ON THE ROLE OF TIME IN READER-BASED TEXT COMPREHENSION[1]

Jean-Pierre Corriveau[2]
Department of Computer Science
University of Toronto
Toronto, CANADA
M5S 1A5

Abstract: Information-processing models for comprehension typically regard text as the depository of a *single determinate* meaning, placed *in* it by the writer. Conversely, a *reader-based* approach views meaning as constituted by the interactions between an *individual* and a text. From a computational standpoint, reader-based understanding suggests abandoning models which depend on *a priori* rules of interpretation and limiting the design of an algorithm to the *quantitative* aspects of text comprehension. I propose that the perception of subject matter be viewed as a *race process* where the generation of *bridging inferences* and *expectations* is partly controlled by quantitative factors (such as the delay for memory retrieval) which emphasize the invisible but omnipresent role of *time* during reading.

1 Introduction

In this paper, I am concerned with the *comprehension* of long, unrestricted, *written text* rather than with the design of single-sentence systems. I assume that *subject matter* is what gives a text a certain unity: if we fail to perceive the subject matter of a text, we find it difficult to understand that text (Bransford and Johnson, 1973).

Researchers in text understanding (*e.g.* Dyer, 1983; Graesser and Clark, 1985) typically assume that conceptual structures are constructed during comprehension: each linguistic element must be *connected* to some of the *cognitive constructs* obtained for the text read so far. Intuitively, one can think of each of such constructs as a set of *links* (or *connections*) that specifies which words and clauses are connected one to the other; each link corresponds to a conceptual path of some sort. In this paper, I shall try to avoid representational issues by working at this basic level of conceptual links.

Local coherence allows the reader to perceive successive clauses of a text as a set of related ideas. The connection between two successive clauses may be explicitly stated in the text (*e.g.* using connectives such as *because, therefore, when*). Such explicit inter-clausal connections are not problematic. The difficulty in perceiving coherence instead involves clauses with no explicit connections. In this case, the reader must *bridge* from one clause to another by means of inferences. Researchers generally assume that there is a *correct* inference path to connect two clauses. Correctness is defined with respect to the *a priori* rules of interpretation. Similarly, it is generally hypothesized that the perception of *global coherence* is also rule-based and depends on certain global patterns of organization called *macrostructures* (Phillips, 1985). The few models which tackle the problem of global coherence assume the existence of a small, correct set of macrostructures, which are specified in cognitive schemata. In other words, the possible gists of a text

[1] Support from the National Science and Engineering Research Council is gratefully acknowledged. I am indebted to Ed Plantinga and Graeme Hirst for several discussions and valuable comments on this research.

[2] UUCP: jpierre@utai.uucp; CSNet: jpierre@ai.toronto.edu

are, in essence, specified *a priori*: the rules of interpretation of these models come to form an 'understanding algorithm' which defines what it is to correctly understand a text.

The idea that a text *contains a single determinate meaning* has been rejected by the proponents of *reader-based understanding* who consider the meaning of a text to be constituted by the interactions between text and reader (see Holub, 1984). From this standpoint, the algorithms of existing models constitute mere mechanical encodings of sets of rules more or less arbitrarily established by the programmers.

In the context of reader-based understanding, comprehension is taken to proceed from the *private response* of a reader to a text. For example, Gadamer (1976) proposes that the act of interpretation be understood as an interaction between the *horizon* provided by the text and the horizon that the interpreter brings to it. From this perspective, the text acts as a *stabilizing* factor in the *idiosyncratic* interpretation of a reader. From a computational viewpoint, reader-based understanding suggests that we abandon attempts at specifying *a priori* rules of comprehension and, in particular, macrostructures. Instead, the design of an algorithm should be limited to the quantitative (*i.e.* non-qualitative) aspects of comprehension. I investigate these aspects in this paper.

2 From Local to Global Coherence

2.1 Fundamental Hypothesis

By definition, one can establish that A *causes* B only after having inferred that B takes place *after* A in the chronology constructed from the text. Unless one assumes that all causal paths are encoded exhaustively in a knowledge base, in which case comprehension reduces to pattern-matching, it is important to notice that there is an order in which certain inferences are produced: certain types of inferences must be preceded by others. Time displays a similar invisible omnipresence for the problems of anaphora (Hirst, 1981; Stevenson, 1986), and ambiguity (Hirst, 1987). The fundamental hypothesis of my work is that time plays a crucial role during comprehension.

2.2 Time-Constrained Comprehension

2.2.1 Time, Working Memory, and Context

Ultimately, text consists of linguistic elements juxtaposed in a linear sequence. At time t_1, a word WORD(t_1) is input. Most aspects of syntax can be viewed as linguistic devices to establish intra-clausal connections, that is, to connect WORD(t_1) to some of the elements of the cognitive construct existing at t_1 for the clause being currently processed, CLAUSE(t_1).

For simplicity, let me gloss over the syntactic analysis process. A few instants later, at time t_2, once WORD(t_1) has been *integrated* with CLAUSE(t_1), the reader has a new construct for the current clause, CLAUSE(t_2). This construct may contain unresolved references such as pronouns or ambiguous words. Following Stevenson's (1986) work, I assume that the reader immediately tries to connect CLAUSE(t_2) with the constructs he has created for previous clauses, rather than waiting for a complete syntactic representation of the current clause.

Typically, a reader only has a limited number of constructs *in focus*, that is, readily accessible. The other constructs obtained for the text read so far require significantly more time to access, if they can still be accessed at all. This corresponds to the usual *heuristic* distinction between a working memory of limited 'capacity' and another partition of storage, the *background*, which is considerably slower to access.

As the reader advances through the text, the constructs in the working memory change, the focus is constantly modified. Let us denote the set of constructs in focus at time t by FOCUS(t), and the set of constructs in the background by BACKGROUND(t). FOCUS and BACKGROUND comprise all the information stored by the comprehender during his reading. Therefore, at time t_2, CLAUSE(t_2) can only be connected to the elements of FOCUS(t_2) and BACKGROUND(t_2). I take *context* to be the union of FOCUS and BACKGROUND at any given point in time. Let us denote the context at time t by CONTEXT(t).

2.2.2 Bridging Inferences Revisited

At a given point t in time, the reader must connect CLAUSE(t) with CONTEXT(t). Ultimately, which connections are made depends on the reader's prior knowledge. Yet I suggest that certain quantitative factors also affect the generation of conceptual connections and thus the perception of subject matter.

Two quantitative factors can be distinguished during the generation of conceptual connections:

- the processing time required for establishing connections.
- the access time of target constructs.

These two factors emphasize a most significant fact about text comprehension, namely, that it is *a real-time process*. The importance of this postulate has been defended by Gigley (1985a, 1985b) in her seminal work in Neurolinguistics from which much of the present research proceeds.

Márkus (1983) makes a crucial observation when he remarks that bridging inferences are generally restricted to what is in focus. For him, the failure to consider the contents of BACKGROUND originates in the *pressure to infer* typically felt by the comprehender who allocates himself but a relatively short amount of time to integrate clauses. Since the cognitive constructs in focus are much faster to access than those in the background, they are the first, and generally, the only ones to be considered. This suggests that the perception of coherence be viewed as a *time-constrained* process: time constitutes the stopping criterion for the generation of conceptual links. More precisely, connecting CLAUSE(t) to CONTEXT(t) should be regarded as a *race process* where the *actual* (as opposed to a theoretically *correct*) set of bridging connections is the one available when this race stops. The order in which bridging inferences are created becomes crucial. For example, given a 'hard' text and a short delay between clauses, most temporal connections, but probably only a few implicit causal ones, would be obtained.

2.3 On the Emergence of Global Coherence

Since CLAUSE is necessarily in focus and consists of the same types of links as the contents of the working memory, CLAUSE is taken to be a subset of FOCUS. Thus, each time a conceptual connection is found between CLAUSE and the rest of FOCUS, it is added to FOCUS, that is, to the working memory. Since FOCUS has limited capacity, it may become filled up at a certain point in time. Each time FOCUS is saturated, the addition of a new connection requires that at least one of the existing connections be moved out of focus, that is, either placed in BACKGROUND or suppressed.

There is a fundamental difference between being moved to BACKGROUND and being suppressed: in the first case, the link is still accessible, in the second case, it is lost. Intuitively, we would like the most 'important' or 'relevant' connections of the text to stay in focus, less important ones to be in the background, and trivial ones to be suppressed. Following such an approach, global coherence can simply be defined as the contents of the working memory at a given point in time. The crucial point is that there can be no *a priori* rules of 'importance' or 'relevance'; the selection of what must be moved out of focus must operate at the quantitative level.

Each conceptual connection links two constructs. The connection itself may be direct or consist of an inference path. Since each element in an inference path is only relevant with respect to the whole path, the conceptual connection between two constructs should be thought of as an indivisible unit. In order to *quantitatively* distinguish conceptual connections, I propose that each one be assigned an *activation level* (or equivalently, *energy*): the links with the greatest energy are kept in focus, the others are moved out of focus. Let us investigate the notion of an activation level.

From a quantitative point of view, when a connection A is added to FOCUS, it may affect the energy of a connection B already existing in FOCUS in one of three ways:

1. *independence*: A and B are independent if the addition of A to FOCUS does not affect B's energy.

2. *additivity*: A and B are additive if the addition of A to FOCUS increases B's energy.

3. *adversity*: A and B are adversative if the addition of A to FOCUS decreases B's energy[3].

Each of these quantitative relations is defined over *conceptual* connections which control the perception of coherence, a semantic phenomenon. In other words, how a link affects another one (when one is added to FOCUS and the other is already in the working memory) is not determined quantitatively but *qualitatively and dynamically*, that is, according to the current context and to the semantic knowledge possessed by the comprehender prior to the reading.

The addition of a new connection to FOCUS is not the only factor that affects the energy of the connection in the working memory. More precisely, time also affects activation levels. My hypothesis is two-fold. Firstly, energy is constantly lost as time goes by. This process is known as *constant decaying*. Secondly, the longer a connection stays in FOCUS, the more difficult it becomes to move out of focus. Since a connection can maintain itself in FOCUS only if it keeps a high level of activation, I take a connection's energy, which is a function defined over time, to be inversely proportional to that connection's decay, which is also a function defined over time. Thus, from a qualitative viewpoint, the energy of a link may be seen as a measure of its 'importance' with respect to coherence.

Let us return to the problem of coherence.

Firstly, if the activation level of a connection drops under a specified minimum, that connection is removed from FOCUS and lost. Intuitively, this connection has been *forgotten* (a quantitative phenomenon) or *contradicted* (a semantic phenomenon which, like all others, carries down to the quantitative level). Since it is seldom the case that the exact wording of a text is memorized (Baddeley, 1976, pp.315–316), it seems reasonable to assume that, for example, syntactic links will be quickly forgotten.

Secondly, upon the arrival of a new connection in FOCUS, the working memory constraints will trigger the selection of a (possibly empty) set of connections to be moved out. Several researchers have suggested numerous kinds of constraints for the working memory. Since there is no general agreement, I propose that there be only one constraint for the working memory: its total activation level at any point in time. This assumption has the advantage of being independent of any type of representation, and of not explicitly restricting the number or the size of elements in FOCUS. This unique constraint on the total energy in FOCUS constitutes the criterion that determines how

[3] Additivity corresponds to the connectionist notion of an *activation* link, and adversity, to the notion of an *inhibitory* link (see Feldman, 1984).

many and which connections must be moved out of focus. Abstracting from details, connections are moved out so that the constraint is *always* respected. Once it has been selected for removal from FOCUS, a connection can either be transferred to the background or forgotten. Since this decision must be quantitative, admission to BACKGROUND should be controlled by a minimum required energy (which would obviously be greater than the threshold used for forgetting). If a connection is above this threshold, it is transferred to the background; otherwise, it is lost.

I claim that the mechanisms described above adequately cover the construction of global coherence from the local connections established within and between clauses. Indeed, since FOCUS is not restricted to local connections but rather maintains a set of 'important' or 'relevant' connections, each new clause can be immediately linked to both its linear neighbours that would have not yet been moved out of focus, and to the 'important' clauses of the text.

3 On the Perception of Subject Matter

According to the proposed model of comprehension, global coherence is restricted to the conceptual connections triggered by the text. As time goes by, only the 'important' bridges (*i.e.* those that account for global coherence) remain in the working memory, the others being moved to the background or forgotten. From this viewpoint, global coherence strictly proceeds from the text. But subject matter is not bound to the text: it also involves what Gadamer (1976) calls the *horizon of the reader*. Three factors, namely *expectations*, *remindings*, and *interestingness*, account for this horizon. Let us briefly discuss each of these.

3.1 On Expectations and Remindings

A schema is a cluster of 'knowledge', that is, a set of conceptual links. These links are bound together in some way or another so that they come to form a conceptual unit. Thus, when one of the elements of a schema is used to bridge two elements of a text, the whole schema is momentarily accessible. An expectation should be regarded as a connection which 'forces its way' in the working memory when another link of one of the schemata it belongs to is required for bridging. In other words, each time a connection of a schema is used during comprehension, it may 'drag' with it other elements of that schema, which constitute the expectations associated with the bridging connection.

Since an expectation is just another connection in the working memory, it obeys all the rules previously introduced. In particular, it has an activation level and decays with time. As with other elements in FOCUS, its initial energy is determined dynamically according to the current context and to the semantic knowledge possessed by the comprehender prior to the reading. Intuitively, we would like to assume that this initial activation level is relatively low, so that expectations would be among the first things to be moved out of focus when necessary. This low initial energy would also emphasize the need for the reading to *confirm* an expectation soon after its addition to the working memory, that is, before it is forgotten. The confirmation of an expectation consists in the addition to FOCUS of connections that are additive with this expectation. From this point of view, expectations play a most important quantitative role in that they *speed up* the bridging of the clause(s) that confirm(s) them: the presence of an expectation in the working memory constitutes a pre-constructed bridge whose explanatory power is grounded in the schema associated with this expectation. Finally, the selection and addition to FOCUS of an expectation is part of the time-constrained race which delimits the bridging of each clause to its context. In other words, time restricts the number of expectations placed in focus.

It was observed earlier that the few information-processing researchers who address the problem of subject matter postulate an *a priori* set of macrostructures that, in essence, specifies exhaustively the gists that their models can recognize. I reject the idea that one can find a correct and complete set of such macrostructures. However, I do not want to deny that, with experience, each reader *acquires* such macrostructures. But each macrostructure merely provides a set of expectations idiosyncratic to the reader: a macrostructure is viewed as a schema for a particular gist. Thus, the expectations of a reader are not limited to the local level but indeed may concern the global coherence of a text. Macrostructures are not required in the proposed model of comprehension; they do not define what the text is or isn't about, but rather they mostly *facilitate* comprehension.

Schemata provide a mechanism to assemble conceptual links into a cognitive unit. Through the generation of expectations, they allow the reader to place in focus elements which are not required for the perception of global coherence.

Schematic membership is not the only device that can introduce non-bridging (and therefore, non-required) connections in the working memory. ***Remindings*** (see Schank, 1982) may also place in focus conceptual links which, in this case, have no direct relation to the text. From this observation, a slight distinction can be made between global coherence and subject matter, which extends beyond the connections generated from the text to include such indirect links. Typically, remindings are even more idiosyncratic than expectations and operate at all levels of comprehension: a word, a sentence, or the entire contents of FOCUS at a given point in time may conjure up such remindings. Their effect varies greatly: one may facilitate comprehension, in which case it essentially acts as a schema which sets up expectations, or it may do nothing more than slow down the reader.

3.2 On Interestingness

The proposed model of comprehension operates at the basic level of conceptual connections. One of the characteristics of this model is that when selected to be added to the working memory, a connection is *qualitatively* assigned an initial activation level. This initial energy constitutes a major factor in determining the future role of the connection itself. I now suggest that this initial activation level corresponds to the *interestingness* of the connection. Hidi and Baird (1986) report that the perception of this interestingness is partly idiosyncratic: "Interest occurs only in the interaction of stimulus and person so that one can never stipulate its origin in one to the exclusion of the other". They also remark that "what is central to the response of interest is that a person is compelled to increase intellectual activity to cope with the greater significance of incoming information". What does it mean to "increase intellectual activity"? The answer to this question has mostly to do with time. Upon the addition of a connection to the working memory, its initial energy may affect several parameters of the model. Consider, for example, a 'very interesting' connection, that is, one with a high initial activation level:

1. working memory capacity: the maximum for the total energy of connections in FOCUS could be increased.

2. decay: all decay functions of elements in FOCUS could be 'slowed down'.

3. racing delay: the time allowed to search for bridges could be significantly increased.

The example intuitively corresponds to a comprehender who, after reading something he considers interesting, starts to concentrate (at least temporarily) on subsequent input. Similarly, a reader who gets bored would probably decrease these parameters. Interestingness is necessary for the perception of subject matter. It cannot be grounded solely in the text, and it affects every aspect of comprehension leading to, in effect, a *personalization*, an *appropriation* of the text. And, in the end, this is what *to comprehend* (from *com* + *prehendere* to grasp) means.

4 Conclusion

In this paper, I have investigated some of the quantitative aspects of text comprehension and suggested that they originate in time. Several, if not all, of the quantitative parameters of the proposed model may vary from one individual to the next; they form a first level of explanation for the idiosyncratic nature of reader-based comprehension.

This paper should be regarded as the preliminary specifications of a model of reader-based text comprehension. All aspects of this proposal must be explicited, explained, and illustrated with respect to a particular representational scheme. The ultimate goal of this research is the detailed specifications of an implementation of a model of comprehension which follows a reader-based design philosophy: a text is processed with respect to *a* reader's knowledge base; there is no quest for the knowledge of a 'competent' reader. In other words, the issue is not *what* should be in the knowledge base but rather *how* every element of this knowledge base should be specified and used.

References

[1] Baddeley, Alan (1976) *The Psychology of Memory*, Basic Books, New York.

[2] Bransford, John and Johnson, Marcia (1973) "Considerations of Some Problems of Comprehension", In: Chase, William (ed.), *Visual Information Processing*, Academic Press, New York.

[3] Dyer, Michael G. (1983) *In-Depth Understanding*, MIT Press, Cambridge, MA.

[4] Feldman, Jerome (1984) "Computational Constraints from Biology", *Proceedings of the Sixth Annual Conference of the Cognitive Science Society*, Boulder, Colorado, p.101.

[5] Gadamer, Hans-Georg (1976) *Philosophical Hermeneutics*, translated by David Linge, University of California Press, Berkeley, CA.

[6] Gigley, Helen (1985a) "Computational Neurolinguistics: What Is It All About?", *Proceedings of the Ninth Joint Conference on Artificial Intelligence*, Los Angeles, California, pp.260–266.

[7] Gigley, Helen (1985b) "Grammar Viewed as A Functional Part of a Cognitive System", *Proceedings of the 22nd Anuual Meeting of the Association for Computational Linguistics*, Chicago, Illinois, pp.324–332.

[8] Graesser, Arthur and Clark, Leslie (1985) *Structures and Procedures of Implicit Knowledge*, Ablex Publishing Corporation, Norwood, NJ.

[9] Hidi, Suzanne and Baird, William (1986) "Interestingness—A Neglected Variable in Discourse Processing", *Cognitive Science*, 10:179–194.

[10] Hirst, Graeme (1981) *Anaphora in Natural Language Understanding*, Lecture Notes in Computer Science 119, Springer-Verlag.

[11] Hirst, Graeme (1987) *Semantic Interpretation and the Resolution of Ambiguity*, Studies in Natural Language Processing, Cambridge University Press, Cambridge, England.

[12] Holub, Robert (1984) *Reception Theory—A Critical Introduction*, Methuen, New York.

[13] Márkus, András (1983) "Shifting the Focus of Attention", *Proceedings of the Eighth International Joint Conference on Artificial Intelligence*, Karlsruhe, West Germany, pp.66–68.

[14] Phillips, Martin (1985) *Aspects of Text Structure*, North Holland Linguistic Series, Amsterdam.

[15] Schank, Roger (1982) *Dynamic Memory*, Cambridge University Press, New York.

[16] Stevenson, Rosemary (1986) "The Time Course of Pronoun Comprehension", *Proceedings of the Eighth Annual Conference of the Cognitive Science Society*, Amherst, Massachusetts, pp.102–109.