Some parameters of literary and news comprehension: Effects of discourse-type perspective on reading rate and surface structure representation

Rolf A. Zwaan *

This paper examines the assumption that there are specific control systems which regulate the comprehension of a particular discourse type, by altering parameters of basic comprehension processes. It focuses on some parameters for the comprehension of literary and news stories. This is done by presenting the same set of texts to two different groups of subjects, once as news stories, and once as literary stories. Two hypotheses are derived from theoretical assumptions concerning literary and news comprehension. The first one states that compared with a news perspective, a literary perspective imposes a slower reading speed. The second states that, compared with newspaper reading, literary reading results in a stronger surface structure representation. Reading-time measurements and memory tasks provide support for both hypotheses. The paper concludes that a combination of on- and off-line methods to study parameters of literary (and news) comprehension is both possible and necessary.

1. Introduction

The idea that the perspective from which a text is read influences the processing of that text is widely held, both in cognitive psychology and the empirical study of literature. Kintsch and Van Dijk (1978), for example, assume that discourse comprehension is always controlled by a specific schema, which contains the reader's knowledge of the structure of conventional text-types. As Kintsch (1980) points out:

'Normally, when a story is read, the comprehension is guided by a "story schema"... just as a scientific report would be read under the control of a "report schema", a news article under a

"newspaper schema", or as any text might be read with some special task-determined schema in control. The point is that ... the same comprehension operations occur when stories, news reports, or scientific texts are being read, but they are applied in different ways and for different purposes and hence have different effects.' (Kintsch 1980: 87).

Van Dijk and Kintsch (1983) extend this idea beyond the knowledge of discourse structures to include also knowledge of the specific social settings in which conventional types of discourse, such as poems, scholarly papers, lectures, and wedding speeches, occur. According to Van Dijk and Kintsch, these conventionalized settings may be systematically linked with schematic discourse structures, in such a way that they enable the language user to make plausible inferences about the actual categories of the schema.

In other words, language users utilize particular cognitive control mechanisms, underlying conventional communicative behavior. These mechanisms not only mobilize declarative conventional knowledge (e.g. about the setting in which a lecture usually takes place, the usual goals of a lecturer, and what is expected of the attendant – e.g., being able to ask a relevant question afterwards), but also procedural knowledge, particular cognitive strategies, such as strategies to deal with typographical deviations in poetry or ads. Some of these knowledge structures and procedures may be acquired via explicit instruction (e.g., a course in sonnet reading), or via incidental learning (learning by doing). Recent evidence for the role of incidental learning in acquiring reading skills is provided by Stanovich and West (1989). Their study demonstrates, in the orthographic domain, how reading itself may develop skills and a knowledge base that then serve to enable more efficient subsequent reading.

In short, discourse-type specific control systems influence the way a text is comprehended: how it is processed on-line, how it is represented in memory, and how the representation is maintained over a longer period of time. Following Kintsch (1980), it seems plausible to assume that the same comprehension operations occur when, for example, news reports or literary stories are read, but that these are used in different ways and for different purposes and thus have different effects. This would mean that the control system, which regulates these processes, alters some basic parameters of the comprehension process, dependent on the type of discourse to be processed. For example, we may assume that in the minds of proficient readers, a specific 'literary control system' is triggered by textual (e.g., the label 'a novel' on the title page of a book) or contextual information (e.g., the instruction 'read this poem'), and then sets the parameters for a conventional literary comprehension process. (Obviously, this is a simplification; a very important factor in triggering a literary control system would be the goal to read a literary text. However, the above account suffices for the present discussion.) To find out more about these parameters is the goal of this paper.

The question, then, is how one can render this idea empirically testable. A first step, quite in accordance with the methodological 'tradition' of the

empirical study of literature, would be to present one and the same text (or preferably: one and the same group of texts) to two groups of subjects and instruct one group that the text is a literary narrative and the other group that the text is of some other type (e.g., a popular scientific text); cf. Meutsch (1987). This makes it possible to obtain data on the influence of perspective, while textual factors are kept constant. Meutsch (1987) distinguished between literary and non-literary comprehension. However, as can be deduced from what has been said above, there is no standard type of 'normal', 'non-literary' comprehension with which literary comprehension is to be compared; mature readers may have evolved control systems for each conventional type of text they frequently encounter (e.g., newspaper articles, fiction, textbooks, manuals, advertisements, and so on). In other words, whatever we learn about literary reading through a design as just described is always relative to some other type of reading with which literary reading is compared. For this reason, it is methodologically important to formulate some assumptions about both types of reading.

The present paper compares the comprehension of literary narratives with the comprehension of another type of narrative for which it can be assumed that experienced readers have developed specific control systems: (eye-witness) news stories. This assumption is plausible in view of the fact that most people read newspaper texts on an almost daily basis; on this basis, one may expect them to have developed a specific news-comprehension control system (NCCS), which enables them to process news stories in a way that they find acceptable to their own standards (e.g. to keep in touch with what happens in the world) and/or to social constraints (e.g. to be able to discuss the world-news with relatives, friends, colleagues, etc).

Now, what hypotheses can be generated about the comprehension of literary versus news stories? A hypothesis that immediately springs to mind is that NCCS favors relatively fast processing: [1] news stories are read relatively fast: an article read on Wednesday morning might be outdated on Thursday morning, when a new edition hits the mat, or even on the same day, whenever radio or television news broadcasts updates. This may lead to a skimming strategy. A control system for comprehending literary narratives, which we will call LCCS, on the other hand, will probably favor a relatively slow processing speed. [2] Most (non-professional) readers will read literary texts for enjoyment, without having to worry about the text becoming outdated: The

---

[1] It is important to observe that very different predictions may be made when non-narrative, e.g. expository, newspaper texts are involved; these are presumably read at a much slower rate than news stories.

[2] The labels NCCS and LCCS are used for ease of communication. It stands to reason that there may be several more CSs involving news or literature: e.g., specific controls systems for the comprehension of televised news or poetry, or for the *production* of news reports and literary narratives. However, these matters will be left aside in this paper.

Iliad is usually not thought of as being superseded by the *Odyssey*, although
the latter is more up to date on the fate of the Greek and Trojan Warriors.
Conventionally, newspaper texts are regarded as disposables, whereas literary
texts are considered our cultural heritage. We may assume that adult readers
in a Western society have acquired this conventional wisdom and these value
judgments through explicit instruction and have internalized them in their
control systems, where they impinge on the reading process: newspaper texts
are skimmed, whereas literary texts are 'savored'.

A further aspect is associated with incidental learning. Readers of literature
are likely to have experienced on several occasions the pitfalls of reading
literature too hastily. Shallow processing of ambiguities, unusual metaphors,
garden-path sentences, puns, changes of perspective, and other features which
– when thought to be purposefully constructed – are considered typical of
literature may result in insufficient comprehension (e.g. by the reader's own
standards). In newspaper texts both author and copy-editor make sure that
potential sources of confusion are removed. And although everyone can come
up with examples where this strategy has turned out counterproductive, it is
plausible to assume that also by way of incidental learning an LCCS imposes a
slower reading speed than an NCCS. Therefore, ideas concerning explicit
instruction as well as inductive learning pertaining to literary and news
comprehension, enable us to launch the hypothesis that a given text is
processed slower when it is presented as a literary text than when it is
presented as a newspaper text.

The second assumption concerns part of the textual representation to be
constructed by readers. If our assumptions about literary and news compre-
hension are valid, literary reading should result in a stronger representation of
how something is expressed in a text than newspaper reading. [3] This idea can,
for example, be derived from Jakobson's (1960) well-known functions of
language. Jakobson suggests that in literature (but not, for instance, in news)
the poetic function is dominant, which implies a focus on the (structure of) the
message itself (rather than on, for instance, the situation described by the
message). Although Jakobson seems to regard the poetic function as a text-im-
manent one, it is possible to translate it into a more cognitively oriented
version, which states that a typical aspect of literary reading is the focus on
stylistic features, which potentially could result in a relatively (i.e., as com-
pared to newspaper reading) strong surface structure representation.

In order to attempt a first test of these assumptions, this paper features two
hypotheses. *Hypothesis 1* predicts that the average reading time for a particu-
lar text will be faster when it is read from a news perspective (NP) than from a

---

[3] In a similar vein, Vipond and Hunt (1984: 274) suggest that readers reading in a 'point-driven'
way, which they more or less equate with literary reading, will have especially accurate memory
for narrative surface (e.g. as opposed to 'information-driven reading').

literary perspective (LP). *Hypothesis 2* predicts that recognition and recall of textual surface features is better when a story is read from a literary than from a news perspective. The focus is on particular aspects of surface structure, namely discourse evaluations (cf. Hunt and Vipond 1986). It is conceivable that discourse evaluations are salient aspects of surface structure, which are probably better remembered than most other aspects of surface structure. This reduces the danger of a floor effect. Hypothesis 1 was tested using a subject-paced reading task and Hypothesis 2 by a combination of a time-based recognition task and a text-completion task. A sceptical objection to the use of this method may be that it creates an 'unnatural' experimental situation which will override any potential effects of discourse-type perspective. While it is known that the experimental situation can impinge upon subjects' behavior, it is certainly possible that the effects of discourse-type perspective are too robust to be overridden by contingencies of the experimental situation. The present study proceeds on this assumption.

It was decided to test Hypothesis 2 by a combination of tasks, because, in the absence of any similar research in the empirical study of literature, it was difficult to decide in advance whether a task would be sensitive enough to any inter-group (LP versus NP) differences; there could be floor or ceiling effects. Therefore, a combination of methods would enhance our chance of finding inter-group differences in surface representation.

As an attempt to rule out potential interfering factors, a pilot study was carried out, in order to empirically select texts that are acceptable both as a news and as a literary story, rather than just selecting texts on an intuitive basis as, for instance, Meutsch (1987) did. The pilot study will be reported in the next section.

To sum up, this study consists of two general phases. First, a pilot study used to select the materials for the reading task. Second, a subject-paced reading task, followed by recognition and recall tasks.

2. The selection of materials: A pilot study

Subjects. The subjects in the pilot study were ten students from the faculties of social sciences and letters of Utrecht University. 8 subjects were female, 2 male. The mean age of the subjects was 20.9 years. They were paid Dfl. 10,00 for participation in the pilot study, which took about 45 minutes.

Materials. Ten text-excerpts were pre-selected for the pilot study, most of them describing mass movements (demonstrations, parades, police charges). Five excerpts stemmed from modern Dutch and Flemish novels (by E. du Perron, *Het Land van Herkomst* (Country of Origin), Willem Frederik Hermans, *Onder Professoren* (Among Professors), Harry Mulisch, *De Aanslag* (The Assault), Jef Geeraerts, *De Zaak Alzheimer* (The Alzheimer Case), and

Jan Wolkers, *De Perzik der Onsterfelijkheid* (The Peach of Immortality). Five other excerpts were taken from newspaper reports in two Dutch quality papers, *De Volkskrant*, and *NRC-Handelsblad*. These excerpts concerned well-covered events from the 1980s, both in the Netherlands and abroad. Two excerpts featured events in the Netherlands: the arrival of Pope John Paul II in the Netherlands, May 1985, and the riots during the coronation of Queen Beatrix of the Netherlands, April 1980. Three excerpts concerned events abroad: an eye-witness description of the chaos after an uprising in South-Yemen, by the *New York Times* service, January 1986; the disaster with the US space-shuttle Challenger, January 1986; and the story of a former Rumanian political prisoner, December 21, 1989. Several excerpts contained explicit references to events which the subjects might recall, such as the demonstration against the stationing of nuclear missiles in the Netherlands, in Amsterdam 1981 (the Mulisch text), or the Challenger-disaster; other excerpts were more opaque as to when or where the reported events took place. Two of the texts were thought to be relatively typical of their genre, the Wolkers excerpt, and the Challenger excerpt. These were included in order to provide the subjects with some points of reference for the genre-classification task.

Procedure. Subjects were tested in groups of five. They were presented with a booklet. Each page contained a text and four questions plus rating scales. The texts were randomly ordered for each subject. The subjects received the following instructions: 'Each text you will be reading is either a passage from a newspaper report, or a passage from a literary novel. It is your task to indicate to which genre each text belongs. You also have to indicate how certain you are of your opinion and how typical you find the text of the genre you have chosen. Finally, please motivate your answer, for instance by underlining passages you found particularly characteristic of the genre you selected. Please remember that you must always choose one genre!' The most important question concerned the binary classification of the texts as either literary or newspaper. The confidence and typicality ratings (on a six-point scale) were included to gain more insight in individual judgments; for instance, if a subject would classify a text as a newspaper text, but gave very low confidence and typicality ratings, the classification was probably more based on guessing than on a well-founded judgment, which could imply that relatively little weight should be given to such a classification. The ratings were not intended for use in quantitative computations. Finally, the underlining task was intended to provide some qualitative data as to the subjects' conceptions of literary and newspaper characteristics, especially concerning textual features.

Results. A first step in the analysis is the computation of the average percentage of correct genre classifications per subject and per text. The average percentage per subject is 57%, which means that the subjects performed little better than when they would be guessing the genres of the texts.

Table 1
Accuracy scores per text

Text	Accuracy score
Wolkers	100%
Challenger	100%
Mulisch	90%
Hermans	70%
Pope John Paul II	70%
Yemen	70%
Coronation	30%
Geeraerts	20%
Rumanian prisoner	10%
Du Perron	10%

The best performance is a 70% accurate classification, the poorest is 50%. This seems to indicate that the texts used are rather opaque as to the genre to which they belong. However, it may well be that some texts are identified correctly by all subjects, whereas other texts are identified incorrectly by a majority of the subjects. Therefore, accuracy scores were also computed for texts. Table 1 shows the results.

Two texts received 100% accuracy scores (as expected, these were the Challenger text and the Wolkers text). This means that these texts should not be used in the reading-time experiment, because they are not credible under both perspectives. The same goes for the Mulisch text, which has an accuracy score of 90%. However, at the other extreme, two texts received only 10% accuracy scores: the Du Perron text and the text about the political prisoner in Rumania. That is, the latter text was found 'novelesque', while the former was regarded as 'newspaperish'. The question arises whether these texts should be selected for further experimentation. One could argue against their selection because they strongly tend towards one genre in the pilot study. However, a counterargument is that this is not the original genre of the excerpts; and since these excerpts have been processed in the original contexts (i.e., as parts of a novel or newspaper text, respectively), there is no reason why this should not be possible in the reading experiment. The accuracy scores for the other texts fall in between these extremes. The Geeraerts text has an accuracy score of 20% and thus tends towards the newspaper side. However, some of the subjects who identified the text as a newspaper excerpt indicated that they doubted their decision, because they noticed also 'literary' characteristics of the text. On the other hand, the Hermans text, which scores 70%, was discarded because the subjects who made incorrect identifications were rather unsure of their judgment and gave low typicality ratings, which implies that the credibility of this text as a newspaper text is questionable.

It is interesting to take a look at the subjects' motivation for their judg-ments. The most typical features of newspaper excerpts, according to the subjects, are: factual, down-to-earth, concise style ('telegram style'), mention-ing of facts and figures, absence of emotions and interpretation (i.e., no discourse evaluations), and 'objective'. The most typical features of the literary excerpts are: discourse evaluations ('adjectives which make the story more vivid and more imaginable', 'subjective emotions', 'opinions'), figurative lan-guage, creative word use, the presence of a subjective instance or narrator. The interesting thing about these criteria is, of course, that they are frequently deployed to back up incorrect classifications. This can mean two things: (1) some of the excerpts are not representative of their genre, (2) the conventional ideas about characteristics of literary and newspaper texts are off the mark and are at times equally applicable to the opposite genre. It seems that the latter is the more plausible account: everyone who studies eye-witness reports in newspapers will notice the many discourse evaluations in almost any text, while there are certainly literary excerpts which are not crowded by discourse evaluations.

In conclusion, the pilot study leads to the selection of six texts for Experiment 1: the literary excerpts by Du Perron and Geeraerts, and the newspaper texts about the visit of Pope John Paul II to Holland, the situation in Yemen, the Rumanian political prisoner, and the coronation riots. [4] With the proviso that the pilot study involved only a small number of subjects it can be said that all of these texts are to some extent acceptable both as newspaper and as literary text, especially when the differences between an explicit genre-classification task (Pilot study) and a subject-paced reading task are taken into account. The former task is in several respects more selective on the texts than the second: subjects are explicitly told that the texts can be either non-fictional or fictional (in other words, their attention is focused on a comparison between genre and text), they have much more time available to read the text, and they are able to interpret it 'holistically' rather than 'sequentially' (i.e., they can look back in the text, rather than moving through the texts on a line-by-line basis). The subjects in the reading task, on the other hand, are much more constrained. They are simply allocated to a perspective from which to read the texts and they are not able to reread some parts of the texts. For these reasons, we may assume that they will be less sensitive to any possible incongruences between induced perspective and text.

---

[4] A possible objection to this uneven distribution of text-types (4 newspaper excerpts and only 2 literary excerpts) is that there are underlying differences between literary and newspaper texts which do not manifest themselves in a text-classification task, but which show up during an on-line task. However, as it does not seem clear beforehand what these differences could be, it seems best to follow the design of the experiment; and if there are underlying differences, they should result in a statistically significant perspective X text interaction.

3. The experiment

3.1. Method

Subjects. Subjects were 40 students (27 female and 13 male) of the faculties of social sciences and humanities at Utrecht University. None of these subjects participated in the pilot study, and all of them were native speakers of Dutch. The mean age of the subjects was 23.6 years ($sd = 4.2$). They were paid Dfl. 10,00 for participation in the experiment, which lasted about 40 minutes.

Materials. The six texts which emerged from the pilot study as being able to pass as both news and literary stories were used as experimental texts in the reading task; two of these texts are presented in English translation in the appendix. The average length of the texts was 216 words. A passage from a travelogue by the Dutch author Cees Nooteboom was included as a practice text for the subjects and as a way to measure subjects' habitual reading rate before allocation to a perspective, because, as previous research has shown (cf. Van Oostendorp 1988), some of the variance in reading-time experiments is attributable to subjects' habitual reading speed. The average reading time per word for the practice text was later used as a covariate in the analyses of variance. The texts were presented line by line on the computer screen, whereby, contrary to most reading-time experiments, line endings did not necessarily correspond to linguistic units (i.e., clauses or sentences). This was done to obtain as natural a 'typography' as possible, as an attempt to optimize ecological validity. However, each new sentence started on a new line. It was decided to present the texts in the same order to all subjects, because this would mean that all subjects had the same time-lag and number of interfering texts between the reading task and the recognition and recall tasks.

For the computer-controlled recognition task, a list of 14 (parts of) sentences was taken from the text about the visit of the pope to Holland, which was the final text for all subjects. The final text was selected, because then there was a short time-lag between reading and recognition task (about 5 minutes). All of these sentences, which were roughly of the same length, contained a discourse evaluation (cf. Hunt and Vipond 1986, Vipond and Hunt 1989), that is to say, (a pair of) substantives, adjectives or adverbs, with which the author invites the reader to share a belief, attitude or value. These evaluations were printed in capitals (needless to say, the evaluations were printed in normal case in the texts!). However, in half of the sentences they were replaced by more neutral synonyms. For instance 'Een enorm leger politie-agenten' (A huge army of policemen) became 'Een groot aantal politie-agenten' (A great number of policemen), and the archaic and slightly pejorative 'Vromen' (Pious people) became 'Gelovigen' (Believers). These synonyms were also printed in capitals, the rest of the line was printed in normal type and served as a retrieval cue. It

was decided to focus on discourse evaluations, because they are considered to be literary both in the theoretical literature (the cited studies by Hunt and Vipond), and in the subject pool of the present experiment, as the results from the pilot study seem to suggest. This enhances our chances of finding differences between groups, although it also entails a restriction to a specific aspect of textual surface structure.

Finally, a booklet was prepared containing materials for further tasks. It contained a short questionnaire inquiring after age, gender, field of study, year of study, and the time spent in an average week reading newspaper articles, textbooks, popular scientific articles, fiction, and other types of text. The booklet further contained a printed version of the first text the subjects read ('about the riots in Amsterdam'), in which 27 words were deleted and replaced by lines; an instruction was printed above the text, asking subjects to supply the original words and write them down on the lines; it was stressed that each line replaced one and only one word.

Apparatus. A LASER 286/2 AT pc was used with a monochrome amber screen; a button box was connected to its serial printer port. The box contained three buttons, marked 'CORRECT', 'NEXT', and 'INCORRECT'.

Design. The independent, between-subjects, variable was the perspective under which the texts were read; there were two levels of this variable, news perspective (NP) and literary perspective (LP). Half of the subjects were randomly assigned to the NP group, and half to the LP group. The dependent variables were reading times for the texts, percentage of correct responses on the recognition task, speed of response, and proportion of accurately recalled words in the recall task. Text order was not varied for reasons explained above.

Procedure. Subjects were tested individually in a sound-proof room. The experimenter made them feel at ease in the room and told them that they were to read seven texts from the computer screen; that the average length of the texts was 216 words ('about two-thirds of a book page'); and that they would read these texts line by line by pressing the button marked 'NEXT' in front of them. It was also pointed out that the subjects could not move backwards in the texts. Finally, the subjects were told that each text would be followed by a statement about its content and that they were to indicate whether this statement was correct or incorrect by pressing the appropriate button. These statements were all rather easy to judge, and were included to keep the subjects on task.

The experimenter then explained that the rest of the instructions, four in all, would appear automatically on the screen, not line by line, like the texts, but in their entirety (with several lines on the screen), and that these instructions would guide them further through the session. Finally, the experimenter told the subjects that the computer session would last about 20 minutes and that as soon as they had finished, a message would appear on the screen

asking the subjects to inform the experimenter of this event, so that he could hand them a booklet containing some other tasks. No mention was made at this point of the types of tasks the subjects were to receive later.

The experimenter then pressed the button marked 'NEXT', which made the first instruction appear on the screen, and left the room. The instruction informed the subjects once again about the general reading procedure and the manipulation of the buttons. After the subjects pressed 'NEXT', a second instruction appeared informing the subjects that the first text they would be reading was a travelogue, which served to familiarize them with the procedure (as mentioned earlier, the reading time for this text was later used as a covariate in analyses of variance). It was stressed that the reading time would be 'globally measured' so that, if subjects wanted to pause, they should do this exclusively between texts.

After the subjects had read the practice text and judged the statement, a third instruction appeared on the screen. By this instruction the subjects were allocated to either the NP or the LP condition. For the NP condition the instruction read: 'The texts you will be reading are all excerpts from reports about important events which occurred in the 1980s, the excerpts come from *De Volkskrant*, *NRC-Handelsblad*, and *De Gazet van Antwerpen*. [5] Roughly speaking, all texts deal with confrontations between police and civilians'. For the LP condition the instruction was: 'The texts you will be reading are all taken from well-known authors writing in the Dutch language, such as Harry Mulisch, Edgar du Perron, and Jef Geeraerts. Roughly speaking, all texts deal with confrontations between police and civilians'. To enhance the effect of the instruction, each subsequent text was announced as either a newspaper excerpt or a novel excerpt, depending on the perspective. Subjects were encouraged to read the texts as they would normally read similar newspaper report/novel passages.

After the subjects had read the six experimental texts and evaluated the statements which followed each text, a fourth instruction appeared, which announced the recognition task. The subjects were told that they would be presented with a series of 14 lines (one at a time) each containing a word in capitals. They were told that these lines came from the last text they read ('about the visit of the Pope to Den Bosch'), except for the word in capitals; this could either be the original word or a synonym of the original. The subjects were to press INCORRECT in case of a synonym and CORRECT in case of the original word. It was stressed that the task involved speed as well as accuracy of response. The subjects were first presented with a series of six well-known proverbs in which a word in capitals was either the original word

---

[5] The name of this Flemish newspaper was mentioned in order to prevent confusion when subjects encountered the Geeraerts excerpt, which would perhaps not be credible as coming from a Dutch newspaper with its references to typically Flemish organizations.

Table 2
Mean reading times (msec/word) for the experimental texts per perspective, adjusted for habitual reading speed (standard deviation in parentheses)

Perspective	Text						Overall
	Text 1	Text 2	Text 3	Text 4	Text 5	Text 6	
NP	300	299	342	285	272	238	289
	(60)	(71)	(75)	(63)	(46)	(56)	
LP	341	342	400	320	315	271	332
	(70)	(72)	(103)	(62)	(53)	(57)	

or a synonym; they were to judge these sentences in order to get used to the procedure.

After completion of this task, a message appeared on the screen which told the subjects to notify the experimenter. The experimenter then handed them the booklet with the questionnaire and cued-recall task.

3.2. Results

Hypothesis 1 predicts that people read faster in the NP condition than in the LP condition. Testing this hypothesis was rather straightforward, since it only involved a comparison of the total reading times for the texts; for ease of comparison between texts these times were converted to reading times per word. The figures obtained this way were subjected to a repeated-measures analysis of variance with the reading times for the texts as within-subject variables, perspective (with the levels NP and LP) as the between-subjects variable, and the reading times for the practice text as a covariate in order to filter out the influence of the subjects' habitual reading rate.

Table 2 shows the mean reading times for the six experimental texts (expressed in msec per word) broken down by perspective; the values are adjusted for the covariate.

The repeated-measures analysis yields a highly significant main effect of perspective: $F(1,37) = 39.08$, $p < 0.0005$, which means that, as far as perspective is concerned, the pattern of reading times of subjects in the NP condition differs significantly from that of the subjects in the LP condition. Also a significant main effect of text was found, $F(5,190) = 53.36$, $p < 0.0005$, which is logical, since the texts differed on several dimensions, such as syntax, use of words, average length of words and topic, narrativity, which are known (cf. Graesser 1981) to influence reading speed. More importantly, though, no significant text X perspective interaction was found ($F < 1$), which means that the perspective effect does not have to be qualified for certain texts. This also is a strong refutation of the potential objection mentioned in footnote 3, which suggested that there might be underlying differences between the two literary

and the four newspaper excerpts, which would possibly show up during on-line reading. If this were the case, a significant statistical interaction would have occurred.

In order to test the robustness of the perspective effect, separate ANOVAs were carried out on the six experimental texts with reading time for the practice text again as a covariate. [6] Evidently, each text displays differences in reading times for the perspectives; moreover, all differences are in the predicted direction: NP subjects read faster than LP subjects. The ANOVAs yielded significant effects for all six texts: Text 1, $F(1,39) = 6.284$, $p < 0.025$; Text 2, $F(1,39) = 6.455$, $p < 0.025$; Text 3, $F(1,39) = 5.508$, $p < 0.025$; Text 4, $F(1,39) = 4.608$, $p < 0.05$; Text 5 $F(1,39) = 9.044$, $p < 0.01$; Text 6, $F(1,39) = 4.299$, $p < 0.05$.

In other words, the finding that people read faster when they think they are reading a newspaper text is valid for all six experimental texts and is thus rather robust. This robustness is even enhanced if we look at the ratios of the reading speeds in the two conditions per text: this ratio is strikingly constant and varies only between 0.86 (Text 3) and 0.89 (Texts 4 and 5), with a mean of 0.88. This means that in the present experimental conditions, newspaper reading is about 12 percent faster than literary reading.

Hypothesis 2 predicts that LP subjects construct a stronger surface structure representation of the text than NP subjects. Since it is well-established that in most cases the surface representation of a text is not retained very long in memory (cf. Graesser and Mandler 1975, and for recent evidence Kintsch et al. 1990), it was decided to measure strength of surface representation directly after the processing of a text, as well as after some time-interval (about 20 minutes) with the interference of some other texts. Also a variation of methods was used: a time-based recognition task, and a text-completion task. The former task concerned the final text the subjects read (since the text order was not varied, this was the same text for all subjects: the one about Pope John Paul II), and took place immediately after the reading of this text (with only the intervention of six practice sentences). It was predicted that subjects in the LP condition would outperform subjects in the NP condition on this task, which means in this case that their responses will be more accurate and/or faster (cf. Massaro 1975). Accuracy was computed simply as the proportion of correct answers: if a subject responded correctly to 8 of the 14 sentences, he/she received an accuracy score of 0.57. Accuracy scores were computed for each subject using this method. Hypothesis 2 predicts that LP subjects attain higher accuracy scores than NP subjects.

In the comparison of the reaction-time data only the correct responses were used; incorrect responses were exempted from analysis. Reaction-time data for

---

[6] Once a significant multivariate effect has been found, it is possible to perform univariate analyses without inflating the error rate (cf. Stevens 1986).

the correct responses contained a few outliers (that is to say, extremely slow reaction times); therefore, it was decided to eliminate reaction times above 2.5 times the standard deviation from the mean; this involved 3.6% of the observations. After this, the mean reaction time for each subject was computed. In short, the above operations resulted in two values for each subject: an accuracy score and a mean reaction time for correct responses. Both accuracy scores and reaction times were subjected to t-tests.

The mean accuracy scores (ACC) were 0.55 ($sd = 0.12$) in the NP condition, and 0.66 ($sd = 0.15$) in the LP condition; this difference is significant in a two-tailed t-test: $t(38) = -2.64$, $p < 0.025$. The reaction times (RT), however, displayed no significant difference: the means were 4402 msec ($sd = 1237$) for the NP and 4441 ($sd = 947$) for the LP condition, $t(38) = -0.11$, n.s. This probably means that there has been some speed-accuracy trade-off on the part of the NP subjects, who performed on chance level, and may have been guessing rather than attempting to answer the questions. On the other hand, because the accuracy scores are rather low (in most reaction-time experiments they are higher than 0.90), the reaction-time measurements become less useful. Therefore, confirmation of Hypothesis 2 is solely obtained through the accuracy of responses.

Our second measure was the text-completion task. This task concerned the first experimental text the subjects read, the one about the riots in Amsterdam. Since this task was not carried out on the computer, but on paper, no reaction-time data were recorded. Again, accuracy scores were computed for each subject, expressed as the proportion of correct completions out of 27 cases. In order to count as a correct response, the supplied word had to be a verbatim copy of the original word; this implied for instance that 'zwaar' ('heavy') was judged incorrect because the original word reads 'loodzwaar' ('leaden'). The mean accuracy score was 0.22 ($sd = 0.11$) for the NP condition, and 0.31 ($sd = 0.16$) for the LP condition. Again, this difference is significant on a two-tailed t-test: $t(38) = -2.04$, $p < 0.05$. This result provides additional

Table 3
Average accuracy scores and reaction times on surface structure recognition and recall tasks, broken down by perspective (sd in parentheses)

Perspective	Task		
	Recognition		Recall
	ACC	RT	ACC
NP	0.55	4402	0.22
	(0.12)	(1237)	(0.11)
LP	0.67	4441	0.31
	(0.15)	(947)	(0.16)

support for Hypothesis 2. It also shows that even after a time interval of approximately 20 minutes and having read five other texts with a similar topic, literary readers have a stronger surface representation of a text than newspaper readers. Recognition and recall data are summarized in table 3.

4. Summary and discussion

The principal aim of this study was to empirically assess some effects of two different discourse-type perspectives (a literary and a news perspective) on the same set of texts. These effects would lend support to the assumption that readers have developed (through instruction and incidental learning) specific control systems for the various types of texts they regularly encounter, and that these control systems, once triggered, set some basic parameters of the comprehension process in order for conventional communicative behavior to occur.

Reading-time measurements show that the same texts (which are, more or less, acceptable both as literary and news stories) are read faster when they are assumed to be from a newspaper. This means that, as predicted, a literary-comprehension control system induces a slower reading speed than a news-comprehension control system. In other words, control systems influence on-line processing, independently of textual features; probably even automatically since it is not very likely that people consciously adapt their reading rate to a certain perspective. One very important caveat must be added here: by no means do these results indicate that differences in processing speed are in general *exclusively* attributable to cognitive factors; on the contrary, work was cited which provides evidence for the effects of several linguistic factors (Graesser 1981, Haberlandt 1988); moreover, a highly significant effect for text was found and reported.

The second assumption stated that literary reading is characterized by a focus on stylistic aspects and should therefore result in a relatively strong surface structure representation of texts. This assumption is supported by the data as well: subjects in the LP condition performed better on surface structure recognition and recall tasks than subjects in the NP condition.

As pointed out above, it has been attempted to optimize the ecological validity of the experiment. However, in some cases ecological validity had to be sacrificed because of methodological considerations. Here are some important examples:

(1) Although news processing usually involves very recent events, the subjects in the present study had to read newspaper texts about not so recent events. This might be 'unnatural'. However, a design in which one and the same text is to be presented as a newspaper text in one condition, and as a literary text in the other makes it impossible to take texts describing very

recent events, because then the text cannot successfully pose as a literary text; it is practically impossible to write a novel about the fall of the Berlin wall, find a publisher and get it published all on the same day as the event takes place. However, since some of the texts were opaque as to when the events took place (the events described by Du Perron for instance took place in the 1930s), these texts might very well be read as describing very recent events. Therefore, the drawback of including non-recent events is perhaps not very severe.

(2) The texts used in the experiment were untitled. This may be considered unnatural, especially for the NP subjects, because the headline is regarded as the most typical aspect of newspaper texts. However, introducing headlines and titles would imply confounding two things: the effect of perspective on reading speed and the effect of title/headline on reading speed. The important question, then, is whether the absence of titles and headlines may have led to artifactual results. This does not appear to be the case, as a sequel to the present study, in which headlines and titles were used with the same texts as in the present experiment, found no effects of title/headline on reading rate, but, on the contrary, mirrors the results reported above (cf. Zwaan, in preparation).

(3) The reading task itself is unnatural: subjects read text from the computer screen instead of paper, and they could not look back and forth in the text. Up to a certain point this remark seems valid. However, since two groups were compared, both should be affected by this 'unnaturalness'. Besides, the large majority of subjects, when asked how reading from the screen compared to natural reading, noted no big differences. On the contrary, most of them expressed surprise about the 'naturalness' of the situation; only two subjects found it a disadvantage that they could not look backwards in the texts. Although part of the subjects' responses may be due to social desirability effects, it seems extremely unlikely that the results found in this experiment are caused by the 'unnaturalness' of the task.

In conclusion, the present study provides support for the idea that there are specific control systems for particular types of discourse. One of the things these systems do, is alter some parameters of the comprehension process. As regards the control systems for literary and news comprehension, it seems that the first imposes a slower reading speed and more emphasis on surface structure representation than the second.

However, it is likely that (many) more parameters determine literary (and news) comprehension. What these parameters are and how their effects can be assessed empirically are important questions for further research. For the moment, it seems that – as Zwaan (1991) suggested – it is fruitful to use on-line tasks (in combination with recognition and recall tasks) to reveal aspects of literary reading. In this sense, a promising line of research has been opened: further research along these lines is both necessary and possible.

Appendix: English translations of two of the experimental texts used in this study

Text 2: Excerpt from the American translation of E. du Perron's *Country of Origin* (translated by F. Bulhof and E. Daverman, UMass Press, 1984), pp. 403–404.

The cabinet resigned, and immediately there was an enormous demonstration on the boulevard. At six o'clock two huge groups started to run into each other, one coming from the Boulevard de la Madeleine and the other from the Boulevard des Italiens. The police left them alone. (...) Suddenly there was five minutes of fierce commotion at the Café de la Paix: windows were broken, tables smashed, a newsstand turned over; and the customers on the terrace got into a fight. A bus was stopped, the passengers forced to get out, and the bus placed diagonally across the street to stop the traffic. (...) Lanterns were bent; flames came out of a newsstand. Finally the police moved in, but carefully. There were more than five thousand demonstrators this time, and their number was growing. A number of policemen were roughed up; a fire engine that happened to be passing by and was suspected of going to hose down the crowd was damaged. The hoses were cut up, and the water sprayed out of the holes. Pickpockets worked diligently during the demonstration. The police arrested three hundred people in all.

N.B. The parentheses indicate places where a sentence and a word from the original text were left out in the experimental text. This was done because the information left out was anachronistic. These were the only changes made in the entire material of the experiment.

Text 5: *Volkskrant*, December 21, 1989. (The story of a former Rumanian political prisoner, now in the Netherlands, translation by R.A.Z.)

His first confrontation with the police dates from winter 1983. He studied to be a stage director at night-school, and worked in a studio during the daytime. The energy shortage, a consequence of megalomaniac investments in the petrochemical industry, was dire. Two measures became simultaneously operative: the energy prize was multiplied and the supply of energy was severely reduced. A propaganda campaign accompanied the cold under the slogan: 50 percent materials, 100 percent performance. Sorin drew a man cut in two, wrote the slogan under it, and sneaked at night to a factory gate. He pinned up the drawing, believing himself unseen. The next day, he was picked up from his work. At first, he was treated in a friendly manner at the police station. He was offered some coffee. During the interrogation, the central question was by whose order Sorin had pinned up that drawing. Sorin remained silent. They hit him. He refused to talk. They threatened to cut his wrists. He denied having anything to do with the drawing. They showed him the door. 'Just go'. As he walked down the corridor, an officer grabbed hold of him and knocked him unconscious. When he came round, he lay in the corner of a cell, his hands and face covered in blood. Two fingers of his right hand were paralyzed; they had cut the tendons. He was then allowed to go. In the bus people wondered at his blood-covered face and hands.

References

Graesser, A.C., 1981. Prose comprehension beyond the word. New York: Springer.

Graesser, A.C. and G. Mandler, 1975. Recognition memory for the meaning and surface structure of sentences. Journal of Experimental Psychology: Human Learning and Memory 104, 238–248.

Haberlandt, K., 1988. Component processes in reading comprehension. In: M. Daneman, G.E. MacKinnon and T.G. Waller (eds.), Reading research: Advances in theory and practice, Volume 6, 67–108. San Diego, CA: Academic Press.

Hunt, R.A. and D. Vipond, 1986. Evaluations in literary reading. Text 6, 53–71.

Jakobson, R., 1960. Linguistics and poetics. In: T.A. Sebeok (ed.), Style in language, 350–377. Cambridge, MA: MIT Press.

Kintsch, W., 1980. Learning from text, Levels of comprehension, or: Why anyone would read a story anyway. Poetics 9, 87–98.

Kintsch, W. and T.A. van Dijk, 1978. Toward a model of text comprehension and production. Psychological Review 85, 363–394.

Kintsch, W., D. Welsch, F. Schmalhofer and S. Zimny, 1990. Sentence memory: A theoretical analysis. Journal of Memory and Language 29, 133–159.

Massaro, D.W., 1975. Experimental psychology and information processing. Chicago, IL: Rand McNally.

Meutsch, D., 1987. Literatur verstehen: Eine empirische Studie. Braunschweig/Wiesbaden: Vieweg.

Meutsch, D. and R. Viehoff, 1989. Comprehension of literary discourse: Results and problems of interdisciplinary approaches. Berlin: De Gruyter.

Stanovich, K.E. and R.F. West, 1989. Exposure to print and orthographic processing. Reading Research Quarterly 24, 402–433.

Stevens, J., 1986. Applied multivariate statistics for the social sciences. Hillsdale, NJ: Erlbaum.

Van Dijk, T.A. and W. Kintsch, 1983. Strategies of discourse comprehension. New York: Academic Press.

Van Oostendorp, H., 1988. Regulatieprocessen bij tekstverwerking (Regulation processes in textprocessing). Dissertation, University of Amsterdam.

Vipond, D. and R.A. Hunt, 1984. Point-driven understanding: Pragmatic and cognitive dimensions of literary reading. Poetics 13, 261–277.

Vipond, D. and R.A. Hunt, 1989. Literary processing and response as transaction: Evidence for the contribution of readers, texts, and situations. In: D. Meutsch and R. Viehoff (eds.), 1989, 155–174.

Zwaan, R.A., 1991. The literary situation-action model and the reading process. In: E. Ibsch, D.H. Schram and G.J. Steen (eds.), Empirical studies in literature: Proceedings of the second international conference (1989), 167–174. Amsterdam: Rodopi.

Zwaan, R.A., in preparation. Levels of representation in literary and news comprehension. University of Utrecht.