inference.

There was, however, a difference between how often the target word was re-read in the 'strong' and 'weak' confirmed conditions. Despite not reaching significance, there was a trend for more regressions back in weak confirmed condition. Moreover, go-past times for the post-target region in this condition were significantly longer than in the 'strong confirmed' condition. This means that even though the weak indication provided enough information for readers to activate the target word concept, they needed additional 'validation' for their inference after they read the target word by looking back into it again from the immediately following region.

Good enough processing

On the basis of presented evidence, it appears that during reading predictions are formed about what might potentially come up next in the text by activating any concepts associated with the current content. However, these predictions appear rather

general and underspecified. If the new concept does not seem to fit immediately with the previous content, readers do not show processing disruptions right away but 'wait' until they get further information. Only then do they carry out corrective processing to maintain cohesion of their representation. In a sense, readers seem to integrate new information in a fast-yet-shallow way. So long as the information is locally coherent, they do not register immediate integration difficulties and only do so when they acquire further information - in our case, when they complete reading of the target word sentence.

Such locally-focused heuristics for integration of new information resonates with 'good enough' accounts of language processing. Proponents of 'good enough' language processing argue against the view that people process language by building a detailed, complete and accurate representation of the input. Instead, they propose that people process language in a 'good enough' way by creating representations that are sufficiently detailed for the current purpose of comprehension; often, underspecified representations are sufficient (Ferreira & Patson, 2007; Sanford & Graesser, 2006). For example, if a mother asks 'Did my kid climb the tree?' and the teacher replies 'Every kid is up a tree right now!', she does not need to create a precise and complete representation of the teacher's utterance to obtain information relevant for her question. People thus process language in a way which allows them to achieve understanding in a most economical way – this means using fast heuristics with the highest possible chance to deliver mostly correct understanding. Inevitably, this will lead to a trade-off between speed and accuracy, and result in a quick-yet-rather shallow processing (Ferreira, Bailey & Ferraro, 2002).

A good illustration of 'good enough' processing is our impaired ability to notice text anomalies if the anomaly is a good fit for the text globally (Erickson & Matteson,

1981; Barton &Sanford, 1993) Imagine you read a question 'How many of each type of animal did Moses take on the ark?' If you responded 'two', you have like most people failed to notice that the question itself is wrong since it was Noah who sailed the ark and not Moses (Erickson & Matteson, 1981). In this case, people understand the question well enough to realize that they are being asked about a biblical figure who played a role in the story about the Great Flood, yet not well enough to realize the difference between Noah and Moses. This is because 'good enough' processing enables people to create a match in their representation of the Great Flood as a biblical event and Moses as a biblical figure, eliminating conflict between existing representation and incoming information.

Similarly, processing of garden-path sentences such as (13) provides support for good enough processing because readers routinely create a representation of Anna bathing the baby, even though they understand the correct meaning, that Anna in fact bathed the baby (Christianson, Hollingworth, Halliwell & Ferreira, 2001).

(13) While Anna bathed the baby played in a crib.

Similarly, readers often wrongly recall standard sentences which present them with unusual situations such as "the dog was bitten by the man" and create a representation of man being bitten by a dog (Ferreira, 2003). Readers in our study might use similar heuristics during their initial reading when they do not immediately realize that new information is difficult to integrate since it does not obviously clash with their 'good enough' representation. Only after the readers fully acquire the concept introduced by the target word and move to subsequent content at which point target word might become more difficult to integrate in their 'good enough' representation, they engage in

deeper re-processing of the text. This would explain why our readers only reacted to increased processing demands for the target word in a delayed fashion.

The strong context effect

Rather surprisingly, strong context produced a higher number of regressions out of the target word, and a spill-over effect of more frequent regressions back into the post-target word. This suggests that target word integration was more demanding in the strong contexts than in the weak context. This makes sense for the 'strong unconfirmed' condition: the previous context gives no cues for the target word and even suggests a different direction; however, it makes little sense for the 'strong confirmed' condition which otherwise appears to make the target word integration easy. However, strong contexts will in general lead readers to create more detailed text representation and such representation may incite them to re-read previous content when they encounter new information. If the text representation is so strong and detailed, readers might be more compelled to check if the new concept is a good fit. If the representation is less detailed, readers might be more willing to incorporate new information without checking and process the text in a good enough way.

A similar effect was seen in Ashby et al. (2005) where readers regressed more often from the low frequency target words in the highly predictable condition than in the unpredictable conditions. They reasoned that highly predictable conditions biased readers towards expecting the target word and the regressions back reflected readers' quick confirmation that their expectations were correct. The fact that regressions out of the target word in the 'strong confirmed' condition in Experiment 4 did not translate into longer overall reading time yet they did in the 'strong unconfirmed' condition supports this explanation for our data too. On this view, in the 'strong confirmed'

condition, readers used regressions from the target word provided a quick check that their prediction was correct and once confirmed, they did not dwell on the target word anymore. In the 'strong unconfirmed' condition, however, readers checked the previous content and found that it did not fit well with the target word. Consequently, they spent more time re-reading the target word itself to complete the integration.

Differences between readers with higher and lower comprehension scores

The rate of re-reading was not the same for all individuals in our sample. Participants were assigned into a Higher or Lower comprehension group depending whether their comprehension accuracy was above the median value. When the two groups were compared in their reading pattern, only the group with a lower comprehension score showed a difference between the 'confirmed' and 'unconfirmed' conditions in target word total reading times. Interestingly, the higher accuracy group showed similar target word reading times across the four conditions. Compared to the lower comprehension group, their first pass reading times were a little longer overall, and significantly longer for the 'confirmed' conditions. Similarly, their total target word reading times were longer in the 'confirmed' conditions yet a little shorter in the 'unconfirmed' conditions, although this difference was not significant. They also showed a strong trend towards making fewer regressions back into the target word in general.

This seems to suggest that those with higher comprehension accuracy process texts more deeply and are less likely to engage in good-enough processing than those with lower comprehension accuracy. Participants in the Lower comprehension group showed a typical good-enough reading pattern; they generated underspecified predictions that in the confirmed conditions resulted in fast integration of the target

word, yet in the unconfirmed conditions caused bigger processing costs as they did not fit with the new content. On the other hand, the reading pattern of Higher comprehension group suggested that they processed text more thoroughly. The differences in their first pass reading times indicated that they did not fall back as much on heuristics of quickly activating associated concepts. Thus, participants from Higher comprehension group might have been less likely to generate predictions about further development and rather integrated information as it arrived, possibly through bridging inferences. Because of that, they found it easier to integrate concepts that did not fit immediately with the preceding context.

We could thus say that skilled comprehenders are less likely to rely on quick processing heuristics and instead monitor both local and global text coherence better (Christianson, Williams, Zacks & Ferreira, 2006; Hannnon & Daneman, 2001; Sanford & Graesser, 2006). Hannon and Daneman (2001) found that skilled readers were more likely to notice simple anomalies such as 'tranquilizing stimulants' during reading than less skilled readers; less skilled readers were more likely to miss these inconsistencies, as long as they were not central to overall meaning. Similarly, Christiansen et al. (2006) found that elderly readers were more prone to mis-interpret garden-path sentences as in (12) to create a representation where Anna bathed the baby than younger readers. Christiansen et al. explained this as a result of compensating their weaker working memory by using processing heuristics. Such heuristics say that when reading phrases such as 'While Anna bathed the baby…' the baby is more likely to be the object of the action rather than an agent of separate action. While such processing might in most cases be good enough to serve the comprehension purpose, it brings on problems in more demanding texts.

Conclusion

I have argued that people generate predictions about potential future developments as they read by activating concepts associated with the current content, on the basis that these might be relevant later. Evidence for this was provided by shorter reading times and fewer regressions back into the target word in the 'confirmed' conditions. When the new concept introduced by the target word did not fit with these predictions, information was instead integrated via more costly bridging inferences. These inferences are computed after the target word has been initially read, as evidenced by the high rate of re-reading in the 'unconfirmed' conditions. Thus, it appears that readers use predictive inference making as a default integration strategy, at least while reading sentences like those in Experiment 4. However, the predictive inferences I described here are different to the fully formed predictions usually referred to in the literature, where violations of predictions manifested during first pass reading times already, as in the case with violations of gender stereotypes for example (Duffy & Keir, 2004).

The predictions generated during this experiment might usefully be classed as underspecified. This is because there was no difference in initial target word processing between the confirmed and unconfirmed conditions. The difference between them only emerged later on in processing when readers corrected their expectations by additional re-reading. Such text processing is akin to good-enough processing where heuristics create representations that are sufficient for a reader's current needs: they are not necessarily detailed and veridical representations of the text. Only when further information was provided in unconfirmed conditions was it necessary to engage in a detailed re-analysis of the text.

Interestingly, comparison of reading pattern between readers with higher and lower comprehension skills suggested that less-skilled readers were more likely to generate underspecified predictions, thus engaging in more 'good enough' style processing heuristics. The higher skilled comprehenders displayed a lower use of predictive inferences and integrated the target word once they arrived to it, probably by generating fast bridging inferences. Importantly, since the skill group distinction was only based on splitting our sample around the median value, it is difficult to truly differentiate two groups on the basis of comprehension skill. Moreover, our readers came from Oxford University student population which has an extensive practice in reading and engaging deeply with texts they read. Future work should thus focus on exploring the use of bridging and predictive inferences in populations selected on the basis of a broader range of comprehension skills.

Chapter 6: General Discussion

This final chapter provides a summary of findings from the research conducted for this thesis and their implications for inference making literature. It discusses how contextual properties affect the choice of inference-making strategy implemented in word-to-text integration, and its time course, focusing on predictive and bridging inferences. Implications for theories of text-processing are drawn; I argue that my data provides support to slightly modified two-stage text processing accounts. The chapter concludes by discussing limitations of the research presented in this thesis and with suggestions for future work.

Research presented in this thesis examined how skilled readers integrate new concepts in their text representation through inference-making. I was particularly interested in bridging and predictive inferences because they are crucial for good reading comprehension (Cain, Oakhill, Barnes, & Bryant, 2001; Cain & Oakhill, 1999; Oakhill, 1984). My experiments explored how prior contextual information affected the type of inferences made and their time course during word-to-text integration. All experiments used eye-tracking as the main method because it provides a window on inference making as it happens during reading. Experiment 1, 2 and 3 directly examined the role of supporting context, manipulating the type and availability of referents provided for target word integration. Experiment 4 aimed to establish whether skilled readers generated predictive inferences in less constraining context and which inference making strategy they prefer to use when reading. My results support a memory-based text processing account as a default text processing mechanism, if readers have good topic knowledge and the text is not demanding.

Recapitulation of Experimental Work and Main Findings

Experiment 1 examined how people integrate a new word into their text representation, building on the results of Yang et al. (2007) who found that readers used bridging inferences to do so. The prior context in Experiment 1 varied in the amount of contextual information it provided for target word integration. As the amount of supporting information decreased in Experiment 1, demands on readers' inferential reasoning increased and they had to draw more substantial inferences to integrate the target word. Since bridging inferences are seen as key to keeping the text coherent I expected that participants would use them to fill in coherence gaps in their mental text

model when integrating the target word. I hypothesized that skilled readers would take the previous context into account during target word integration; as such, target word reading times would lengthen as it became more difficult to link them with the previous content.

Participants in Experiment 1 read sentence pairs in three conditions that varied availability of target word referents in the previous context. The easiest integration condition was titled the 'control' condition because it was supposed to provide a baseline for very easy integration. The context-giving sentence in this condition provided a referent that shared its lexical and surface form with the target word, just like the word 'exploded' acts as a referent for the target word 'explosion in (1):

(1) The bomb hit the ground and exploded. The explosion was reported to the commander.

The 'paraphrase' condition provided a semantically related referent such as 'blew up', and the 'inference' condition only informed readers that the bomb hit the ground.

The results of Experiment 1 showed that readers integrated the target word significantly faster in the control condition than in the other two conditions. This difference was visible early in processing, evident in first fixations and first pass reading times. As expected, the differences in target word reading times were the most pronounced in longer overall word reading times and in participants' tendency to re-read the target word. In line with my prediction that skilled readers use previous context to make their word integration more effective, participants took less time to generate bridging inferences in contextually-supportive control condition although I failed to find clear benefits associated with semantically related referents in the paraphrase condition. The rapid integration seen the control condition sparked my further interest because it

suggested that skilled readers might integrate words into text representation and monitor the difficulty of this integration very early in processing, as evidenced by reductions in initial fixation duration.

However, it was impossible to tell from the results of Experiment 1 whether rapid processing seen in the control condition reflected rapid inference when the context was very explicit, or whether lexical repetition might be driving the effect. Because the referent and the target word in the control condition matched in their lexical and surface form (e.g. 'exploded' - 'explosion'), fast processing in early looking times could just reflect an already activated lexical entry, or perceptual priming from seeing the same visual stimulus again. Furthermore, some items in the paraphrase condition in Experiment 1 presented readers with synonyms rather than referents that were contextually-related yet not synonymous with the target word. Because of these two problems, it was necessary to treat the result of Experiment 1 with caution.

To verify the effect of supporting context on generation of bridging inferences and to address these two issues, I added an extra condition to Experiment 2. This condition was titled the 'synonym' condition and it presented readers with a synonym as a target word referent instead of the lexically identical referent used in the control condition. The paraphrase condition was altered to only provide contextually related referents and not synonyms as was the problem in Experiment 1. I reasoned that if the fast target word processing seen in the early processing measures in Experiment 1 resulted from support from prior context, the synonym and the control condition should show the same pattern. If however it resulted from lexical repetition, target word reading times in the synonym condition should be longer than in the control condition.

The results of Experiment 2 did not show a difference between the synonym and the control condition. However, it was impossible to interpret this finding with regard to the initial predictions because Experiment 2 did not replicate any effects seen in the Experiment 1: there was no difference in the target word reading times across conditions. There are two ways to interpret this null effect: either participants did not draw inferences while they read the target word, or they made inferences equally easily in all conditions, meaning there was no measurable effect of facilitation from the availability of additional context. Both interpretations are inconsistent with data from Experiment 1, and previous literature (McKoon and Ratcliffe, 1992; Yang et al. 2006, 2007). The results of Experiment 2 were puzzling especially because out of the 36 items, 25 items were identical to those used in Experiment 1 in the inference and control conditions. Extensive comparative analysis was carried out to identify reasons for the differences between the two experiments but left no clear explanation. When target word reading times for these 25 items were compared across the two experiments, it transpired that reading times in the inference condition were significantly longer in Experiment 1 compared to Experiment 2 despite the fact that the sentences and target words were identical.

Thus, it is clear that at least one of the two experiments provided a spurious view of how people use previous context to generate inferences when they integrate new words into their text representation. A number of methodological problems were identified during the comparison of Experiments 1 and 2. First, the stimuli in both experiments showed a number of problems; items in the inference condition did not always require an inference related to the context-giving sentence to integrate the target word; the number of words separating the explicit referent and the target word varied to an extent that made it difficult to distinguish the effects of lexical repetition, and for

some items in the synonym condition, some referents were rated as unnatural in the given contexts. Second, I did not use filler items in either experiment to disguise the manipulation yet I increased the number of experimental paragraphs by a third. With the introduction of an extra condition in Experiment 2, participants might have become aware of the experimental manipulation and strategically changed their reading behaviour. Third, comprehension questions were presented on the screen together with the experimental paragraphs in both experiments. This might have resulted in different re-reading pattern as readers attempted to reanalyse the text with the question in mind (Magliano, Trabasso & Graesser, 1999; Van Der Schoot, Vasbinder, Horsley & Van Lieshout, 2008). Finally, the stimuli were rather easy in both experiments and it was possible that readers manifested the equivalent of a ceiling effect, making any differences in sensitivity to the supporting context impossible to identify.

With these limitations in mind, Experiment 3 set out to address these problems and clarify which reading pattern accurately reflects the effect of varying contextual support on inference making during reading. Apart from creating a new set of stimuli, a distance manipulation separating the target word from the context-giving sentence by additional text was added. The purpose of this was to increase processing difficulty by placing higher working memory demands on the integration. Participants might have been sensitive to the amount of supporting context provided for the word integration via bridging inferences, however, the stimuli in Experiments 1 and 2 might have been too easy and participants might have manifested a ceiling effect. In such case, all conditions would provide information sufficient to create a detailed text representation and target word reading times would be similar in all conditions.

To provide an additional test of whether participants incorporated inferred information in their text representation, an off-line instantiation task was included in

Experiment 3. The task explored if people were more likely to falsely recall seeing inferred information explicitly stated in the text in more supportive contexts.

The results of the Experiment 3 mirrored those of Experiment 2 in showing no effect of varying contextual support on target word reading times. This suggests that the pattern of results in Experiment 2 holds even after the methodological issues were addressed. In addition, there was no effect of manipulating the distance between the context-giving sentence and the target word. Interestingly, however, the instantiation task revealed that readers did not instantiate the inferred information equally strongly in all conditions, as would be expected from the on-line results. The more contextual support was available, the more likely people were to instantiate inferred information. Despite showing sensitivity to the amount of supporting context in the off-line instantiation memory task, participants' reading times on the target word did not differ across conditions during on-line reading.

There were two potential causes for this phenomenon; first, supportive information might have proved more harmful than helpful because of the 'reverse cohesion' effect. It had been observed that overly cohesive texts which provide a high amount of argument overlap and connectives between adjacent sentences can serve to decrease comprehension in skilled readers with good knowledge of the topic, relative to less cohesive texts (O'Reilly and McNamarra, 2007; McNamara & Kintsch, 1996; McNamara et al., 1996). The 'control' and 'synonym' conditions were constructed to be highly cohesive texts. Potentially therefore, they might have confused participants by emphasizing information that is not normally contained in text (the rationale being that people easily fill in the information themselves). This confusion might have consequently prolonged reading times on the target word in control and synonym conditions, leading to the pattern of data obtained in Experiment 2 and 3.

This suggestion however is difficult to reconcile with there being no effect of distance: as the intervening text in the distant condition made paragraphs less cohesive, this should have mitigated the reverse cohesion effect yet it did not. Thus, a more likely explanation was the second explanation: instead of drawing bridging inferences to integrate the target word, readers made predictive inferences. As there was no effect of context on reading times across conditions, it follows that even the least supportive 'inference' condition provided cues sufficient enough to generate a predictive inference.

Experiment 4 thus aimed to distinguish whether readers make predictive or bridging inferences when they read stimuli like those in the previous three experiments. Participants read paragraphs that varied how strongly prior context indicated the future direction of the text. The target word then either confirmed this direction or not. In conditions where prior context provided cues in a direction not invoking the target word, readers could not have generated the necessary predictive inferences; instead, integration would require a bridging inference to be made.

I predicted that if people integrated the target word through predictive inferences, target word reading times will be shorter in conditions where it confirmed the direction suggested in the context-giving sentence. To test out whether participants in previous experiments integrated the target words via predictive inferences even in the absence of strong contextual cues (i.e. the inference condition of Experiments 1-3), I predicted that conditions where target word confirmed the set out direction would manifest similar target word reading times regardless of how strong were the contextual cues in the context-giving sentence. On the other hand, if participants used bridging inferences to integrate the target word, reading times would increase in a stepwise pattern as a function of how easy it was to establish the link with the previous context.

As predicted in the scenario supporting the use of predictive inferences, the results of Experiment 4 showed shorter reading times for the target words in the conditions where target word confirmed the direction suggested in the context-giving sentence. Furthermore, both strong and weak contextual cues in the confirmed conditions elicited similar reading times. This finding, together with a lack of stepwise increase in reading times for the target words across conditions supports the proposition that readers used predictive inferences to integrate the target words in these sentences. When these could not be computed, they engaged in the construction of bridging inferences as evidenced by longer reading times for the target word and the regions preceding it, and an increased proportion of regressions back into the target word. As the confirmed conditions in Experiment 4 modelled the synonym and the inference conditions from Experiment 1-3, these results are consistent the view that readers used predictive inferences in the previous experiments.

Interestingly, the results also revealed that the predictive inferences seen in Experiment 4 were rather underspecified. This is because target word reading times were the same during initial reading across all conditions regardless of whether it confirmed the suggested text direction or not; reading times only differed in later reading measures and re-reading rate. If the inferences generated after reading of the context-giving sentence were fully specified, even initial target word reading times in the unconfirmed conditions would be longer, as it shall be clear that the predictions were not met.

Altogether, data presented in this thesis suggests that readers used predictive inferences as a default integration mechanism, at least for rather simple texts used here. Current findings thus pose an interesting challenge for the literature which questions the occurrence and use of predictive inferences, and established bridging inferences as a

primary means for word-to-text integration when people read sentences of the type tested in this thesis. I will now discuss the implications of my findings with regard to the theories of inferences making during reading.

Relationship to the theories of inference making

The inference-making literature takes a rather conservative position to predictive inferences. While bridging inferences are well established (Klin & Myers, 1993; McKoon & Ratcliff, 1992; Singer, Halldorson, Lear, & Andrusiak, 1992a), there is more debate regarding the nature of predictive inferences, with some studies failing to find them (Keefe & McDaniel, 1993; Magliano, Bagget, Johnson, & Graesser, 1993; Potts, Keenan, & Golding, 1988). As discussed in Chapter 1, the primary role of bridging inferences is to maintain coherence between subsequent sentences in a text, but they also maintain global text coherence, especially if sentence level coherence is lost (McKoon & Ratcliffe, 1992). Thus, it has been argued that bridging inferences are the crucial mechanism for word-to-text integration. While I do not contest this view, I propose that predictive inferences are also a means of word-to-text integration and that under some circumstances, they are might have priority over bridging inferences.

If readers generate predictive inferences in a fashion proposed here – automatic, underspecified and rapidly pruned if unconfirmed – the methodology used in many previous studies is not fit to capture them. Returning to the studies cited in the General Introduction, classic bridging inference experiments such as Singer et al. (1992a) can also be interpreted as showing predictive inferences. Singer et al. (1992a) compared readers' sentence reading times on the second sentence in causally linked sentences such as (2a) and control sentences (2b):

(2a) Dorothy poured the bucket of water on the fire. The fire went out.

(2b) Dorothy placed the bucket of water by the fire. The fire went out.

People were faster to read the second sentence in (2a) than the identical sentence in (2b). Singer et al. reasoned this was because participants could "causally bridge" in (2a). However, their results can be equally interpreted through prediction rather than bridge. It is highly possible that reading about water being poured on the fire caused readers to generate a predictive inference that the fire went out. Thus, faster reading times on the second sentence in (2a) could have reflected prediction rather than a causal bridge.

Another study (Singer, Andrusiak, Reisdorf & Black, 1992b) presented readers with sentences such as (3):

(3) Carol found that the milk was three weeks old. The smell turned her stomach

They argued that since the first sentence in (3) offered no referent for the word 'smell' , readers had to construct a bridge back to the previous content by drawing on their world knowledge that three-week old milk turns bad, and bad milk smells. Evidence for this bridging inference being drawn came from the observation that people were as fast to answer an inference question ('Did the milk go bad?') as a factual or literal question such as 'Did Carol find the milk?' However, this explanation does not rule out a possibility that readers generated a prediction for the concept of smell when reading about the milk being three weeks old in the first sentence. Similarly, St George et al. (St George, Mannes, & Hoffman, 1997) recorded ERP potentials for the sentence *'It was too bad the turkey burned'* when it was preceded either by (4a) or (4b). They found that the N400 peaks were smaller following sentences like (4a):

(4a) Pam forgot about the turkey in the oven.

(4b) Pam put the turkey in the oven.

They explained the reduction in N400 by the ease of making the bridge between the sentences afforded by (4a). Yet, the result can also be explained by readers pre-activating their world knowledge that things get burned when forgotten about in the oven.

Methodological issues also cloud how to interpret some of the previous findings about predictive inference making as already mentioned in Chapter 1. For example, Potts, Keenan and Golding (1988, Experiment 1) asked, participants to read what they titled predictive sentences such as (5a) and coherence sentences such as (5b).

(5a) No longer able to control his anger, the husband threw the delicate porcelain vase against the wall. He had been feeling angry for weeks, but had refused to seek help.

(5b) No longer able to control his anger, the husband threw the delicate porcelain vase against the wall. It cost him well over one hundred dollars to replace the vase.

Subsequently, participants performed a naming task for the target word 'broke'. Potts et al. found that participants were to name targets faster in the coherent condition (5b) than in the predictive condition (5a). The problem with this methodology was that naming times were measured after the second sentence was presented. While the second sentence further reinforced the target word action ('broke') in the coherence condition (5b), it did not reinforce it in the predictive condition (5a). Since predictive inferences are pruned quickly if the following context does not confirm them, it is not surprising

that Potts et al. did not find speeded responses to the target word when the text took a different direction, as in (5a). Had they measured naming times immediately after the target sentence, they might have found evidence for predictive inferences as Keefe and McDaniel (1993) did when they replicated Potts et al.'s study.

Thus, the prominent position of bridging inferences over predictive inferences for word-to-text integration is therefore less straightforward than would seem initially. If readers are given a chance to activate associated concepts in the form of predictive inferences, they will do so automatically provided their topic knowledge is high and the text is sufficiently easy. Data in this thesis are consistent with this view and suggest that even relatively weak indications are sufficient for predictive inferences to be generated. I will now discuss the implications of my findings with respect to the theoretical accounts of general text processing.

Relationship to general text processing accounts

Such a mechanism for predictive inferences as described in the previous paragraph fits well with theoretical accounts of text processing that emphasize automatic activation of concepts through resonance of contents currently in reader's active memory with concepts in episodic and semantic long-term memory (Myers & O'Brien, 1998). When reading text, skilled readers activate concepts associated with the current content, as manifested in the results presented in this work. If the current content provides indications of how the scenario described in the text might develop, skilled readers may generate predictions about what might come later in the text. Such a mechanism would allow them to integrate new concepts efficiently as the concepts would be pre-activated in their mental representations through predictive inferences.

However, this mechanism seems to be automatic and rather dumb (Cook et al., 2014). It appears from Experiment 4 that predictions readers generate are underspecified, short-lived and pruned quickly if they are not confirmed by the future context. Such mechanism for word-to-text integration would provide support to heuristics-driven and memory-based text processing accounts.

Two of the most influential views on text processing are the memory-based view (Gerrig & McKoon, 1998; Myers & O'Brien, 1998) and the explanation-based view (Greasser, Singer & Trabasso, 1994). I will first review the memory-based text processing account. The memory-based view of text processing focuses on the influence of bottom-up components in reading. It proposes that much of comprehension relies on lower-level automatic memory retrieval processes such as resonance (Myers & O'Brien, 1998). According to the resonance model, when readers encode new concepts in working memory these concepts generate a signal to long-term memory, reaching both memory of the previous context and general world knowledge. Long-term memory concepts which share features in common with the current content of working memory will "resonate", and those with the highest resonance will be activated and briefly incorporated into the text representation.

This activation process is considered to be unguided, nonstrategic, and in a sense dumb because any information with strong enough resonance will be activated, regardless of its global fit with the text (Cook et al., 2014; O'Brien, Rizzella, Albrecht, & Halleran, 1998). Because of this, readers will activate a large amount of information, much of which may not be useful to the reader. Thus, this information will be quickly extinguished unless future content confirms it (Cook, Limber & O'Brien, 2001; Keefe & McDaniel, 1993). In a memory-based processing account, therefore many inferences

that readers will generate are passive as they arise as a result of an automatic memory resonance process.

In contrast to the memory-based approach, the constructionist or explanation-based view (Graesser, Singer & Trabasso, 1994) argues that people read under a principle of 'search after meaning'. Thus, both automatic and controlled strategic processes interact to influence on-line processing. In the explanation-based view, inferential processes are active and strategic rather than automatic and passive and involve a directed, resource-demanding search of long-term memory to locate specific information (Long, Oppy & Seely, 1996). On this view, when a new concept enters an individual's working memory, it will be integrated in their text representation by checking how this piece of information fits with actions, events and states mentioned in the text, rather than by the automatic resonance process described in the memory-based approach outlined above. While the memory-based view focuses on the bottom-up components involved in reading, the explanation-based view accentuates the top-down components. A large body of evidence supports this view and shows that readers engage in directed and strategic processes such as tracking the causes of events (Trabasso & Van den Broek, 1985), attending to goal information (Long, Oppy & Seely, 1996; Lutz & Radvansky, 1997) and holding different standards of coherence and normality depending on the context and purpose of reading situation (Filik, 2008; Van den Broek, Lorch, Linderholm, & Gustafson, 2001).

Data from this thesis support the memory-based processing account provided certain conditions are met, which I will discuss in more detail later. Because readers integrated the target word equally quickly regardless of the previous context in Experiments 2 and 3, and in relevant parts of Experiment 4, it appears they used predictive inferences to do so; that is, they generated these inferences even when prior

context provided few direct referents to the target word. As the scenarios used in our stimuli were quite common, it seems that even weak indications were sufficient to create adequate resonance with the world knowledge of readers, such that the target word concepts were activated (i.e., 'ill baby taken away' activated 'hospital') and thus generated as predictive inferences.

Had readers integrated the target word strategically using all information available to them prior to the reading of the target word, I should have seen a stepwise increase in target word reading times as the amount of supporting context decreased. Since there were no differences, it seems that readers did not use constructive, strategic processing when integrating the target words in Experiments 2 and 3. The case is somewhat different however for Experiment 4. Here I saw clear evidence of re-processing and additional reading in the unconfirmed conditions. Such a pattern suggests strategic reading as readers spent longer time reading the target word and the following region and returned to these regions repeatedly when they did not automatically fit into readers' text representation.

Thus, participants in my experiments engaged in constructionist processing too, but it appears that their default processing was more in line with memory-based accounts of text processing. On the basis of this evidence it seems that readers use both memory-based processing and explanation-based processing, but use them at different times and for different purposes, potentially as a function of text difficulty and comprehension skill; I return to these points later. With regard to each text processing strategy being used at distinct time, the data in this thesis support Cook's framework of two-stage processing (Cook & Myers, 2004; Cook & Gueraud, 2005). Cook and colleagues proposed that comprehension is achieved via a two-stage process, similar to the bonding and resolution model proposed by Sanford and Garrod (2005). In the first

stage, the contents of active memory become linked with information automatically reactivated from long-term memory. This information may be semantic or episodic (relating to earlier portions of the text), depending on the degree of featural overlap between just read active concepts and concepts in long-term memory, and both types can become activated (Rizzella & O'Brien, 2002). The concepts which become activated will serve as antecedents in rapid automatic bridging inferences, such as 'car' for the pronoun 'it' in (6a), or as predictive inferences, such as 'stain' in (6b).

(6a) A bright red car was parked in the backyard. It looked brand new.

(6b) Not paying attention to what she was doing, Mary spilled red wine on her wedding gown. The stain ruined her dress.

Because concepts generated as predictive inferences are already active in a reader's mental representation, they will be rapidly processed when confirmed in the text. However, when simple resonance fails to reactivate information sufficient for comprehension to happen, readers will engage in active problem-solving and explicit 'search after meaning' processes as proposed by the explanation-based view (Graesser et al., 1994).

When a text is easy; describing a common scenario and using straightforward syntax and vocabulary, and when readers have detailed and deep knowledge of the topic (stemming for instance from their world knowledge), a memory-based approach is sufficient to create a good enough text representation for the comprehension purposes. However, when the text is difficult, describing for example an unusual scenario that gives little indication of future development, uses demanding syntax or rare vocabulary, poses high demands on attention and memory, and readers' knowledge of the background is low, memory-based processing will be insufficient to deliver a rich and

coherent text representation. To achieve satisfactory comprehension, readers would need to engage in constructionist processing. This pattern fits well with data in this thesis. I suggest that when texts were easy, readers integrated the target word mostly through predictive inferences brought about by resonance processes. However, when integration became more difficult, as in the 'unconfirmed' conditions of Experiment 4 where prior context did not provide indication for the target word, readers displayed substantial additional re-reading and reanalysis in line with explanation-based processing.

The two-stage account of text processing is consistent with the idea that readers often process texts in a shallow way and create an underspecified text representation. As suggested in the discussion section of Experiment 4, normal understanding of language, whether through reading or a dialogue, often involves incomplete processing. People rarely construct precise, complete semantic analyses of texts or utterances, either because information offered explicitly is incomplete or because a precise and complete interpretation would be too effortful given the payoff (Sanford & Graesser, 2006). Instead, they complete representations which are 'good-enough'; this means they are sufficiently detailed for person's current comprehension purpose. The term 'good-enough' processing was coined especially by Christianson, Ferreira and colleagues (Christianson, Hollingworth, Halliwell, & Ferreira, 2001; Ferreira, Bailey and Ferraro, 2002) to denote quick-yet-shallow text processing which aims to deliver the most likely meaning with the lowest processing costs by using general heuristics.

Good-enough processing thus fits well with the memory-based text processing stage. In fact, the good-enough processing heuristics might use the same mechanisms as memory resonance processes, such as building on featural overlap. A typical example of good-enough processing is the 'Moses illusion' (Erickson & Mattson, 1981) where

readers do not detect an anomaly in the question "how many animals of each kind did Moses take on the Ark" because Moses and Noah share many features (Biblical figures, Old Testament, old men…). Readers notice problems with their good-enough representations only when content just read no longer fits with their representation. It is at that point when they start to actively reprocess previous information in the way described in the explanation-based view; they focus on global coherence and overall reading goal, evaluate previous information and strategically re-read previous content. This is the picture I saw in Experiment 4; the initial target word reading times in Experiment 4 did not differ across conditions, regardless of whether the target word confirmed the direction suggested by the earlier content or not. Only when readers acquired further information by reading the target word sentence to the end did they engage in additional re-processing of the target word.

Thus, readers seemed to use fast-yet-shallow processing as the default during reading and when it failed to deliver sufficient comprehension, readers reverted to explanation-based type of processing. Data in this thesis supported this view although there are limitations which should be acknowledged and addressed in future work to firmly establish memory-based processing and generation of predictive inferences as a default.

Limitations and future work

There are several factors which make it difficult to confidently apply my findings to reading in general; the texts used my experiments were easy, readers had very good general world knowledge and were highly practiced. All participants were Oxford University undergraduate and graduate students who are generally well read and their work requires them to deeply engage with written materials. Thus, it is difficult to

say how well the processing patterns captured in the data might generalize to other readers and other texts. Good comprehension skill (Cain, Oakhill & Bryant, 2000; Rayner, Chace, Slattery & Ashby, 2006; Perfetti, Schmalhoffer & Yang, 2008), good background knowledge of the topic (Chiesi et al.,1979; McNamara & O'Reilly, 1996) and text difficulty (Weaver & Bryant, 1995) are all factors that can strongly influence text processing. It is thus possible that predictive inferences are only used as a default word integration mechanism in specific instances where readers are skilled, they have good knowledge of the topic and texts they are reading are easy.

As mentioned earlier, text difficulty might modify how much readers use memory-based processing. My stimuli were very short and described common scenarios. Such texts in general do not place high demands on comprehension resources inlcuding working memory, topic knowledge, or attention and goal orientation. It is possible that readers are very economical in choosing their processing style and would thus only employ fast-yet-shallow text processing when they know the texts are easy enough to do so while still achieving reasonably good comprehension. Thus, texts presenting readers with more difficult scenarios or topics where their background knowledge is less good might promote bridging inferences rather than predictive as default for word integration.

Having good knowledge of the text background is important. It can even compensate for worse general reading skill. When less skilled readers read about topics they know well they can show comparable comprehension to skilled comprehenders. Chiesi et al. (1979) found that less skilled readers (in this case children) who knew more about baseball showed better recall of a text about a baseball game than adults who were skilled readers yet lacked baseball knowledge. Comprehension depends on a reader's activation of appropriate related knowledge while reading and linking it with

what is read through inference generation. If prior knowledge is good, inference generation becomes easier because concepts are richly connected meaning that currently active concepts will easily achieve strong resonance with concepts in long-term memory. This way, many associated concepts will be available to the reader making it easier to keeping various concepts in active memory. If a reader knows little about the topic he is reading about, they need to create the knowledge base at the same time as they are reading. Creating this knowledge base adds extra processing costs for memory, since it not only needs to store and process what is read, but use this information to build up a memory base for the topic. Thus, future work should study whether readers would still engage in memory-based processing when presented with texts from areas they are not familiar with.

Similarly, testing participants with lower comprehension skills would provide important insights into whether they use two-stage processing similarly to skilled readers. If memory-based and good-enough processing relies on the ability to quickly activate associated concepts through memory resonance, less skilled comprehenders might not show integration via predictive inferences, simply because their resonance would not be strong enough to achieve the activation threshold. Future work should thus present less skilled comprehenders with the same stimuli to see whether they would use resonance-based processing like skilled readers in our sample, or whether they would use bridging inferences instead to integrate the target words. Furthermore, I saw that skilled readers in Experiment 4 realized well when they needed to switch from memory-based text processing to explanation-based text processing. It is also possible that less skilled comprehenders might in fact overly rely on memory-based processing and they might have problems with adjusting their strategy to explanation-based processing when necessary.

Another problem of the presented work is the lack of a proper baseline condition which would establish target word reading times in a neutral setting without any previous contextual facilitation or reference, particularly in Experiments 1- 3. This is an important measure which should have been included in the experiments because it is otherwise difficult to firmly establish that any inference on the target word had been drawn. If target word reading times in neutral settings were the same as reading times in our experimental conditions, it would suggest that previous context did not have an effect on the target word integration and inferences to integrate the target word were not drawn on-line. If reading times on the target word differed across conditions this itself would be an indication that some inference had been drawn on-line to integrate the target word even though a baseline condition would still be desirable to establish this firmly. However, since the target word reading times did not differ across conditions in Experiments 2 and 3, and the results of Experiment 1 seem to be unreliable, there was no definite evidence that readers drew bridging inferences while reading target words.

All four studies presented in this thesis looked at higher-order discourse processing via measuring reading times on one word. While this is a usual method in many eye-tracking studies which examine discourse processing, it is not unproblematic. Discourse processing is a complicated process and integration of a specific concept may involve other portions of the text not just that particular word, or the one following in case of the post-target word region. However, the eye-tracking analyses used here, in common with those in the literature, focused on fixation durations and regressions to the target and post-target word and did not describe processing which happens beyond these two regions. Analysis of regressive eye-movements tell us whether readers re-read the target word or whether they regressed back to the earlier portions of the text, but it was not possible to find out which portions of text they regressed to. Similarly, I do not

know where the regressions back to the target word came from. Potentially, this information is very important. Did readers only realise the need to re-process the target word once they had finished the sentence? Or did they notice the difficulty with integration a few words down already?

Similarly, I could not identify which earlier portions of text served as cues to activate associated concepts. It would therefore be interesting to establish what this activation process looks like on-line as there is currently little information available. Think aloud protocols (Suh & Trabasso, 1993) provide some insight into when readers might activate predictive inferences, but this is rather limited because it reflects off-line processing and only captures active predictive inferences – those that readers' attention was pointed to. However, the mechanism behind predictive inferences as suggested by this thesis should operate largely via good-enough processing and passive resonance of active concepts with concepts in long term memory. Both of these processes are rapid and passive, and as such would normally escape reasoning in think aloud protocols. Eye-tracking could capture the activation process and following integration of associated concepts at it happens, but it would need to look beyond fixation durations and probabilities of regressing at the actual eye-movements *patterns* (scanpaths) across the whole paragraph. Recent analytic advances such as scanpath analysis (van der Malsburg & Vasishth, 2011; Vasishth, von der Malsburg, & Engelmann, 2013) might offer valuable insights beyond the traditional methods adopted here.

Scanpaths analyses examine readers' overall sentential processing pattern without focusing on one word only. It is thus an ideal methodology to capture reading behaviour. It examines sequences of fixations across the whole sentence and searches for similar fixation sequences in the data. However, scanpaths analysis is problematic in that it is very complex and difficult to describe. Finding the mean and variance for a set

of scanpaths is difficult and requires advanced statistical procedures. While there are R -software packages to facilitate this (van der Malsburg, Vasishth & Kliegl, 2012), the eye-tracking software used to record data in this thesis (MiniEye and EyeTrack) did not allow the scanpaths analysis to be carried out. Moreover, scanpaths analysis requires that all texts used in the experiment have uniform spatial layout. Despite technical demands, scanpaths could provide valuable insights into the time course of detecting the lack of fit between the text representation and the target word, and how readers proceed to solve such problems.

Standard eye-tracking studies aggregate data across trials in a condition and the individual fixation sequence patterns are lost. If there were certain distinctive fixation patterns reflecting different cognitive processing, it would not be possible to identify them in the aggregate data. Since readers may use a number of different re-processing strategies (van der Malsburg & Vasishth, 2011), scanpaths are an ideal tool to describe individual differences and distinct strategies adopted to complete word integration through inference making. Furthermore, comprehension skill is known to influence reading strategy (Schroeder, 2011; Stine-Morrow, Loveless, & Soederberg, 1996; Van Der Schoot et al. 2008) and scanpaths analysis could provide detailed insight how this is manifested. Work presented in this thesis, in particular Experiment 4, suggested that skilled comprehenders might be less likely to engage in fast-yet-shallow processing styles; scanpaths analysis has the potential to provide a detailed account of differences between skilled and less skilled comprehenders. Thus, future work should use scanpaths analysis to examine how readers generate predictions with relation to previous context.

Finally, the experiments in this thesis, especially Experiment 1-3, suffered from lacking in statistical power. The repeated measures design with three and especially four factor levels requires a large number of items or subjects, which is laboursome and time

intensive. Eye movements are very sensitive to fatigue which invariably ensues from reading a high number of experimental items. Thus, future research in this area needs account for these problems when coming up with experimental design as within-subject factors with many levels require a high number of participants to achieve desired power. It is also necessary to note that the effect sizes in experiments such as presented here are small: I was looking at reading times on only one word which was kept the same across conditions. Thus, the expected differences in reading times across conditions would be rather small to start with. Power differences might partially explain why Yang et al. (2006, 2007) observed different results: their participants read 120 paragraphs in the course of the experiment whereas mine read up to 40. However, they do not explain why Experiment 1 produced such different results than Experiment 2 and 3 since all experiments were similarly underpowered and looking at the effects of similar size.

Overall Conclusion

Successful integration of words into a reader's mental representation of test is the backbone of good comprehension. Without it, a mental model might be disjointed, impoverished and lacking in coherence, and as a consequence, fail to deliver good comprehension. This thesis studied word-to-text integration by investigating readers' ability to draw inferences. I was specifically interested in how previous context influenced readers' choice of integration strategy and its time course; whether they made bridging or predictive inferences, and whether the integration happened during the reading of the target word, afterwards or even before the target word had been read. Results of the experiments presented in this thesis suggested that if text is easy and readers have good background knowledge of the topic (general world knowledge in the case of my experiments), even rather weak, indirect indications of future development

are sufficient to activate associated concepts. Because the target words tapped the activated concepts, readers could integrate them easily via predictive inference-making.

These findings fit with the idea that as people read, concepts they had just read send signals to their long term memory. Any concepts in their long-term memory which share features with concepts currently active in working memory start to resonate and if the resonance reaches a threshold, these concepts become activated and available to readers in the form of predictive inferences. When the target word confirms the predictions, integration should happen rapidly and automatically, without reader's active involvement. However, if the target word concept does not confirm the predictive inference, the inference would need to be eliminated. Under these circumstances, target word integration would then became strategic and goal-directed as readers actively tried to establish the target word's relevance and relationship to the prior content. Thus, predictive inference making would serve word integration in texts which are easy, or where there is extensive prior background knowledge, but when this automatic processing fails, readers turn to process text more actively via bridging inferences.

This thesis opened with stating that a good story can only be appreciated if readers are able to comprehend it. To do so, people need to create a rich and coherent representation of what they read. Even if part of this work seems passive – through memory-based processing – it can only be completed if concepts in people's long-term memory and their current active memory connect and communicate. Otherwise, resonance will be too weak to automatically activate associated concepts. I hope this thesis provided my readers with everything they needed to create a good text representation, and that despite its length, it has been an enjoyable read.

References

Anderson, A., Garrod, S. C., & Sanford, A. J. (1983). The accessibility of pronominal antecedents as a function of episode shifts in narrative text. *Quarterly Journal of Experimental Psychology, 35*(3), 427-440.

Anderson, J. E., & Holcomb, P. J. (2005). An electrophysiological investigation of the effects of coreference on word repetition and synonymy. *Brain and language, 94*(2), 200-216.

Anderson, R. C., & Ortony, A. (1975). On putting apples into bottles—A problem of polysemy. *Cognitive Psychology, 7*(2), 167-180.

Anderson, R. C., Pichert, J. W., Goetz, E. T., Schallert, D. L., Stevens, K. V., & Trollip, S. R. (1976). Instantiation of general terms. *Journal of Verbal Learning and Verbal Behavior, 15*(6), 667-679.

Ashby, J., Rayner, K., & Clifton, C. (2005). Eye movements of highly skilled and average readers: Differential effects of frequency and predictability. *The Quarterly Journal of Experimental Psychology Section A, 58*(6), 1065-1086.

Barton, S. B., & Sanford, A. J. (1993). A case study of anomaly detection: Shallow semantic processing and cohesion establishment. *Memory & cognition, 21*(4), 477-487.

Bloom, C. P., Fletcher, C. R., Van Den Broek, P., Reitz, L., & Shapiro, B. P. (1990). An on-line assessment of causal reasoning during comprehension. *Memory & Cognition, 18*(1), 65-71.

Branigan, H. P., Pickering, M. J., & Cleland, A. A. (1999). Syntactic priming in written production: Evidence for rapid decay. *Psychonomic Bulletin & Review, 6*(4), 635-640.

Cain, K., & Oakhill, J. (2006). Profiles of children with specific reading comprehension difficulties. *British Journal of Educational Psychology, 76*(4), 683-696.

Cain, K., & Oakhill, J. V. (1999). Inference making ability and its relation to comprehension failure in young children. *Reading and writing, 11*(5-6), 489-503.

Cain, K., Oakhill, J. V., Barnes, M. A., & Bryant, P. E. (2001). Comprehension skill, inference-making ability, and their relation to knowledge. *Memory & Cognition, 29*(6), 850-859.

Cain, K., Oakhill, J., & Bryant, P. (2000). Phonological skills and comprehension failure: A test of the phonological processing deficit hypothesis. *Reading and Writing, 13*(1-2), 31-56.

Cain, K., Oakhill, J., & Bryant, P. (2004). Children's reading comprehension ability: Concurrent prediction by working memory, verbal ability, and component skills. *Journal of educational psychology, 96*(1), 31.

Calvo, M. G., & Castillo, M. D. (1996). Predictive inferences occur on-line, but with delay: Convergence of naming and reading times. *Discourse Processes, 22*(1), 57-78.

Calvo, M. G., Castillo, M. D., & Schmalhofer, F. (2006). Strategic influence on the time course of predictive inferences in reading. *Memory & cognition, 34*(1), 68-77.

Calvo, M. G., Meseguer, E., & Carreiras, M. (2001). Inferences about predictable events: Eye movements during reading. *Psychological research, 65*(3), 158-169.

Camblin, C. C., Gordon, P. C., & Swaab, T. Y. (2007). The interplay of discourse congruence and lexical association during sentence processing: Evidence from ERPs and eye tracking. *Journal of Memory and Language, 56*(1), 103-128.

Chiesi , H . L. , Spillich , G. J., & Voss , J. F. (1979) . Acquisition of domain-related knowledge in relation to high and low domain knowledge . *Journal of Verbal Learning and Verbal Behavior, 18*, 257-272

Christianson, K., Hollingworth, A., Halliwell, J. F., & Ferreira, F. (2001). Thematic roles assigned along the garden path linger. *Cognitive psychology, 42*(4), 368-407.

Christianson, K., Williams, C. C., Zacks, R. T., & Ferreira, F. (2006). Younger and older adults'" good-enough" interpretations of garden-path sentences. *Discourse Processes, 42*(2), 205-238.

Cook, A. E., & Guéraud, S. (2005). What have we been missing? The role of general world knowledge in discourse processing. *Discourse Processes, 39*(2–3), 265–278.

Cook, A. E., & Myers, J. L. (2004). Processing discourse roles in scripted narratives: The influences of context and world knowledge. *Journal of Memory and Language, 50*(3), 268-288.

Cook, A. E., Lassonde, K. A., Splinter, A. F., Guéraud, S., Stiegler-Balfour, J. J., & O'Brien, E. J. (2014). The role of relevance in activation and instantiation of predictive inferences. *Language, Cognition and Neuroscience, 29*(2), 244-257.

Cook, A. E., Limber, J. E., & O'Brien, E. J. (2001). Situation-based context and the availability of predictive inferences. *Journal of Memory and Language, 44*(2), 220-234.

Daneman, M., & Carpenter, P. A. (1980). Individual differences in working memory and reading. *Journal of verbal learning and verbal behavior, 19*(4), 450-466.

der Malsburg, T., & Vasishth, S. (2011). What is the scanpath signature of syntactic reanalysis? *Journal of Memory and Language, 65*(2), 109–127.

der Malsburg, T., Vasishth, S., & Kliegl, R. (2012, December). Scanpaths in reading are informative about sentence processing. In *24th International Conference on Computational Linguistics (p. 37*).

DeSchepper, B., & Treisman, A. (1996). Visual memory for novel shapes: implicit coding without attention. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 22*(1), 27.

Dimigen, O., Sommer, W., Hohlfeld, A., Jacobs, A. M., & Kliegl, R. (2011). Coregistration of eye movements and EEG in natural reading: analyses and review. *Journal of Experimental Psychology: General, 140*(4), 552.

Duffy, S. A., & Keir, J. A. (2004). Violating stereotypes: Eye movements and comprehension processes when text conflicts with world knowledge. *Memory & Cognition, 32*(4), 551-559.

Duffy, S. A., & Rayner, K. (1990). Eye movements and anaphor resolution: *Effects of antecedent typicality and distance. Language and Speech, 33*(2), 103-119.

Duffy, S. A., Morris, R. K., & Rayner, K. (1988). Lexical ambiguity and fixation times in reading. *Journal of memory and language, 27*(4), 429-446.

Duke, N. K., & Pearson, P. D. (2002). Effective practices for developing reading comprehension. *What research has to say about reading instruction,* 3, 205-242.

Egeth, H. E., & Yantis, S. (1997). Visual attention: Control, representation, and time course. *Annual review of psychology, 48*(1), 269-297.

Einstein, G. O., McDaniel, M. A., Owen, P. D., & Cote, N. C. (1990). Encoding and recall of texts: The importance of material appropriate processing. *Journal of Memory and Language, 29*(5), 566-581.

Erickson, T. D., & Matteson, M. E. (1981). From words to meaning: A semantic illusion. *Journal of Verbal Learning and Behaviour*, 20, 540–551.

Ferreira, F. (2003). The misinterpretation of noncanonical sentences. *Cognitive psychology, 47*(2), 164-203.

Ferreira, F., & Patson, N. D. (2007). The 'good enough'approach to language comprehension. *Language and Linguistics Compass, 1*(1-2), 71-83.

Ferreira, F., Bailey, K. G., & Ferraro, V. (2002). Good-enough representations in language comprehension. *Current Directions in Psychological Science, 11*(1), 11-15.

Filik, R. (2008). Contextual override of pragmatic anomalies: Evidence from eye movements. *Cognition, 106*(2), 1038-1046.

Forster, K. I. (1981). Frequency blocking and lexical access: One mental lexicon or two? *Journal of Verbal Learning and Verbal Behavior, 20*(2), 190-203.

Forster, K. I., & Davis, C. (1984). Repetition priming and frequency attenuation in lexical access. *Journal of experimental psychology: Learning, Memory, and Cognition, 10*(4), 680.

Frazier, L., & Rayner, K. (1982). Making and correcting errors during sentence comprehension: Eye movements in the analysis of structurally ambiguous sentences. *Cognitive psychology, 14*(2), 178-210.

Garnham, A. (2001). *Mental models and the interpretation of anaphora*. Psychology Press.

Garnham, A., Oakhill, J., & Johnson-Laird, P. N. (1982). Referential continuity and the coherence of discourse. *Cognition, 11*(1), 29-46.

Garrod, S. C., & Sanford, A. J. (1982). The mental representation of discourse in a focussed memory system: Implications for the interpretation of anaphoric noun phrases. *Journal of semantics, 1*(1), 21-41.

Garrod, S., & Sanford, A. (1977). Interpreting anaphoric relations: The integration of semantic information while reading. *Journal of Verbal Learning and Verbal Behavior, 16*(1), 77-90.

Garrod, S., & Terras, M. (2000). The contribution of lexical and situational knowledge to resolving discourse roles: Bonding and resolution. *Journal of Memory and Language, 42*(4), 526-544.

Garrod, S., O'Brien, E. J., Morris, R. K., & Rayner, K. (1990). Elaborative inferencing as an active or passive process. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 16*(2), 250.

George, M. S., Mannes, S., & Hoffman, J. E. (1997). Individual differences in inference generation: An ERP analysis. *Journal of Cognitive Neuroscience, 9*(6), 776-787.

Gerrig, R. J., & McKoon, G. (1998). The readiness is all: The functionality of memory-based text processing. *Discourse Processes, 26*(2-3), 67-86.

Gordon, P. C., & Hendrick, R. (1998). The representation and processing of coreference in discourse. *Cognitive Science, 22*(4), 389-424.

Gordon, P. C., Grosz, B. J., & Gilliom, L. A. (1993). Pronouns, names, and the centering of attention in discourse. *Cognitive science, 17*(3), 311-347.

Goswami, U., & Bryant, P. (1990). *Phonological skills and learning to read* (p. 1). London: Lawrence Erlbaum.

Graesser , AC & Kreuz , RJ (1993). A theory of inference generation during text comprehension.*Discourse Processes, 16*, 145-160.

Graesser, A. C., Singer, M., & Trabasso, T. (1994). Constructing inferences during narrative text comprehension. *Psychological review, 101*(3), 371.

Hannon, B., & Daneman, M. (2001). Susceptibility to semantic illusions: An individual-differences perspective. *Memory & cognition, 29*(3), 449-461.

Haviland, S. E., & Clark, H. H. (1974). What's new? Acquiring new information as a process in comprehension. *Journal of verbal learning and verbal behavior, 13*(5), 512-521.

Hoover, W. A., & Gough, P. B. (1990). The simple view of reading. *Reading and writing, 2*(2), 127-160.

Hulme, C., & Snowling, M. (1992). Deficits in output phonology: An explanation of reading failure?. *Cognitive Neuropsychology, 9*(1), 47-72.

Jacoby, L. L., & Dallas, M. (1981). On the relationship between autobiographical memory and perceptual learning. *Journal of Experimental Psychology: General, 110*(3), 306.

Just, M. A., & Carpenter, P. A. (1980). A theory of reading: from eye fixations to comprehension. *Psychological review, 87*(4), 329.

Just, M. A., Carpenter, P. A., & Woolley, J. D. (1982). Paradigms and processes in reading comprehension. *Journal of experimental psychology: General, 111*(2), 228.

Kalyuga, S., Ayres, P., Chandler, P., & Sweller, J. (2003). The expertise reversal effect. *Educational psychologist, 38*(1), 23-31.

Keefe, D. E., & McDaniel, M. A. (1993). The time course and durability of predictive inferences. *Journal of Memory and Language, 32*(4), 446-463.

Keenan, J. & Kintsch, W. (1974) The identification of explicitly and implicitly presented information in W. Kintsch (Ed.), *The representation of meaning in memory*, Erlbaum, Hillsdale, NJ, pp. 153–176

King, J., & Just, M. A. (1991). Individual differences in syntactic processing: The role of working memory. *Journal of memory and language, 30*(5), 580-602.

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

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

Klin, C. M., & Myers, J. L. (1993). Reinstatement of causal information during reading. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 19*(3), 554.

Koornneef, A. W., & Van Berkum, J. J. (2006). On the use of verb-based implicit causality in sentence comprehension: Evidence from self-paced reading and eye tracking. *Journal of Memory and Language, 54*(4), 445-465.

Kuperberg, G. R., Paczynski, M., & Ditman, T. (2011). Establishing causal coherence across sentences: An ERP study. *Journal of Cognitive Neuroscience, 23*(5), 1230-1246.

Landi, N., & Perfetti, C. A. (2007). An electrophysiological investigation of semantic and phonological processing in skilled and less-skilled comprehenders. *Brain and language, 102*(1), 30-45.

Lassonde, K. A., & O'Brien, E. J. (2009). Contextual specificity in the activation of predictive inferences. *Discourse Processes, 46*(5), 426-438.

Liversedge, S. P., & Findlay, J. M. (2000). Saccadic eye movements and cognition. *Trends in cognitive sciences, 4*(1), 6-14.

Long, D. L., Oppy, B. J., & Seely, M. R. (1994). Individual differences in the time course of inferential processing. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 20*(6), 1456.

Lutz, M. F., & Radvansky, G. A. (1997). The fate of completed goal information in narrative comprehension. *Journal of Memory and Language, 36(*2), 293-310.

Magliano, J. P., Baggett, W. B., Johnson, B. K., & Graesser, A. C. (1993). The time course of generating causal antecedent and causal consequence inferences. *Discourse Processes, 16*(1-2), 35-53.

Magliano, J. P., Trabasso, T., & Graesser, A. C. (1999). Strategic processing during comprehension. *Journal of Educational Psychology, 91*(4), 615.

Magliano, JP, & Mills, KK (2003). Assessing reading skill with a think-aloud procedure and latent semantic analysis. *Cognition and Instruction, 21*, 251–283

Mak, W. M., & Sanders, T. J. (2013). The role of causality in discourse processing: Effects of expectation and coherence relations. *Language and Cognitive Processes, 28*(9), 1414-1437.

Marian, V., Spivey, M., & Hirsch, J. (2003). Shared and separate systems in bilingual language processing: Converging evidence from eyetracking and brain imaging. *Brain and language, 86*(1), 70-82.

McKoon, G., & Ratcliff, R. (1986). Inferences about predictable events. *Journal of Experimental Psychology: Learning, memory, and cognition, 12*(1), 82.

McKoon, G., & Ratcliff, R. (1992). Inference during reading. *Psychological review, 99*(3), 440.

McNamara, D. S., & Kintsch, W. (1996). Learning from texts: Effects of prior knowledge and text coherence. *Discourse processes, 22*(3), 247-288.

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.

Meuter, R. F., & Allport, A. (1999). Bilingual language switching in naming: Asymmetrical costs of language selection. *Journal of memory and language, 40(*1), 25-40.

Mitchell, D. C., Shen, X., Green, M. J., & Hodgson, T. L. (2008). Accounting for regressive eye-movements in models of sentence processing: A reappraisal of the Selective Reanalysis hypothesis. *Journal of Memory and Language, 59*(3), 266-293.

Morris, R. K. (1994). Lexical and message-level sentence context effects on fixation times in reading. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 20*(1), 92.

Morris, R. K., & Folk, J. R. (1998). Focus as a contextual priming mechanism in reading. *Memory & Cognition, 26*(6), 1313-1322.

Murray, J. D., Klin, C. M., & Myers, J. L. (1993). Forward inferences in narrative text. *Journal of Memory and Language, 32*(4), 464-473.

Myers, J. L., & O'Brien, E. J. (1998). Accessing the discourse representation during reading. *Discourse processes, 26*(2-3), 131-157.

Myers, J. L., Shinjo, M., & Duffy, S. A. (1987). Degree of causal relatedness and memory. *Journal of Memory and Language, 26*(4), 453-465.

Nation, K., & Snowling, M. (1997). Assessing reading difficulties: The validity and utility of current measures of reading skill. *British Journal of Educational Psychology, 67*(3), 359-370.

Nation, K., & Snowling, M. J. (1998). Individual differences in contextual facilitation: Evidence from dyslexia and poor reading comprehension. *Child development, 69*(4), 996-1011.

Nation, K., & Snowling, M. J. (1999). Developmental differences in sensitivity to semantic relations among good and poor comprehenders: Evidence from semantic priming. *Cognition, 70*(1), B1-B13.

Noordman, L. G., Vonk, W., & Kempff, H. J. (1992). Causal inferences during the reading of expository texts. *Journal of Memory and Language, 31*(5), 573-590.

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

Oakhill, J. (1982). Constructive processes in skilled and less skilled comprehenders' memory for sentences. *British Journal of Psychology, 73*(1), 13-20.

Oakhill, J. (1983). Instantiation in skilled and less skilled comprehenders. Quarterly *Journal of Experimental Psychology, 35*(3), 441-450.

Oakhill, J. (1984). Inferential and memory skills in children's comprehension of stories. *British Journal of Educational Psychology, 54*(1), 31-39.

Oakhill, J., Garnham, A., & Reynolds, D. (2005). Immediate activation of stereotypical gender information. *Memory & Cognition, 33*(6), 972-983.

Oakhill, J., Yuill, N., & Parkin, A. (1986). On the nature of the difference between skilled and less-skilled comprehenders. *Journal of Research in Reading, 9*(2), 80-91.

O'Brien, E. J., Rizzella, M. L., Albrecht, J. E., & Halleran, J. G. (1998). Updating a situation model: a memory-based text processing view. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 24*(5), 1200.

O'Brien, E. J., Shank, D. M., Myers, J. L., & Rayner, K. (1988). Elaborative inferences during reading: Do they occur on-line?. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 14*(3), 410.

O'Reilly, T., & McNamara, D. S. (2007). Reversing the reverse cohesion effect: Good texts can be better for strategic, high-knowledge readers. *Discourse processes, 43*(2), 121-152.

Ozuru, Y., Dempsey, K., & McNamara, D. S. (2009). Prior knowledge, reading skill, and text cohesion in the comprehension of science texts. *Learning and instruction, 19*(3), 228-242.

Paris, S. G., & Myers, M. (1981). Comprehension monitoring, memory, and study strategies of good and poor readers. *Journal of Literacy Research, 13(*1), 5-22.

Peracchi, K. A., & O'Brien, E. J. (2004). Character profiles and the activation of predictive inferences. *Memory & cognition, 32*(7), 1044-1052.

Perfetti, C. A. (1985). *Reading ability*. Oxford University Press.

Perfetti, C. A., & Hart, L. (2001). The lexical basis of comprehension skill. In D. Gorfien (Ed.), *On the consequences of meaning selection* (pp. 67-86), Washington, DC: American Psychological Association

Perfetti, C. A., & Hogaboam, T. (1975). Relationship between single word decoding and reading comprehension skill. *Journal of Educational Psychology, 67*(4), 461.

Perfetti, C., Yang, C. L., & Schmalhofer, F. (2008). Comprehension skill and word-to-text integration processes. *Applied Cognitive Psychology, 22*(3), 303-318.

Petten, C. V., Kutas, M., Kluender, R., Mitchiner, M., & McIsaac, H. (1991). Fractionating the word repetition effect with event-related potentials. *Cognitive Neuroscience, Journal of, 3*(2), 131-150.

Pollatsek, A., Reichle, E. D., & Rayner, K. (2006). Tests of the EZ Reader model: Exploring the interface between cognition and eye-movement control. *Cognitive Psychology, 52*(1), 1-56.

Potts, G. R., Keenan, J. M., & Golding, J. M. (1988). Assessing the occurrence of elaborative inferences: Lexical decision versus naming. *Journal of Memory and Language, 27*(4), 399-415.

Raney, G. E., & Rayner, K. (1995). Word frequency effects and eye movements during two readings of a text. *Canadian Journal of Experimental Psychology/Revue canadienne de psychologie experimentale, 49*(2), 151.

Raney, G. E., Therriault, D. J., & Minkoff, S. R. (2000). Repetition effects from paraphrased text: Evidence for an integrated representation model of text representation. *Discourse Processes, 29*(1), 61-81.

Ratcliff, R., Hockley, W., & McKoon, G. (1985). Components of activation: Repetition and priming effects in lexical decision and recognition. *Journal of Experimental Psychology: General, 114*(4), 435.

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

Rayner, K., & McConkie, G. W. (1976). What guides a reader's eye movements?. *Vision research, 16*(8), 829-837.

Rayner, K., Chace, K. H., Slattery, T. J., & Ashby, J. (2006). Eye movements as reflections of comprehension processes in reading. *Scientific Studies of Reading, 10*(3), 241-255.

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: Section A, 53*(4), 1061-1080.

Rayner, K., Warren, T., Juhasz, B. J., & Liversedge, S. P. (2004). The effect of plausibility on eye movements in reading. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 30*(6), 1290.

Rizzella, M. L., & O'Brien, E. J. (2002). Retrieval of concepts in script-based texts and narratives: The influence of general world knowledge. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 28*(4), 780.

Rugg, M. D., & Nagy, M. E. (1987). Lexical contribution to nonword-repetition effects: Evidence from event-related potentials. *Memory & Cognition, 15*(6), 473-481.

Sanford, A. J. , & Garrod S.C. (2005). Memory-based approaches and beyond. *Discourse Processes, 39*(2-3), 205-224.

Sanford, A. J., & Garrod, S. C. (1980). Memory and attention in text comprehension: The problem of reference. *Attention and performance VIII*, 459-474.

Sanford, A. J., & Graesser, A. C. (2006). Shallow processing and underspecification. *Discourse Processes, 42*(2), 99-108.

Scarborough, D. L., Cortese, C., & Scarborough, H. S. (1977). Frequency and repetition effects in lexical memory. *Journal of Experimental Psychology: Human perception and performance, 3*(1), 1.

Schmalhofer, F., McDaniel, M. A., & Keefe, D. (2002). A unified model for predictive and bridging inferences. *Discourse Processes, 33*(2), 105-132.

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, 1218-1226.

Schroeder, S. (2011). What readers have and do: Effects of students' verbal ability and reading time components on comprehension with and without text availability. *Journal of Educational Psychology, 103*(4), 877.

Singer, M. (1993). Causal bridging inferences: Validating consistent and inconsistent sequences. *Canadian Journal of Experimental Psychology/Revue canadienne de psychologie experimentale, 47*(2), 340.

Singer, M., Andruslak, P., Reisdorf, P., & Black, N. L. (1992b). Individual differences in bridging inference processes. *Memory & Cognition, 20*(5), 539-548.

Singer, M., Halldorson, M., Lear, J. C., & Andrusiak, P. (1992a). Validation of causal bridging inferences in discourse understanding. *Journal of Memory and Language, 31*(4), 507-524.

Stanovich, K. E. (1982). Individual Differences in the Cognitive Processes of Reading I. Word Decoding. *Journal of Learning Disabilities, 15*(8), 485-493.

Staub, A., & Rayner, K. (2007). Eye movements and online comprehension processes. *The Oxford handbook of psycholinguistics,* 327, 342.

Stine-Morrow, E. A., Loveless, M. K., & Soederberg, L. M. (1996). Resource allocation in on-line reading by younger and older adults. *Psychology and Aging, 11*(3), 475.

Stine-Morrow, E. A., Milinder, L. A., Pullara, O., & Herman, B. (2001). Patterns of resource allocation are reliable among younger and older readers. *Psychology and Aging, 16*(1), 69.

Sturt, P. (2003). The time-course of the application of binding constraints in reference resolution. *Journal of Memory and Language, 48*(3), 542-562.

Suh, S. U., & Trabasso, T. (1993). Inferences during reading: Converging evidence from discourse analysis, talk-aloud protocols, and recognition priming. *Journal of memory and language*, *32*(3), 279-300.

Trabasso, T., & Van Den Broek, P. (1985). Causal thinking and the representation of narrative events. *Journal of memory and language, 24*(5), 612-630.

Traxler, M. J., & Pickering, M. J. (1996). Plausibility and the processing of unbounded dependencies: An eye-tracking study. *Journal of Memory and Language, 35*(3), 454-475.

Traxler, M. J., Foss, D. J., Seely, R. E., Kaup, B., & Morris, R. K. (2000). Priming in sentence processing: Intralexical spreading activation, schemas, and situation models. *Journal of Psycholinguistic Research, 29*(6), 581-595.

Traxler, M. J., Morris, R. K., & Seely, R. E. (2002). Processing subject and object relative clauses: Evidence from eye movements. *Journal of Memory and Language, 47*(1), 69-90.

Van den Broek, P., Lorch, R. F., Linderholm, T., & Gustafson, M. (2001). The effects of readers' goals on inference generation and memory for texts. *Memory & Cognition, 29*(8), 1081-1087.

Van Der Schoot, M., Vasbinder, A. L., Horsley, T. M., & Van Lieshout, E. C. (2008). The role of two reading strategies in text comprehension: An eye fixation study in primary school children. *Journal of Research in Reading, 31*(2), 203-223.

van der Schoot, M., Vasbinder, A. L., Horsley, T. M., & Van Lieshout, E. C. D. M. (2008). The role of two reading strategies in text comprehension: An eye fixation study in primary school children. *Journal of Research in Reading, 31*(2), 203–223.

Vasishth, S., von der Malsburg, T., & Engelmann, F. (2013). What eye movements can tell us about sentence comprehension. *Wiley Interdisciplinary Reviews: Cognitive Science*, *4*(2), 125-134.

Weaver, C. A., & Bryant, D. S. (1995). Monitoring of comprehension: The role of text difficulty in metamemory for narrative and expository text. *Memory & Cognition, 23*(1), 12-22.

Whitney, P., Ritchie, B. G., & Clark, M. B. (1991). Working-memory capacity and the use of elaborative inferences in text comprehension. *Discourse Processes, 14*(2), 133-145.

Yang, C. L., Perfetti, C. A., & Schmalhofer, F. (2005). Less skilled comprehenders' ERPs show sluggish word-to-text integration processes. *Written Language & Literacy, 8*(2), 157-181.

Yang, C. L., Perfetti, C. A., & Schmalhofer, F. (2007). Event-related potential indicators of text integration across sentence boundaries. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 33*(1), 55.

Yuill, N., & Oakhill, J. (1988). Understanding of anaphoric relations in skilled and less skilled comprehenders. *British Journal of Psychology, 79*(2), 173-186.

Yuill, N., Oakhill, J., & Parkin, A. (1989). Working memory, comprehension ability and the resolution of text anomaly. *British journal of psychology, 80*(3), 351-361.

Zwaan, R. A., & Radvansky, G. A. (1998). Situation models in language comprehension and memory. *Psychological bulletin, 123*(2), 162.

Appendix A

Below is a list of all eye-movements measured and analyzed in this thesis with the description of on-line processes the eye-movements are thought to reflect (following Liversedge and Findlay,2000; Rayner, 1998).

Eye tracking measurement	*Definition*	*Cognitive processes reflected by the measure*
First fixation duration	The duration of the first fixation in a region.	The earliest processing measure, reflecting the immediate processing of that word, such as lexical retrieval and cognitive processes that are very fast.
First pass time/ Gaze duration	The sum of all fixations made in a region prior to the eyes moving to the right or left.	A later measure reflecting the initial processing of a region involving fast processes such as lexical access but also more complex integrative process.
Probability of a regression out of region	Probability of making regressive eye movements to preceding regions.	A later effect that reflects difficulty in comprehending the region being fixated.
Go-past reading time	The sum of all the temporally contiguous fixations in a region, including any regressive eye movements to the left of the region, until the point of fixation progresses to the region to the right.	Reflects both early measures reflecting initial integrative processing of a region as well as later measures such as regressing out due to difficulties integrating a word.
Number of total fixations	The total number of fixations made in a particular region.	A late processing measure that can reflect higher level comprehension processes whereby more fixations corresponds to increased comprehension difficulty. Reflects both first pass and second pass reading.
Total reading time	The sum of all fixations made in the region (i.e., first-pass fixations plus any regressive fixations).	A late processing measure that can reflect higher level reading comprehension processes. Reflects both first pass and second pass reading.
Probability of regression into region	The returning to a word that has been read previously.	Reflects difficulty in integrating the region being regressed to with the current text.
Second pass reading time	The time spent re-reading a region having previously encountered it.	A late processing measure that reflects difficulty integrating the word or region being re-read.

Appendix B

Stimulus set for Experiment 1

Target words are marked in *italics.*

1.

Control: Allen's baby became very ill, so Allen put the baby in the car and rushed off to the hospital. The *hospital* had a long waiting queue.

Paraphrase: Allen's baby became very ill, so Allen put the baby in the car and rushed off to casualty. The *hospital* had a long waiting queue.

Inference: Allen's baby became very ill, so Allen put the baby in the car and rushed off. The *hospital* had a long waiting queue.

Comprehension Question (CQ): Where did Allen rush off to? The car/ The hospital

---

2.

Control: John couldn't see anything in the dark until he found the box of matches on the table and lit one. The weak *light* filled the room with a warm, yellow glow.

Paraphrase: John couldn't see anything in the dark until he found a box of matches on the table and struck one. The weak *light* filled the room with a warm, yellow glow.

Inference: John couldn't see anything in the dark until he found a box of matches on the table in front of him. The weak *light* filled the room with a warm, yellow glow.

CQ: What did John do with the matches? He lit one/ He found them

---

3.

Control: Mary packed her drink into her schoolbag and set off for school, but the bottle leaked and her books got soaked. At school, she found that her books were *soaked* in orange juice.

Paraphrase: Mary packed her drink into her schoolbag and set off for school, but the bottle leaked and her books got really wet. At school, she found that her books were *soaked* in orange juice.

Inference: Mary packed her drink into her schoolbag and set off for school, but the bottle leaked on the way there. At school, she found that her books were *soaked* in orange juice.

CQ: Why were the books soaked in orange juice? Mary's bottle leaked/ Mary packed a drink to school

---

4.

Control: A little kitten was climbing an apple tree when a branch broke and the kitten landed on the ground. The kitten *landed* safely on a pile of old leaves.

Paraphrase: A little kitten was climbing an apple tree when a branch broke and the kitten fell down to the ground. The kitten *landed* safely on a pile of old leaves.

Inference: A little kitten was climbing an apple tree when a branch under its paws broke suddenly . The kitten *landed* safely on a pile of old leaves.

CQ: Why was the kitten on a pile of old leaves? It fell there/ It climbed there

---

5.
Control: Cathy had a beautiful blue lamp on her table when suddenly an earthquake hit the city and the lamp broke. Tiny pieces of *broken* blue glass lay on the floor.

Paraphrase: Cathy had a beautiful blue lamp on her table when suddenly an earthquake hit the city and the lamp fell down. Tiny pieces of *broken* blue glass lay on the floor.

Inference: Cathy had a beautiful blue lamp on her table when suddenly an earthquake hit the city where she lived. Tiny pieces of *broken* blue glass lay on the floor.

CQ: What happened to Cathy's lamp? It fell and broke/It lay on the floor

---

6.
Control: After playing with his pony, Ben picked up his bike and rode away. He *rode off* at a very high speed to get home before his dad returned from work.

Paraphrase: After playing with his pony, Ben picked up his bike and pedalled off. He *rode off* at a very high speed to get home before his dad returned from work.

Inference: After playing with his pony, Ben picked up his bike lying nearby. He *rode off* at a very high speed to get home before his dad returned from work.

Comp question: How did Ben get home? He rode a pony/ He rode a bike

---

7.
Control: Jenny had just finished the last batch of her amazing cookies and put them in the cooker. A mouth-watering smell came from the *cooker* and filled the kitchen.

Paraphrase: Jenny had just finished the last batch of her amazing cookies and left them to bake. A mouth-watering smell came from the *cooker* and filled the kitchen.

Inference: Jenny had just finished the last batch of her amazing cookies that her son loved so much. A mouth-watering smell came from the *cooker* and filled the kitchen.

Comp Q: What smelled nice? The cooker/ The cookies

8.
Control: Annie rubbed her eyes, got out of bed and left her room to go to the wash basin in the bathroom. She filled the *basin* with hot water and splashed it all over her face.

Paraphrase: Annie rubbed her eyes, got out of bed and left her room to go to the bathroom. She filled the *basin* with hot water and splashed it all over her face.

Inference: Annie rubbed her eyes, got out of bed and left her room to get ready for school. She filled the *basin* with hot water and splashed it all over her face.

CQ: Where did Annie wash her face? In her room/ in the bathroom

---

9.
Control: Julie had had a hard day at school so Mum decided to cheer her up before dinner by buying her favourite chocolate cheesecake. A yummy chocolate *cheesecake* waited on the table for Julie to come home.

Paraphrase: Julie had had a hard day at school so Mum decided to cheer her up before dinner by buying her favourite chocolate dessert. A yummy chocolate *cheesecake* waited on the table for Julie to come home.

Inference: Julie had had a hard day at school so Mum decided to cheer her up before dinner as she had to study more in the evening. A yummy chocolate *cheesecake* waited on the table for Julie to come home.

CQ: What did Mum get to cheer Julie up? A nice cake/ A nice dinner

---

10.
Control: Shelly's room was very hot and the flowers in the vase had no more water left so they started to turn yellow. The flowers all turned *yellow* by the end of the day.

Paraphrase: Shelly's room was very hot and the flowers in the vase had no more water left so they slowly started to die. The flowers all turned *yellow* by the end of the day.

Inference: Shelly's room was very hot and the flowers in the vase had no more water left for the rest of the day. The flowers all turned *yellow* by the end of the day.

CQ: Why did the flowers turn yellow? Because they had no water/ Because it was the end of the day

---

11.
Control: Jenny had been running for two hours when she tripped on a tree root and injured her ankle. Jenny's *ankle* hurt so much that she had to stop running.

Paraphrase: Jenny had been running for two hours when she tripped on a tree root and injured herself. Jenny's *ankle* hurt so much that she had to stop running.

Inference: Jenny had been running for two hours when she tripped on a tree root hidden in the grass. Jenny's *ankle* hurt so much that she had to stop running.

CQ: Why did Jenny's ankle hurt? She ran for too long/ She tripped on a tree root

---

12.
Control: The fire destroyed everything in the forest, including the grizzly bear's house. The *grizzly bear* had to look for a new home in another forest.

Paraphrase: The fire destroyed everything in the forest, including the homes of many animals. The *grizzly bear* had to look for a new home in another forest.

Inference: The fire destroyed everything in the forest and the whole place was black and deserted. The *grizzly bear* had to look for a new home in another forest.

CQ: Why did the bear had to look for a new home? He liked a different forest /His home burnt

---

13.
Control: The artist left a beautiful ice swan in his yard, forgetting that the midday sun would melt the statue. Later that day, there was nothing left of *the statue* but a watery puddle.

Paraphrase: The artist left a beautiful ice swan in his yard, forgetting that the midday sun would melt the ice. Later that day, there was nothing left of *the statue* but a watery puddle.

Inference: The artist left a beautiful ice swan in his yard, forgetting about the midday sun that filled the yard. Later that day, there was nothing left of *the statue* but a watery puddle.

CQ: What happened to the statue? It melted/ It disappeared

14.
Control: A little boy stole the baby's lollipop so the baby began to cry. The baby's mum bought the baby another *lollipop* to make her happy again.

Paraphrase: A little boy stole the baby's sweet, so the baby began to cry. The baby's mum bought the baby another *lollipop* to make her happy again.

Inference: A little boy stole from a baby, so the baby began to cry. The baby's mum bought the baby another *lollipop* to make her happy again.

CQ Why did Mom need buy a lollipop for the baby? The boy took the first one/ The baby lost the first one

---

15.
Control: Tammy's Mum was having a surprise birthday party for Tammy and she decorated the whole house for her with balloons. Lots of colourful *balloons* were lying everywhere and party bags with sweets were ready for kids.

Paraphrase: Tammy's Mum was having a surprise birthday party for Tammy and she decorated the whole house for her. Lots of colourful *balloons* were lying everywhere and party bags with sweets were ready for kids.

Inference: Tammy's Mum was having a surprise birthday party for Tammy and she decorated the whole house with party decorations. Lots of colourful *balloons* were lying everywhere and party bags with sweets were ready for kids.

CQ What did Mum get to decorate the house for the party? Balloons/ Party Bags

---

16.
Control: It was the day of the band's biggest performance and Jane had been nervous since she woke up. Jane was so *nervous* she couldn't even finish her breakfast.

Paraphrase: It was the day of the band's biggest performance and Jane had been tense since she woke up. Jane was so *nervous* she couldn't even finish her breakfast.

Inference: It was the day of the band's biggest performance. Jane was so *nervous* she couldn't even finish her breakfast.

Q: Why was Jane nervous? Her band had a big performance/ She didn't finish breakfast

---

17.
Control: Annie was making tea for her friend Bill when suddenly her dog jumped on her and the tea spilled from the cup. The *spilled tea* stained the carpet.

Paraphrase: Annie was making tea for her friend Bill when suddenly her dog jumped on her and the cup slipped from her hands. The *spilled tea* stained the carpet.

Inference: Annie was making tea for her friend Bill when suddenly her dog jumped on her. The *spilled tea* stained the carpet.

CQ: Why was the carpet stained? The dog jumped on Annie/ Bill spilled the tea

---

18.
Control: Lucy's favourite restaurant was closed because it was being painted. The *painting* was supposed to last for next five months.

Paraphrase: Lucy's favourite restaurant was closed because it was being decorated. The *painting* was supposed to last for next five months.

Inference: Lucy's favourite restaurant was closed when she came there with her dad. The *painting* was supposed to last for next five months.

CQ: When will the restaurant open again? In five months/ It was closed for good.

19.
Control: Alex was roasting marshmallows in the fire place when a spark flew out and burnt a hole in his sleeve. There was now a big *black hole* on his new shirt.

Paraphrase: Alex was roasting marshmallows in the fire place when a spark flew out and his sleeve caught fire. There was now a big *black hole* on his new shirt.

Inference: Alex was roasting marshmallows in the fire place to have some ready for his friends when a spark flew out. There was now a big *black hole* on his new shirt.

Q.C.: What made a hole in Alex's shirt? A spark from the fire/ A roasted marshmallow

---

20.
Control: Tom's little boat was swaying back and forth on the big waves and Tom started to feel unwell. Tom felt so *unwell* he had to throw up.

Paraphrase: Tom's little boat was swaying back and forth on the big waves so that Tom started to feel poorly. Tom felt so *unwell* he had to throw up.

Inference: Tom's little boat was swaying back and forth on the big waves which splashed water into the boat. Tom felt very *unwell* and had to throw up.

CQ: What made Tom sick? The boat was swaying too much/ The waves were too big

---

21.
Control: It was snowing very hard and the pavements in town were covered with slippery snow. People had to walk very slowly to stop themselves *slipping* on the wet pavestones.

Paraphrase: It was snowing very hard and the pavements in town were covered with freezing snow. People had to walk very slowly in order to prevent themselves *slipping* on the wet pavestones.

Inference: It was snowing very hard and the whole town looked like winter wonderland. People had to walk very slowly to stop themselves *slipping* on the wet pavestones.

CQ: Why were the stones slippery? They were wet from the rain/ they were wet from the snow

---

22.
Control: Annie was very tired after basketball training and quickly fell asleep. The training session had been very hard, because the team was preparing for an important match.

Paraphrase: Annie was very tired after basketball practice and quickly fell asleep. The training session had been very hard because the team was preparing for an important match.

Inference: Annie was very tired and quickly fell asleep when she arrived home. The training session had been very hard because the team was preparing for an important match.

CQ: Why was Annie so tired? She had a hard training session/ She didn't get enough sleep

---

23.
Control: Jeff was cooking spaghetti when the show started and he left the pot on the hob so long the smoke started to rise from it. Dark s*moke* started to fill the whole kitchen.

Paraphrase: Jeff was cooking spaghetti when the show started and he left the pot on the hob for such a long time, it started burning. Dark s*moke* started to fill the whole kitchen.

Inference: Jeff was cooking spaghetti when the show started and he left the pot on the hob for a really long time. Dark s*moke* started to fill the whole kitchen.

CQ: Why was the kitchen full of smoke? The spaghetti burnt/ Dark smoke filled the kitchen

---

24.

Control: John was watching a film with his friends when his mum brought them some pop-corn. The whole sofa was covered with *pop-corn* as the boys ate messily.

Paraphrase: John was watching a film with his friends when his mum brought them some snacks. The whole sofa was covered with *pop-corn* as the boys ate messily.

Inference: John was watching a film with his friends as they always did on Sunday nights. The whole sofa was covered with pop-corn as the boys ate messily.

CQ: When did the sofa get messy? When they were watching a film/ when they spilled pop-corn on it

---

25.

Control: The queue in the supermarket was very long, and so Jim decided to read a magazine as he was waiting. By the time Jim got to the cashier, he had finished reading his *magazine* completely.

Paraphrase: The queue in the supermarket was very long, and so Jim decided to read as he was waiting. By the time Jim got to the cashier, he had finished reading the *magazine* completely.

Inference: The queue in the supermarket was very long. By the time Jim got to the cashier, he had finished reading the *magazine* completely.

CQ: Why did Jim read the whole magazine? He had to wait for a long time / It was a long walk to the cashier

---

26.

Control: Jane was lying in bed with a fever and a headache, so her Mum prepared some medicine to make her better.Mum put the *medicine* and some tea on the table next to the bed

Paraphrase: Jane was lying in bed with a fever and a headache, so her Mum prepared some tablets to make her better. Mum put the *medicine* and some tea on the table next to the bed.

Inference: Jane was lying in bed with a fever and a headache and felt very poorly. Mum put the *medicine* and some tea on the table next to the bed.

CQ: Why did Mum put the medicines on the table? Because Jane was ill/ Because Mum was ill

---

27.

Control: A slippery banana peel lay on the pavement in front of the shop when Jerry slipped on it. Jerry *slipped* very suddenly as he was coming out with his shopping.

Paraphrase: A banana peel lay on the pavement in front of the shop when Jerry stepped on it. Jerry *slipped* very suddenly as he was coming out with his shopping.

Inference: A banana peel lay on the pavement in front of the shop where Jerry always shopped. Jerry *slipped* very suddenly as he was coming out with his shopping.

CQ: What did Jerry slip on? The pavement/ The banana peel

Appendix C

Stimulus set for Experiment 2

Target words are marked in *italics.*

1.

Control: Allen's baby became very ill, so Allen put the baby in the car and rushed off to the hospital. The ***hospital*** had a long waiting queue.

Synonym: Allen's baby became very ill, so Allen put the baby in the car and rushed off to casualty. The *hospital* had a long waiting queue.

Paraphrase: Allen's baby became very ill, so Allen put the baby in the car and rushed off to get medical help. *The hospital* had a long waiting queue.

Inference: Allen's baby became very ill, so Allen put the baby in the car and rushed off. The *hospital* had a long waiting queue.

CQ: Where did Allen rush off to? The car/ The hospital

2.

Control: John couldn't see anything in the dark until he found the box of matches on the table and lit one. The weak ***light*** filled the room with a warm, yellow glow.

Synonym: John couldn't see anything in the dark until he found the box of matches on the table and ignited one. The weak *light* filled the room with a warm, yellow glow.

Paraphrase: John couldn't see anything in the dark until he found a box of matches on the table and struck one. The weak *light* filled the room with a warm, yellow glow.

Inference: John couldn't see anything in the dark until he found a box of matches on the table in front of him. The weak *light* filled the room with a warm, yellow glow.

CQ: What did John do with the matches? He lit one/ He found them

---

3.

Control: Mary packed her drink into her schoolbag and set off for school, but the bottle leaked and her books got soaked. At school, she found that her books were ***soaked*** in orange juice.

Synonym: Mary packed her drink into her schoolbag and set off for school, but the bottle leaked and her books got drenched. At school, she found that her books were *soaked* in orange juice.

Paraphrase: Mary packed her drink into her schoolbag and set off for school, but the bottle leaked and her books got covered in water. At school, she found that her books were *soaked* in orange juice.

Inference: Mary packed her drink into her schoolbag and set off for school, but the bottle leaked on the way there. At school, she found that her books were *soaked* in orange juice.

CQ: Why were the books soaked in orange juice? Mary's bottle leaked/ Mary packed a drink to school

---

4.

Control: A little kitten was climbing an apple tree when a branch broke and the kitten landed on the ground. The kitten ***landed*** safely on a pile of old leaves.

Synonym: A little kitten was climbing an apple tree when a branch broke and the kitten alighted on the ground. The kitten *landed* safely on a pile of old leaves.

Paraphrase: A little kitten was climbing an apple tree when a branch broke and the kitten tumbled to the ground. The kitten *landed* safely on a pile of old leaves.

Inference: A little kitten was climbing an apple tree when a branch under its paws broke suddenly. The kitten *landed* safely on a pile of old leaves.

CQ: Why was the kitten on a pile of old leaves? It fell there/ It climbed there

---

5.

Control: Cathy had a beautiful blue lamp on her table when suddenly an earthquake hit the city and the lamp broke. Tiny pieces of ***broken*** blue glass lay on the floor.

Synonym: Cathy had a beautiful blue lamp on her table when suddenly an earthquake hit the city and the lamp smashed. Tiny pieces of *broken* blue glass lay on the floor.

Paraphrase: Cathy had a beautiful blue lamp on her table when suddenly an earthquake hit the city and the lamp fell down. Tiny pieces of *broken* blue glass lay on the floor.

Inference: Cathy had a beautiful blue lamp on her table when suddenly an earthquake hit the city where she lived. Tiny pieces of *broken* blue glass lay on the floor.

CQ: What happened to Cathy's lamp? It fell and broke/It lay on the floor

---

6.

Control: After playing with his pony, Ben picked up his bike and rode away. He ***rode off*** at a very high speed to get home before his dad returned from work.

Synonym: After playing with his pony, Ben picked up his bike and cycled away. He *rode off* at a very high speed to get home before his dad returned from work.

Paraphrase: After playing with his pony, Ben picked up his bike and sped off . He *rode off* at a very high speed to get home before his dad returned from work.

Inference: After playing with his pony, Ben picked up his bike lying nearby. He *rode off* at a very high speed to get home before his dad returned from work.

CQ: How did Ben get home? He rode a pony/ He rode a bike

7.

Control: Jenny had just finished the last batch of her amazing cookies and put them in the oven. A mouth-watering smell came from the *stove* and filled the kitchen.

Synonym: Jenny had just finished the last batch of her amazing cookies and put them in the oven. A mouth-watering smell came from the ***stove*** and filled the kitchen.

Paraphrase: Jenny had just finished the last batch of her amazing cookies and left them to bake. A mouth-watering smell came from the *stove* and filled the kitchen.

Inference: Jenny had just finished the last batch of her amazing cookies that her son loved so much. A mouth-watering smell came from the *stove* and filled the kitchen.

CQ: What smelled nice? The stove / The cookies

8.

Control: Annie rubbed her eyes, got out of bed and left her room to go to the wash basin in the bathroom. She filled the ***basin*** with hot water and splashed it all over her face.

Synonym: Annie rubbed her eyes, got out of bed and left her room to go to the washbowl in the bathroom. She filled the *basin* with hot water and splashed it all over her face.

Paraphrase: Annie rubbed her eyes, got out of bed and left her room to go to the bathroom. She filled the *basin* with hot water and splashed it all over her face.

Inference: Annie rubbed her eyes, got out of bed and left her room to get ready for school. She filled the *basin* with hot water and splashed it all over her face.

CQ: Where did Annie wash her face? In her room/ in the bathroom

---

9.

Control: Julie had had a hard day at school so Mum decided to cheer her up before dinner by buying her favourite olive loaf. Fresh and crispy olive ***bread*** waited on the table for Julie to come home.

Synonym: Julie had had a hard day at school so Mum decided to cheer her up before dinner by buying her favourite olive loaf. Fresh and crispy olive *bread* waited on the table for Julie to come home.

Paraphrase: Julie had had a hard day at school so Mum decided to cheer her up before dinner by stopping at her favourite bakery to get Julie a treat. Fresh and crispy olive bread waited on the table for Julie to come home.

Inference: Julie had had a hard day at school so Mum decided to cheer her up before dinner as she had to study more in the evening. Fresh and crispy olive bread waited on the table for Julie to come home.

CQ: What did Mum get to cheer Julie up? A nice bread loaf/ A nice dinner

---

10.
Control: Shelly's room was very hot and the flowers in the vase had no more water left so they started to wilt. The flowers had ***wilted*** by the end of the day.

Synonym: Shelly's room was very hot and the flowers in the vase had no more water left so they started to wither. The flowers had *wilted* by the end of the day

Paraphrase: Shelly's room was very hot and the flowers in the vase had no more water left so they slowly started to die. The flowers had *wilted* by the end of the day.

Inference: Shelly's room was very hot and the flowers in the vase had no more water left for the rest of the day. The flowers had *wilted* by the end of the day.

CQ: Why did the flowers wilt? Because they had no water/ Because it was the end of the day

---

11.
Control: Jenny had been running for two hours when she tripped on a tree root and injured her ankle. Jenny's ***ankle*** hurt so much that she had to stop running.

Synonym: Jenny had been running for two hours when she tripped on a tree root and injured her lower leg. Jenny's *ankle* hurt so much that she had to stop running.

Paraphrase: Jenny had been running for two hours when she tripped on a tree root and injured herself. Jenny's *ankle* hurt so much that she had to stop running.

Inference: Jenny had been running for two hours when she tripped on a tree root hidden in the grass. Jenny's *ankle* hurt so much that she had to stop running.

CQ: Why did Jenny's ankle hurt? She ran for too long/ She tripped on a tree root

---

12.
Control: The fire destroyed everything in the forest, including the rabbit's house. The sad ***rabbit*** had to look for a new home in another forest.

Synonym: The fire destroyed everything in the forest, including the hare's house. The sad *rabbit* had to look for a new home in another forest.

Paraphrase: The fire destroyed everything in the forest, including the homes of many animals. The sad *rabbit* had to look for a new home in another forest.

Inference:The fire destroyed everything in the forest and the whole place was black and deserted. The sad *rabbit* had to look for a new home in another forest.

CQ Why did the rabbit have to look for a new home? He liked a different forest /His home burnt

---

13.
Control: The artist left a beautiful ice swan in his yard, forgetting that the midday sun would melt the statue. Later that day, there was nothing left of ***the statue*** but a watery puddle.

Synonym: The artist left a beautiful ice swan in his yard, forgetting that the midday sun would melt the sculpture. Later that day, there was nothing left of *the statue* but a watery puddle.

Paraphrase: The artist left a beautiful ice swan in his yard, forgetting that the midday sun would melt the ice. Later that day, there was nothing left of *the statue* but a watery puddle.

Inference: The artist left a beautiful ice swan in his yard, forgetting about the midday sun that filled the yard. Later that day, there was nothing left of *the statue* but a watery puddle.

CQ: What happened to the statue? It melted/ It disappeared

14.
Control: A rude little boy stole the baby's plaything so the baby began to cry. The baby's mum bought the baby another *plaything* to make her happy again.

Synonym: A rude little boy stole the baby's toy so the baby began to cry. The baby's mum bought the baby another *plaything* to make her happy again.

Paraphrase: A rude little boy stole the baby's cuddly bear, so the baby began to cry. The baby's mum bought the baby another *plaything* to make her happy again.

Inference: A rude little rude boy stole from a baby, so the baby began to cry very loud. The baby's mum bought the baby another *plaything* to make her happy again.

CQ Why did Mom need buy a plaything for the baby? The boy took the first one/ The baby lost the first one

---

15.
Control: The flat was much sunnier than he liked, so Arthur walked over and closed the blinds. With the blinds ***closed***, the room felt a lot darker and Arthur was pleased.

Synonym: The flat was much sunnier than he liked, so Arthur walked over and shut the blinds. With the blinds closed, the room felt a lot darker and Arthur was pleased.

Paraphrase: The flat was much sunnier than he liked, so Arthur walked over and reached for the blinds. With the blinds closed, the room felt a lot darker and Arthur was pleased.

Inference: The den was much sunnier than he liked, so Arthur walked over to the window. With the blinds closed, the room felt a lot darker and Arthur was pleased.

CQ: Why did Arthur walk over to the window? To see the sun/ To darken the room

---

16.
Control: It was the day of the band's biggest music performance and Jane had been nervous since she woke up. Jane was so ***nervous*** she couldn't even finish her breakfast.

Synonym: It was the day of the band's biggest music performance and Jane had been anxious since she woke up. Jane was so *nervous* she couldn't even finish her breakfast.

Paraphrase: It was the day of the band's biggest music performance and Jane had been worrying since she woke up. Jane was so *nervous* she couldn't even finish her breakfast.

Inference: It was the day of the band's biggest music performance. Jane was so *nervous* she couldn't even finish her breakfast.

CQ: Why was Jane nervous? Her band had a big performance/ She didn't finish breakfast

---

17.
Control: Annie was making tea for her friend Bill when suddenly her dog jumped on her and the tea spilled from the cup. The ***spilled tea*** stained the carpet.

Synonym: Annie was making tea for her friend Bill when suddenly her dog jumped on her and the tea dribbled from the cup. The *spilled tea* stained the carpet

Paraphrase: Annie was making tea for her friend Bill when suddenly her dog jumped on her and the cup slipped from her hands. The *spilled tea* stained the carpet.

Inference: Annie was making tea for her friend Bill when suddenly her dog jumped on her. The *spilled* tea stained the carpet.

CQ: Why was the carpet stained? The dog jumped on Annie/ Bill spilled the tea

---

18.
Control: Lucy's favourite restaurant was closed because it was being decorated. The ***decoration*** was supposed to last for next five months.

Synonym: Lucy's favourite restaurant was closed because it was being painted. The *decoration* was supposed to last for next five months.

Paraphrase: Lucy's favourite restaurant was closed because it was being completely renovated. The *decoration* was supposed to last for next five months.

Inference: Lucy's favourite restaurant was closed when she came there with her dad. The *decoration* was supposed to last for next five months.

CQ: When will the restaurant open again? In five months/ It was closed for good.

19.
Control: Alex was roasting marshmallows in the fire place when a spark flew out and burnt a hole in his sleeve. There was now a big ***black hole*** on his new shirt.

Synonym: Alex was roasting marshmallows in the fire place when a spark flew out and burnt an opening his sleeve. There was now a big *black hole* on his new shirt.

Paraphrase: Alex was roasting marshmallows in the fire place when a spark flew out and his sleeve caught fire. There was now a big *black hole* on his new shirt.

Inference: Alex was roasting marshmallows in the fire place to have some ready for his friends when a spark flew out. There was now a big *black hole* on his new shirt.

CQ.: What made a hole in Alex's shirt? A spark from the fire/ A roasted marshmallow

20.

Control: Tom's little boat was swaying back and forth on the big waves and Tom started to feel unwell. Tom felt so ***unwell*** he had to throw up.

Synonym: Tom's little boat was swaying back and forth on the big waves and Tom started to feel queasy. Tom felt so *unwell* he had to throw up.

Paraphrase: Tom's little boat was swaying back and forth on the big waves so that Tom started to turn green. Tom felt so *unwell* he had to throw up.

Inference: Tom's little boat was swaying back and forth on the big waves which splashed water into the boat. Tom felt very *unwell* and had to throw up.

CQ: What made Tom sick? The boat was swaying too much/ The waves were too big

21.

Control: It was snowing very hard and the pavements in town were covered with slippery snow. People had to walk very slowly to stop themselves ***slipping*** on the wet pavestones.

Synonyms: It was snowing very hard and the pavements in town were covered with icy snow. People had to walk very slowly to stop themselves *slipping* on the wet pavestones

Paraphrase: It was snowing very hard and walking on pavements was getting very dangerous. People had to walk very slowly in order to prevent themselves *slipping* on the wet pavestones.

Inference: It was snowing very hard and the whole town looked like winter wonderland. People had to walk very slowly to stop themselves *slipping* on the wet pavestones.

CQ: Why were the stones slippery? They were wet from the rain/ they were wet from the snow

22.

Control: Annie was very tired after basketball training and quickly fell asleep. The ***training*** session had been very hard, because the team was preparing for an important match.

Synonym: Annie was very tired after basketball practice and quickly fell asleep. The training session had been very hard, because the team was preparing for an important match.

Paraphrase: Annie was very tired after a difficult basketball game and quickly fell asleep. The training session had been very hard because the team was preparing for an important match.

Inference: Annie was very tired and quickly fell asleep when she arrived home. The training session had been very hard because the team was preparing for an important match.

CQ: Why was Annie so tired? She had a hard training session/ She didn't get enough sleep

23.
Control: Jeff was cooking spaghetti when the show started and he left the pot on the hob so long the smoke started to rise from it. Dark ***smoke*** started to fill the whole kitchen.

Synonym: Jeff was cooking spaghetti when the show started and he left the pot on the hob so long fumes started to rise from it. Dark s*moke* started to fill the whole kitchen

Paraphrase: Jeff was cooking spaghetti when the show started and he left the pot on the hob for such a long time, it started burning. Dark s*moke* started to fill the whole kitchen.

Inference: Jeff was cooking spaghetti when the show started and he left the pot on the hob for a really long time. Dark s*moke* started to fill the whole kitchen.

CQ: Why was the kitchen full of smoke? The spaghetti burnt/ Dark smoke filled the kitchen

---

24.
Control: John was watching a film with his friends when his mum brought them some crisps. The whole sofa was covered with ***crisps*** as the boys ate messily.

Synonym: John was watching a film with his friends when his mum brought them some Pringles. The whole sofa was covered with *crisps* as the boys ate messily.

Paraphrase: John was watching a film with his friends when his mum brought them some nibbles. The whole sofa was covered with *crisps* as the boys ate messily.

Inference: John was watching a film with his friends as they always did on Sunday nights. The whole sofa was covered with crisps as the boys ate messily.

CQ: Why did the sofa get messy? Because they were watching a film/ Because they spilled crisps on it

---

25.
Control: The queue in the supermarket was very long, and so Jim decided to read a magazine as he was waiting. By the time Jim got to the cashier, he had finished reading his ***magazine*** completely.

Synonym: The queue in the supermarket was very long, and so Jim decided to read a periodical as he was waiting. By the time Jim got to the cashier, he had finished reading his *magazine* completely.

Paraphrase: The queue in the supermarket was very long, and so Jim decided to read as he was waiting. By the time Jim got to the cashier, he had finished reading his *magazine* completely.

Inference: The queue in the supermarket was very long as Jim went shopping just before dinner time. By the time Jim got to the cashier, he had finished reading his *magazine* completely.

CQ: Why did Jim read the whole magazine? He had to wait for a long time / It was a long walk to the cashier

---

26.
Control: Jane was lying in bed with a fever and a headache, so her Mum fetched her medicine to make her better. Mum put the ***medicines*** and some tea on the table next to the bed.

Synonym: Jane was lying in bed with a fever and a headache, so her Mum fetched her antibiotics to make her better. Mum put the *medicines* and some tea on the table next to the bed.

Paraphrase: Jane was lying in bed with a fever and a headache, so her Mum fetched her tablets to make her better. Mum put the *medicine* and some tea on the table next to the bed.

Inference: Jane was lying in bed with a fever and a headache and felt very poorly. Mum put the *medicine* and some tea on the table next to the bed.

CQ: Why did Mum put the medicines on the table? Because Jane was ill/ Because Mum was ill

---

27.
Control: A banana peel lay on the pavement in front of the shop when Jerry slipped on it. Jerry ***slipped*** very suddenly as he was coming out with his shopping.

Synonym: A banana peel lay on the pavement in front of the shop and it made Jerry skid. Jerry *slipped* very suddenly as he was coming out with his shopping.

Paraphrase: A banana peel lay on the pavement in front of the shop when Jerry stepped on it. Jerry *slipped* very suddenly as he was coming out with his shopping.

Inference: A banana peel lay on the pavement in front of the shop where Jerry always shopped. Jerry *slipped* very suddenly as he was coming out with his shopping.

CQ: What did Jerry slip on? The pavement/ The banana peel

28.
Control: No longer able to control his fury, Alice's boyfriend hurled the porcelain vase against the wall and broke it into pieces. It took Alice a long time to clean up the ***pieces*** from the floor.

Synonym: No longer able to control his fury, Alice's boyfriend hurled the porcelain vase against the wall and broke it into bits. It took Alice a long time to clean up the *pieces* from the floor.

Paraphrase: No longer able to control his fury, Alice's boyfriend hurled the porcelain vase against the wall and broke it. It took Alice a long time to clean up the *pieces* from the floor.

Inference: No longer able to control his fury, Alice's boyfriend hurled the porcelain vase against the wall with his full strength. It took Alice a long time to clean up the pieces from the floor.

Q: What happened with the vase? Boyfriend broke it/ Alice broke it

---

29.
Control: Mary was having severe labour pains and so she went to the hospital to give birth. Her baby's ***birth*** was a joyous occasion for the whole family.

Synonym: Mary was having severe labour pains and so she went to the hospital to deliver. Her baby's birth was a joyous occasion for the whole family.

Paraphrase: Mary was having severe labour pains and so she went to the hospital to see a midwife. Her baby's *birth* was a joyous occasion for the whole family.

Inference: Mary was having severe labour pains and so she went to the hospital in a hurry. Her baby's birth was a joyous occasion for the whole family.

Q: Why did Mary go to the hospital? To give birth/ To stop the pain

---

30.

Control: The drunk driver was speeding on the highway when he lost control and crashed his car. The car got completely destroyed in ***the crash*** but the driver survived.

Synonym: The drunk driver was speeding on the highway when he lost control and smashed his car. The car got completely destroyed in *the crash* but the driver survived.

Paraphrase: The drunk driver was speeding on the highway when he lost control of his car and steered towards a tree. The car got completely destroyed in *the crash* but the driver survived.

Inference: The drunk driver was speeding on the highway when he lost control of his car all of a sudden. The car got completely destroyed in *the crash* but the driver survived.

Q: Why did the driver crash? He was drunk/ He was speeding

31

Control: Nina noticed a mistake in her test just before she handed it in so she ran back to her desk and corrected it. She got an A in the test thanks to ***correcting*** the mistake.

Synonym: Nina noticed a mistake in her test just before she turned it in so she ran back to her desk and fixed it. She got an A in the test thanks to correcting the mistake.

Paraphrase: Nina noticed a mistake in her test just before she turned it in so she ran back to her desk and erased it. She got an A in the test thanks to correcting the mistake.

Inference: Nina noticed a mistake in her test just before she turned it in so she ran back to her desk as the time was running out. She got an A in the test thanks to correcting the mistake.

Q: Why did Nina ran to her desk? To correct the mistake/ To hand her paper in

---

32.

Control: When Holly noticed a stranger in her flat, she sat still in the dark corner of the laundry room and felt frightened. She felt so ***frightened*** her heart was racing like crazy.

Synonym: When Holly noticed a stranger in her flat, she sat still in the dark corner of the laundry room and felt terrified. She felt so *frightened* her heart was racing like crazy.

Paraphrase: When Holly noticed a stranger in her flat, she sat still in the dark corner of the laundry room and trembled with fear. She felt so *frightened* her heart was racing like crazy.

Inference: When Holly noticed a stranger in her flat, she sat still in the dark corner of the laundry room and made no sound. She felt so *frightened* her heart was racing like crazy.

Q: Why was Holly frightened? The corner was dark/ There was a stranger

---

33.
Control: Robin got new a painting that her friend Maria made for her, so she looked for a nail and a hammer to hang it on the wall. After ***hanging*** the painting she stood in front of it and admired it for a long time.

Synonym: Robin got new a painting that her friend Maria made for her, so she looked for a nail and a hammer to attach it to the wall. After *hanging* the painting she stood in front of it and admired it for a long time.

Paraphrase: Robin got new a painting that her friend Maria made for her, so she looked for a nail and a hammer to put it up. After *hanging* the painting she stood in front of it and admired it for a long time.

Inference: Robin got new a painting that her friend Maria made for her, so she looked for a nail and a hammer in her tool box. After *hanging* the painting she stood in front of it and admired it for a long time.

Q: What did Robin need the hammer for? For hanging the painting/ For making the painting

---

34.
Control: After spotting the cockroach in the bathtub, Nick reached over to grab an old shoe and killed it. Nick didn't like ***killing*** bugs, but he did it for his girlfriend.

Synonym: After spotting the cockroach in the bathtub, Nick reached over to grab an old shoe and murdered? it. Nick didn't like killing bugs, but he did it for his girlfriend.

Paraphrase: After spotting the cockroach in the bathtub, Nick reached over to grab an old shoe and swung it. Nick didn't like killing bugs, but he did it for his girlfriend.

Inference: After spotting the cockroach in the bathtub, Nick reached over to grab an old shoe and got ready. Nick didn't like killing bugs, but he did it for his girlfriend.

Q: Why did Nick grab the shoe? To kill the cockroach/ To help his girlfriend

---

35.
Control: Fiona had ripped her favourite skirt and thought she ruined it, but then she found some matching thread and sewed it up. Having ***sewed*** the skirt she wore it for another summer.

Synonym: Fiona had ripped her favourite skirt and thought she ruined it, but then she found some matching thread and stitched it up. Having sewed the skirt she wore it for another summer.

Paraphrase:. Fiona had ripped her favourite skirt and thought she ruined it, but then she found some matching thread and started fixing it. Having sewed the skirt she wore it for another summer.

Inference: Fiona had ripped her favourite skirt and thought she ruined it, but then she found some matching thread and needles. Having sewed the skirt she wore it for another summer.

Q: What did Fiona do with the matching thread? She found it/She sewed the skirt with it

---

36

Control: The policeman took out his gun and aimed it at the fleeing murderer, and then shot him. The ***shooting*** was reported on the news.

Synonym: The policeman took out his gun and aimed it at the fleeing murderer, and then fired. The shooting was reported on the news.

Paraphrase: The policeman took out his gun and aimed it at the fleeing murderer, and then pulled the trigger. The shooting was reported on the news.

Inference: The policeman took out his gun and aimed it at the fleeing murderer as he ran away. The shooting was reported on the news.

Q: What happened to the murderer? He got shot/ He fled

Appendix D

Stimulus set for Experiment 3

Target words are written in capitals at the beginning of each paragraph. The sentences in italics were used in the Long Distance condition to separate the context-giving sentence and the target word sentence. Only experimental items are included.

---

1. BIRTH

Control

Mary was having severe stomach cramps and so she went to the hospital to give birth. *Many friends and relatives came to see her. They brought lots of flowers and cards.* Mary's whole family celebrated the **birth** of her first baby.

Synonym

Mary was having severe stomach cramps and so she went to the hospital to deliver. *Many friends and relatives came to see her. They brought lots of flowers and cards.* Mary's whole family celebrated the **birth** of her first baby.

Paraphrase

Mary was having severe stomach cramps and so she went to the hospital to see a midwife. *Many friends and relatives came to see her. They brought lots of flowers and cards.* Mary's whole family celebrated the **birth** of her first baby.

Inference

Mary was having severe stomach cramps and so she went to the hospital in a hurry. *Many friends and relatives came to see her. They brought lots of flowers and cards.* Mary's whole family celebrated the **birth** of her first baby.

---

2. ACCIDENT

Control

Anne looked up and saw the bus was heading in her direction, but it was too late to stop the **accident** from happening. *Both Anne and the bus driver were very scared. Luckily, neither of them were going very fast.* The accident was later reported on the news.

Synonym

Anne looked up and saw the bus was heading in her direction, but it was too late to stop the collision*. Both Anne and the bus driver were very scared. Luckily, neither of them were going very fast.* The accident was later reported on the news.

Parahrase

Anne looked up and saw the bus was heading in her direction, but it was too late to avoid it on a narrow road. *Both Anne and the bus driver were very scared. Luckily, neither of them were going very fast.* The accident was later reported on the news.

Inference
Anne looked up and saw the bus was heading in her direction while honking at her and trying to brake. *Both Anne and the bus driver were very scared. Luckily, neither of them were going very fast*. The accident was later reported on the news.

---

3. HOSPITAL
Control
Allen's baby became very ill, so he put the baby in the car and rushed off to the hospital that was nearby. *The baby was crying due to her fever. Allen's wife tried to comfort her, but she kept crying*. The hospital had a long waiting queue.

Synonym
Allen's baby became very ill, so he put the baby in the car and rushed off to the infirmary that was nearby. *The baby was crying due to her fever. Allen's wife tried to comfort her, but she kept crying*. The hospital had a long waiting queue.

Parahrase
Allen's baby became very ill, so he put the baby in the car and rushed off to quickly get medical help. *The baby was crying due to her fever. Allen's wife tried to comfort her, but she kept crying*. The hospital had a long waiting queue.

Inference
Allen's baby became very ill, so he put the baby in the car and rushed off, speeding on the way. *The baby was crying due to her fever. Allen's wife tried to comfort her, but she kept crying*. The hospital had a long waiting queue.

---

4. CRASHED
Control
The drunk driver was speeding on the highway when he lost control and crashed his car. *He told himself he will never drink before driving again. It was a stupid thing to have done.* The car got completely destroyed when he crashed but he survived.

Synonym
The drunk driver was speeding on the highway when he lost control and smashed his car. *He told himself he will never drink before driving again. It was a stupid thing to have done*. The car got completely destroyed when he crashed but he survived.

Parahrase
The drunk driver was speeding on the highway when he lost control of his car and steered into a tree. *He told himself he will never drink before driving again. It was a stupid thing to have done*. The car got completely destroyed when he crashed but he survived.

Inference
The drunk driver was speeding on the highway when all of a sudden he lost control of his car. *He told himself he will never drink before driving again. It was a stupid thing to have done.* The car got completely destroyed when he **crashed** but he survived.

---

5. PARENTS

Control

After a long time away, Lisa booked a flight to visit her parents*. Her boss did not want her to go anywhere. They really needed her in the company.* The visit came as a great surprise to her parents as she hadn't told them she was coming.

Synonym

After a long time away, Lisa booked a flight to see her mum and dad. *Her boss did not want her to go anywhere. They really needed her in the company.* The visit came as a great surprise to her parents as she hadn't told them she was coming.

Parahrase

After a long time away, Lisa booked a flight home to see her family. *Her boss did not want her to go anywhere. They really needed her in the company.* The visit came as a great surprise to her parents as she hadn't told them she was coming.

Inference

After a long time away, Lisa decided to book a flight one rainy Monday afternoon. *Her boss did not want her to go anywhere. They really needed her in the company.* The visit came as a great surprise to her parents as she hadn't told them she was coming.

---

6. CRYING

Control

After discovering he had failed his drug test and lost the medal, Victor started crying in front of everyone. Victor regretted bitterly giving in to the temptation to cheat. He had practiced so hard for this race. His crying annoyed the other athletes who thought he shouldn't have cheated.

Synonym

After discovering he had failed his drug test and lost the medal, Victor started sobbing in front of everyone. Victor regretted bitterly giving in to the temptation to cheat. He had practiced so hard for this race. His crying annoyed the other athletes who thought he shouldn't have cheated.

Parahrase

After discovering he had failed his drug test and lost the medal, Victor's eyes started welling up. Victor regretted bitterly giving in to the temptation to cheat. He had practiced so hard for this race. His crying annoyed the other athletes who thought he shouldn't have cheated.

Inference

After discovering he had failed his drug test and lost the medal, Victor was overwhelmed with emotion. *Victor regretted bitterly giving in to the temptation to cheat. He had practiced so hard for this race.* His crying annoyed the other athletes who thought he shouldn't have cheated.

---

7. CORRECTED

Control

Nina noticed a mistake in her test just before she handed it in so she ran back to her desk and corrected it. It was a very important test. It determined whether Nina would be accepted to college or not. She got an A in the test because she corrected the mistake.

Synonym

Nina noticed a mistake in her test just before she handed it in so she ran back to her desk and fixed it. It was a very important test. It determined whether Nina would be accepted to college or not. She got an A in the test because she corrected the mistake.

Parahrase

Nina noticed a mistake in her test just before she handed it in so she ran back to her desk and altered it. It was a very important test. It determined whether Nina would be accepted to college or not. She got an A in the test because she corrected the mistake.

Inference

Nina noticed a mistake in her test just before she handed it in so she ran back to her desk as time was running out. *It was a very important test. It determined whether Nina would be accepted to college or not.* She got an A in the test because she **corrected** the mistake.

---

8. STOLE

Control

A suspicious man stuck a hanger in the window of Tom's car, opened the door, and stole the car in no time at all. *Nobody saw the man opening the car. It was dark and everything happened so fast.* In the morning, Tom found out that somebody **stole** his car at night.

Synonym

A suspicious man stuck a hanger in the window of Tom's car, opened the door, and pinched the car in no time at all. Nobody saw the man opening the car. It was dark and everything happened so fast. In the morning, Tom found out that somebody stole his car at night.

Parahrase

A suspicious man stuck a hanger in the window of Tom's car, opened the door, and drove off with the car in no time at all. Nobody saw the man opening the car. It was dark and everything happened so fast. In the morning, Tom found out that somebody stole his car at night.

Inference

A suspicious man stuck a hanger in the window of Tom's car, opened the door, and got in the car in no time at all. *Nobody saw the man opening the car. It was dark and everything happened so fast.* In the morning, Tom found out that somebody **stole** his car at night.

---

9. DEFEATED

Control

The team started the game badly and were defeated five nil by the end. Nobody expected the

game to be like this. The team was the tournament's favourite. After getting defeated in this match, there was no hope for getting into the finals anymore.

Synonym
The team started the game badly and were beaten five nil by the end. Nobody expected the game to be like this. The team was the tournament's favourite. After getting defeated in this match, there was no hope for getting into the finals anymore.

Parahrase
The team started the game badly and didn't score enough goals by the end. Nobody expected the game to be like this. The team was the tournament's favourite. After getting defeated in this match, there was no hope for getting into the finals anymore.

Inference
The team started the game badly despite their great initial prospects. *Nobody expected the game to be like this. The team was the tournament's favourite*. After getting defeated in this match, there was no hope for getting into the finals anymore.

---

10. SEWED
Control
Fiona had ripped her favourite skirt and thought she had ruined it, but then she found some matching thread and sewed it up. Fiona had always received many compliments whenever she wore this skirt. It was a favourite for a reason. Having sewed the skirt she wore it for another summer.

Synonym
Fiona had ripped her favourite skirt and thought she had ruined it, but then she found some matching thread and stitched it up. Fiona had always received many compliments whenever she wore this skirt. It was a favourite for a reason. Having sewed the skirt she wore it for another summer.

Parahrase
Fiona had ripped her favourite skirt and thought she had ruined it, but then she found some matching thread and started fixing it. Fiona had always received many compliments whenever she wore this skirt. It was a favourite for a reason. Having sewed the skirt she wore it for another summer.

Inference
Fiona had ripped her favourite skirt and thought she had ruined it, but then she found some matching thread and needles. *Fiona had always received many compliments whenever she wore this skirt. It was a favourite for a reason*. Having sewed the skirt she wore it for another summer.

---

11. FIRED
Control
Rick had lost one customer after another when his supervisor called him to his office and fired

him. Rick had never got along well with his supervisor. He was a very unpleasant man to work for. Rick was not upset as being fired from a lousy job didn't hurt him all that much.

Synonym
Rick had lost one customer after another when his supervisor called him to his office and sacked him. Rick had never got along well with his supervisor. He was a very unpleasant man to work for. Rick was not upset as being fired from a lousy job didn't hurt him all that much.

Parahrase
Rick had lost one customer after another when his supervisor called him to his office and told him to leave. Rick had never got along well with his supervisor. He was a very unpleasant man to work for. Rick was not upset as being fired from a lousy job didn't hurt him all that much.

Inference
Rick had lost one customer after another when his supervisor called him to his office and gave him a talking to. Rick had never got along well with his supervisor. He was a very unpleasant man to work for. Rick was not upset as being **fired** from a lousy job didn't hurt him all that much.

---

12. DONATION
Control
After hearing the charity talk, the banker felt sorry so for the homeless that he reached for his cheque book and made a large **donation**. The whole event turned out to be very successful. The talk especially moved a lot of people. This was the largest donation the charity had ever received.

Synonym
After hearing the charity talk, the banker felt so sorry for the homeless that he reached for his cheque book and made a large contribution. The whole event turned out to be very successful. The talk especially moved a lot of people. This was the largest donation the charity had ever received.

Parahrase
After hearing the charity talk, the banker felt so sorry for the homeless that he reached for his cheque book and gave them money. The whole event turned out to be very successful. The talk especially moved a lot of people. This was the largest donation the charity had ever received.

Inference
After hearing the charity talk, the banker felt so sorry for the homeless that he reached for his pen without hesitation. *The whole event turned out to be very successful. The talk especially moved a lot of people*. This was the largest donation the charity had ever received.

---

13. BEATEN

Control
When school was over, the bully waited for Joey in a small alley and beat him repeatedly. Joey was very scared of the bully, but he never reported him. He didn't want the teachers to intervene. Joey got beaten so badly his body was covered in bruises.

Synonym
When school was over the bully waited for Joey in a small alley and hit him repeatedly. Joey was very scared of the bully, but he never reported him. He didn't want the teachers to intervene. Joey got beaten so badly his body was covered in bruises.

Parahrase
When school was over, the bully waited for Joey in a small alley and attacked him fiercely. Joey was very scared of the bully, but he never reported him. He didn't want the teachers to intervene. Joey got beaten so badly his body was covered in bruises.

Inference
When school was over, the bully waited for Joey in a small alley as he was coming home. *Joey was very scared of the bully, but he never reported him. He didn't want the teachers to intervene.* Joey got beaten so badly his body was covered in bruises.

---

14. FIRE
Control
Mila was smoking her cigarette and carelessly dropped it during her walk in the woods causing a fire to flare up. Mila decided this was her last cigarette. From now on, she was a non-smoker. It took two days to extinguish the fire and it engulfed many trees.

Synonym
Mila was smoking her cigarette and carelessly dropped it during her walk in the woods causing flames to flare up. Mila decided this was her last cigarette. From now on, she was a non-smoker. It took two days to extinguish the fire and it engulfed many trees.

Parahrase
Mila was smoking her cigarette and carelessly dropped it during her walk in the woods where it set the leaves alight. Mila decided this was her last cigarette. From now on, she was a non-smoker. It took two days to extinguish the fire and it engulfed many trees.

Inference
Mila was smoking her cigarette and carelessly dropped it during her walk in the woods without noticing. *Mila decided this was her last cigarette. From now on, she was a non-smoker*. It took two days to extinguish the fire and it engulfed many trees.

---

15. BREAK UP
Control
When Helen found out that her boyfriend was not faithful to her she called him to say she

wants to break up with him. She could not believe he had really done this to her. She felt betrayed and humiliated. The break up was very emotional and they haven't spoken since.

Synonym
When Helen found out that her boyfriend was not faithful to her she called him to say she wants to split up with him. She could not believe he had really done this to her. She felt betrayed and humiliated. The break up was very emotional and they haven't spoken since.

Parahrase
When Helen found out that her boyfriend was not faithful to her she called him to say that she is leaving him. She could not believe he had really done this to her. She felt betrayed and humiliated. The break up was very emotional and they haven't spoken since.

Inference
When Helen found out that her boyfriend was not faithful to her she was so upset she called him immediately. *She could not believe he had really done this to her. She felt betrayed and humiliated*. The break up was very emotional and they haven't spoken since.

---

16. SMELL
Control
A week after his fishing trip, Charles accidentally forgot about the fish he left in his boot leaving a horrible smell. He was afraid she'd be annoyed at him as he was running late. He tried to hurry up. Charles opened the car and the smell made him gag.

Synonym
A week after his fishing trip, Charles accidentally forgot about the fish he left in his boot leaving a horrible stench. He was afraid she'd be annoyed at him as he was running late. He tried to hurry up. Charles opened the car and the smell made him gag.

Parahrase
A week after his fishing trip, Charles accidentally forgot about the fish he left in his boot stinking out the whole car. He was afraid she'd be annoyed at him as he was running late. He tried to hurry up. Charles opened the car and the smell made him gag.

Inference
A week after his fishing trip, Charles accidentally forgot about the fish he left in his boot until he needed to go to pick up his wife. *He was afraid she'd be annoyed at him as he was running late. He tried to hurry up.* Charles opened the car and the smell made him gag.

---

17. DINNER
Control
Jane was out of town for the evening and Jim was left to take care of the kids and make dinner for them all. The kids loved staying with Jim. He'd always play video games with them and watch TV. The dinner he made turned out surprisingly well given he had not cooked much before.

Synonym
Jane was out of town for the evening and Jim was left to take care of the kids and make tea for them all. The kids loved staying with Jim. He'd always play video games with them and watch TV. The dinner he made turned out surprisingly well given he had not cooked much before.

Parahrase
Jane was out of town for the evening and Jim was left to take care of the kids and prepare some food for them all. The kids loved staying with Jim. He'd always play video games with them and watch TV. The dinner he made turned out surprisingly well given he had not cooked much before.

Inference
Jane was out of town for the evening and Jim was left to take care of the kids all by himself which made him nervous. *The kids loved staying with Jim. He'd always play video games with them and watch TV*. The dinner he made turned out surprisingly well given he had not cooked much before.

---

18. SHOOTING
Control
The policeman took out his gun and aimed it at the fleeing murderer, shooting at him as he ran. The neighbours were shocked. It used to be a safe neighbourhood but none of them felt safe anymore. The shooting was reported on the news.

Synonym
The policeman took out his gun and aimed it at the fleeing murderer, firing at him as he ran. The neighbours were shocked. It used to be a safe neighbourhood but none of them felt safe anymore. The shooting was reported on the news.

Parahrase
The policeman took out his gun and aimed it at the fleeing murderer, and then pulled the trigger. The neighbours were shocked. It used to be a safe neighbourhood but none of them felt safe anymore. The shooting was reported on the news.

Inference
The policeman took out his gun and aimed it at the fleeing murderer as he ran away. *The neighbours were shocked. It used to be a safe neighbourhood but none of them felt safe anymore.* The shooting was reported on the news.

---

19. WROTE
Control
Mike had not been able to solve a tough maths problem until he spoke to his friend Jim and wrote Jim's ideas down. Jim was extraordinarily good at maths. Some people even thought he was a genius. Excited, he wrote down everything Jim mentioned in their conversation.

Synonym

Mike had not been able to solve a tough maths problem until he spoke to his friend Jim and jotted Jim's ideas down. Jim was extraordinarily good at maths. Some people even thought he was a genius. Excited, he wrote down everything Jim mentioned in their conversation.

Parahrase
Mike had not been able to solve a tough maths problem until he spoke to his friend Jim and took notes while they talked. Jim was extraordinarily good at maths. Some people even thought he was a genius. Excited, he wrote down everything Jim mentioned in their conversation.

Inference
Mike had not been able to solve a tough maths problem until he spoke to his friend Jim about it in a cafe. *Jim was extraordinarily good at maths. Some people even thought he was a genius.* Excited, he wrote down everything Jim mentioned in their conversation.

---

20. LIGHT

Control
John couldn't see anything in the dark until he found the box of matches on the table and tried to light one. Constant electricity cuts were getting really annoying. He would need to speak to the landlord soon. The weak light filled the room with a warm, yellow glow.

Synonym
John couldn't see anything in the dark until he found the box of matches on the table and tried to ignite one. Constant electricity cuts were getting really annoying. He would need to speak to the landlord soon. The weak light filled the room with a warm, yellow glow.

Parahrase
John couldn't see anything in the dark until he found a box of matches on the table and tried to strike one. Constant electricity cuts were getting really annoying. He would need to speak to the landlord soon. The weak light filled the room with a warm, yellow glow.

Inference
John couldn't see anything in the dark until he found a box of matches on the table just in front of him. *Constant electricity cuts were getting really annoying. He would need to speak to the landlord soon*. The weak light filled the room with a warm, yellow glow.

---

21. COLLAPSED
Control
Near the end of his first marathon, fatigued Sean barely finished the race and then collapsed. Many athletes around seemed to have problems with exhaustion after the run. The novices were the worst. Nobody was surprised that Sean collapsed as it was a marathon after all.

Synonym
Near the end of his first marathon, fatigued Sean barely finished the race and then fell to the ground. Many athletes around seemed to have problems with exhaustion after the run. The

novices were the worst. Nobody was surprised that Sean collapsed as it was a marathon after all.

Parahrase
Near the end of his first marathon, fatigued Sean barely finished the race and then broke down. Many athletes around seemed to have problems with exhaustion after the run. The novices were the worst. Nobody was surprised that Sean collapsed as it was a marathon after all.

Inference
Near the end of his first marathon, fatigued Sean barely finished the race he was so exhausted by the run. *Many athletes around seemed to have problems with exhaustion after the run. The novices were the worst.* Nobody was surprised that Sean collapsed as it was a marathon after all.

---

22. STOPPED
Control
With his petrol tank half empty, Frank kept driving his car further until it stopped all of a sudden. It was getting very cold. Frank was worried as he hadn't packed enough warm clothes. Having stopped in the middle of nowhere caused Frank serious problems.

Synonym
With his petrol tank half empty, Frank kept driving his car further until it came to a halt all of a sudden. It was getting very cold. Frank was worried as he hadn't packed enough warm clothes. Having stopped in the middle of nowhere caused Frank serious problems.

Parahrase
With his petrol tank half empty, Frank kept driving his car further until it ran out of fuel. It was getting very cold. Frank was worried as he hadn't packed enough warm clothes. Having stopped in the middle of nowhere caused Frank serious problems.

Inference
With his petrol tank half empty, Frank kept driving his car further into the mountains. *It was getting very cold. Frank was worried as he hadn't packed enough warm clothes.* Having stopped in the middle of nowhere caused Frank serious problems.

---

23. STUDY
Control
Since she was a child, Sally dreamed of being a scientist so she decided to study Physics at university. The universe had always fascinated her. Sally would often daydream about working on important space exploration projects. She loved to study Physics so much she spent all her free time working on problems.

Synonym
Since she was a child, Sally dreamed of being a scientist so she decided to read Physics at university. The universe had always fascinated her. Sally would often daydream about working

on important space exploration projects. She loved to study Physics so much she spent all her free time working on problems.

Parahrase
Since she was a child, Sally dreamed of being a scientist so she decided to take a Physics course at university. The universe had always fascinated her. Sally would often daydream about working on important space exploration projects. She loved to study Physics so much she spent all her free time working on problems.

Inference
Since she was a child, Sally dreamed of being a scientist and wanted to work for NASA straight after graduation. *The universe had always fascinated her. Sally would often daydream about working on important space exploration projects*. She loved to study Physics so much she spent all her free time working on problems.

---

24. ANSWERING
Control
The telephone had already rung five times before Rob reached over, answering it. Rob didn't like getting disturbed from work. However, the phone was ringing too loud to ignore it. He immediately regretted answering the phone as it turned out to be an angry client.

Synonym
The telephone had already rung five times before Rob reached over, picking it up. Rob didn't like getting disturbed from work. However, the phone was ringing too loud to ignore it. He immediately regretted answering the phone as it turned out to be an angry client.

Parahrase
The telephone had already rung five times before Rob reached over and started talking. Rob didn't like getting disturbed from work. However, the phone was ringing too loud to ignore it. He immediately regretted answering the phone as it turned out to be an angry client.

Inference
The telephone had already rung five times before Rob walked across the office. *Rob didn't like getting disturbed from work. However, the phone was ringing too loud to ignore it.* He immediately regretted answering the phone as it turned out to be an angry client.

---

25. CLOSED
Control
The flat was much sunnier than he liked, so Arthur walked over and closed the blinds. He was tired as he'd been up all night. He was looking forward to some rest. With the blinds closed, the room felt a lot darker and Arthur was pleased.

Synonym
The flat was much sunnier than he liked, so Arthur walked over and shut the blinds. He was tired as he'd been up all night. He was looking forward to some rest. With the blinds closed, the room felt a lot darker and Arthur was pleased.

Parahrase
The flat was much sunnier than he liked, so Arthur walked over to adjust the blinds. He was tired as he'd been up all night. He was looking forward to some rest. With the blinds closed, the room felt a lot darker and Arthur was pleased.

Inference
The flat was much sunnier than he liked, so Arthur walked over to the window. *He was tired as he'd been up all night. He was looking forward to some rest.* With the blinds closed, the room felt a lot darker and Arthur was pleased.

---

26. CRIMINAL
Control
Emily's husband had been very secretive about his work but Emily suspected nothing until the police told her he was a criminal. Her husband was a very responsible and peaceful man. He had always taken good care of her. After discovering that her husband was a criminal, Emily wasn't sure what to do.

Synonym
Emily's husband had been very secretive about his work but Emily suspected nothing until the police told her he was a gangster. Her husband was a very responsible and peaceful man. He had always taken good care of her. After discovering that her husband was a criminal, Emily wasn't sure what to do.

Parahrase
Emily's husband had been very secretive about his work but Emily suspected nothing until the police charged him with fraud. Her husband was a very responsible and peaceful man. He had always taken good care of her. After discovering that her husband was a criminal, Emily wasn't sure what to do.

Inference
Emily's husband had been very secretive about his work but Emily suspected nothing until the police called one day. *Her husband was a very responsible and peaceful man. He had always taken good care of her.* After discovering that her husband was a criminal, Emily wasn't sure what to do.

---

27. RUNNING

Control
Having bought a wonderful new dress which was a size too small, Melissa decided to start running to lose weight. The dress was a good motivation as it was very expensive. She had bought it for her brother's wedding. Melissa never enjoyed running but she was determined to fit into the dress.

Synonym
Having bought a wonderful new dress which was a size too small, Melissa decided to start jogging to lose weight. The dress was a good motivation as it was very expensive. She had

bought it for her brother's wedding. Melissa never enjoyed running but she was determined to fit into the dress.

Parahrase
Having bought a wonderful new dress which was a size too small, Melissa started training for half marathon to lose weight. The dress was a good motivation as it was very expensive. She had bought it for her brother's wedding. Melissa never enjoyed running but she was determined to fit into the dress.

Inference
Having bought a wonderful new dress which was a size too small, Melissa decided she needed to lose weight. *The dress was a good motivation as it was very expensive. She had bought it for her brother's wedding.* Melissa never enjoyed running but she was determined to fit into the dress.

---

28. PRESENT
Control
Ben had a crush on the pretty girl from his class so he stopped at a department store on the way to school and bought her a present. Ben had wanted to ask the girl out for a long time. That day, he finally decided to do it. The girl loved the present and thanked Ben with a happy smile.

Synonym
Ben had a crush on the pretty girl from his class so he stopped at a department store on the way to school and bought her a gift. Ben had wanted to ask the girl out for a long time. That day, he finally decided to do it. The girl loved the present and thanked Ben with a happy smile.

Parahrase
Ben had a crush on the pretty girl from his class so he stopped at a department store on the way to school to get her something nice. Ben had wanted to ask the girl out for a long time. That day, he finally decided to do it. The girl loved the present and thanked Ben with a happy smile.

Inference
Ben had a crush on the pretty girl from his class and so he stopped at a department store on the way to school. *Ben had wanted to ask the girl out for a long time. That day, he finally decided to do it.* The girl loved the present and thanked Ben with a happy smile.

---

29. RESTORATION

Control
The old Victorian house needed a lot of work so the Historical Society came in to do a restoration. The Historical Society had been doing a fantastic job protecting old buildings. Their large endowment made everything easier. The results of the four-week restoration were impressive.

Synonym
The old Victorian house needed a lot of work so the Historical Society came in to do a

renovation. The Historical Society had been doing a fantastic job protecting old buildings. Their large endowment made everything easier. The results of the four-week restoration were impressive.

Parahrase
The old Victorian house needed a lot of work so the Historical Society came in to fix the crumbling facade. The Historical Society had been doing a fantastic job protecting old buildings. Their large endowment made everything easier. The results of the four-week restoration were impressive.

Inference
The old Victorian house needed a lot of work so the Historical Society came in to help. *The Historical Society had been doing a fantastic job protecting old buildings. Their large endowment made everything easier.* The results of the four-week **restoration** were impressive.

---

30. DISCOVERING

Control
Finally the hours of lab work paid off with Joseph discovering an important gene involved in cancer. His academic career looked very promising. His papers were widely cited and his work was respected by his peers. After discovering the gene for tumour growth he got a lot of media attention.

Synonym
Finally the hours of lab work paid off with Joseph finding an important gene involved in cancer. His academic career looked very promising. His papers were widely cited and his work was respected by his peers. After discovering the gene for tumour growth he got a lot of media attention.

Parahrase
Finally the hours of lab work paid off with Joseph stumbling upon an important gene involved in cancer. His academic career looked very promising. His papers were widely cited and his work was respected by his peers. After discovering the gene for tumour growth he got a lot of media attention.

Inference
Finally the hours of lab work paid off and Joseph was happy that his cancer research was going well. *His academic career looked very promising. His papers were widely cited and his work was respected by his peers.* After discovering the gene for tumour growth he got a lot of media attention.

---

31. SOAKED
Control
Mark was walking in the park when the sky darkened and a sudden downpour soaked him. Gusts of wind blew leaves into his face. Mark deeply regretted not checking the weather properly. Without an umbrella, he got soaked to the bone by the time he got back home.

Synonym
Mark was walking in the park when the sky darkened and a sudden downpour drenched him. Gusts of wind blew leaves into his face. Mark deeply regretted not checking the weather properly. Without an umbrella, he got soaked to the bone by the time he got back home.

Parahrase
Mark was walking in the park when the sky darkened and a sudden downpour caught him unprepared. Gusts of wind blew leaves into his face. Mark deeply regretted not checking the weather properly. Without an umbrella, he got soaked to the bone by the time he got back home.

Inference
Mark was walking in the park when the sky darkened and he saw menacing clouds above. *Gusts of wind blew leaves into his face. Mark deeply regretted not checking the weather properly.* Without an umbrella, he got soaked to the bone by the time he got back home.

---

32. TRIPPED

Control
Paul was walking very quickly and didn't notice how uneven the pavement was until he tripped accidentally. It was still quite early in the morning, but the street was already very busy. Paul hated the rush hour. He felt very embarrassed that he tripped like a small boy.

Synonym
Paul was walking very quickly and didn't notice how uneven the pavement was until he stumbled accidentally. It was still quite early in the morning, but the street was already very busy. Paul hated the rush hour. He felt very embarrassed that he tripped like a small boy.

Parahrase
Paul was walking very quickly and didn't notice how uneven the pavement was until his foot got caught in a hole. It was still quite early in the morning, but the street was already very busy. Paul hated the rush hour. He felt very embarrassed that he tripped like a small boy.

Inference
Paul was walking very quickly and didn't notice how uneven the pavement was as he was rushing to a meeting. *It was still quite early in the morning, but the street was already very busy. Paul hated the rush hour.* He felt very embarrassed that he tripped like a small boy.

---

33. PURCHASED

Control
Erica decided to reward herself so she went on-line and purchased plenty of clothes she didn't need. It always made her feel good. After a week like the last one, she needed some cheering up. She had purchased four dresses and three pairs of shoes in the first hour alone.

Synonym
Erica decided to reward herself so she went on-line and bought plenty of clothes she didn't

need. It always made her feel good. After a week like the last one, she needed some cheering up. She had purchased four dresses and three pairs of shoes in the first hour alone.

Parahrase
Erica decided to reward herself so she went on-line and spent a lot of money on clothes she didn't need. It always made her feel good. After a week like the last one, she needed some cheering up. She had purchased four dresses and three pairs of shoes in the first hour alone.

Inference
Erica decided to reward herself as she felt down after a tough week and so she went on-line. *It always made her feel good. After a week like the last one, she needed some cheering up.* She had purchased four dresses and three pairs of shoes in the first hour alone.

---

34. CONCEAL

Control
Michael had been in love with his best friend's girlfriend but he made sure to conceal his feelings. She was the best girl he'd ever met. Ashamed, he envied his best friend's good fortune. It took him a lot of energy and effort to conceal his feelings well.

Synonym
Michael had been in love with his best friend's girlfriend but he made sure to hide his feelings. She was the best girl he'd ever met. Ashamed, he envied his best friend's good fortune. It took him a lot of energy and effort to conceal his feelings well.

Parahrase
Michael had been in love with his best friend's girlfriend but he did everything to never show this. She was the best girl he'd ever met. Ashamed, he envied his best friend's good fortune. It took him a lot of energy and effort to conceal his feelings well.

Inference
Michael had been in love with his best friend's girlfriend and it had troubled him deeply. *She was the best girl he'd ever met. Ashamed, he envied his best friend's good fortune.* It took him a lot of energy and effort to conceal his feelings well.

---

35. PERSONALITY

Control
Having survived the war, Esther then became a successful entrepreneur thanks to her strong **personality**. *She never gave up. It took a lot of effort to build a successful business like hers.* Her friends thought she had the strongest personality of anyone they knew.

Synonym
Having survived the war, Esther then became a successful entrepreneur thanks to her strong character. *She never gave up. It took a lot of effort to build a successful business like hers.* Her friends thought she had the strongest personality of anyone they knew.

Parahrase
Having survived the war, Esther then became a successful entrepreneur as her strong will

never gave up. *She never gave up. It took a lot of effort to build a successful business like hers.* Her friends thought she had the strongest personality of anyone they knew.

Inference
Having survived the war, Esther then became a successful entrepreneur despite all the hardships. *She never gave up. It took a lot of effort to build a successful business like hers.* Her friends thought she had the strongest personality of anyone they knew.

---

36. PERFORMANCE

Control
The precision and the perfect synchrony with which the dancers moved was unbelievable and the audience loved the **performance**. It was tremendously difficult to get through auditions for this dance company. Hundreds of dancers tried every year. It was by far the best performance that year.

Synonym
The precision and the perfect synchrony with which the dancers moved was unbelievable and the audience loved the show. It was tremendously difficult to get through auditions for this dance company. Hundreds of dancers tried every year. It was by far the best performance that year.

Parahrase
The precision and the perfect synchrony with which the dancers moved was unbelievable and the audience enjoyed watching them. It was tremendously difficult to get through auditions for this dance company. Hundreds of dancers tried every year. It was by far the best performance that year.

Inference
The precision and the perfect synchrony with which the dancers moved was unbelievable and only the most talented could dance like this. *It was tremendously difficult to get through auditions for this dance company. Hundreds of dancers tried every year.* It was by far the best performance that year.

---

37. HURRY

Control
Having slept in, the couple was in a huge hurry to get to the train station in time. They threw their belongings in a suitcase without checking. The room looked terribly messy after they left. They were in such a hurry that they forgot to take their camera with them.

Synonym
Having slept in, the couple was in a huge rush to get to the train station in time. They threw their belongings in a suitcase without checking. The room looked terribly messy after they left. They were in such a hurry that they forgot to take their camera with them.

Parahrase
Having slept in, the couple was running frantically to get to the train station in time. They threw their belongings in a suitcase without checking. The room looked terribly messy after they left. They were in such a hurry that they forgot to take their camera with them.

Inference
Having slept in, the couple really struggled to get to the train station in time. *They threw their belongings in a suitcase without checking. The room looked terribly messy after they left*. They were in such a hurry that they forgot to take their camera with them.

---

38. AFFECTION

Control
Sitting in the corner of the café, the writer watched a mother attending with affection to her disabled son. It must have been very difficult to take care of him. The mother seemed happy though. There was so much affection in the way she helped her little boy.

Synonym
Sitting in the corner of the café, the writer watched a mother attending with love to her disabled son. It must have been very difficult to take care of him. The mother seemed happy though. There was so much affection in the way she helped her little boy.

Parahrase
Sitting in the corner of the café, the writer watched a mother attending with gentleness to her disabled son. It must have been very difficult to take care of him. The mother seemed happy though. There was so much affection in the way she helped her little boy.

Inference
Sitting in the corner of the café, the writer watched a mother attending to her disabled son as the boy was sipping tea*. It must have been very difficult to take care of him. The mother seemed happy though.* There was so much affection in the way she helped her little boy.

---

39. DISTRESSED

Control
Henry felt very distressed before meeting his most important client for the first time. He spent hours preparing materials for this meeting. It could be a turning point for his career. As the meeting began, he was so distressed he could barely speak.

Synonym
Henry felt very worried before meeting his most important client for the first time. *He spent hours preparing materials for this meeting. It could be a turning point for his career*. As the meeting began, he was so distressed he could barely speak.

Parahrase
Henry's heart was racing and his palms sweaty before meeting his most important client for the first time. *He spent hours preparing materials for this meeting. It could be a turning point for his career*. As the meeting began, he was so distressed he could barely speak.

Inference
Henry was meeting his most important client for the first time and he knew the client had very high expectations. *He spent hours preparing materials for this meeting. It could be a turning point for his career*. As the meeting began, he was so distressed he could barely speak.

---

40. ENEMIES

Control
During his lifetime, Mozart had many enemies due to his unconventional and rude behaviour. How he could be so talented yet behave like this was a surprise to all. Some felt it was unfair. However, even his enemies had to praise his musical genius.

Synonym
During his lifetime, Mozart had many foes due to his unconventional and rude behaviour. How he could be so talented yet behave like this was a surprise to all. Some felt it was unfair. However, even his enemies had to praise his musical genius.

Parahrase
During his lifetime, Mozart had alienated many people due to his unconventional and rude behaviour. How he could be so talented yet behave like this was a surprise to all. Some felt it was unfair. However, even his enemies had to praise his musical genius.

Inference
During his lifetime, Mozart was known for his unconventional and rude behaviour. *How he could be so talented yet behave like this was a surprise to all. Some felt it was unfair.* However, even his enemies had to praise his musical genius.

Comprehension questions for Experiment 3

Participants needed to decide whether the presented statement was true or false based on the information they had read in the paragraph.

Allen's baby had to be taken to the hospital.

Victor won the race but lost his medal.

The team was likely to get into the finals.

The charity was raising money for poor children.

Helen forgave her boyfriend that he cheated.

The policeman shot the murderer.

Novices were the most likely to collapse after the marathon.

Rob regretted answering the phone as he had to speak to his angry girlfriend

Melissa wanted to lose weight by before going to her brother's wedding

Joseph's scientific career was stagnating.

Historical society offered no help with restoring the old Victorian building.

Erica had a tough week and decided to make herself feel better by spending money.

The dancers in the dance company were only moderately good.

The mother seemed happy to be helping her disabled son.

Henry was woefully unprepared for the meeting with his client.

Fillers based statements:

Jennifer was shocked when she won the Academy award.

Leila was very happy with her landlady and they got along well.

The customer left the waitress no tip because she was rude to him.

Julia had so much art she could open her own art gallery easily.

William got into his college's first team despite his illness.

Appendix 4

Stimulus set for Experiment 4

Target words are written in capitals at the beginning of each paragraph. Only experimental items are included.

---

1. HOSPITAL

Strong / Confirmed (SC)

Allen's baby became very ill so he put the baby in the car and rushed off to the nearby infirmary. The hospital had a long queue.

Weak / Confirmed (WC)

Allen's baby became very ill so he put the baby in the car and rushed off, speeding on the way. The hospital had a long queue.

Strong/ Unconfirmed (SU)

Allen's baby became very unhappy so he put the baby in the car and rushed off to his work that was nearby. The hospital had a long queue.

Weak/ Unconfirmed (WU)

Allen's baby became very unhappy so he put the baby in the car and rushed off, speeding on the way. The hospital had a long queue.

---

2. ACCIDENT

S/C

Anne looked up and saw the bus was heading in her direction but it was too late to stop the collision. The accident was later reported on the news.

W/C

Anne looked up and saw the bus was heading in her direction while honking at her. The accident was later reported on the news.

S/U

Anne looked up and saw the bus was passing in her direction but it was too late to stop it and get on. The accident was later reported on the news.

W/U

Anne looked up and saw the bus was passing in her direction while honking at her. The accident was later reported on the news.

---

3. BIRTH

S/C

Mary's big belly hurt as she went to the hospital to deliver. Mary's whole family awaited the **birth** of her first baby impatiently.

W/C
Mary's big belly hurt as she went to the hospital in a hurry. Mary's whole family awaited the **birth** of her first baby impatiently.

S/U
Mary's tender belly hurt as she went to the hospital after eating her dinner. Mary's whole family awaited the **birth** of her first baby impatiently.

W/U
Mary's tender belly hurt as she went to the hospital in a hurry. Mary's whole family awaited the **birth** of her first baby impatiently.

---

4. CRASHED

S/C
Eric was driving on the motorway when he lost control and wrecked his car. The car was brand new when he crashed it but Eric survived.

W/C
Eric was driving on the motorway when he lost control of his car all of a sudden. The car was brand new when he crashed it but Eric survived.

S/U
Eric was driving on the motorway when he tried to stop and check his car. The car was brand new when he crashed it but Eric survived.

W/U
Eric was driving on the motorway when he tried to stop his car all of a sudden. The car was brand new when he crashed it but Eric survived.

---

5. PARENTS

S/C
After a long time away from home, Lisa booked a flight to see her family. The visit came as a great surprise to her parents as she hadn't told them she was coming.

W/C
After a long time away from home, Lisa booked a flight one rainy Monday afternoon. The visit came as a great surprise to her parents as she hadn't told them she was coming.

S/U
After a long time away from her favourite people, Lisa booked a flight to see her friends again. The visit came as a great surprise to her parents as she hadn't told them she was coming.

W/U
After a long time away from her favourite people, Lisa booked a flight one rainy Monday afternoon. The visit came as a great surprise to her parents as she hadn't told them she was coming.

---

6. CRYING

S/C
After discovering he had failed his drug test and lost everything, Victor started sobbing in front of everyone. His crying annoyed the other athletes who thought he shouldn't have cheated.

W/C
After discovering he had failed his drug test and lost everything, Victor was overwhelmed with emotion. His crying annoyed the other athletes who thought he shouldn't have cheated.

S/U
After discovering he had failed his drug test again, Victor denied it in front of everyone. His crying annoyed the other athletes who thought he shouldn't have cheated.

W/U
After discovering he had failed his drug test again, Victor was overwhelmed with emotion. His crying annoyed the other athletes who thought he shouldn't have cheated.

---

7. CORRECTED
S/C
As she was handing her test in, Nina noticed a mistake in it and she ran back to her desk to fix it. Because she corrected the mistake, she got an A in the end.

W/C
As she was handing her test in, Nina noticed a mistake in it and she ran back to her desk as time was running out. Because she corrected the mistake, she got an A in the end.

S/U
As she was handing her test in, Nina noticed a mistake in it and she was afraid there were more to discover. Because she corrected the mistake, she got an A in the end.

W/U
As she was handing her test in, Nina noticed a mistake in it and she was afraid as time was running out. Because she corrected the mistake, she got an A in the end.

---

8. VANISHED
S/C
A suspicious man opened Tom's car and stole the car in no time at all. Tom found out his car vanished overnight as he came to the parking lot in the morning.

W/C
A suspicious man opened Tom's car and got in the car in no time at all. Tom found out his car vanished overnight as he came to the parking lot in the morning.

S/U
A tired man opened Tom's car and checked the engine in no time at all. Tom found out his car vanished overnight as he came to the parking lot in the morning.

W/U
A tired man opened Tom's car and got in the car in no time at all. Tom found out his car vanished overnight as he came to the parking lot in the morning.

---

9. DEFEATED

S/C

The team started the game badly and were beaten five nil by the end. After being defeated in this match, there was no hope for getting into the finals anymore.

W/C

The team started the game badly but were trying their best to win the game. After being defeated in this match, there was no hope for getting into the finals anymore.

S/U

The team started the game well and were leading five nil before the end. After being defeated in this match, there was no hope for getting into the finals anymore.

W/U

The team started the game well and were trying their best to win the game. After being defeated in this match, there was no hope for getting into the finals anymore.

---

10. SEWED

S/C

Fiona had ripped her favourite skirt and thought she had ruined it, but she found a matching thread and stitched it up. Having sewed the skirt she wore it for another summer.

W/C

Fiona had ripped her favourite skirt and thought she had ruined it, but she found a matching thread and needles. Having sewed the skirt she wore it for another summer.

S/U

Fiona had ripped her favourite skirt and thought she had ruined it, but she found a matching petticoat and fitted it underneath. Having sewed the skirt she wore it for another summer.

W/U

Fiona had ripped her favourite skirt and thought she had ruined it, but she found a matching petticoat. Having sewed the skirt she wore it for another summer.

---

11. FIRED

S/C

Rick had lost one customer after another when his supervisor called him to his office and sacked him. Rick did not mind being fired from a lousy job he hated and he wasn't upset at all.

W/C

Rick had lost one customer after another when his supervisor called him to his office to chat. Rick did not mind being fired from a lousy job he hated and he wasn't upset at all.

S/U

Rick had served one customer after another when his supervisor called him to his office and

offered him a drink. Rick did not mind being fired from a lousy job he hated and he wasn't upset at all.

W/U
Rick had served one customer after another when his supervisor called him to his office to chat. Rick did not mind being fired from a lousy job he hated and he wasn't upset at all.

---

12. DONATED
S/C
After hearing the charity talk, the banker felt so sorry for the homeless that he reached for his cheque book and made a large **contribution**. He donated the largest amount the charity had ever received.

W/C
After hearing the charity talk, the banker felt so sorry for the homeless that he reached for his cheque book without thinking. He donated the largest amount the charity had ever received.

S/U
After hearing the charity talk, the banker felt so sorry for the homeless that he reached for his handkerchief and wiped his eyes. He donated the largest amount the charity had ever received.

W/U
After hearing the charity talk, the banker felt so sorry for the homeless that he reached for his handkerchief without thinking. He donated the largest amount the charity had ever received.

---

13. BRUISED
S/C
When school was over, the bully waited for Joey in a park and hit him repeatedly. Joey was bruised all over the next day.

W/C
When school was over, the bully waited for Joey in a park on the way home. Joey was bruised all over the next day.

S/C
When school was over, Joey's classmate waited for him in a park to play with him. Joey was bruised all over the next day.

S/C
When school was over, Joey's classmate waited for him in a park on the way home. Joey was bruised all over the next day.

---

14. FIRE
S/C
Mila lit up her cigarette and carelessly dropped it during her walk, causing flames to flare up. It took two days to control the fire and it engulfed many trees.

W/C

Mila lit up her cigarette and carelessly dropped it during her walk. It took two days to control the fire and it engulfed many trees.

S/U
Mila lit up her cigarette and carelessly exhaled during her walk, spoiling the fresh air. It took two days to control the fire and it engulfed many trees.

W/U
Mila lit up her cigarette and carelessly exhaled during her walk. It took two days to control the fire and it engulfed many trees.

---

15. SEPARATION
S/C
When Helen heard that her boyfriend was not faithful to her, she called him to split up with him. The separation was very emotional and they haven't spoken since.

W/C
When Helen heard that her boyfriend was not faithful to her, she called him immediately. The separation was very emotional and they haven't spoken since.

S/U
When Helen heard that her boyfriend was complaining about her, she called him to make peace with him. The separation was very emotional and they haven't spoken since.

W/U
When Helen heard that her boyfriend was complaining about her, she called him immediately. The separation was very emotional and they haven't spoken since.

---

16. SMELL
S/C
After his fishing trip, Charles accidentally forgot about the fish he left in his car, leaving a horrible stench. The strong smell made Charles gag when he opened the car.

W/C
After his fishing trip, Charles accidentally forgot about the fish he left in his car he was so busy with work. The strong smell made Charles gag when he opened the car.

S/U
After his fishing trip, Charles accidentally forgot about the pizza he left in his car, leaving a horrible stain. The strong smell made Charles gag when he opened the car.

W/U
After his fishing trip, Charles accidentally forgot about the pizza he left in his car he was so busy with work. The strong smell made Charles gag when he opened the car.

---

17. COOKING

S/C
With Jane out of town, Jim was left to take care of the kids and keep them well fed. His cooking was very tasty despite not having much experience with it.

W/C
With Jane out of town, Jim was left to take care of the kids without help. His cooking was very tasty despite not having much experience with it.

S/U
With Jane out of town, Jim was left to entertain the kids and keep them busy all day. His cooking was very tasty despite not having much experience with it.

W/U
With Jane out of town, Jim was left to entertain the kids without help. His cooking was very tasty despite not having much experience with it.

---

18. SHOOTING

S/C
The policeman took out his weapon and pursued the fleeing murderer, firing at him as he ran. The shooting was reported on the news.

W/C
The policeman took out his weapon and pursued the fleeing murderer as he ran away. The shooting was reported on the news.

S/U
The policeman took out a club and pursued the fleeing murderer, hitting him as he ran. The shooting was reported on the news.

W/U
The policeman took out a club and pursued the fleeing murderer as he ran away. The shooting was reported on the news.

---

19. ABSCOND

S/C
The customers quickly took to the new barman who was a trustworthy guy and nobody expected he would run away with the profits. When he absconded with monthly takings, everyone was shocked.

W/C
The customers quickly took to the new barman who was a trustworthy guy and nobody expected what he'd do. When he absconded with monthly takings, everyone was shocked.

S/U
The customers quickly took to the new barman who was a trustworthy guy and owner expected he would increase the profits. When he absconded with monthly takings, everyone was shocked.

W/U

The customers quickly took to the new barman who was a trustworthy guy, as the owner expected they would. When he absconded with monthly takings, everyone was shocked.

---

20. LIGHT

S/C
John couldn't see anything in the dark when he found the box of matches on the table and tried to ignite one. The weak light filled the room with a warm, yellow glow.

W/C
John couldn't see anything in the dark when he found a box of matches on the table just in front of him. The weak light filled the room with a warm, yellow glow.

S/U
John couldn't see anything in the dark when he found a small paper box on the table and tried to open it. The weak light filled the room with a warm, yellow glow.

S/U
John couldn't see anything in the dark when he found a small paper box on the table just in from of him. The weak light filled the room with a warm, yellow glow.

---

21. COLLAPSED

S/C
Near the end of his first marathon, fatigued Sean successfully finished the race and then fell down. Nobody was surprised that Sean collapsed as it was a marathon after all.

W/C
Near the end of his first marathon, fatigued Sean successfully finished the race he had trained so hard for. Nobody was surprised that Sean collapsed as it was a marathon after all.

S/U
Near the end of his first marathon, enthusiastic Sean successfully finished the race and then cheered. Nobody was surprised that Sean collapsed as it was a marathon after all.

W/U
Near the end of his first marathon, enthusiastic Sean successfully finished the race he had trained so hard for. Nobody was surprised that Sean collapsed as it was a marathon after all.

---

22. STOPPED

S/C
With his petrol tank half empty, Frank kept driving his car further when it came to a sudden halt. Having stopped in the middle of nowhere caused Frank serious problems.

W/C
With his petrol tank half empty, Frank kept driving his car further into the mountains. Having stopped in the middle of nowhere caused Frank serious problems.

S/U
With his petrol tank full, Frank kept driving his car further when he came to a mountain hut. Having stopped in the middle of nowhere caused Frank serious problems.

W/U
With his petrol tank full, Frank kept driving his car further into the mountains. Having stopped in the middle of nowhere caused Frank serious problems.

---

23. PHYSICS

S/C
Since she was a child, Sally dreamed of being a scientist so she decided to do science at university. She loved physics so much she spent all her free time working on problems.

S/C
Since she was a child, Sally dreamed of being a scientist and wanted to have her own lab. She loved physics so much she spent all her free time working on problems.

S/U
Since she was a child, Sally dreamed of being an astronaut so she decided to get very fit at university. She loved physics so much she spent all her free time working on problems.

W/U
Since she was a child, Sally dreamed of being an astronaut and wanted to explore space. She loved physics so much she spent all her free time working on problems.

---

24. ANSWERING

S/C
The telephone had rung five times before Rob walked across the office to pick it up. He immediately regretted answering the phone as it turned out to be an angry client.

W/C
The telephone had rung five times before Rob walked across the office that morning. He immediately regretted answering the phone as it turned out to be an angry client.

S/U
The telephone had rung five times before Rob arrived in the office to start a new project. He immediately regretted answering the phone as it turned out to be an angry client.

W/U
The telephone had rung five times before Rob arrived in the office that morning. He immediately regretted answering the phone as it turned out to be an angry client.

---

25. CLOSED

S/C
Sunshine poured in to the flat and heated the air as Arthur walked over to shut the blinds. With the blinds closed the room felt a lot darker and Arthur was pleased.

W/C
Sunshine poured in to the flat and heated the air as Arthur walked over to the window. With the blinds closed the room felt a lot darker and Arthur was pleased.

S/U
Sunshine poured in to the flat and heated the air as Arthur leaned out to clean the blinds. With the blinds closed the room felt a lot darker and Arthur was pleased.

W/U
Sunshine poured in to the flat and heated the air as Arthur leaned out of the window. With the blinds closed the room felt a lot darker and Arthur was pleased.

---

26. CRIMINAL

S/C
Emily never knew much about her husband's job when the police called one day and charged him with an offence. Hearing that her husband was a criminal shook Emily badly.

W/C
Emily never knew much about her husband's job when the police called one day to speak to him. Hearing that her husband was a criminal shook Emily badly.

S/U
Emily never knew much about her husband's job when his boss called one day and charged him with a new project. Hearing that her husband was a criminal shook Emily badly.

W/U
Emily's husband had been secretive about his job when his boss called one day to speak to him. Hearing that her husband was a criminal shook Emily badly.

---

27. RUNNING

S/C
Having bought a wonderful new dress which was a size too small, Melissa decided to start jogging. Melissa never enjoyed running but she was determined to fit into the dress.

W/C
Having bought a wonderful new dress which was a size too small, Melissa decided to lose weight. Melissa never enjoyed running but she was determined to fit into the dress.

S/U
Having bought a wonderful new dress which was a size too big, Melissa decided to start eating. Melissa never enjoyed running but she was determined to fit into the dress.

W/U
Having bought a wonderful new dress which was a size too big, Melissa decided to fix it. Melissa never enjoyed running but she was determined to fit into the dress.

---

28. PRESENT

S/C

Ben had a crush on a girl from his class so he stopped in a shop on the way from school and bought her a gift. The girl saw the present and thanked Ben with a happy smile.

W/C

Ben had a crush on a girl from his class and so he stopped in a shop on the way from school. The girl saw the present and thanked Ben with a happy smile.

S/U

Ben held a grudge against a girl from his class so he stopped in a shop on the way from school and bought himself chocolate. The girl saw the present and thanked Ben with a happy smile

W/U

Ben held a grudge against a girl from his class so he stopped in a shop on the way from school. The girl saw the present and thanked Ben with a happy smile

---

29. RESTORATION

S/C

The old museum needed a lot of work so the Historical Society came in to do a renovation. The intensive **restoration** yielded impressive results.

W/C

The old museum needed a lot of work so the Historical Society came in to help. The intensive **restoration** yielded impressive results.

S/U

The old museum needed more attention so the Historical Society came in to provide materials. The intensive **restoration** yielded impressive results.

W/U

The old museum needed more attention so the Historical Society came in to help. The intensive **restoration** yielded impressive results.

---

30. DISCOVERED

S/C

Finally the hours of lab work paid off with Joseph finding an important gene involved in cancer. After Joseph discovered the gene for tumour growth he got a lot of media attention.

W/C

Finally the hours of lab work paid off and Joseph was happy that his research was successful. After discovering the gene for tumour growth he got a lot of media attention.

S/U

Finally the hours of lab work were over with Joseph mutating an important gene involved in cancer. After Joseph discovered the gene for tumour growth he got a lot of media attention.

W/U
Finally the hours of lab work were over and Joseph was happy that his research was done. After Joseph discovered the gene for tumour growth he got a lot of media attention.

---

31. SOAKED

S/C
Mark was walking in the park when the sky darkened and a sudden downpour drenched him. With a light jacket on, he got soaked to the bone by the time he got back home.

W/C
Mark was walking in the park when the sky darkened and the wind started blowing strongly. With a light jacket on, he got soaked to the bone by the time he got back home.

S/U
Mark was walking in the park when the sky lightened and a sudden rainbow appeared. With a light jacket on, he got soaked to the bone by the time he got back home.

W/U
Mark was walking in the park when the sky lightened and the wind started blowing strongly. With a light jacket on, he got soaked to the bone by the time he got back home.

---

32. TRIPPED

S/C
Paul was walking very quickly and didn't notice how uneven the pavement was until he stumbled accidentally. He felt very embarrassed that he tripped like a small boy.

W/C
Paul was walking very quickly and didn't notice how uneven the pavement was as he was rushing to a meeting. He felt very embarrassed that he tripped like a small boy.

S/U
Paul was walking very quickly and didn't notice how dirty the pavement was until he stepped on a dead rat. He felt very embarrassed that he tripped like a small boy.

W/U
Paul was walking very quickly and didn't notice how dirty the pavement was. He felt very embarrassed that he tripped like a small boy.

---

33. PURCHASED

S/C
Erica decided to reward herself so she went online and bought plenty of useless things. She had purchased four dresses and three pairs of shoes in the first hour alone.

W/C
Erica decided to reward herself after a tough week at work and so she went online. She had purchased four dresses and three pairs of shoes in the first hour alone.

S/U
Erica decided to give herself a break so she went online and read about plenty of useless things. She had purchased four dresses and three pairs of shoes in the first hour alone.

W/U
Erica decided to give herself a break after a tough week at work and so she went online. She had purchased four dresses and three pairs of shoes in the first hour alone.

---

34. CONCEAL

S/C
Michael was in love with his best friend's girlfriend but he made sure to hide it. It took him a lot of energy and effort to conceal his feelings well.

W/C
Michael was in love with his best friend's girlfriend and it had troubled him deeply. It took him a lot of energy and effort to conceal his feelings well.

S/C
Michael was upset with his best friend's girlfriend but he made sure she knew it. It took him a lot of energy and effort to conceal his feelings well.

W/U
Michael was upset with his best friend's girlfriend and she had troubled him deeply. It took him a lot of energy and effort to conceal his feelings well.

---

35. STRONGEST

S/C
Having survived the war, Esther became a successful entrepreneur thanks to her strong personality. Her friends thought she had the strongest character of anyone they knew.

W/C
Having survived the war, Esther became a successful entrepreneur despite all the hardships. Her friends thought she had the strongest character of anyone they knew.

S/U
Having survived the war, Esther became a nursery teacher thanks to her kind personality. Her friends thought she had the strongest character of anyone they knew.

W/U
Having survived the war, Esther became a nursery teacher despite all the hardships. Her friends thought she had the strongest character of anyone they knew.

---

36. TALENTED

S/C
The dancers moved with unbelievable precision and synchrony that only a natural could produce. They were s by far the most talented group in the academy that year.

W/C
The dancers moved with unbelievable precision and synchrony that the choreographer saw only rarely. They were by far the most talented group in the academy that year

S/U
The dancers moved with imperfect precision and synchrony that only a lack of practice could produce. They were by far the most talented group in the academy that year.

W/U
The dancers moved with imperfect precision and synchrony that the choreographer saw only rarely. They were by far the most talented group in the academy that year

---

37. HURRY

S/C
Having overslept, the couple was in a huge rush to get to the train station in time. They were hurrying so much they forgot to take their camera with them.

W/C
Having overslept, the couple was packing light to get to the train station in time. They were hurrying so much they forgot to take their camera with them.

S/U
Having slept little, the couple was in a bad mood when they got to the train station in time. They were hurrying so much they forgot to take their camera with them.

W/U
Having slept little, the couple was packing light to get to the train station in time. They were hurrying so much they forgot to take their camera with them.

---

38. AFFECTION

S/C
Sitting in the corner of the café, the writer watched a mother attending with love to her disabled son. There was so much affection in the way she helped her little boy.

W/C
Sitting in the corner of the café, the writer watched a mother attending to her disabled son as the boy was sipping tea. There was so much affection in the way she helped her little boy.

S/C
Sitting in the corner of the café, the writer watched a mother talking with resignation to her disabled son. There was so much affection in the way she helped her little boy.

S/C
Sitting in the corner of the café, the writer watched a mother talking to her disabled son as the boy was sipping tea. There was so much affection in the way she helped her little boy.

---

39. DISTRESSED

S/C
Henry felt anxious before meeting his most important client for the first time as they were known for being difficult. As the meeting began, he was so distressed he could barely speak.

W/C
Henry was meeting his most important client for the first time and he knew they were known for being difficult. As the meeting began, he was so distressed he could barely speak.

S/U
Henry felt excited before meeting his most important client for the first time as they were known for powerful connections. As the meeting began, he was so distressed he could barely speak

W/U
Henry was meeting his most important client for the first time and he knew the client had powerful connections. As the meeting began, he was so distressed he could barely speak

---

40. ENEMIES

S/C
During his lifetime, Mozart had many foes due to his unconventional and rude behaviour. However, even his enemies had to praise his musical genius.

W/C
During his lifetime, Mozart was known for his unconventional and rude behaviour. However, even his enemies had to praise his musical genius.

S/U
During his lifetime, Mozart had many friends due to his unconventional and cheerful behaviour. However, even his enemies had to praise his musical genius.

W/U
During his lifetime, Mozart was known for his unconventional and cheerful behaviour. However, even his enemies had to praise his musical genius.

Comprehension statements for Experiment 4

Participants needed to decide whether the presented statement was true or false based on the information they had read in the paragraph. The following statements were presented:

Mary had her first baby.

Victor won the race but lost his medal.

The team was likely to get into the finals.

The charity was raising money for poor children.

Helen forgave her boyfriend for mistreated her.

The policeman ran after the murderer.

Novices were likely to collapse after the marathon.

Rob regretted answering the phone as he had to speak to his angry girlfriend.

Melissa bought the dress even though it did not fit her well.

Joseph's scientific career was stagnating.

Erica decided to make herself feel better by spending money.

The dancers were not a very promising group within the academy.

Henry was really enjoying the meeting with his client.

Fillers based statements:

Jennifer thought the nominees were not good.

Leila was very happy with her landlady and they got along well.

George thought there was no way his cabin would get flooded.

The customer left the waitress a tip even though she was rude to him.

Julia loved collecting art but she did not enjoy the auctions.

The workers' supervisor was a good poker player.

William got into his college's first team despite his illness.