Identifying Context-dependent Modes of Reading

Miho Fuyama[1] *, Shohei Hidaka[2], and Masaki Suwa[1]

[1] Keio University, 5322 Endo, Fujisawa-shi, Kanagawa 252-0882 Japan,
[2] Japan Advanced Institute of Science and Technology,
1-1 Asahidai, Nomi, Ishikawa 923-1292, Japan
{miho02}@sj9.so-net.ne.jp,
{shhidaka}@jaist.ac.jp
{suwa}@sfc.keio.ac.jp

Abstract. Reading time is commonly used as a basis for the study of reading. It has been observed in previous studies that unexpected sentences initiating a new context in a body of text considerably increases its reading time. There are two general accounts of the mechanism behind this slowdown. The first views reading, both slow and fast, to be a single, continuous process. The second views reading as consisting of multiple, distinct modes of processing. We consider in this paper the problem of distinguishing between these two accounts. To do so, we present a new kind of statistical analysis of data obtained from two different experiments on reading time. Our analysis supports the two-mode account of reading slowdown.

Keywords: Reading; Comprehension; Gamma Distribution

1 Introduction

Reading has been studied in various levels of cognitive processes, such as word identification, syntactic parsing, discourse processing [12]. At the levels of the comprehending narratives such as novels, readers have rich experiences; in the beginning they imagine the characters and the world, then devoutly read page after page with fun, and catch their breath at the unexpected unfoldment.

Past studies have focused to discuss the reading of shorter and controlled stories, and more rapid processing of reading than those broad changes with the contexts of literary works as described above. In those studies, reading time has been used to identify factors which may affect reading processes[8, 15, 4]. It has been reported that reading slows down considerably at the beginning of a clause, sentence, paragraph, or story – constructs which indicate a change in context – and subsequently speeds up again [1, and see also Gernsbacher (1997) for the

* Please note that the LNCS Editorial assumes that all authors have used the western naming convention, with given names preceding surnames. This determines the structure of the names in the running heads and the author index.

review]. These findings have been interpreted using theoretical models of reading processes.

For example, Gernsbacher(1990)[3] have proposed the *Structure building framework* which suggests that reading involves three component process - *foundation-laying*, *mapping*, and *shifting*. Foundation-laying is the first process of comprehension and in which readers lay a foundation for their mental structure. When the incoming information is consistent with previous information, readers *mapping* the new information onto their foundation. When there are inconsistencies between the incoming information and the existing information, readers *shift* and build a new structure.

In this study, we are interested in readers' cognitive processes during reading longer and real literary works. The question is whether readers' reading processes changes with contexts of stories in a qualitative manner, like *Structure building framework*, instead of only their speed. We also analyzed the reading time, but developed a statistical technique discussed in 1.1 which allows us to detect qualitative differences of cognitive processes depended on the contexts of these stories.

1.1 Statistical analysis of reading time

We assume the reading processes are composed of several subprocesses, and we call such qualitatively distinct collection of subprocesses involved in reading a *reading mode*. Then different collection of subprocess are regarded as different mode in this definition.

The question is, given reading data, how we can infer the number of reading modes reflected in the data? In the previous studies about short texts and more rapid processes, differences in reading time alone have been interpreted as reflection of two qualitatively distinct processes [2]. However, reading time may vary depending on multiple factors like frequency, familiarity, and the lengths of words [7, 14]. We cannot, therefore, naively interpret reading time alone as an indicator of multiple reading modes.

This observation motivates the development of new analysis technique for reading time. The analysis we present as an alternative is based on a statistical theory of processing time [6]. In this theory, the presence of multiple different modes of processing can be detected by the statistical distribution of the processing time.

If the reading process consisted of n subprocesses with the same constant processing rate over time, in other words the process finished only when all these subprocesses have finished, the reading time would follow a gamma distribution with shape parameter n (Figure 1 (a)). If, on the other hand, the reading process consisted of one subprocess with process rate t^k as a function of the process time t, in other words the process finished when at least one subprocess has finished, reading time would follow a Weibull distribution with shape parameter k (Figure 1 (b)).

Setting $n = 1$ in a gamma distribution or $k = 1$ in a Weibull distribution yields an exponential distribution. There is, therefore, a statistical relationship

between the types of distributions that processing time follows even as the number of Poisson processes involved with process vary.

This statistical analysis allows us to distinguish processes which have a same average speed of processes but have different number of subprocesses (Figure 2, A and B), and to distinguish processes which have same number of subprocesses but have different average speed of processes. This subprocess estimation gives an advantage over the previous studies analyzing differences in the reading time alone.

We adopt this statistical account of processing time in evaluating the number of reading modes based on reading time. If each observation in a reading time dataset follows essentially the same distribution as the others, we would treat this as an indicator of a single reading mode. If, on the other hand, the data set appears to have been generated by sampling from a mixture of distributions, we treat it as an indicator of multiple reading modes (Figure 3). Each dataset in question is composed of observations about a single subject. This technique therefore removes overall reading speed as a factor in the analysis.

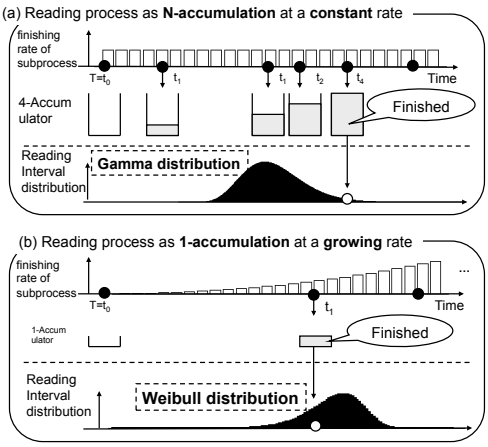

Fig. 1. Schematic illustration of the different types of reading processes and corresponding statistical distributions.

1.2 Approach

When reading, one is generally also engaged in many other processes – eye movements, posture management, etc. If one were only lightly engaged in reading and more heavily preoccupied with a number of these other activities, it is entirely possible that their preoccupation could appear as distinct reading modes in our statistical analysis. To prevent the detection of such false modes, it would be valuable to have a measure of reading engagement independent from reading time. We could then test the results of our statistical analysis based on their correlation with that measure.

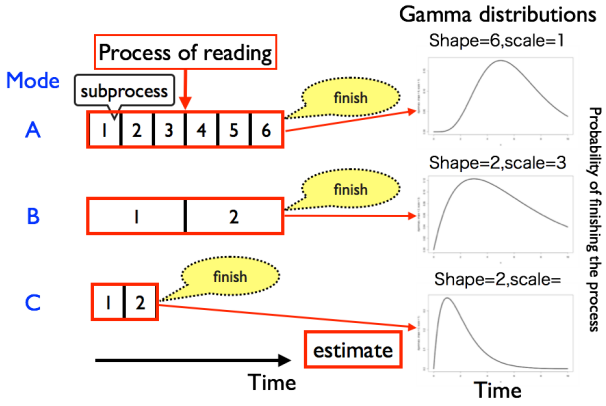

Fig. 2. Processes which have different numbers of subprocess and average speeds are estimated different distributions. These examples illustrate the processes which finish only when all subprocesses are finished. Mode A: 6 subprocesses, each takes short time on average, Mode B: 2 subprocesses, each takes long time, Mode C: 2 subprocesses, each takes short time. The overall average of both Mode A and B is the same, but their distributions (on the right hand side of the figure) are different.

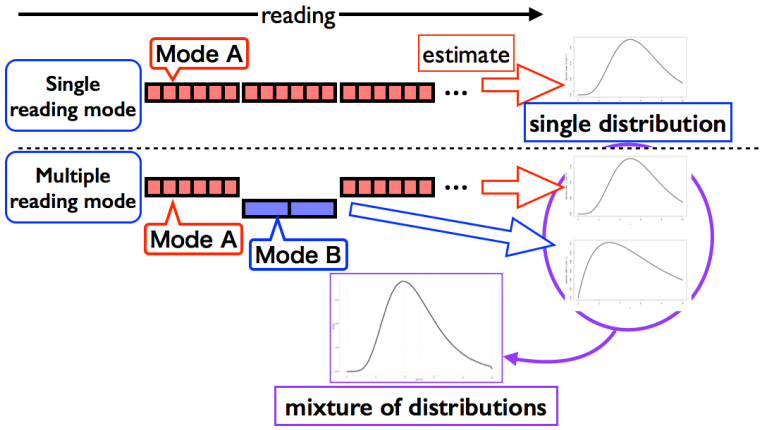

Fig. 3. (top) A single reading mode is repeated across multiple pages results in a single gamma distribution.
(bottom) Two reading modes, A and B, appear across different pages result in mixture of two gamma distributions.

Since the analytic technique we will use is statistical in nature, it requires relatively large datasets in order to produce meaningful results. To this end, and although this is not typical of existing studies of reading, we use entire novels as the texts in our experiments.

Given the burden that reading such long texts places on the subjects of our experiments, our first experiment consisted of only one subject – Miho Fuyama, the first author of this paper. She is an avid reader, which suggests that she is generally easily engaged in reading as an activity. In Experiment 1, we studied her reading time and the degree of engagement in reading across two books in order to empirically establish the validity of our analysis. We then analyzed data generated during her readings of 18 additional novels in order to test whether her reading process has a single or multiple reading modes.

Having validated our statistical analysis, we adopted it in our second experiment to a cross sectional study of multiple subjects. In Experiment 2, we asked 5 subjects to read a short novel. The subjects were also asked to evaluate their degrees of reading engagement each two pages after the reading session. This experiment was designed to evaluate whether our findings from Experiment 1 hold in general. We also evaluated changes in reading modes could be related to the semantic structure of the text itself. To do so, we analyze the consistency of the dynamics governing the change of reading modes across subjects and treat the consistent dynamics as text-specific semantic effects in reading.

2 Experiment 1

The first author was the sole subject of several high-load reading tasks. We asked her to read 20 Japanese novels. Each session tool one day including breaks. The set of samples from these 20 sessions of 20 novels was submitted to statistical analysis using the scheme described in the previous section, and we estimated the statistical distribution of her reading time for each two pages. For two of the novels (novels 17 and 18 in Table 1), she evaluated her degrees of absorption each two pages as an indicator of her engagement to reading. Specifically we asked her how absorbed she was in reading every pair of pages in these novels. These absorption ratings were used to validate the statistical analysis.

2.1 Participant

The subject was the first author, Miho Fuyama, who was 30 years old when the experiment was conducted. She is a native Japanese speaker, is a regular reader, and has normal vision.

2.2 Material

We used 20 Japanese novels, which the first author read for the first time in this experiment. The titles, authors, and page lengths of books are listed in Table 1. We selected as texts books witten by authors who have won Japan's prestigious literature prizes, such as the Naoki Prize or Akutagawa Prize.

Table 1. The novels read in Experiment 1.

No.	Title (Abbreviated)	Author	Page length
1	Shikisai	H. Murakami	370
2	Kamisama	H. Mori	314
3	Nameraka	H. Kawakami	189
4	Tenchi	T. Ubukata	474
5	Chinmoku	Y. Ogawa	308
6	Hikari	S. Miura	297
7	Kuchi	M. Banto	309
8	Mizuumi	B. Yashimoto	206
9	Kogoeru	A. Shino	401
10	Self-Reference	T. Enjo	308
11	Shi no izumi	H, Minagawa	427
12	Kisetsu no kioku	K. Hosaka	316
13	Eien no deguchi	E. Mori	313
14	Hokanaranu hito he	K. Shiraishi	295
15	Shorou tomurai dou	N. Kyogoku	498
16	Kodoku no utagoe	A, Tendo	312
17	Neko	Y. Ogawa	359
18	Ruto 225	C. Fujino	282
19	Yasashii uttae	Y. Ogawa	260
20	Burahuman	Y. Ogawa	146

2.3 Procedure

In each session of the experiment, the subject was asked to read a novel. Each session lasted several hours (including breaks), but was completed in one day. The subject reported her degrees of absorption for every two pages read in novels number 17 and 18. These reports were made approximately 100 days after the reading sessions. Her degree of absorption was measured on a five-level scale – "extremely bored", "bored", "normal", "absorbed", and "deeply absorbed". This scale was coded using the numbers $-2, -1, 0, 1$, and 2 respectively for each of the states. As the experiment required her to focus on and to become absorbed in such long texts, the subject was allowed to perform her readings at her home in order to minimize her tension. She was also allowed to have breaks whenever she wanted. The breaks were typically 5 to 15 minutes long, but there were also several hour-long lunch breaks. While reading, she sat at her desk and was videotaped with two small web cameras.

2.4 Analysis

From the videos, we transcribed the reading time for each pair of pages. These reading times were measured as the lengths of time between page turns, excluding time spent on break. The statistical analysis is performed on these transcribed reading times. We analyzed the aggregate of the data gathered across all

the sessions of the experiment in order to increase the statistical power of our analysis.

In our analysis, we fitted mixtures of exponential distributions, those of Weibull distributions, and those of gamma distributions to the aggregate data. For each mixture distribution, ranging from 1 to 5 components, we estimated the parameters by maximizing likelihood. As these statistical models have different numbers of parameters, we chose the model with the smallest Bayesian Information Criterion statistic [13] as the one which best explains the data.

2.5 Results and discussion

We found that a mixture of two gamma distributions provides the best fit to the aggregate data amongst all the distributions considered. Figure 4 illustrates the differences between these various classes of distributions in explaining our data. It shows the hazard function $H(t)$ of the page-turn interval t. The hazard function $H(t)$ is the conditional probability (density) to finish reading given the reading being unfinited until t. Exponential distributions in general exhibit a constant $H(t)$, which means this random process has "no memory" – a constant rate of reading interval regardless of time. Weibull and gamma distributions, in contrast, have increasing hazard functions. This means that the reading becomes more and more likely to be finished as time goes on. The two class of distributions, however, exhibit differences in the shape of their hazard functions.

The exponential distribution, with a constant hazard function, did not fit the data well in Figure 4 (BIC = 29421.71). Likewise, the Weibull distribution has large deviation from the data at the tails of distribution ($t < 30$ and $140 < t$) (BIC = 26146.06). The single gamma distribution fits better than the exponential and Weibull distributions (BIC = 25722.64), but the mixture of two gamma distributions provides the best fit (BIC = 25655.29). In addition, mixtures of three gamma distributions (BIC = 25677.24) or more did not provide better fits than the two-component gamma distribution.

Figure 5 shows the probability density function of empirical reading intervals and the estimated probability density function, which is a mixture of two gamma distributions. One subcomponent, Distribution 1, has shape 13.80 and scale 4.24. The other subcomponent, Distribution 2, has shape 7.58 and scale 10.67. This result suggests that the subject shows of two distinct modes in her reading, with each mode involving different reading subprocesses. It is worth noting that, at this point, we have not established the relationship between the two statistically estimated modes and the putative cognitive processes for reading.

Correlation to reading engagement We now address the question of whether the two distinct modes identified in our analysis are actually reflective of the text being read. In order to test this, we analyzed the correlation between the temporal change in mode and the degree of absorption reported by the reader. We obtained the reader's post-hoc report on engagement for each two pages of the books No. 17 and 18.

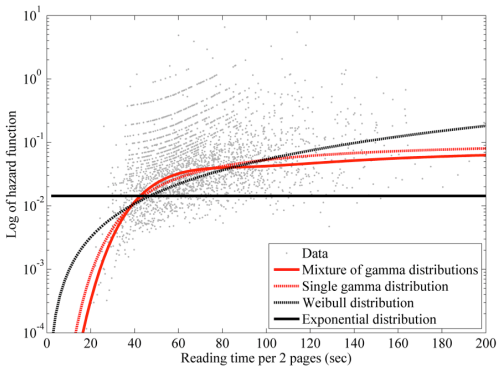

Fig. 4. The hazard function for the sample (dots) and for the estimated probability distributions (lines) of reading time per two pages.

Taking the book No. 17 as a representative case, Figure 6 shows the temporal profile of the weighted-average of shape parameters (black dots) and the reader's degrees of absorption (red dots). The weights was given by the mixture of the two gamma distributions for each reading time of two pages. The corresponding moving average of the two over 5 data points are shown as black and red line, respectively.

We performed correlation analysis for a pair of the estimated shape parameters and the degrees of absorption. For the book No. 17 across 141 pairs of pages, we had correlation -0.284 ($p < 0.001$). For the book No. 18 across 118 pairs of pages, we had correlation -0.283 ($p < 0.01$). This indicates that the temporal changes in the modes identified from our reading time analysis (Figure 6) does indeed reflect changes in reading engagement.

Remember that the shape parameter can be interpreted as the number of subprocesses involved in processing a given text, and the scale parameter can be interpreted simply as inverse of reading speed (Figure 1). Taking this theory into account, we conclude that the two modes estimated in this analysis are likely to represent a fast reading mode (Distribution 1) with a larger number of subprocesses and a slow reading mode (Distribution 2) with a smaller number of subprocesses.

3 Experiment 2

In Experiment 1, our statistical analysis detected two different modes of behavior in the reading data generated by the experiment. We further showed that the change in mode over time had a statistically significant correlation to the levels of engagement with the text reported by the subject. Our goal for Experiment 2 was to establish whether or not these findings are consistent across multiple subjects and, if so, to identify the various factors governing the reading modes detected in Experiment 1. In order to answer these questions, we design a short

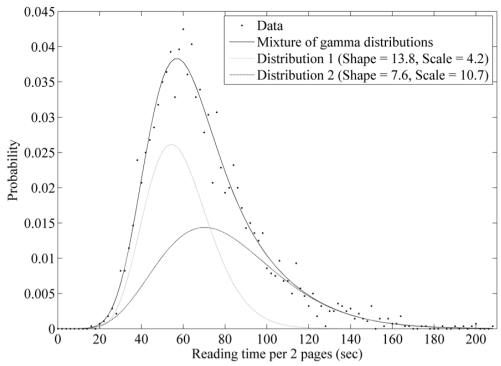

Fig. 5. Sample (dots) and estimated (solid) probability distribution of reading time per two pages. The two curves under the fitting curve shows subcomponents of the gamma mixture distribution.

experiment for the other subjects. In our second experiment, we asked different subjects to read a short novel but kept the rest of the procedure the same as it was in No.17 and No.18 of Experiment 1. Namely, subjects were asked to read a short novel or a part of it, and then they are asked to report their degrees of absorption for each two pages. The novel itself took less than an hour to read.

We expect two possible cases: We may have individual variance in reading time across subjects which would reflect that different subjects exhibited very different ways of processing the text. Second, the reading time may depend on the contextual structure of the text, and different subjects show similar mode changes in reading the same text.

The major factor dictating reading modes would be a subject's reading strategy in the first case. In the second case, it would be the contextual structure of the text itself.

3.1 Participants

In Experiment 2, which is on-going, we employed five participants to read one short story, and employed an additional subject to read another story. As this data collection is on-going, the numbers of participants for the two conditions were not balanced at the moment. The subjects were 4 male and 2 female undergraduate and graduate students at Keio University. Most of these subjects were not regular readers.

3.2 Procedure

The procedure was the same except for the length of the text and the environment in which the reading took place. During each session, one participant read a 49-page long short story in a room reserved specifically for the experiment.

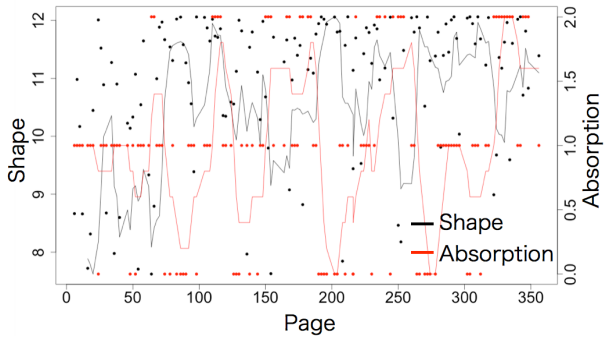

Fig. 6. Page-based temporal profile of the statistical property (shape parameter) of reading time and the absorption ratings of the case No.17.

Right after the reading session, the participant was asked to report their degrees of absorption in the same scale as Experiment 1 for each two pages. The short story they read is entitled "Onna no inai otoko tachi", which was commercially available as part of an omnibus and is authored by Haruki Murakami. In a controlled situation, a sixth participant read a part of another book "Chinmoku Hakubutukan" written by Yoko Ogawa. As the introductory part, from page 3 to 40, of the second book was chosen, this particular text does not include any major change of context. After their reading session, each subject was asked to report his/her degrees of absorption for each two pages using the same five point scale used in Experiment 1.

3.3 Analysis

For consistent comparison, we analyzed the aggregate of the reading time data across subjects by fitting to it a two-component mixture of gamma distributions. We fixed the class of distributions, instead of identifying it from data. This is largely due to the small sample size of our data at this point. Each participant provides reading time data for only 23 pairs of pages, and it did not sufficient statistical power to be conclusive even as its aggregation across subjects. Thus, we employed the statistical distribution estimated in Experiment 1.

3.4 Results and Discussion

Each panel of Figure 7 shows the page-based temporal profile of the modes estimated from reading time. In each panel, a dot shows estimated shape parameter for each reading time data point, and the line indicates its moving average. As in Experiment 1, we found that the temporal changes in modes were significantly correlated to the reported degree of reading engagement ($R = -0.238, p < 0.01$). This result replicates and generalizes the findings of Experiment 1.

Secondly, the result shown in Figure 7 exhibit inter-subject consistency in temporal changes in reading modes. Subjects A, B, C, D, and E read the same text, while subject F read another one. We found a similar inverted U-shape profile in the data of each of the first five subjects, whereas the data generated by subject F did not clearly show a similar pattern.

We performed correlation analysis on each pair of the subjects in order to test whether two groups of subjects have similar temporal profile of shape parameters within the group A, B, C, D, and E and dissimilar with the subject F. The average of correlations between pairs of subjects A, B, C, D, and E was 0.647 ($p < 0.01$), but the average correlations between subject F and the other subjects was 0.040 ($p > 0.693$).

As smaller shape and larger scale parameters suggest that fewer but slow subprocesses are involved in reading a text, high probability of this reading mode in the beginning of reading supports the structure-building account. In addition, this particular story has a "twist" about two-thirds of the way in. This twist is apparent in the upward trend of the probabilities in the first five panels of Figure 7. Our interpretation of this trend is that it reflects the structure-building which is like the *Structure building framework's laying a foundation* or *shifting* taking place due to the sudden contextual change in the storyline.

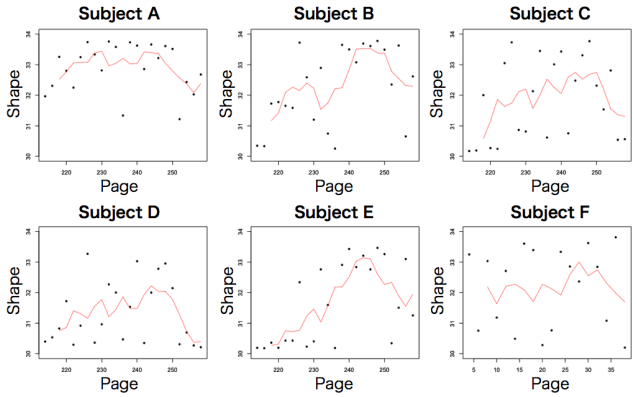

Fig. 7. The page-based temporal profile of the estimated shape parameters for each subject.

4 General Discussion

Reading is an essentially mental and subjective experience. Its cognitive underpinnings have been difficult to characterize directly, and reading time is a major tool for drawing inferences about the underlying cognitive mechanism behind reading. This study offers a new approach to the analysis of reading time, an

approach capable of identifying different modes of reading behavior from reading time data.

In Experiment 1, we collected an analyzed reading time data generated by a single subject reading several full novels in a natural situation. We observed significant correlation between subject's report of her engagement in reading and her reading modes inferred from the estimated reading time distribution. This experiment has three major implications:

1. In contrast to conventional studies on controlled, short readings, it is perhaps the first study on involving reading entire books in a more natural situation.
2. It establishes a new analytical technique for reading time data by associating the estimated modes with the subject's engagement in reading.
3. It provides evidence in favor of theroretical models which assume the existence of qualitatively distinct reading modes.

A clear limitation of Experiment 1 is that we could not employ many subjects due to the intense nature of the experiment. In Experiment 2, each session was designed to be as minimally demanding as possible. This allowed us to perform the experiment using a number of different subjects. We once again observed two distinct reading modes, and found that the mode switches across different subjects reading the same story were consistent with each other. This suggests that, to a large extent, the reading modes are dictated by the contextual structure of the text being read.

In the context of past discussions of reading, the implications of our findings may be related to the issue of *global coherence* in text, which is supposedly necessary to comprehend a text as a whole [8, 5]. In contrast to local inference on the level of words and sentences, global coherence refers to context- level smoothness over long passages which facilitates updates to textual knowledge and incorporating such knowledge with existing common knowledge [11]. Although there are a few experimental findings on global coherence using short texts [9], there is no well-established method of evaluating the global coherence of long texts such as those used in Experiment 1. Our new experimental procedure offers a way to access the relationship between temporal changes in the mechanism by which a reader processes text and the structure of the text itself. We expect that further studies along these lines would characterize global coherence.

In past studies, there are various theoretical model of reading processes such as *Structure building framework*, *Construction-integration model*, *Event indexing model* and so on. Some discussion have been done to integrate these models and provide the more comprehensive model of reading comprehension[10]. Mcnamara(2009)[10] concluded that each models explain different spectrum of comprehension processes. Our results of two distinct modes of reading can consist with his suggestion. It can be possible that several reading models of previous studies explain reading comprehension in one mode, and the other models explain that in another mode. If so, it should be reasonable to discuss our broad changes of mode with long text and previous models which focus on the rapid processes at the short text together for more understanding of reading processes.

5 Acknowledgements

The authors are grateful to Dr. Neeraj Kashyap for his proofreading of this manuscript. The first author was supported by Keio University Doctorate Student Grant-in-Aid Program and Mori Grants. The second author was supported by Grant-in-Aid for Scientific Research B KAKENHI No. 23300099.

References

1. Anderson, A., Garrod, S.C., Sanford, A.J.: The accessibility of pronominal antecedents as a function of episode shifts in narrative text. The Quarterly Journal of Experimental Psychology Section A: Human Experimental Psychology 35(February), 427–440 (1983)
2. Gernsbacher, M.A.: Two Decades of Structure Building. Discourse processes 23(3), 265–304 (1997)
3. Gernsbacher, M.A., Varner, K.R., Faust, M.E.: Investigating differences in general comprehension skill. Journal of Experimental Psychology: Learning, Memory, and Cognition 16(3), 430–445 (1990)
4. Graesser, A.C., Millis, K.K., Zwaan, R.A.: Discourse comprehension. Annual review of psychology 48, 163–89 (1997)
5. Graesser, A.C., Singer, M., Trabasso, T.: Constructing inferences during narrative text comprehension. Psychological review 101(3), 371–95 (1994)
6. Hidaka, S.: A computational model associating learning process, word attributes, and age of acquisition. PLOS ONE 8(10), e76242 (2013)
7. Inhoff, A.W., Rayner, K.: Parafoveal word processing during eye fixations in reading: Effects of word frequency. Perception & psychophysics 40(6), 431–439 (1986)
8. Kintsch, W.: The role of knowledge in discourse comprehension: A construction-integration model. Psychological Review 95(2), 163–182 (1988)
9. Magliano, J.P., Millis, K.K.: Assessing reading skill with a think-aloud procedure and latent semantic analysis. Cognition and Instruction 21(3), 251–283 (2003)
10. Mcnamara, D.S., Magliano, J.: Toward a Comprehensive Model of Comprehension. The Psychology of Learning and Motivation 51, 297–384 (2009)
11. Rapp, D.N., Kendeou, P.: Revising what readers know: Updating text representations during narrative comprehension. Memory & cognition 35(8), 2019–2032 (2007)
12. Rayner, K., Reichle, E.D.: Models of the Reading Process. Wiley Interdisciplinary Reviews: Cognitive Science 1(6), 787–799 (2010)
13. Schwarz, G.: Estimating the Dimension of a Model. The Annals of Statistics 6(2), 461–464 (1978)
14. White, S.J.: Eye movement control during reading: Effects of word frequency and orthographic familiarity. Journal of experimental psychology: Human perception and performance 34(1), 205–223 (2008)
15. Zwaan, R.A., Radvansky, G.A., Hilliard, A.E., Curiel, J.M.: Constructing Multidimensional Situation Models During Reading. Scientific Study of Reading 2(3), 199–220 (1998)