Abstract

This article studies the role of the perceptual dimension in text comprehension processes. The current researches in the domain of reading deal with the elaboration of cognitive representation of the text content without taking into account the text format. However, linguists and computer scientists confronted with the problems related to automatic reading notice that the text format cannot be dissociated from its linguistic content. We theorize that the Text Format affects comprehension processes in reading. The experiment looks at the effect of text format on the eye behaviour of readers as they process a narrative containing a perspective shift. The results showed that the presence of a paragraph cue (indentation) on the line where the perspective shift took place enabled early integration of this shift. The results of this experiment are discussed in relation to Gernsbacher’s (1989, 1990, 1997) Structure Building Framework model and work by Lorch and colleagues (Lorch, & Chen, 1986; Lorch, Lorch, Ritchey, McGovern, & Coleman, 2001).

Key words: Text format, comprehension, perspective shift, eye movements
Understanding a text requires building a coherent mental representation that integrates information from the text and from the reader's knowledge base (Kintsch, 1988, 1998; Gernsbacher, 1989, 1995, 1997; Gernsbacher, Varner & Faust, 1990). Integration starts as soon as reading begins, in such a way that the mental representation is built gradually, getting richer and richer as new information is supplied by the text. According to Structure Building Framework (Gernsbacher, 1997), building this mental representation, or structure, involves laying a foundation, at the beginning of the text, mapping coherent information onto the developing structure, and shifting to initiate a new sub-structure when incoming information is less coherent. Thus, most representations comprise several branching substructures. The building blocks of these mental structures are memory nodes activated by incoming stimuli. Initial activation forms the foundation of mental structures. Once a foundation is laid, subsequent information is often mapped on because the more coherent the incoming information is with the previous information, the more likely it is to activate the same or connected memory nodes. In contrast, the less coherent the incoming information is, the less likely it is to activate the same or connected memory nodes. In this case, the incoming information activates a different set of nodes, and the activation of these other nodes forms the foundation for a new substructure. In addition, once memory nodes are activated, they transmit processing signals to enhance (increase) or suppress (dampen or decrease) other nodes' activation (Gernsbacher & Faust, 1991).

Coherence is a key element in text comprehension. A large body of research has taken an interest in the processes used by readers to maintain coherence as they build an integrated mental representation of the information drawn from the text and from their own knowledge base. Many discourse features contribute to maintaining coherence, including referential (Kintsch, 1988, Van Dijk & Kintsch,1983) and causal relations (Fletcher, Hummel & Marsolek, 1990). Research has shown that a break in the coherence of a narrative brought about by the introduction of a new theme (Hyönä, 1994, 1995; Hyönä & Niemi, 1990; Lorch, Lorch & Morgan, 1987) or a perspective shift (Baccino & Pynte, 1998; Millis, 1995) increases reading time on the line in question. A longer reading time is generally considered to reflect the occurrence of an integration process. The extra processing time at topic shift or coherence break is spent on encoding that new piece of information, integrating it with the previous representation of text by building a new sub-structure connected with the structure-base or foundation. Evidence for these processes come from a variety of experiments of reading and comprehension of texts with perspective shift, topic shift, use of adverb informing readers about the narrative structure (Gernsbacher, 1990, 1996, 1997). Nevertheless, paragraph marks have not been considered as shift cues in these experiments. Text format is not taken into account in the elaboration of the coherent mental representation.

Even though text format is granted an informative and intentional role in the linguistic research, just like the words the text contains (Virbel, 1986). For example, centering a title has the following discursive equivalent: "The title of this book is ...". But current models of text comprehension do not include title centering, or any text mark despite the fact that its discursive equivalent is part of the mental representation of the text. A text's format does indeed have an intentional informative value, i.e., the author formats a text to transmit some particular information to the reader. In the field of electronic document processing, research on automatic text generation and reading has shown that the discursive equivalent of formatting must be taken into account in order to arrive at the meaning of a text (Pascual, 1991). In psycholinguistics, work by Lorch and colleagues (Lorch & Lorch, 1995; Lorch, & al., 2001) has shown that the semantic structure of a text is highlighted by organizing marks. Readers remember what is said in the text in a hierarchical way by defining an access order based on the importance of the information. When structural cues are lacking (no numbering), readers recall information in its order of appearance in the text. Structural markers that enhance recall may already take effect during encoding, i.e., at information intake time.
Perspective shift and text format

(Meyer & Rice, 1989; Mayer, Dyck & Cook, 1984). Longer reading times on numbered segments support this hypothesis (Lorch & Chen, 1986). Formatting (paragraphs, titles, etc.) in procedural texts plays the same structuring role as numbering because it enables encoding based on predefined markers of hierarchical structure (Schmid & Baccino, 2001). If the hierarchical organization of text information is taken into account at the onset of reading, encoding takes more time on procedural texts that are formatted than when no organizational cues are present (unformatted presentation). Inversely, when time is spent searching for how information is organized at the beginning of the reading process, processing is faster near the end of the text. With an unformatted text, the reader can only read the text "as it comes", i.e., as if it was a list of words to interpret as the reading progresses. This type of reading leads to shorter reading times within the text, but clearly slows down the process at the end, when the reader has to integrate all of the information read and fit the pieces together to form a coherent representation.

The study presented in this article deals with information encoding during reading, and more specifically, with the integration phase that follows a break in text coherence. Our goal is to determine the role played by a text's visual form in the integration phase involved in the construction of a coherent mental representation. Perspective shifts were used to materialize breaks in coherence, as shown by Millis (1995) for narrative perspective encoding. Initial reading time appears to be greater on lines with a perspective shift, but not on subsequent lines, suggesting that readers assess the coherence of the shift with respect to the preceding passage and integrate it into the mental representation being built. A perspective shift thus constitutes a break in coherence that necessitates an additional integration process (Black, Turner, & Bower, 1979; Millis, 1995). The additional time taken on the text segments with a perspective shift is used by the reader to spatially encode the position of the break in the narrative, like any other linguistic difficulty encountered in a text (Baccino & Pynte, 1994, 1998). The spatial encoding process is generally assessed using pointing tasks where a text is displayed on a computer display and the reader has to quickly find the text segments where a perspective shift occurs. Perceptual factors enter into the integration of narrative coherence breaks, and format is a crucial part of the spatial encoding process (underlining, indentations, etc.). An eye-tracking technique was used to measure information intake by readers confronted with a coherence break generated by a narrative perspective shift. The break was or was not signalled by indentation at the beginning of the line. A pointing task after the reading phase was run to see how subjects memorize the spatial location of a perspective shift when it was introduced by a visual mark or not. Research on text structure has demonstrated its impact on text recall, but for methodological reasons studies on this topic offer no information about the encoding process during reading. The purpose of this experiment was precisely to show that the physical format of a text being read has an effect on the information integration process. It was postulated that text format interacts with high-level processing of perspective shifts as a result of the spatial encoding strategy implemented during reading. A text's format provides additional cues to the reader and facilitates the cognitive processing needed for comprehension.

METHOD

Participants

The participants were 23 unpaid adult students at the Literature, Arts, and Humanities School of the University of Nice at Sophia Antipolis, France. They were all native speakers of French.

Eye movements data recording

The experiment was driven by two interconnected computers. One computer displayed the texts to be read while the other recorded the ocular data. Black characters were presented
on a white background, on a high-resolution screen. Eye movements were tracked using the corneal reflection technique from the Dr. Bouis\(^1\) oculometric system, and sampled every millisecond by a digital board (National Instruments). The spatial resolution of this system is less than one degree of visual angle (about one character space). To keep the ocular data free of parasitic noise from unexpected head movements, the subjects read while biting a bar that held their upper jaw stationary. Text presentation was preceded by a two-dimensional calibration phase.\(^2\) Before each text was read, the calibration was checked using the infrared beam by having the subject fixate an \(X\) located in the upper left-hand corner of the screen. If the subject's gaze was properly positioned on the \(X\), the text appeared; if not, the calibration phase was repeated. During reading, the X-Y coordinates of the eye movements were automatically stored in the computer for later processing, aimed at determining both the spatial pathway (fixation location) and the time course (fixation duration) of the subject's gaze. This processing method reduced the data to a collection of saccades (eye movements) and fixations (pauses) for statistical analysis.

**Linguistic materials and design**

Subjects saw 24 experimental texts, 16 filler texts, and 2 practice texts. All the texts contained four lines. The experimental texts had a particular construction: the first two lines described a scene and introduced a main character, and always presented the scene from an external perspective. Perspective is defined as the position of the narrator in the story (Genette, 1972). When the narrator describes a situation and a character in the third person singular, the narrative perspective is said to be external. An internal perspective is expressed in the first person singular. A shift to the internal perspective occurred in 50% of the cases on the experimental line (line 3). When this happened, the narrator’s position in the story changed (the narrator and the main character were the same person; see example).

**Text Example**

\[\text{Unformatted Presentation}\\
A\ dull\ sound\ ran\ through\ the\ walls\ of\ the\ building\ and\ gradually\ tapered\ off.\\
An\ athlete\ with\ a\ muscular\ body\ and\ a\ tanned\ chest\ was\ lifting\ weights.\\Slowly, / he(I) let out / a\ sigh,\ wiping\ his(my)\ brow.\\It was hot, the gym was not ventilated.\\
\text{Formatted Presentation}\\
A\ dull\ sound\ ran\ through\ the\ walls\ of\ the\ building\ and\ gradually\ tapered\ off.\\An\ athlete\ with\ a\ muscular\ body\ and\ a\ tanned\ chest\ was\ lifting\ weights.\\Slowly, / he(I) let out /a\ sigh,\ wiping\ his(my)\ brow.\\It was hot, the gym was not ventilated.\\
\]

(The texts were presented in French, see the original example in appendix)

Narrative perspective was the first experimental factor manipulated in the experiment (within-subject). It had two categories (no perspective shift versus perspective shift). Presentation format was the second manipulated factor (within-subject): half of the texts were

---

\(^1\) This apparatus detects eye movements as deviations in an infrared light beam projected onto the cornea of the eye. The cornea acts as a mirror, and a photoelectric cell placed on the other side of the eye detects the reflection of the beam.

\(^2\) The procedure consisted of having the subject consecutively fixate 9 points displayed at an equal distance from each other on the screen. The spatial coordinates of these points were recorded. The 9 coordinates defined the plane in which the eye movements were tracked. For the data processing, the coordinates were used to geometrically project the eye positions onto the stimulus plane.
presented with a 5-space indentation on the experimental line, thereby marking the beginning of a second paragraph (see example,). This factor also had two categories (indentation versus no indentation). Each participant read 24 different texts, six in each of the four experimental conditions.

The experimental line (line 3) was always the same length and had the same structure: adverb / personal pronoun + verb / complement. It was divided into three segments (separated by a slash, in the text example) to allow for a finer analysis of the eye movements.

Experimental procedure

The experiment was run in two phases: a reading phase and a pointing phase. During the reading phase, eye saccades and fixations were recorded. Progressive fixations, which followed progressive saccades (rightward), were then distinguished from regressive fixations (leftward), which preceded an eye movement towards a word already read in the line. The experimental line was analyzed as a whole and in segments. The eye movements were broken down into several variables: the number of fixations and their total duration, and the number and duration of progressive and regressive fixations. The analysis of regressive fixation time was not possible on all segments of the line because there were not enough regressive fixations on the short segments (Adverb and Personal-Pronoun+Verb).

When the subjects were done reading, they left-clicked on the mouse. This caused an X and a box to appear at the bottom of the screen. Clicking on the X displayed a question in the box. The second phase consisted of searching in the text for the word that answered the question and pointing to it with the mouse. This task was designed to inform us about the spatial encoding done during reading. Pointing time and the mouse's path were recorded. In target pointing tasks with a mouse, pointing deviation is taken to be the area between the optimal path (here, the straight line connecting the X at the bottom of the screen to the center of the target word) and the path actually followed by the mouse (Baccino & Kennedy, 1995).

Analyses of variance with subjects (F1) and items (F2) as the random variables were conducted on all dependent variables.

RESULTS

Reading task: analysis of the experimental line

For the entire experimental line, the analysis of variance yielded a main effect of format on total fixation time, $F_1(1,22) = 7.31 \ p < .05$, $F_2(1,20) = 3.39 \ p = .08$. The readers looked at this line longer when it was indented (With indentation $m = 2470$, without indentation $m = 2348$). Broken down by type of fixation, this effect was only found on progressive fixations, which lasted longer when there was indentation, $F_1(1,22) = 18.59 \ p < .01$, $F_2(1,20) = 9.64 \ p < .01$. Evidence that text format is taken into account was provided by this increase in eye fixations.

Reading task: analysis by segment

The division of the experimental line into three segments (adverb - personal pronoun + verb - complement) allowed us to analyze the reader's eye movements on the portion of text where the shift took place. For the Adverb segment, no effect on any of the dependent variables considered was observed.

1) Analysis of the Personal Pronoun + Verb segment. When a perspective shift existed, the reader discovered it on the Personal Pronoun + Verb segment. For this segment of the sentence, there was a main effect of format on total fixation time, $F_1(1,22) = 9.96 \ p < .01$, $F_2(1,20) = 8.92 \ p < .01$. The eye fixated longer on the Personal Pronoun + Verb segment when this phrase was encountered with indentation than when there was not any paragraph
mark. The interaction between the perspective shift factor and the indentation factor was marginally significant, $F_1(1,22) = 3.62, p = .07$, $F_2(1,20) = 3.49, p = .07$ (Figure 1). Note that the perspective-shift effect in this interaction was found on formatted texts only, $F_1(1,22) = 4.72, p < .05$, $F_2(1,20) = 6.77, p < .05$. In the presence of indentation, the perspective shift caused an immediate increase in fixation time (with perspective shift, $m=367$ ms, without shift $m=329$). With no spatial markers, reading time was not affected by the change in point of view ($F < 1$).

![Figure 1: Fixation Time on the segment Personal Pronoun + Verb as a function of Perspective Shift and Indentation](image)

The main indentation effect found on all fixations was observed again on progressive fixation time, $F_1(1,22) = 4.67, p < .05$, $F_2(1,20) = 4.59, p < .05$. In contrast, there was no difference in the regressive fixation pattern.

2) Analysis of the complement segment. Indenting increased fixation time on the complement, $F_1(1,22) = 9.56, p < .01$, $F_2(1,20) = 4.85, p < .05$. The same effect was found on progressive fixation time $F_1(1,22) = 19.39, p < .01$, $F_2(1,20) = 8.20, p < .01$. Progressive fixation time was longer on this segment when the line was indented line than when the line was formatted. The regressive-fixation analysis yielded a significant interaction between the two experimental factors, indentation and perspective, $F_1(1,22) = 4.80, p < .05$, $F_2(1,20) = 1.5$ ns (see Figure 2).
More regressive fixations were found when there was a perspective shift on a non-indented line, $F_1(1,22) = 7.54\ p < .05$, $F_2(1,20) = 3.14\ p = .09$.

This effect indicates deferred processing of the coherence break, since the perspective shift occurred on the Personal Pronoun + Verb segment, but no effect was noted on this part of the sentence when it was not indented. When no visual cues signalling the perspective shift in the text were given (no indentation), the readers did not process the shift on the Personal Pronoun+Verb segment until the end of the line. When the line was indented, the perspective shift had an immediate impact (see analysis of the Personal Pronoun + Verb segment). The perspective shift was encoded by the reader immediately on the segment where it occurred (Personal Pronoun + Verb) only when the line was indented. Thus, the indented text format was conducive to the immediate integration of new information generated by the perspective shift. When the paragraph change was not marked (no indentation), the integration process was deferred at the end of the line. Formatting thus seems to help readers with the cognitive processing of breaks in narrative coherence.

**Pointing task**

The pointing task required the reader to use the mouse to click on the antecedent of the pronoun located in the experimental line. This task was used to maintain the reader's level of vigilance throughout the reading phase of the experiment. Attentive reading of all texts presented was required to carry out this task. Analyses of variance were computed on pointing time and on the deviation of the mouse's path (see experimental procedure).

**Temporal analysis of pointing**. The perspective shift significantly increased pointing time, $F_1(1,22) = 22.29\ p < .01$, $F_2(1,20) = 11.04\ p < .01$. (Without perspective shift, $m=324$ ms; with perspective shift, $m= 446$ ms). There was no effect of indentation on total pointing time.

**Spatial analysis of pointing**. The analysis of the area between the optimal path and the path actually taken by the mouse did not yield any main effect, but one nearly significant
interaction between our two factors, F(1,22) = 4.19 p < .06, F2(1,20) = 2.96 p = .10. The perspective factor had a nearly significant effect when the line was indented: F1(1,22) = 3.59 p = .07, F2(1,20) = 2.14 p = .16 (Without perspective shift, m= 6082 px²; with perspective shift, m= 7192 px²). When there was indentation, pointing deviation was greater with a perspective shift than without. Thus, with a paragraph marker, the search for the antecedent of the personal pronoun "I" in the first part of the text disrupted the readers, who apparently assumed the pronoun introduced a new character.

On the pointing task, there was no main format effect, even though indentation offered a spatial clue to readers. Moreover, the perspective shift increased pointing time. This result can be explained in terms of the particularities of the pointing task: subjects were supposed to answer a question that referred to the anaphor "I" or "he" by pointing to a noun, not a pronoun. This task thus forced them to consider the character mentioned in the second line of the text as the referent of the anaphor. Some readers regarded the pronoun “I” as a new character (the narrator) that was introduced into the story on the third line and was not connected to the preceding character. In this case, they hesitated before pointing, rereading the text to search for another possible antecedent. The "separateness" of the character on line 2 and the personal pronoun "I" was enhanced by the presence of indentation, which divided the text into two paragraphs with potentially different themes. With another pointing task (on a word of the line), Baccino & Pynte (1998) have obtained the opposite effect, the perspective shift facilitate pointing on the line because it was encoded spatially by the reader.

DISCUSSION

The first aim of this article was to study the effects of text format on the cognitive processes underlying the reading of narrative texts with perspective shifts. Changing the point of view (perspective) in a text triggers high-level cognitive operations that we wanted to test with different text presentation formats. Our analyses yielded an interaction between how the reader integrates the perspective shift and how the text is presented. Indenting the line where the perspective shift occurred led to more fixation on the words that triggered the shift (personal pronoun + verb), which means that the readers immediately took the coherence break into account. When a paragraph marker was lacking, the integration of the shift was postponed until the end of the line (increase in regressive fixations during complement reading). The indented text format thus enabled immediate integration of the new information conveyed by the viewpoint change. When no visual cues warn the reader that new information will be introduced on the current line, integration takes place at the end of the sentence. But when the paragraph is marked, the reader organizes the encoding of text information on the basis of formatting cues. Paragraph marks inform readers they will probably encounter a piece of new information in the first line. As soon as it appears, it will be integrated in the mental text representation by creating a new sub-structure, according to Structure Building Framework Model (Gernsbacher, 1990, 1997). As the reading progresses, then, readers use the structural information supplied by the text's format in the same way as they use semantic information taken from the text and from their own knowledge base. The mental representation of a text thus seems to include physical or architectural parameters. Despite this fact, the perceptual dimension of written texts is totally neglected in current theories of reading comprehension (Albrecht & O'Brien, 1993; Gernsbacher, 1989, 1997; Kintsch, 1988, 1998; Just & Carpenter, 1987; O'Brien, Rizella, Albrecht, & Halleran, 1998; Van Dijk & Kintsch, 1983). The first level of representation considered in these models is the surface structure, which includes words, sentences, and sentence syntax, but does not take into account any formatting features into account like blank lines, indentations, typographical changes, etc. A spatial or organizational level of representation should be added to the three commonly accepted levels (surface structure, text
base, and coherent mental representation). According to Kintsch (1988), text format cannot be integrated into the surface structure, because one characteristic of this level of representation is that it is not retained in the reader's memory beyond two sentences, unlike the two levels above it (Kintsch, Welsch, Schmalhofer, & Zimny 1990). A text's format serves to structure linguistic information, just like the numbering and overviews studied by Lorch (Lorch & Chen, 1986; Lorch & Lorch, 1995, 1996, Lorch & al., 2001). Lorch's results suggest that a mental schema of a text's format is constructed in order to organize encoding and recall in a hierarchical way. Unlike the text's syntactic structure, this schema seems to be critical to recall. Readers appear to use spatial cues supplied by formatting to devise a reading strategy that acts as a cognitive processing structure for integrating text information (Schmid & Baccino, 2001); this structure also organizes information in memory (Lorch & Lorch, 1996). When the text contains no spatial marking, new information is integrated at the end of the line. The increase in end-of-line regressive fixations observed in this study can be taken as an indication that information is being integrated at that time, when encoding operations (wrap-up processes) aimed at achieving a coherent mental representation of the text are taking place (Hyönä, 1994).

The results obtained here support the idea that the narrative perspective is incorporated into the reader's mental representation (Millis, 1995). Perspective shifts change the narrator's viewpoint from that of an all-knowing narrator, to that of a story character or a person describing events from the outside (Genette, 1972). Our results suggest that where there is a transition from the third person singular (used to describe a character from the outside) to the first person singular, readers do indeed interpret it as a shift in the narrator's point of view. But this integration process varies with the text format.

Irrespective of the representation level at which text format first takes effect, the results presented in this article show that the visuospatial dimension of a text enters into play in the comprehension process, just like the words the text contains. When readers see a paragraph mark they anticipate the need of sub-structure elaboration in the line (Gernsbacher, 1997) like they do it when they read some adverbs (then, further more, moreover…). In the construction of meaning, presentation format interacts with linguistic information in the text. These results are compatible with Virbel's theory (1986) and raise the question of the existence of a spatial representation level for texts that supplies meta-textual information to be integrated into the mental representation.

REFERENCES
Genette, G. (1972). Discours du récit, Figure III, Seuil Paris, 203-211.


**APPENDIX**

**Text Example in French**

*Unformatted Presentation*

Un bruit sourd parcourt les murs du bâtiment et s’estompe peu à peu. Un athlète au corps musclé et au torse bronzé soulève des haltères. Longuement, il(la) lance un soupir en s’y essuyant le front.

Il fait chaud, le gymnase manque d’aération.

*Formatted Presentation*

Un bruit sourd parcourt les murs du bâtiment et s’estompe peu à peu. Un athlète au corps musclé et au torse bronzé soulève des haltères. Longuement, il(la) lance un soupir en s’y essuyant le front.

Il fait chaud, le gymnase manque d’aération.