

## Coherence is also in the Eye of the Beholder

Max M. Louwerse (mlouwerse@memphis.edu)  
Danielle S. McNamara (d.mcnamara@mail.psyc.memphis.edu)  
Arthur C. Graesser (a.graesser@mail.psyc.memphis.edu)  
Patrick Jeuniaux (pjeuniaux@mail.psyc.memphis.edu)  
Fang Yang (fangyang@mail.psyc.memphis.edu)

Department of Psychology / Institute for Intelligent Systems  
Memphis, TN 38152 USA

### Abstract

Many eye tracking studies have investigated fixations in text comprehension tasks. Experimenters thereby manipulated text aspects (e.g., cohesion) in order to determine how texts are processed in different conditions. Stimuli tend to be carefully counterbalanced, and both texts and subjects are considered as random variables in analyses of variance. Often ignored are the variations of knowledge and reading skills between the readers. This paper argues that it is essential that readers' individual differences be taken into account. One of a series of experiments will be presented that show how coherence, measured in terms of fixation times, depends on both text cohesion and individual reading ability.

### Introduction

Text comprehension requires the construction of a coherent mental representation of the situations described in the text. Two constructs are of particular importance here: cohesion and coherence. We have used the term 'coherence' to refer to representational relationships formed in the mind of the reader and 'cohesion' to refer to properties of the text. Cohesion is the consistency of elements in the text, whereas coherence is the consistency of elements in the mental representation in the mind of the comprehender. Cohesion does not always guarantee coherence, and the absence of cohesion can still allow for coherence. However, a text containing cohesion devices (e.g., connectives like *because*) will generally be perceived as being more coherent than a text without them, while a text without cohesion devices may still be perceived as coherent.

For instance, in the sentence *The motor cyclist got a ticket, because he was driving too fast*, the connective *because* facilitates building a causal relation between the two events. However, without the connective, the two sentences (*The motor cyclist got a ticket. He was driving too fast*) can still be interpreted causally.

There are different types of cohesion, local and global, grammar-driven and vocabulary-driven, with cohesion strands varying from referential, temporal, spatial, causal, additive and intentional (Louwerse, 2004). Recently, a tool has been developed, called 'Coh-Metrix', that measures cohesion ([www.coh-metrix.memphis.edu](http://www.coh-metrix.memphis.edu)). The modules of Coh-Metrix use lexicons, part-of-speech classifiers, syntactic parsers, templates, corpora, latent semantic

analysis, and other components that are widely used in computational linguistics (Graesser, McNamara, Louwerse, & Cai, 2004; McNamara, Louwerse & Graesser, 2004).

Many psycholinguistic studies have shown that cohesion generally facilitates coherence (see Gernsbacher, 1994; Graesser, Goldman & Gernsbacher, 2003; Louwerse, 2002; Louwerse & Graesser, 2004; Zwaan & Radvansky, 1998). That is, the presence of cohesion devices in the text facilitates the construction of a coherent mental representation in terms of reading time, fixation time and recall. For instance, the presence of the connective *because* cues the comprehender that a causal relation needs to be built. Such a cue reduces processing time (Louwerse, 2001), facilitates recall and improves summarization (Louwerse, 2002). Similar findings have been reported for referential, temporal, spatial, and intentional cohesion (Zwaan & Radvansky, 1998).

But the literature has also found that cohesion does not always lead to coherence and comprehension benefits. McNamara and colleagues have shown strong interactions in cohesion and coherence in low- and high-knowledge readers (McNamara, 2001; McNamara & Kintsch, 1996; McNamara, Kintsch, Songer & Kintsch, 1996). In a series of experiments they found that low-knowledge readers benefit more from high-cohesive text both in comprehension and recall, but high-knowledge readers actually perform better on low-cohesive texts. This can be explained by the high-knowledge reader being induced reading the text more actively in order to bridge the cohesive gaps. High-knowledge readers are able to do this, because they have the needed background information. Low-knowledge readers, on the other hand, desperately need the cohesive cues to build a coherent mental representation.

Though there is a wealth of literature on cohesion and coherence (Graesser, Gernsbacher, & Goldman, 2004) and there is recent evidence that reading ability plays an important factor in the comprehension data for cohesion and genre variables, little is known of the interactions between cohesion and reading ability (McNamara, et al., 1996). Even less is known about these variables on online processes, with hardly any empirical evidence collected through highly precise methods like eye tracking. Our eye tracking labs have done exactly that: investigate the interactions of text genre, cohesion, and reading ability on online processes in

children. This study is in part instigated by the provocative findings of McNamara et al. (1996) showing that children with low domain knowledge benefit from a high cohesion text, but those with high domain knowledge benefit from a low cohesion text. McNamara et al. (1996) however used offline measures of coherence (free-recall, comprehension questions), children's domain knowledge, and expository (biology) texts. The question is whether the interaction between reading ability and cohesion can be, 1) extended to texts other than expository texts, 2) whether reading ability rather than domain knowledge can account for the individual difference effects, and 3) whether these findings also hold for online measures of coherence.

## Experiment

### Participants

Forty-four children from the Memphis City Schools participated in this experiment for payment. Children ranged in age from 8-11 years and were in their 3<sup>rd</sup> to 5<sup>th</sup> grade ( $M = 4.24$ ,  $SD = .52$ ). All children had normal or corrected-to-normal vision (no glasses). Prior to participating in the experiment, the child's parents or caregivers were informed about the nature of the study and signed an informed consent.

### Materials

Each child read four texts (which we call Plants, Heat, Orlando, and Moving). Two of these texts (Plants and Heat) are expository texts, the other two (Orlando and Moving) narrative texts. The original versions of these texts were modified such that the cohesion was higher in their manipulated versions. Version and presentation order of the texts was counterbalanced, such that each participant read one version of each of the four texts (two high and two low cohesion versions). A cohesion analysis by Coh-Metrix showed that for the far majority of the indices high cohesion texts have higher scores than the low cohesion texts.

### Apparatus

An SMI iView X High-Speed eye tracker stabilized the participants' head using an ergonomic chinrest. Participants were calibrated both before they began the experimental session and throughout the session to ensure reliable data. During calibration, participants viewed nine points on a 1024 x 768 computer monitor and the tracker recorded corresponding x-y coordinates. The eye tracker had a temporal resolution of 240Hz. The horizontal viewing angle was  $\pm 30^\circ$  and vertical view angle was  $30^\circ$  up and  $45^\circ$  down. The initial calibration process required approximately five minutes.

### Procedure

Participants arrived with their parents or caregivers in the research lab and were introduced to the equipment and procedure. As soon as the child felt comfortable,

parents/caregivers left the experiment room and waited outside for the duration of the experiment. Participants were first asked to answer questions from the Woodcock-Johnson Psychoeducational Battery (Reading Mastery Test, Word Attack) in order to determine individual differences between participants. This was followed by the eye tracking component of the experiment, whereby the participant read each of the four stories from a computer screen while their eye gaze was recorded. After each text, participants were asked to recall the story and to answer some comprehension questions. The results presented in this study will focus on the eye tracking data only.

## Results

The participants were divided in low ability and high ability readers on the basis of a median-split of performance on the Woodcock-Johnson Psychoeducational Battery (Reading Mastery Test, Word Attack). Two participants were removed from the analyses because over 40% of their fixation times were missing, potentially due to an error in calibration. Of the remaining 42 participants an analysis of variance was carried out on the total fixation time per word (in milliseconds). This fixation time measure is considered as a good reflection of reading comprehension at a text level.

A main effect was found for the average total fixation time per word for text genre ( $F(1, 80) = 6.95$ ,  $p = .01$ ). Narrative texts were read faster ( $M = 261.51$ ,  $SD = 150.01$ ) than expository texts ( $M = 297.37$ ,  $SD = 190.07$ ). No main effect was found for cohesion, though the observed pattern suggested that high cohesion texts were read faster than low cohesion texts. Not surprisingly, reading ability yielded a significant difference ( $F(1, 80) = 5.70$ ,  $p = .02$ ), with high ability readers reading faster than low ability readers.

A significant interaction was found for high and low ability readers, expository and narrative text genre, and low and high cohesion ( $F(1, 80) = 6.14$ ,  $p = .02$ ). The nature of this interaction is presented in Fig. 1a (high ability readers) and Fig. 1b (low ability readers).

These two figures show that while for low ability readers high cohesion texts facilitate processing – more so for expository than for narrative texts – for high ability readers this is not the case. Though the same patterns for narrative texts can be found for both types of readers, processing for expository text is considerably different. Instead of high cohesion facilitating processing, high cohesion hinders processing in high ability readers.

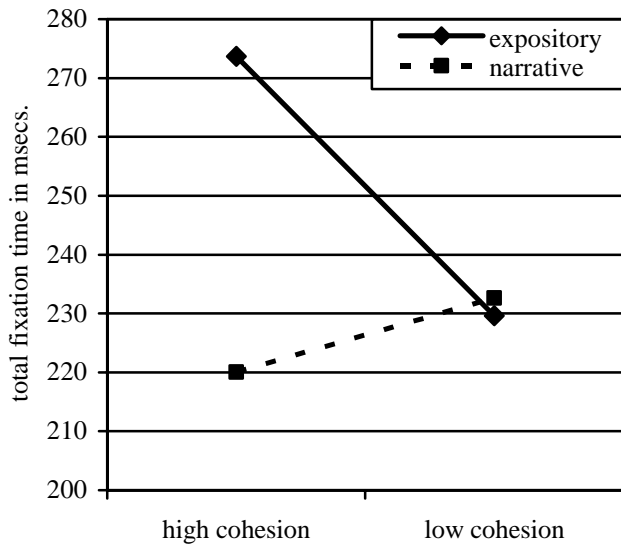


Fig. 1a. Total fixation time by genre and cohesion for high ability readers.

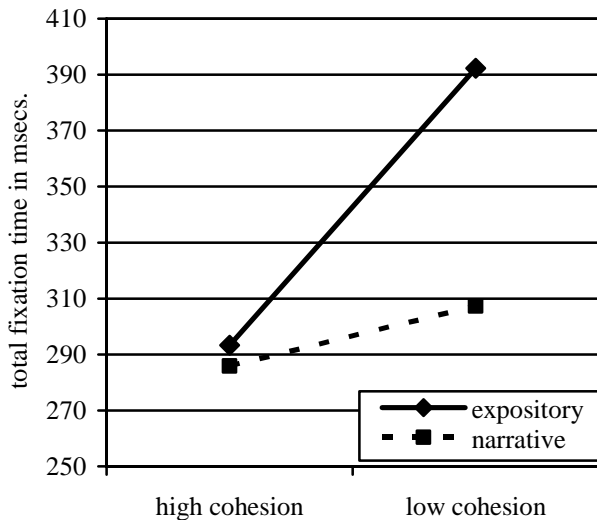


Fig. 1b. Total fixation time by genre and cohesion for low ability readers.

One question arises as to whether the three-way interaction is attributable to artifacts of text attributes, such as a greater number of difficult words in expository texts, or a greater number of easy words in high cohesion texts. Instead of the cohesive glue that this experiment was aimed to investigate, it may have been the case that word level information caused the interactive effects shown in Figures 1a and 1b. To rule out this possibility, the gaze duration on the word was considered as dependent variable, that is, the fixation time of the first time a reader looks at a word. If the same effects are found for gaze duration, the effects may be attributed to word information, rather than cohesion

information. Neither significant main effects nor interactions were found for gaze duration (all  $F$ s < 1.01,  $p$ s > .3), suggesting that the source of the interactive effects must be attributable to factors beyond the word level. That hypothesis is further confirmed by examining the number of regressions a participant makes. Instead of the time spent on a word, regressions measure the number of times a participant moves back and forth between words. Identical patterns were found for regressions as for total fixation time. Expository texts yielded more regressions than narrative texts ( $F(1, 80) = 9.53, p < .01$ ), high ability readers read faster than low ability readers ( $F(1, 80) = 7.18, p < .01$ ), and the same interaction was found for genre, cohesion and reading ability ( $F(1, 80) = 7.73, p < .01$ ). These results thus collectively suggest that when studies consider cohesion, they should also take reading ability into account.

Nonetheless, there is one other possibility for the effects found: in the manipulated texts (high cohesion version): extra words were necessarily added. For instance, causal, temporal and additive connectives were inserted between sentences. These added grammatical items may have contributed to an average lower total fixation time and lower number of regressions for the high cohesion texts. To allow for a comparison *ceteris paribus*, we should compare the average fixation time and regressions for only those words that can be found both in the low and high cohesion versions. Accordingly, 5.6% of the data was removed from Heat, 6.3% from Plants, 4.2% from Orlando and 8.5% from Moving. Main effects were found for text genre ( $F(1, 80) = 6.58, p = .01$ ) as well as for reading ability ( $F(1, 80) = 5.65, p = .02$ ). An interaction was also found between genre, cohesion, and reading ability ( $F(1, 80) = 6.52, p = .01$ ).

Reader	Expository	
	High cohesion	Low cohesion
high ability	266.64 (119.99)	229.65 (122.60)
low ability	283.75 (194.24)	392.74 (241.66)

Reader	Narrative	
	High cohesion	Low cohesion
high ability	215.62 (104.71)	232.68 (94.65)
low ability	280.42 (154.72)	307.34 (199.89)

Table 1. Total fixation times in milliseconds (and SD) for words shared between high and low cohesion texts

## Conclusion

The results of this eye tracking experiment thus support the findings in McNamara et al. (1996). The interaction between reading ability and cohesion however only holds up for expository texts; in narrative texts cohesion does not have an impact on processing for readers with either high or low reading ability. However, for the expository texts the McNamara et al. findings were replicated using 1) reading ability rather than domain knowledge, and 2) online

(fixation and regression) measures rather than offline measures.

The studies by McNamara et al. as well as the one reported here use manipulated texts. Recently, however, we have replicated the findings reported here using natural texts. Subjects read nine expository texts that varied in levels of cohesion (high, medium and low) as identified by Coh-Metrix. Again, those participants with a low reading ability benefited from the high cohesion texts, but those with a high reading ability did not.

These findings suggest that in eye tracking experiments where text comprehension is a dependent variable, not only text genre and text cohesion should be taken into account, as is usually done, but also the reader. Measuring coherence in the mind of the reader using the eye requires adding into the equation individual differences of its beholder.

### Acknowledgments

This research was supported by grants from National Science Foundation (IIS-0416128) and IES (R3056020018-02). Any opinions, findings, and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of NSF or IES. Correspondence concerning this article should be addressed to Max M. Louwerse, Institute for Intelligent Systems / Department of Psychology, University of Memphis, 202 Psychology Building, Memphis, Tennessee 38152-3230.

### References

- Gernsbacher, M.A. (1994). *Handbook of Psycholinguistics*, New York: Academic Press.
- Graesser, A.C., Gernsbacher, M.A., & Goldman, S. (2003) (Eds.). *Handbook of discourse processes*. Mahwah, NJ: Erlbaum.
- Graesser, A.C., McNamara, D.S., Louwerse, M.M. & Cai, Z. (2004). Coh-Metrix: Analysis of text on cohesion and language. *Behavior Research Methods, Instruments, and Computers*, 36, 193-202
- Louwerse, M.M. & Graesser, A.C. (2004). Coherence in discourse. In Strazny, P. (ed.), *Encyclopedia of linguistics*. Chicago: Fitzroy Dearborn.
- Louwerse, M.M. (2001). An analytic and cognitive parameterization of coherence relations. *Cognitive Linguistics*, 291-315.
- Louwerse, M.M. (2002). Computational retrieval of themes. In M. M. Louwerse & W. van Peer (Eds.), *Thematics: Interdisciplinary Studies* (pp. 189-212). Amsterdam/Philadelphia: John Benjamins.
- Louwerse, M. M. (2004). Un modelo conciso de cohesion en el texto y coherencia en la comprension. *Revista Signos*, 37, 41-58.
- McNamara, D.S. (2001). Reading both high and low coherence texts: Effects of text sequence and prior knowledge. *Canadian Journal of Experimental Psychology*, 55, 51-62.
- McNamara, D.S., & Kintsch, W. (1996). Learning from text: Effects of prior knowledge and text coherence. *Discourse Processes*, 22, 247-287.
- McNamara, D.S., Kintsch, E., Songer, N.B., & Kintsch, W. (1996). Are good texts always better? Text coherence, background knowledge, and levels of understanding in learning from text. *Cognition and Instruction*, 14, 1-43.
- McNamara, D.S., Louwerse, M.M., & Graesser, A.C. (2002). *Coh-Metrix: Automated cohesion and coherence scores to predict text readability and facilitate comprehension*. Technical report, Institute for Intelligent Systems, University of Memphis, Memphis, TN.
- Zwaan, R.A. & Radvansky, G.A. (1998). Situation models in language comprehension and memory. *Psychological Bulletin*, 123, 162-185.