Cognition 175 (2018) 53–68

Contents lists available at ScienceDirect

Cognition

journal homepage: www.elsevier.com/locate/cognit

Original Articles

Gaze patterns reveal how situation models and text representations contribute to episodic text memory

Roger Johansson[a,*], Franziska Oren[b], Kenneth Holmqvist[c,d]

[a] Department of Psychology, Lund University, Sweden
[b] Department of Psychology, University of Copenhagen, Denmark
[c] UPSET, North-West University Vaal, South Africa
[d] Faculty of Arts, Masaryk University, Brno, Czech Republic

ARTICLE INFO

Keywords:
Memory
Eye movements
Situation models
Text memory
Episodic memory
Internal simulation

ABSTRACT

When recalling something you have previously read, to what degree will such episodic remembering activate a situation model of described events versus a memory representation of the text itself? The present study was designed to address this question by recording eye movements of participants who recalled previously read texts while looking at a blank screen. An accumulating body of research has demonstrated that spontaneous eye movements occur during episodic memory retrieval and that fixation locations from such gaze patterns to a large degree overlap with the visuospatial layout of the recalled information. Here we used this phenomenon to investigate to what degree participants' gaze patterns corresponded with the visuospatial configuration of the text itself versus a visuospatial configuration described in it. The texts to be recalled were scene descriptions, where the spatial configuration of the scene content was manipulated to be either *congruent* or *incongruent* with the spatial configuration of the text itself. Results show that participants' gaze patterns were more likely to correspond with a visuospatial representation of the described scene than with a visuospatial representation of the text itself, but also that the contribution of those representations of space is sensitive to the text content. This is the first demonstration that eye movements can be used to discriminate on which representational level texts are remembered and the findings provide novel insight into the underlying dynamics in play.

1. Introduction

Imagine a scenario where you are absorbed by an exciting mystery novel and you are reading a text passage starting at the top of the page. The novel is written in a first-person perspective where the protagonist observes a tower block from a distance. It is vividly described how she inspects what can be discerned in each window, starting with the bottom flat and systematically moving her attention towards the top. The description ends at the bottom of the page. The physical layout of this text passage is thus spatially incongruent with how space unfolds in the story. At a later occasion, in a book discussion club, you are then required to recall information from this particular text passage. To what degree will such remembering activate a visuospatial representation of the text itself versus a visuospatial representation of what was described in it, and what is the consequence of spatial inconsistencies between those representations?

From previous research, there is evidence that a visuospatial representation analog to the physical layout of the text (such as a page) gets reactivated more or less automatically when information is recalled from it (Kennedy, Brooks, & Flynn, 2003). On the other hand, there is compelling evidence that a situation model of what a text passage describes, such as a visuospatial arrangement, is the primary level of representation for ongoing text comprehension as well as for subsequent recollections (e.g., Garnham, 1981; Rinck & Bower, 2000; Rinck, Hähnel, Bower, & Glowalla, 1997). In current models of memory retrieval, it is recognized that multiple forms of knowledge representations, such as sensory input from the external environment and internal thoughts and ideas, may combine in different ways and at different levels (e.g., Brainerd, Reyna, & Mojardin, 1999; Schacter & Madore, 2016) and it is possible that both representations of space may contribute to the information that is finally retrieved. For instance, it has been shown that memory for a text's situation model is superior to memory for the wording and syntactic structure (Kintsch, Welsch, Schamalhofer, & Zimmy, 1990), but also that this relationship may change due to different task goals (Curiel & Radvansky, 2002; Zwaan, 1994). Nonetheless, to our knowledge, no previous research on text memory has directly targeted to what degree space from the encoded text itself and space contained in a corresponding situation model

* Corresponding author at: Department of Psychology, Lund University, Box 213, SE-221 00 Lund, Sweden.
E-mail addresses: roger.johansson@psy.lu.se (R. Johansson), fo@psy.ku.dk (F. Oren), kenneth.holmqvist@ownit.nu (K. Holmqvist).

https://doi.org/10.1016/j.cognition.2018.02.016
Received 25 April 2017; Received in revised form 12 February 2018; Accepted 13 February 2018

contribute to episodic remembering. The present study aims to address this fundamental issue by tracking participants' eye movements when they retold previously read texts from memory.

1.1. Memory for the text structure

The situation where you are confronted with a question to which you know you have read the answer in a particular location of a text is probably familiar to most of us. For instance, this has happened to all three of the present authors. This notion is in line with the "spatial coding" hypothesis (Kennedy, 1992), which holds that readers construct a mental representation of the text in a visuospatial format analog to the physical layout of it (Kennedy, 1992; Kennedy et al., 2003). The spatial coding hypothesis is primarily based on Kennedy and Murray's seminal finding that readers are capable of launching accurate regressive saccades onto a target word as far as 40–50 characters back into the text (Kennedy & Murray, 1987; see also Baccino & Pynte, 1994; Pynte, Kennedy, Murray, & Courrieu, 1988). It has been claimed that the level of such a representation involves the computation of spatial coordinates in a two-dimensional space, such as a page or a display screen. Each piece of the text will then get associated with particular coordinates in that spatial representation (Kennedy, 1992; Kennedy & Murray, 1987). Furthermore, it has been suggested that such a mental representation of the text functions as an "external memory" store, which is permanently present and can be used to off-load working memory demands (Kennedy et al., 2003). In effect, this enables readers to selectively re-inspect previously read words through a spatial indexing mechanism that is far more efficient and economic than ab-initio re-readings (Kennedy et al., 2003). According to the spatial coding hypothesis, a mental representation that spatially corresponds with an encoded text should thus more or less automatically become activated when one recalls information from it.

In support of this claim, it has been demonstrated that the location of numbers in a text is spontaneously coded for (Fischer, Mills, & Shaki, 2010) and several studies have reported that participants can retain spatial representations of previously read texts. For instance, it has been shown that readers are capable of indicating in which corner of a page specific information was presented (e.g., Christie & Just, 1976; Rothkopf, 1971; Zechmeister & McKillip, 1972), and that participants who are confronted with a blank sheet perform well above chance levels when they are instructed to point toward target words from an encoded text (e.g., Le Bigot, Passerault, & Olive, 2011; O'Hara, Sellen, & Bentley, 1999; Rawson & Miyake, 2002). Moreover, in a study by Piolat, Roussey, and Thunin (1997), it was reported that texts encoded in a page-by-page manner produced superior retrieval performance on word location compared to texts that were encoded as scrolling text. This result indicates that the spatial reference frame provided by a page may act as an important retrieval cue and one's memory for word locations would thus be diminished when such a reference frame is unavailable. Also, it has been shown that writers frequently produce regressive saccades several sentences back into their compositions and such look-backs appear to serve a functional role for planning ongoing production (Torrance, Johansson, Johansson, & Wengelin, 2016).

However, while more recent studies have replicated Kennedy and Murray's (1987) original finding that readers are capable of producing regressive saccades far back into the texts, the claim that such look-backs are very precise has not been supported. For instance, Inhoff and Weger (2005) found that readers' distal regressions tend to end up only in a rough vicinity of the target word, and that several corrective saccades are needed before the target word becomes fixated. Moreover, Fischer (1999) conducted a thorough investigation of readers' memory for word location, where spatiotemporal characteristics of the presented sentences were systematically manipulated, as well as the delay between encoding and retrieval. The overall outcome of that study was that a precise memory for word location does indeed exist, but it has a limited spatial span and is only available for a short period of time (see also, Radach & McConkie, 1998). Based on such findings, it has been suggested that readers do not primarily rely on a page-based coordinate system but rather reconstruct the text by retrieving the temporal sequence of its content, which in turn can be used to deduce if particular information was presented in the beginning or towards the end of a text (e.g., Inhoff & Weger, 2005; Therriault & Raney, 2002). In line with such ideas, Fischer (2000) argued that an imprecise memory of word location is ultimately a consequence of the reader's goal, which is to understand what the text is about rather than to remember its spatial layout.

To summarize, there is considerable evidence that mental representations of the text can be retained in a visuospatial format analog to the physical layout of it. But the spatial precision of such representations appears to be quite coarse when considering specific word locations.

1.2. Memory for what the text is about

To understand a specific state of affairs described by a text, readers are required to go beyond particular words, grammar and propositions and construct mental representations in a format that instead is based on inferences from associated memories and general world knowledge (Zwaan & Radvansky, 1998). In language processing, this level of representation is often referred to as a *situation model* (Van Dijk & Kintsch, 1983; Zwaan & Radvansky, 1998). The ability to generate such situation models has been shown to play a significant role for ongoing text comprehension as well as for subsequent memory retrieval (Gambrell & Jawitz, 1993; Garnham, 1981; McKoon & Ratcliff, 1992). A prominent view is that situation models of specific events are grounded in corresponding sensorimotor simulations of how we perceive and act upon the world (Barsalou, 1999, 2008; Glenberg & Kaschak, 2002; Zwaan, 2009). For instance, it has been found that text comprehension often involves automatic activation of mental imagery (Bergen, Chang, & Narayan, 2004; Stanfield & Zwaan, 2001; Zwaan, Stanfield, & Yaxley, 2002) and it has been suggested that words can activate associated perceptual and motor experiences, which allow readers to vicariously experience what is being described (Zwaan, 2004; Zwaan & Madden, 2009). In support of this view, neurocognitive studies have demonstrated modality-specific brain activation during the comprehension of individual words (e.g. Hauk, Johnsrude, & Pulvermüller, 2004) and phrases (e.g. Aziz-Zadeh, Wilson, Rizzolatti, & Iacoboni, 2006), as well as during more extensive narratives (Speer, Reynolds, Swallow, & Zacks, 2009).

A major function of constructing situation models is to enable readers to predict content from future text segments (see Altmann & Mirković, 2009, for a review) and such predictions are critically dependent upon established relationships in a spatiotemporal framework (Kurby & Zacks, 2008). Relevant changes in the cause of events within and between such spatiotemporal frameworks are then prone to induce readers to update their current situation model (Glenberg, 1997; Kurby & Zacks, 2008; Zwaan & Radvansky, 1998). For instance, when reading a story where the main character is described to "take a cab from her office to the zoo", such a change in location would require the reader to establish a new set of narrative expectations. Several studies have investigated the influence of spatial change in text comprehension with the typical finding that the accessibility of an object within a described space is modulated by the distance between that object and the location that currently occupies the reader's focus of attention (e.g., Morrow, Greenspan, & Bower, 1987; O'Brien & Albrecht, 1992; Rinck & Bower, 2000; Rinck et al., 1997).

In sum, there is substantial evidence that situation models are fundamentally grounded in sensorimotor simulations of how the narrated events unfold in time and space. A mental representation of a visuospatial arrangement established in a previously read scene description would thus correspond to an internal simulation of the visuospatial circumstances contained in the corresponding situation model.