A Cognitive Basis for a Computer Writing Environment

TR87-032

June 1988

John B. Smith
Marcy Lansman

The University of North Carolina at Chapel Hill
Department of Computer Science
CB#3175, Sitterson Hall
Chapel Hill, NC 27599-3175

A TextLab Report.

Portions of this research were supported by the National Science Foundation, Grant# IRI-8519517 and the Army Research Institute Contract# MDA903-86-C-0345.

Introduction

During the past ten years, our understanding of writing has changed significantly. It was in 1980 that Dick Hayes and Linda Flower first outlined what has since become the standard model for both composition theorists as well as cognitive psychologists studying writing. As a result, the focus of research has shifted from the *products* of *writing* to the processes of writers.

During that same period, a revolution also took place in computers. The first Apple microcomputer was delivered in 1978. Before that time, virtually all access to computing was through mainframe or mini-mainframe machines operated from a central location. These machines provided a highly technical, generally unfriendly, computing environment. To use the machine, one either went to the central computing facility or accessed it remotely via telephone line. All that changed with the microcomputer. The computer became a personal, rather than public, instrument that could be used wherever electricity was available. And the complex interface of the mainframe was replaced by the more inviting, often graphic interface we have come to associate with the micro.

These changes in computers also produced changes in computing. One of the most important was a shift away from numerical to symbolic computing, particularly word processing. Writers immediately saw that the user-friendly microcomputer was superior to the typewriter or pen as a tool for writing. They could rearrange sections, format the document, or produce a complete new draft at will. This new breed of computer writer came not just from scientific and technical fields but from the humanities, the ranks of students, and from managers and other professionals in business and industry. The computer emphasized that writing was a common denominator for many different jobs and activities and that becoming a better writer helped the individual become a better scholar, student, or professional.

Not surprisingly, this rapid growth in computer writing led to more advanced writing tools. Spelling checkers became an expected part of word processing programs. Recognizing that the structure of a document is separable from the text or content that fits within that structure, system developers offered writers programs to help them outline their ideas and then write their documents within that framework. Even programs that analyze – albeit rather crudely – the writer's style are beginning to appear.

While these programs offer writers new tools, they do so piecemeal and with minimum concern for the large-scale structure of the writing task. Their designs often seem driven more by what the computer can be easily programmed to do rather than what will help writers most. Badly needed are tools designed from the outset to closely match and to augment the inherent cognitive processes human beings use to perform the complex, multifaceted task of writing.

The nature of the interaction between tool and tool user for computer writing invites, perhaps demands, a reconciliation between cognitive research and system design. Computer writing systems are examples of "intelligence amplification" systems. This type of program is intended to help the user think better or more efficiently. Thus, they don't work with extrinsic data, such as payroll information or observed data from an experiment, but with intrinsic data, data that are part of the thought processes of the human being using the system. The design of such a system must closely match the mental processes of the users performing the supported task. If it does not, the system will intrude on the user's thinking, perhaps distorting as well as slowing down those mental processes.

The research of cognitive psychologists and composition theorists offers important insights that can guide development of more compatible computer systems. In the sections that follow, we first review some of their more important theories and experimental results in order to establish a cognitive basis for a computer writing environment. We then show how those insights influenced key design decisions for a system we are developing. While our system could be used by a variety of writers for many different purposes, it is intended primarily for professionals who write as a part of their job. Nevertheless, we believe it illustrates the important relation between cognitive theory and system design and the necessity to consider them together. Our discussion ends with a brief description of our efforts to test both the theoretical basis and the system we have developed in accord with it.

Research on Written Communication

Introduction

Research dealing with written communication is extensive and can be found within several disciplines. The group most directly concerned with writing, *per se*, are the composition theorists. While the emphasis they place on students writing within an academic setting sometimes limits the generality of their work, their research has provided many important insights, especially the role of planning in the overall writing process.

A second major body of research is that of cognitive psychologists. Important for our concerns is their work on the different cognitive processes used by writers, the different intermediate products on which those processes operate, and the succession of subgoals writers must set for themselves in order to produce a document. Research on reading comprehension is also relevant for identifying the characteristics of written documents that make them easier to read and comprehend.

Reading Comprehension

Comprehending a written text involves cognitive processes ranging from decoding individual words to abstracting the 'gist' of the text as a whole. As a result of these various cognitive processes, readers create a memory representation of the text that is usually quite different from the linear sequence of words that they read. This mental representation may be similar or dissimilar to the meaning the writer intended to communicate. Consequently, if writers want to produce texts that can be read and understood easily and accurately, they must understand the cognitive processes used for reading and the textual features that facilitate those processes.

While decoding individual words is complex activity and a subject of continuing research, we will not consider that work here, since writers can do little to affect that process other than selecting words that will be known by their readers. Rather, we focus on research that addresses the active construction of meaning from, first, combinations of words and, then, larger segments, ranging from sentences and paragraphs up to the entire text.

Readers rarely recall text verbatim [Bransford & Franks, 1977; Sachs, 1967]. Instead, they combine the meanings of groups of words to form more abstract mental representations that are stored and later recalled. Many theorists have suggested that text meaning is represented as a series of propositions [Anderson, 1983; Kintsch & van Dijk, 1978], where a proposition is an elemental unit of meaning, composed of concepts rather than words, that makes an assertion about an event or state. Thus, a proposition posits a relationship between two or more concepts. A sentence may be broken into more than one proposition, but a given proposition may also be expressed by several alternative sentences.

The meaning of connected text is also transmitted through relationships between sentences and their underlying propositions. These relationships, called "coherence relations," are conveyed by a number of rhetorical devices, the most well-studied being common referents. The mental representation of such relationships can be symbolized by a 'coherence graph', which shows the links among a number of propositions [Kintsch & van Dijk, 1978].

Coherence graphs indicate that many texts can be represented as hierarchical structures in which key propositions are linked to subordinate propositions. Thus, by selecting a major superordinate idea and then relating subordinate ideas to it, one can construct a tree-diagram or "text base" that indicates the content structure of the text. The psychological reality of such a representation is supported by the fact that recall of a proposition is significantly affected by the position of that proposition in the hierarchy: propositions high in the tree structure are recalled by experimental subjects better than propositions lower in the structure [Meyer, 1975; Kintsch & Keenan, 1973; Britton, Meyer, Hodge, & Glynn, 1980].

The process by which the individual links in the hierarchy are constructed has been examined in detail by Kintsch & van Dijk [1978]. In order to build a text base, the reader follows a step-by-step process in which the propositions in a sentence are related to referents in adjacent sentences. Since short-term memory can retain only a few propositions at a time, the reader first attempts to connect a new proposition to one already in short-term memory. If the link is made, the new text being processed is perceived as coherent with the text just read. If not, an inferential bridging process is initiated to locate a similar proposition in long-term memory and place it in short-term memory. But in this last case, comprehension is slowed considerably [Kintsch & van Dijk, 1978]. Thus, textual features that highlight relations among propositions facilitate comprehension.

The structure of a written text is not limited to the relationships between adjacent sentences. Recent theories of reading comprehension deal with the more global structure of the text as well as lower level structures. Van Dijk, [1980] in particular has been concerned with the "macrostructure" of the text. Beginning with the first phrase in the first sentence, readers form and test hypotheses as to the overall point of the paragraph. Subsequent sentences cause them to revise their hypotheses. As readers proceed through the text, they abstract from the paragraphs generalizations and hypotheses concerning the main points of sections, chapters, even the entire piece. The resulting mental representation of the text is a hierarchical macrostructure with the main point(s) of the piece at the top and successively more detailed summary propositions or "macrofacts" at lower levels.

Thus, as readers comprehend texts they analyze those texts at several levels simultaneously. At a local level, they integrate individual propositions by establishing common referents, conditional relations, etc. At a global level, they form hypotheses as to the higher level meaning structure of the text, i.e., the main point of each paragraph, the superordinate point of each section, etc.

The simultaneous demands of local and global analysis place a tremendous cognitive burden on the reader. These demands are somewhat lightened by the fact that readers often approach a text with some knowledge of what the global structure of that text will be. For example, readers of an experimental article expect the introduction to provide a rationale for the experiment. Readers of a fairy tale expect the initial sentences to provide a setting for the story. These preconceived ideas about the structures of various types of texts have been labelled "schemata" by cognitive scientists, and their importance in text comprehension has been amply demonstrated (See Bower & Cirilo [1985] for a brief review).

The schema for a certain type of text may be activated either by the context in which the text is found (e.g., one expects to read an experimental article when it is published in a certain type of journal) or by characteristics of the text itself. Once a particular schema is activated, readers expect the text to have a certain structure, and they search the text for the propositions that can fill pre-established positions in that structure. If the text is structured as the schema suggests, comprehension is facilitated. If not, comprehension is impaired [Kintsch & Greene, 1978; Thorndyke, 1977].

However, even when the general structure of a text is dictated by a relatively fixed schema, the more detailed structure is not. For example, although readers expect the introduction of an experimental paper to provide the rationale for the experiment, that rationale may be structured in many different ways. The reader depends on the text itself to reveal the particular structure for that particular case. Furthermore, for many types of technical prose, no schema exists, i.e., there is no set form that all documents follow. In these cases, the reader is completely dependent on cues provided by the writer in order to successfully comprehend the macrostructure of the text.

Whether or not the reader has a pre-existing schema, the process of abstracting structure is not foolproof. Success is gauged by the extent to which the reader derives from the text the main points the writer wished to communicate. All of us have had the experience of discussing an article with a colleague who derived an entirely different message than we did. In the case of aesthetic literature, such ambiguity may be tolerable, even desirable. But for technical prose, it represents a failure on the part of the writer.

What strategies, then, can we recommend to writers to increase the probability that readers will comprehend the macrostructure of their texts? First, the writer must have a clear idea of what that structure is. Second, that structure should be made explicit in the document. If van Dijk is right in claiming that readers formulate hypotheses as to the main point of a paragraph or sections as soon as they begin to read the first sentence, then the writer can lighten readers' cognitive load by making those points as accessible as possible. Third, the writer should keep in mind readers' pre-existing expectations (schemata) concerning the structure of the text. If the text violates expectations, the writer must be particularly clear in indicating the intended structure of the text.

Hierarchical structure is particularly important in text organization. Various theories of reading comprehension agree that at both local and global levels, readers attempt to abstract the hierarchical structure of text, i.e., they constantly try to locate the main point of a paragraph, section, or entire text. Once identified, the main point can then be represented in long-term memory while subordinate or irrelevant points are allowed to be forgotten.

Research indicates that specific features which signal the structure of the text facilitate comprehension. For example, thematic titles presented prior to a well-structured text significantly increase free recall of the content of that text [Schwartz & Flammer, 1981]. Within a text, advance organizers – passages containing the main concepts of a text or section of text but at a higher level of abstraction – positively affect comprehension [Ausubel, 1963]. Hierarchical texts in which the structure is signaled or cued are comprehended more effectively than texts in which the structure is not signaled [Meyer, Brandt,& Bluth, 1980]. And at the paragraph level, inclusion of a topic- or theme-sentence in the initial position, rather than in an internal position or not at all, results in more accurate comprehension [Kieras, 1980; Williams, Taylor, & Ganger, 1981]. Thus, clear signaling of the author's intended hierarchical structure of concepts through typographic and rhetorical conventions strongly influences the reader's comprehension of a text.

Guidelines for Effective Documents

These results offer clear advice for writers. That advice can be consolidated and restated as the following guidelines:

- Structured documents are more easily comprehended than unstructured ones.

- Hierarchical structure is a particularly effective, perhaps optimal, form.
- Textual features that signal or cue the hierarchical structure of a document increase its comprehensibility. These include:
  - –Descriptive titles
  - –Advance organizers, or summaries
    - —for the document as a whole
    - —for major sections
    - —for individual paragraphs (particularly topic-sentences in initial positions).

While these guidelines do not guarantee success, they suggest that a document that is hierarchically structured should be understood more easily and more accurately than one that is not. Since the individual points made by a document are understood as they relate to one another, their aggregate impact is likely to be more convincing when these relations culminate in a single high-level concept as opposed to the same points taken individually or related in non-hierarchical ways. Consequently, writers that follow these guidelines should produce documents that are more efficient and more effective than those who do not.

These guidelines can also serve as a target for developers who wish to build more effective computer writing environments. The functions and organization of such systems should help writers, naturally and unobtrusively, construct documents with these features. Critical questions for research, then, are the strategies writers use to transform loosely connected networks of ideas into coherent, tightly-structured hierarchical documents and the architecture of computer systems that can assist them in this process. We will return to these questions, below, when we describe our attempt to develop such a system.

The Cognitive Processes of Writers

So far, in identifying some of the more important characteristics that make a document readable, we have been concerned primarily with the products of writing. Here, we consider the processes writers use to produce those products.

Cognitive psychologists have been slow to turn their attention to writing, perhaps because drawing generalizations about the mental processes that underlie an activity that is so open-ended is difficult. Psychologists feel more comfortable studying situations in which a specific stimulus is presented, a specific response is requested, and the response is then analyzed to infer the cognitive processes that intervened between stimulus and response. Writing does not fit this general paradigm. The environmental variables that lead a writer to write are not usually well-specified; the response – the written product – is complex and difficult to analyze objectively; and the processes that intervene between stimulus and response vary immensely from individual to individual.

In spite of these difficulties, an increasing number of cognitive psychologists and composition theorists are becoming interested in the cognitive processes that go on while a person is writing. In reviewing their work, we will focus on research dealing with the cognitive strategies used by writers since our goal is to develop better computer tools to enhance those strategies. We will be particularly concerned with the strategies writers use to generate and modify the structure of their documents, rather than strategies that underlie the composition of individual sentences.

In much of the early literature on composition, producing a document was assumed to involve three consecutive stages: *planning*, *writing*, and *revising*. During the first stage, writers gathered and organized their ideas. During the second stage, they translated these ideas in coherent text. During the third stage, they revised that text to produce the final document. As most of us can testify from our experiences as writers, the process of writing is much more complex than indicated by this simple three-stage model. Indeed, the model seems more prescriptive than descriptive: It says more about how some teachers think we should write than how we actually do write. Recent research on the cognitive processes of writers has indicated that the three-stage sequential model is indeed a gross over-simplification of what goes on during writing. At the same time, that research also suggests that the recommendation to isolate the various phases of writing, thereby reducing cognitive load, is valid. In the remarks that follow, we will look at research that describes the strategies writers use to manage these various phases.

Research on the role of planning in writing has taken many forms. Populations ranging from elementary school children to professional writers have been studied. Methods have ranged from formal studies, in which instructions to outline were experimentally evaluated, to observational studies, in which a single professional author recorded his thoughts throughout the process of writing an article. The results of such a broad range of studies are hard to summarize, especially since few of those studies were motivated by a comprehensive model of writing. However, the research does seem to converge on the conclusion that skilled and mature writers, when compared to unskilled and immature writers, plan what they are going to write and often separate the planning phase of writing from the composing phase.

Developmentally, the strategy of planning a document, in contrast to simply writing whatever comes to mind, emerges fairly late in childhood. This point has been made most clearly by Bereiter and Scardamalia [1987], who asked children of various ages to produce a written plan for a paper they were going to write. They found that children under the age of 14 produced "plans" that were nothing more than rough drafts of the papers themselves. This result is consistent with the general finding that when writing, children often simply tell all they know about a given topic, as they would in a conversation [Bereiter & Scardamalia, 1987]. As children learn to express themselves in written as well as spoken language, they only gradually acquire the strategy of planning what they want to say. Bereiter and Scardamalia [1987] found that older students, when asked, can produce plans that are distinct from the text itself. But other investigators have shown that even high school and college students devote little time to planning before they begin to write and that few produce written outlines [Humes, 1983].

Given that the ability to produce a written plan increases with age, one might ask whether written plans actually improve the quality of the final document. Research on adult writers indicates that they do. Kellog [1983] hypothesized that writing an outline before beginning to compose a draft would reduce both the capacity demands and the memory load associated with composing. He compared two groups of college students, one that was asked to produce an outline before beginning to compose a complex letter and one that was not. Using the method of 'trained introspection,' he asked all subjects to indicate once per minute whether they were planning, translating (i.e., composing sentences), or revising. Results indicated that the subjects who outlined spent more of their actual writing time translating and producing text judged to be more effective and better developed than those that did not. In a survey of faculty members, Kellog also found that those who were the most productive used outlines.

In the studies reported above, writers were instructed to produce written plans before beginning to write. Clearly not all planning results in a written plan. Nor does all planning take place before the writer begins to compose a draft. A number of researchers have asked what planning strategies writers adopt when they are not explicitly instructed to produce written plans. Matsuhashi [1981] assumed that whenever writers pause during the act of writing they must be planning. She studied videotapes of writers to determine exactly when planning takes place. Results based on one skilled high school writer indicated that planning took place throughout composition, both within and between sentences. Furthermore, a project that required the subjects to generalize rather than simply narrate required more planning time.

Unfortunately, the fact that a writer pauses during writing does not tell us much about what mental processes were taking place during the pause. The writer may have been planning the next sentence or simply daydreaming. To address this issue requires more powerful observational and analytic techniques. It also requires a broader orientation in which planning is viewed in the context of the overall writing process and writers' strategic movement between the different phases of that process. The researchers who have addressed these issues most directly are Linda Flower and John Hayes.

While the work of Flower and Hayes as been far-ranging, we will be concerned here with three major contributions. The first is method. Flower and Hayes were the first to use thinking aloud protocols extensively as a method for looking into the writer's mind during the writing process. The second contribution has been a number of informal observations on the writing task, observations that indicate how varied the plans, mental representations, and goals generated by the writer are. Third is their formal model. Their model goes well beyond the earlier three-stage model by indicating how alternative writing strategies might be represented formally. Although it falls short of capturing the richness suggested by their informal observations, it is an important first step toward a more rigorous understanding of writers' cognitive processes.

As noted above, some researchers have assumed that when writers pause, they are planning. Common sense tells us, however, that this is not always the case. Rather than make such assumptions, we need a more informative way to study what is going on in the writer's head. One way is to ask writers to tell the experimenter what they are thinking. The resulting record of verbalized thoughts is called a "think-aloud protocol." Such protocols have been widely used to study problem solving. John Hayes and Linda Flower, who view writing as a type of problem-solving, imported the technique for studying writing.

The technique is certainly not perfect and has generated considerable debate [Nisbett & Wilson, 1977; Ericsson & Simon, 1980; Ericsson & Simon, 1984]. Not all cognitive processes that go on during writing or any other mental activity are accessible to the writer's conscious awareness. Furthermore, requiring writers to think aloud may change the writing process. Nevertheless, analysis of such protocols has provided important clues as to how writers work. We will discuss Flower and Hayes use of this method in more detail below when we describe their attempts to verify their model of the writing process.

A second major accomplishment of Flower and Hayes has been to show the complexity and diversity of the cognitive processes that go on during writing. They have convincingly argued that the three-stage model is a vast oversimplification and that any realistic model must provide for many different strategies for combining the various subprocesses involved in writing.

In their informal observations, Flower and Hayes have looked at the planning process for writing from several points of view. From one perspective, writing is a goal-directed process. Starting with the overall goal of producing a document with certain characteristics, writers develop a hierarchy of subgoals. Thus, for example, if the overall goal is to write a publishable experimental paper in a psychological journal, the writer may set a subgoal to review the literature in such a way as to highlight the need for a particular study, and, perhaps, a sub-subgoal to discuss the shortcomings of a pertinent study. Flower and Hayes have also shown that expert writers develop more elaborate goal structures than novice writers [Hayes & Flower, 1986].

From another point of view, the writer is seen as juggling a set of constraints [Hayes & Flower, 1980]. The final document must integrate the writer's knowledge of the subject, must be expressed in syntactically correct sentences, and must accomplish a certain purpose. Since meeting all these constraints simultaneously places too large a cognitive load on writers, they develop strategies to lighten the load by relaxing one or another of the constraints during different phases of writing. For example, during brainstorming, the writer relaxes the requirement that ideas be integrated. During organization, the writer relaxes the constraint that ideas be expressed in sentences but increases the requirement that ideas be integrated.

From a third point of view, writing requires that information be transformed through a series of representations, in which each successive representation is a closer approximation to formal language [Flower & Hayes, 1984]. Some of the intermediate forms typically produced by writers include words and phrases, visual images, loosely organized semantic networks, outlines, and verbatim segments.

Flower's and Hayes' informal observations on the nature of the writing process are filled with perceptive insights as to why writing is such a frustrating and at the same time satisfying activity. Their formal model attempts to go further by providing a systematic description of writers' cognitive processes and their strategies for managing those processes. To express the model, Flower and Hayes use the three types of representation most common in cognitive psychology: the box model, the flow chart, and the production system.

Their box model, shown in Figure 1, has three major components: the task environment, which consists of everything outside the writer's head; the writer's long-term memory; and the "monitor," a kind of homunculus which directs the actual cognitive processes of writing. The monitor is shown as directing three types of processes, reminiscent of the three phases in the stages model: planning, translating, and reviewing.

The difference between the Flower and Hayes model and the stage model is that the three processes do not take place in a fixed order. The range of possible sequences is described by the production systems shown in Figure 2. The system at the top is general to all writers. It indicates that under certain circumstances, "edit" processes (rule 1) and "generate" processes (rule 2) can interrupt other ongoing processes, but otherwise the active goal dictates the activity. That is, when the goal is to generate, writers "generate" (rule 7), when the goal is to oranize, they "organize "(rule 8), etc.

The system at the bottom of Figure 2 shows four possible writing strategies, which Hayes and Flower refer to as "configurations." Each of these can be inserted as rules 3-6 in the general system shown at the top. For example, in Strategy (Configuration) 4, writers follow the conventional three-stage model: they generate all the ideas to be included in the text (rule 3), organize them (rule 4), translate them all into text (rule 5), and then review the text (rule 6). On the other hand, in Strategy 1, writers generate an idea, organize it (it's not clear how one idea can be organized), translate it into text, review that text, and begin again. In other words, one idea is completely processed before the next is generated.

The subprocesses involved in the three major types of writing activities are represented by flow charts. As an example, the flow chart for the "generate" process is shown in Figure 3. It shows that ideas are generated in chains, that the previous idea was considered useful enough to include in the plan, and that the goal is still to generate. The flow chart allows for the possibility that the writer will either write down or not write down the ideas generated.

These models attempt to bring the modeling techniques of cognitive psychology to bear on the process of writing. But the question is, what does this formalization provide that less formal descriptions do not? Typically, cognitive psychologists justify formal models by arguing that they alone are sufficiently explicit to be testable. Ideally, a formal model generates predictions that can be matched against empirical data. Discrepancies between model and data lead to modifications in the model. But exactly what type of observation would cause Hayes and Flower to modify or reject their model? What kind of think-aloud protocol would disconfirm some feature of the model?

Figure 1:
Flower and Hayes Box Model

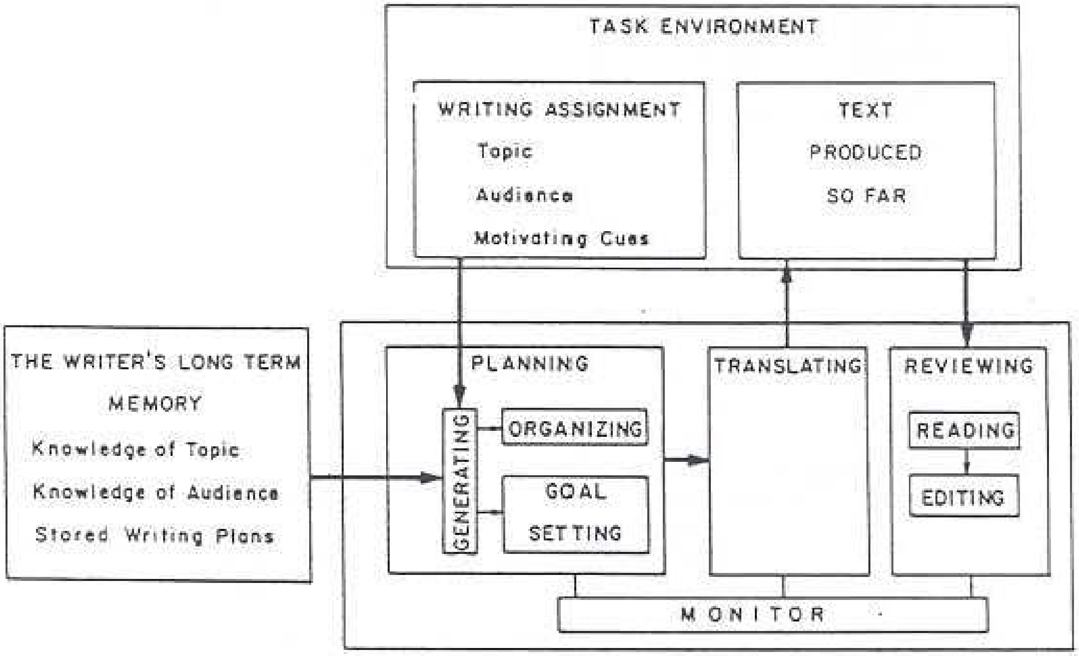
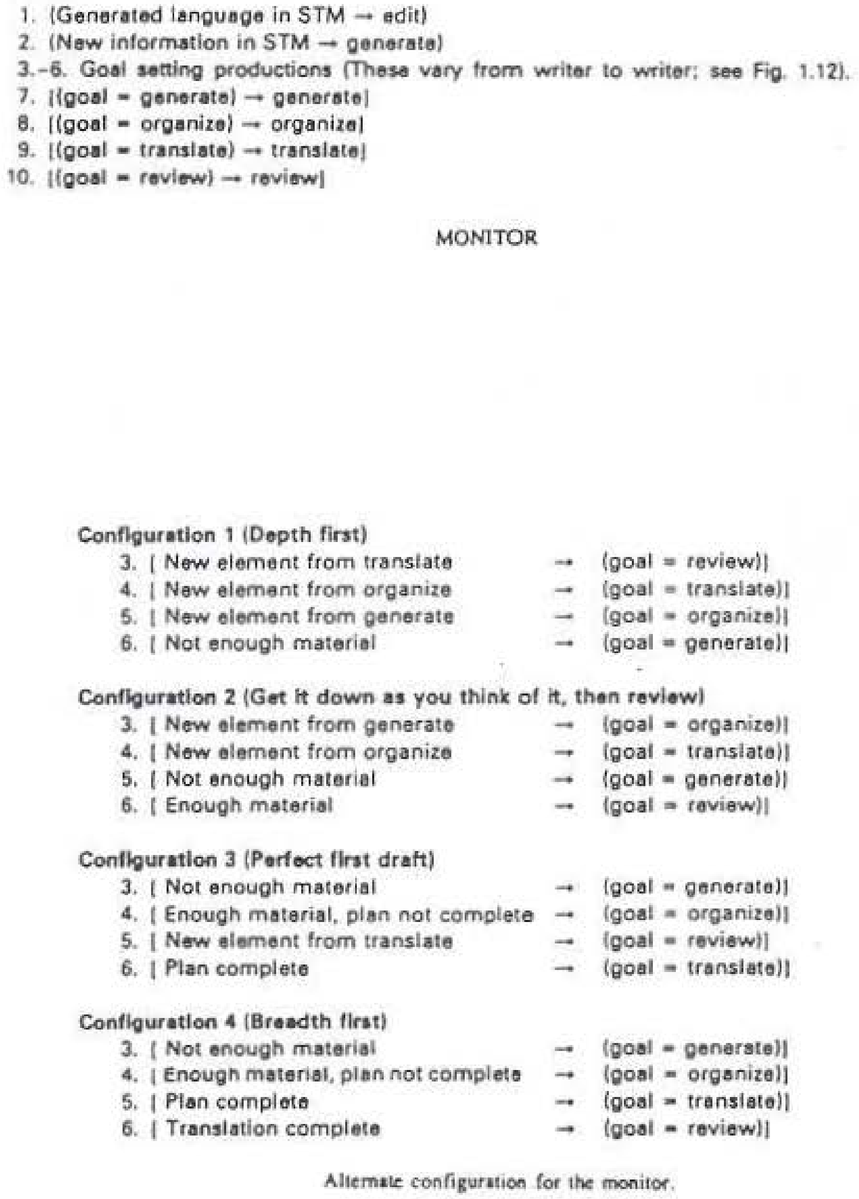


Figure 2:
Flower and Hayes Production Model



1. (Generated language in STM → edit)
2. (New information in STM → generate)
3.–6. Goal setting productions (These vary from writer to writer; see Fig. 1.12).
7. [(goal = generate) → generate]
8. [(goal = organize) → organize]
9. [(goal = translate) → translate]
10. [(goal = review) → review]

MONITOR

Configuration 1 (Depth first)

3. [New element from translate → (goal = review)]
4. [New element from organize → (goal = translate)]
5. [New element from generate → (goal = organize)]
6. [Not enough material → (goal = generate)]

Configuration 2 (Get it down as you think of it, then review)

3. [New element from generate → (goal = organize)]
4. [New element from organize → (goal = translate)]
5. [Not enough material → (goal = generate)]
6. [Enough material → (goal = review)]

Configuration 3 (Perfect first draft)

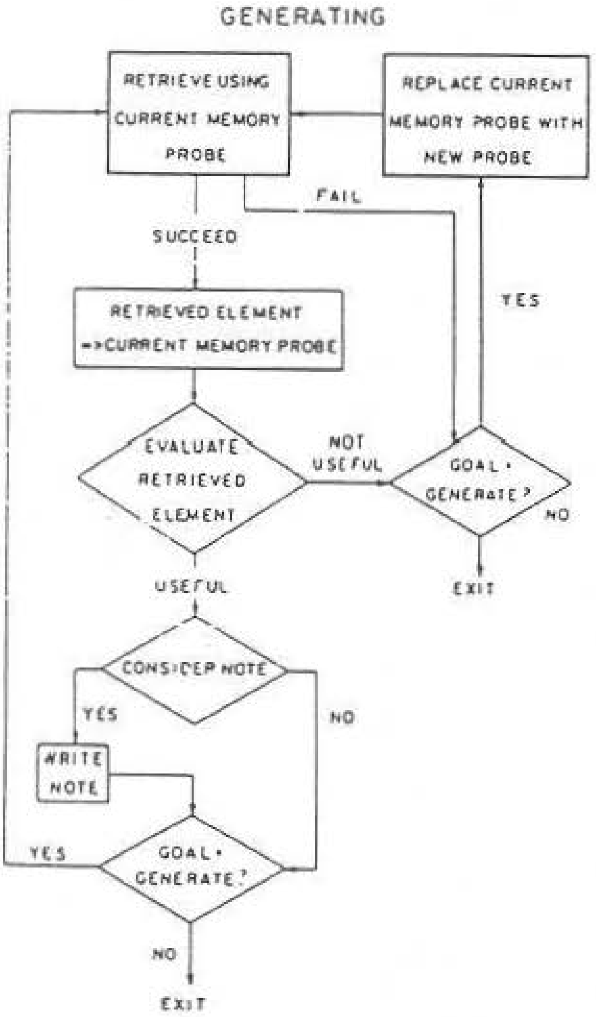
3. [Not enough material → (goal = generate)]
4. [Enough material, plan not complete → (goal = organize)]
5. [New element from translate → (goal = review)]
6. [Plan complete → (goal = translate)]

Configuration 4 (Breadth first)

3. [Not enough material → (goal = generate)]
4. [Enough material, plan not complete → (goal = organize)]
5. [Plan complete → (goal = translate)]
6. [Translation complete → (goal = review)]

Alternate configuration for the monitor.

Figure 3:
Flower and Hayes Flow Chart Model



Looking first at the box model that represents the overall structure of the process, we might ask what features are open to question. The most likely flaw in the box model is that it omits factors that are important in the writing process, such as time constraints. A second could be that some processes that go on during writing may not be categorizable as "planning," "translating," or "reviewing." Third, some protocol statements may combine two processes (see, for example, Berkenkotter's [1983] analysis of an experienced writer's thinking processes). And some may not fit into any of the three categories. Otherwise, its is hard to see how the structure could be shown to be inaccurate.

Turning to the flow chart for the generation process, it is again difficult to see how data from think-aloud protocols could show it to be incorrect, although many protocols might lack sufficient detail to test the model. The model specifies that if the writer's present goal is to generate ideas, he or she will use the current memory probe to search memory and either succeed or fail in generating an idea. A writer might easily fail to report in the protocol that his or her current intention was to generate and might also fail to report which, if any, memory probe was used to search memory. In other words, matching the flow chart against the protocol data might be very difficult. On the other hand, the protocol might disconfirm the model by suggesting that writers use a single probe again and again to generate a series of ideas.

The production system showing the interaction between generating, organizing, translating, and revising seems the most susceptible to revision on the basis of protocol analysis. For example, it seems likely that many writers would fail to follow any of the four strategies suggested by the model and that a hybrid version would be found in the protocols.

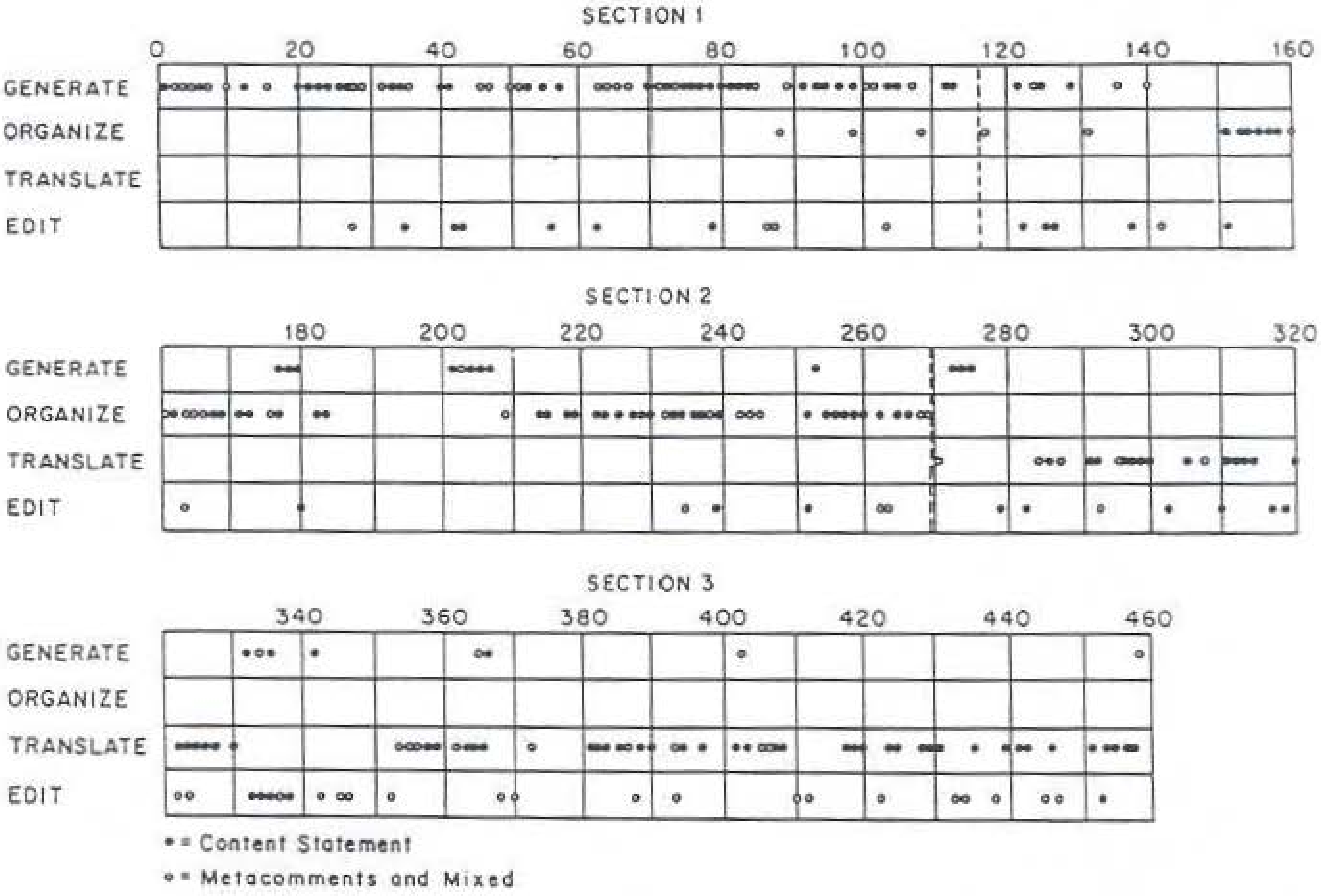
While analysis of protocols could raise problems such as these and, in turn, lead to refinement of the model, they have not. Hayes and Flower [1980] have published only one preliminary attempt to test the model. The data they present is a single protocol, characterized in Figure 4, produced by a single writer. In making their case, they assumed that the output of the generation process was words and sentence fragments, the output of the organization process was indented fragments, and the output of translation was complete sentences. On the basis of comments from the protocol, such as "And what I'll do now is jot down random thoughts," they concluded that the writer was best described by Strategy 4, i.e., the goals of generating, planning, and translating were adopted sequentially. They then divided the protocol into three segments: a generate segment (interrupted occasionally by editing), an organize segment (interrupted by generating and editing), and a translate segment (also interrupted by generating and editing). (At the time of publication, they had not analyzed the section of the protocol dealing with revision.) They then tested the hypothesis that written output generated during the three protocol segments would be of the appropriate types, e.g., that words and fragments would be produced during the "generate" segment. The hypothesis was confirmed: the majority of the written output was of the appropriate type.

According to Hayes and Flower [1980], analysis of this one protocol provided a "rigorous test" of the model. In fact, the analysis showed that when a writer said that he was going to generate ideas, he produced output that looked like ideas; when he said he was going to organize, the output looked like an organized plan; and when he said he was ready to write, he produced output that looked like written text. Thus, they concluded, the protocol supports Productions 7-9 in Figure 2.

But one must ask what kind of protocol would have caused them to revise their productions? Suppose, for example, the writer had said, "Now I'll jot down some ideas," and then proceeded to write down complete, connected sentences. Would Hayes and Flower have modified Production 7 to read:

$$[(\text{goal} = \text{generate}) \rightarrow \text{translate}]?$$

Figure 4:
Flower and Hayes Characterization of Protocol



Or, as seems more likely, would they reinterpret what looks like a goal statement – i.e., decide that by 'ideas' the writer had really meant text – or conclude that the writer had changed the goal without saying so?

A more severe criticism of this protocol analysis as a "rigorous test" of the Hayes and Flower model is that it involves a single subject who, they note, had "especially clear indications of ongoing writing processes." What of other writers? Did their protocols support the model, disconfirm it, or were they simply not clear enough to support a judgment?

Testing a formal model against think aloud protocols is extremely difficult. Few of the elements of the model can be observed directly and accurately. For example, writers may not articulate their goals. In fact, they may not even be conscious of them. Even the intermediate products of writing, e.g., ideas to be included or an organization plan, may not be mentioned in the protocol or observed in the output. Perhaps for these reasons or for others, in the eight years since it was published, Hayes and Flower have not refined their model of the planning process in response to actual protocols. While it may provide an intuitive sense of writers' strategies, it has been less successful at predicting actual, observable patterns or providing a rigorous, systematic understanding of writing.

Flower and Hayes have been more successful at abstracting informal observations from their protocols. They identified a number of different cognitive processes used by writers. They showed in their research on multiple representations that writers produce not just text but a variety of different information forms. And they showed that writing is not a simple process involving three sequential stages but, rather, is a complex task that involves multiple goals and recursive invocation of one cognitive process from another. While their formal model has not been completely successful, their work as a whole represents the largest single contribution to our understanding of the writing process.

Cognitive Modes

The work of Flower and Hayes, Bereiter and Scardamalia, and the other researchers cited above provides a rich body of concepts with which tounderstand the writing process. In this section, we will draw on that material in an attempt to build a cognitive basis for a computer writing environment. Most important are the concepts of *cognitive processes*, *intermediate products*, *goals*, and *constraints*. While each of these constituents is important, they take on added significance in combination. For example, to achieve a particular goal, writers use particular mental processes to produce particular intermediate products; however, both processes and products are constrained in ways consistent with that goal. In the remarks that follow, we examine the relations among these four elements. To clarify these interdependencies, we introduce the concept of *cognitive mode*.

Intuitively, a cognitive mode is a particular way of thinking that writers adopt in order to accomplish some part of the overall writing task. For example, early in the process, writers frequently engage in an exploratory mode of thinking. The goals for this activity are not to produce a draft of the document or even an organizational plan, but rather to externalize ideas and to consider various relations among them. Consequently, this way of thinking often carries with it a particular mood – relaxed, open to different possibilities, perhaps even playful. These goals and the accompanying relaxation of constraints are inherent in the mode, part of what makes exploratory thinking *exploratory* rather than *organizational* or some other form. Similarly, certain forms are appropriately produced during exploration while others are not. For example, words or phrases are typically jotted down to represent an idea; sustained prose is usually not. To produce these preliminary working products, writers emphasize particular cognitive processes and not others. For example, recall, representation, clustering, associating, and noting superordinate/subordinate relations are favored during exploration; sustained linguistic encoding, large-scale abstraction, and close analysis of text generally are not. Thus, a mode of thinking integrates particular sets of goals, constraints, products, and cognitive processes into a *complex whole*.

Looking more precisely at each of the constituents, we mean by *product* the symbolization of a concept or relation among concepts. While one can experience an amorphous thought, to relate that idea to other ideas, to recall it later, or to communicate it, one must transform it into symbolic form. Different cognitive modes provide different options for representation, such as words, notes and other jottings, outlines, and other forms. Thus, different forms tend to prevail in different modes. Some representations eventually become part of the final written document. Some do not. Those that do not are considered intermediate products that serve as stepping stones on the path from early, inchoate thinking to the final, refined document.

Processes act on products. In one mode, the processes might be perceiving an associative relation between two ideas or noting that one is subordinate or superordinate to the other. In another mode, the process might be constructing a large, integrated hierarchical structure composed of many such subordinate/superordinate relations. In still another, an encoding process might transform a word or phrase that represents an idea into a sentence that expresses it. Thus, different cognitive processes operate on different cognitive products to define them or to transform one form into another.

The *goals* for a mode represent the writer's intentions in adopting that particular way of thinking. While goals may be abstract, they are manifest in the target or final product the writer aims to produce. Goals are, thus, linked to the specific forms available in a given mode and, consequently, are implicit within that mode. For example, the goals for exploration are to externalize ideas and to consider various possible relations among small groups of ideas. But they are realized in particular concrete forms: words, phrases, or other symbols; clusters of such symbols; and small relational structures represented in various ways.

The *constraints* for a mode determine the choices available. Constraints are relaxed or tightened in accord with writers' large scale strategies in electing different modes of thinking for different purposes. For example, during exploration, constraints are relaxed to encourage spontaneity and flexibility and to increase the pool of potential ideas. During organization, constraints are tightened in order to build a coherent organizational plan. During writing, they are tightened still further as the writer produces continuous prose.

While products, processes, goals, and constraints can be discussed individually, they form a unified whole. Thus, specific interdependencies are inherent within the various modes. When writers enter a particular mode of thinking, they do so in order to achieve a particular goal. That goal will be represented as a product of a particular type and will be produced by a specific set of cognitive processes in accord with constraints appropriate for that mode. These *combinations* determine the kinds of objects that can be conceptualized, the kinds of relations that can be formulated among them, and the end product that can be produced in that mode of thought. The cognitive modes and their constituents that we believe are most important for writers are shown in Figure 5.

Experienced writers are likely to use these various modes in accord with conscious strategies. Strategies may be global, corresponding, for example, to the large-scale shifts from planning, to writing, to revising. Or they may be local, as in the case of recursive reapplication of planning mode during writing. Thus, writers shift cognitive modes in order to focus on one set of activities at a time and avoid dealing with all phases of the writing process at once – an impossible task. They also shift modes in response to specific problems in the structure of ideas they are currently working on.

The use of cognitive modes in accord with a global strategy should produce a progression of cognitive products that, in general, is orderly and predictable. As we noted

Figure 5:
Cognitive Modes for Writing

Constituents

Modes

	Processes	Products	Goals	Constraints
Exploration	•Recalling •Representing •Clustering •Associating •Noting subordinate superordinate relations	•Individual concepts •Clusters of concepts •Networks of related concepts	•To externalize ideas •To cluster related ideas • To gain general sense of available concepts •To consider various possible relations	•Flexible •Informal •Free expression
Situational Analysis	•Analyzing objectives •Selecting •Prioritzing •Analyzing audiences	•High-level summary statement •Prioritized list of readers(types) •List of (major) actions desired	•To clarify rhetorical intentions •To identify & rank potential readers •To identify major actions •Consolidate realization •To set high-level strategy for document	•Flexible •Extrinsic prespective
Organization	•Analyzing •Synthesizing •Building abstract structure •Refining structure	•Hierarchy of concepts •Crafted labels	•To transform network of concepts into coherent hierarchy	•Rigorous •Consistent •Hierarchical •Not sustained prose
Writing	•Linguistic encoding	•Coherent prose	•To tranform abstract representation of concepts & relations into prose	•Substained expression •Not (necessarily) refined

Figure 5 (cont.): Cognitive Modes for Writing

Constituents

Modes	Processes	Products	Goals	Constraints
Editing: Global Organization	•Noting large scale relations •Noting & correcting inconsistencies •Manipulating large scale structural components	•Refined text structure •Consistent structural cues	• To verify & revise large-scale organizational components	• Focus on large-scale features and components
Editing: Coherence Relations	•Noting coherence relations between sentences & paragraphs •Restructing to make relations coherent	•Refined paragraphs and sentences •Coherent logical relations between sentences and paragraphs	•To verify & revise coherence relations within intermediate sized components	•Focus on structural relations among sentences & paragraphs •Rigorous logical and structural thinking
Editing: Expression	•Reading •Linguistic analysis •Linguistic trasformation •Linguistic encoding	•Refined prose	•To verify & revise text of document	•Focus on expression •Close attention to linguistic detail

above, concepts are externalized, clustered, and linked into a loose network of associations during exploration. During organization, that loose network of ideas is transformed into a coherent structure for the document, which for expository writing is normally a hierarchy. During writing, the individual concepts and relations in the organizational plan are transformed into continuous prose, graphic images, or other developed forms. Editing is the process of refining the structure and expression of the document produced during writing.

However, this flow is not one-way and continuous, as suggested by the stages model. Rather, modes may be engaged recursively to solve specific problems. As a result, the flow of intermediate products may be reversed or restarted. For example, writers may find while organizing that they do not have critical information needed for a particular section. Rather than interrupt the current mode in order to get that information, they may elect to continue and leave the section in question undeveloped. Later, when the missing data is available, they would interrupt their writing, revert to organization or perhaps even exploratory mode, and build the missing branch of the document's structure. When the missing piece has been filled in, they would then resume writing. Thus, the general pattern in the transformation of intermediate products is predictable, but it may be interrupted for a specific, local reason.

In describing cognitive modes, we have suggested a number of predictions raised by the concept. For example, different modes should be preferred at different times in the overall writing process. Recursive invocation of one mode from another should be traceable to specific features or problems in the product currently being developed. Specific sets of cognitive processes should be used in conjunction with one another and with specific cognitive products. Thus, the general concept of cognitive mode as well as the specific modes shown in Figure 5 both generate hypotheses that can and should be tested experimentally. We return to this issue in section three of this paper when we describe several new techniques we have developed for protocol analysis and the particular hypotheses we are examining.

Implications for System Design

Introduction

In the previous section, we reviewed research in written communication in order to synthesize principles for developing a computer writing environment that would closely match the cognitive processes of writers. Here, we examine several key design decisions we made in light of those principles in our attempt to build an advanced Writing Environment (called WE).

Most important is the question of a single-mode system versus a multimodal design. Should all functions always be available to the user or should they be divided so that only certain combinations can be used at any one time? We also consider the dynamics of the system. As the writer transforms information expressed in one form into another, how can this flow of intermediate products best be managed and supported? This discussion is interleaved with our consideration of modes. The section ends with a brief description of features that might have been included in WE but were not.

Multimodal Design

In the previous section, we suggested that writing can be viewed as a complex process involving different cognitive modes. A key question for system design, then, is how best to support these different cognitive modes and the flow of intermediate products among them?

Two approaches are possible. In a single mode system, all system functions would always be available. For a writing environment, the set of functions would be the union of those required to support all of the cognitive processes for the different cognitive modes. A multimodal approach would divide the environment into separate system modes, each corresponding to one of the cognitive modes. If the second approach was followed, each system mode would include only the functions appropriate for its corresponding cognitive mode.

We adopted a multimodal system design for several reasons. As we discussed in the previous section, writers seem to manage the overall writing task by dividing that process into phases in which they engage different cognitive modes. Each mode is unique in terms of its particular combination of processes, products, goals, and constraints. Consequently, supporting these large-grained "chunks" of activity, each with its own unique requirements, in separate system modes seemed both natural and efficient: natural, in that system architecture would both mirror and reinforce cognitive strategy; efficient, in that specific system operations could be matched closely with specific cognitive processes. Also, specific rules for the objects that can be created and manipulated in each system mode could be matched with the specific intermediate products that writers define and transform in the corresponding cognitive mode, in accord with the goals and constraints for that mode.

Consequently, WE provides four system modes, each represented in a different window on the computer screen. We label these network mode, tree mode, editor mode, and text mode. They correspond to the exploratory, organizational, writing, and editing Modes of writing, respectively. They are initially displayed on the screen as shown in Figure 6. However, the screen can be reconfigured so that any single mode or combination of modes can be enlarged to occupy the entire screen. We did not include a mode for situational analysis, and we included only one mode for editing. Our reasons for both decisions are explained, below.

Network Mode

Network mode, shown in the upper left quadrant of Figure 6 and expanded in Figure 7, provides an environment tailored to the exploratory mode. The cognitive processes emphasized during exploration include retrieving potential concepts from long-term memory and/or from external sources, representing these concepts in symbolic form, clustering them, and noting specific relations among small groups of concepts, such as association or superordinate/subordinate relations. The intermediate products that are usually produced include individual concepts, clusters of associated ideas, and small relational structures. Since constraints are minimal in this cognitive mode, the emphasis is on flexibility and freedom so that the writer can consider various relational possibilities. These conditions can be met by a system mode that conforms to an underlying set of rules consistent with those for a network - or, more specifically, a directed graph - embedded in a two-dimensional space. To see why these rules are appropriate and to give a feel for the actual operation of the system, we describe, below, how the writer creates each form of intermediate product normally produced during exploration.

The system permits the writer to represent an idea by creating a small box (*node* in graph theory terminology) that contains a word or phrase signifying that concept. The writer creates the node simply by pointing with a mouse to the place on the screen where it is to be placed, selecting the "create" option from a menu, and then typing a word or phrase to represent the concept.

To cluster two nodes or ideas, the writer selects one of them and then points to the place on the screen where it should be placed.

To define a relationship between a pair of nodes that is stronger than simple spatial

Figure 6:
WE Initial Display

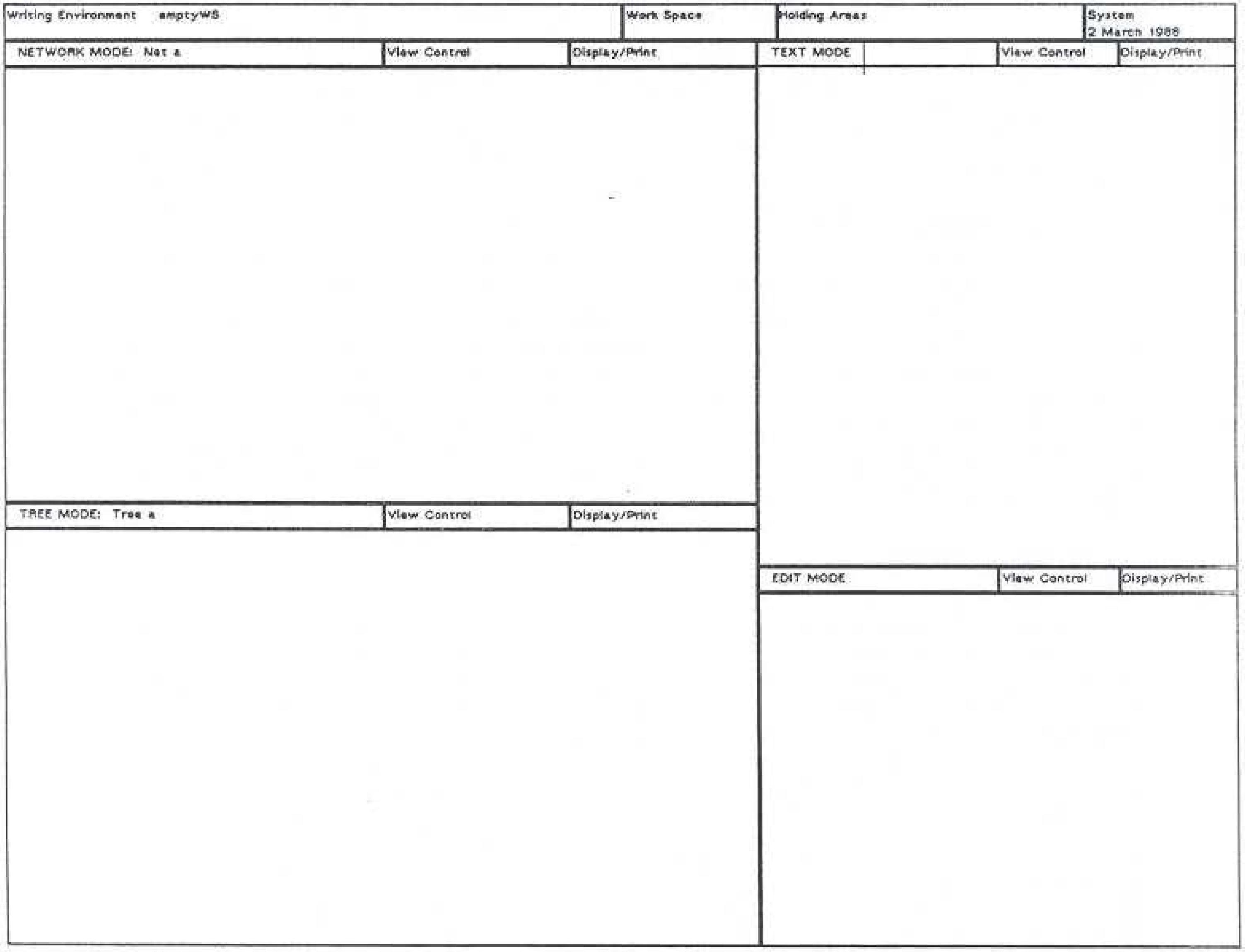
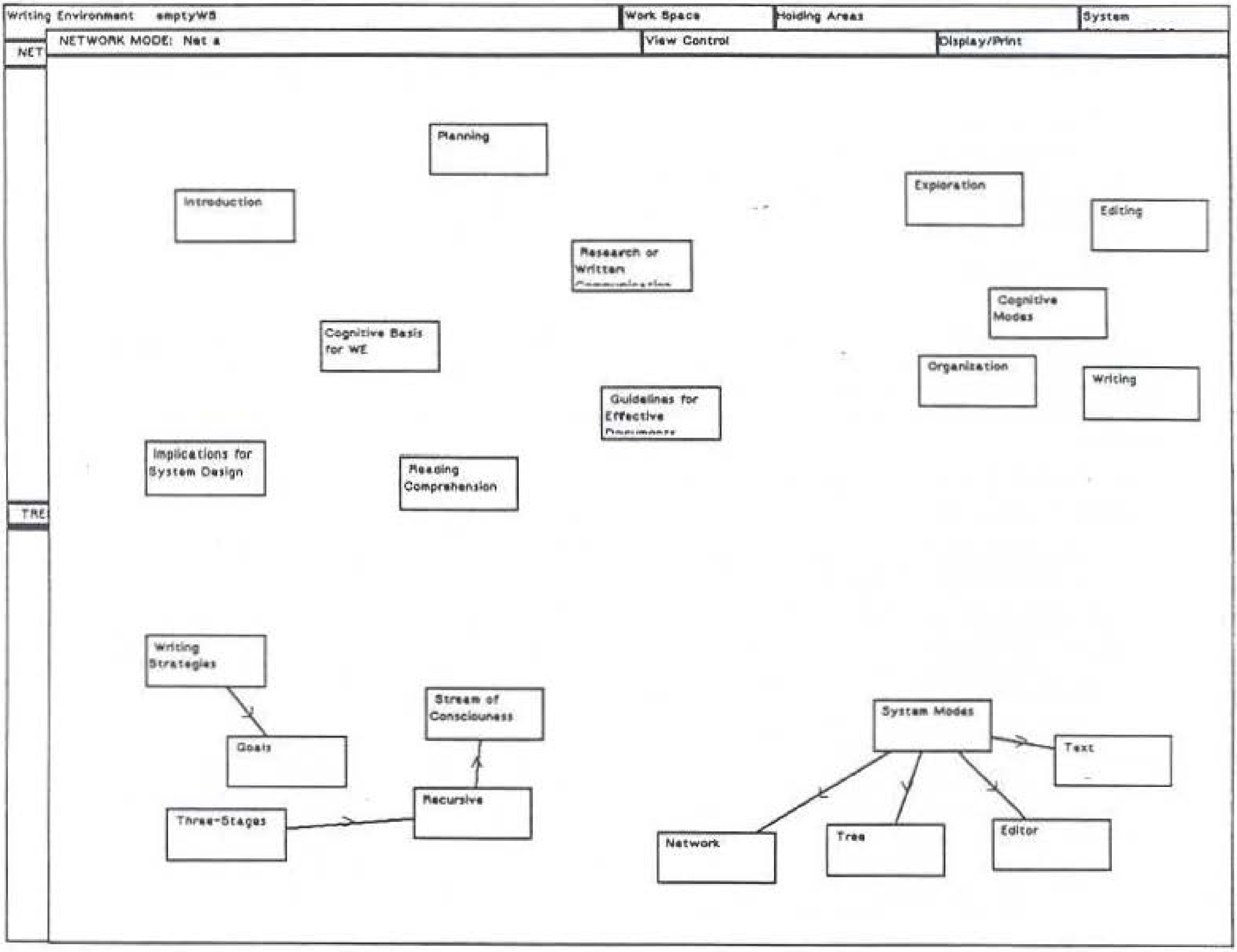


Figure 7:
WE Network Mode



proximity, the writer can create a directed link between them. Links, as well as nodes, can then be named, such as "is part of" as in "Associating is *part* of Exploring". Again, the manual operations for this process require little cognitive overhead and distract minimally from the conceptual task at hand.

To produce a hierarchical relation among a small group of nodes,the user simply constructs directed links from the superordinate node to each of the subordinate ones. Thus, in Figure 7, the writer linked a node labelled "System Modes" to nodes labelled "Network", "Tree", "Editor", and "Text." However, since the rules of network mode are those of a directed graph, the system does not "know" that these relations formed a hierarchy. Consequently, the system does not protect the writer from turning a hierarchy into a cyclic graph.

Thus, Network Mode provides a set of system operations that facilitate the cognitive processes normally used during Exploration. It provides concrete representations of concepts, clusters, relations, and structures. And it permits easy transformation of one well-defined intermediate product into another. Figure 7, in which network mode has been resized to fill the screen, shows examples of these various intermediate products.

Tree Mode

Tree mode, which appears in the lower left quadrant of Figure 6, provides an environment tailored to the organizational mode. The primary goal of this cognitive mode is to construct a coherent hierarchical structure for the document. The rationale for organizing the document as a hierarchy is found in the guidelines for effective documents, described above:

- structured documents are more easily comprehended than nonstructured ones
- hierarchy is a particularly effective, perhaps optimal, structure
- signaling the hierarchical structure through various typographical and rhetorical cues increases comprehension

Although writers can construct trees or hierarchies in network mode, we elected to support exploration and organization in separate system modes because the two are quite different. In exploration, constraints are lowered to emphasize flexibility; in organization, constraints are tightened to emphasize coherence and consistency.

The cognitive processes for the two are also different. While noting superordinate and subordinate relations during exploration is a natural act, organization is a much more deliberate activity that requires a different set of cognitive processes. Writers must think on a broader scale, noting relations among not just small groups of concepts, as during exploration, but whole substructures of ideas. They must note parallel relations among corresponding sections of the tree and balance the overall structure. Organization is, thus, a *building* task in which the parts must be fitted together with care and consistency to produce a coherent structure for the document.

The intermediate product that can be defined and manipulated in tree mode is, of course, a hierarchical structure, represented as a tree. Each node may have several links that leave it but each (except the root) can have one and only one link coming to it. This last restriction precludes cycles that would violate the integrity of the hierarchy. Thus, in tree mode the system "knows" that the structure is hierarchical and *insures* its integrity.

All operations within tree mode apply to a single tree. They include functions to define, develop, and edit a hierarchical structure represented as a tree. Users begin by constructing a root node for the tree. They can then construct a new superordinate node that becomes the new root or a subordinate node, referred to as a "child" of the "parent" node to which it is subordinate. Nodes as well as branches (a node and all of its descendants) can also be moved from one location to another in the tree. Figure 8, in which tree mode has been resized to fill the screen, shows a tree that has been constructed using these operations.

Although WE separates exploration and organization into two separate system modes, the two are closely related. Nodes as well as small hierarchical structures can be moved from network mode into tree mode. Thus, work done during the exploratory process is not lost when writers shift from network to tree mode since intermediate products flow naturally from one mode to the other, as suggested in the discussion of cognitive modes, above.

Finally, while the architecture of the system encourages writers to first use network mode for exploration before going to tree mode for organization, it does not require them to do so. If writers believe some structure other than a hierarchy is more appropriate, they can continue to work in network mode to develop an alternative organization plan. For example, they could use network mode to construct a long string of nodes, a highly interconnected network, even a single all-encompassing node that represents the entire (reductive) structure and then write the document accordingly. With this approach, they could skip tree mode entirely. Thus, the system encourages strategies that have been shown to be effective, but does not require them.

Editor Mode

Editor mode, shown in the lower right quadrant of Figure 9, provides a standard text editor for expanding the concept represented by a node into prose. Thus, it supports the cognitive process of linguistic encoding. The intermediate product, of course, is a block of conventional text that is associated with a particular node. The underlying system rules are those of a linear sequence of characters divided into words, lines, paragraphs, etc. In future extensions of WE, the system will support editors for other kinds of data, such as graphics, sound, and video.

Since the editor can be invoked from either network mode or tree mode, writers do not have to wait until the hierarchy for the document is complete to begin writing. They can expand a concept into text at any time after the node is created. Thus, the system can be used with a variety of writing strategies, including a pure three-stage approach, a recursive pattern, or a stream of consciousness in which the entire text is written within a single node.

Text Mode

Text mode, shown in the upper right quadrant of Figure 10, provides an environment for editing the document. However, it has a different relation to the editing process than the other system modes have with their corresponding cognitive modes. As indicated in Figure 5, editing is a complex activity that involves three different cognitive modes. The first addresses the global organization of the document and involves verifying large-scale features and, possibly, moving and refitting large units, such as paragraphs and sections. The second focuses on coherence relations among smaller segments, such as sentences, within an intermediate-scale frame of reference, such as a paragraph or section. Using a third cognitive mode, writers edit the actual linguistic expression to clarify sentences, to shift their meaning or emphasis, and to make them more graceful.

No single system mode supports all three editing modes. Rather, we presume that large-scale organizational editing will be done in tree or possibly network mode, where the whole (hierarchical) structure for the document can be seen and manipulated directly. At the other end of the spectrum, linguistic editing will be done in editor mode. Text mode supports the intermediate editing mode that focuses on coherence relations within and between paragraphs and sections.

Figure 8:
WE Tree Mode

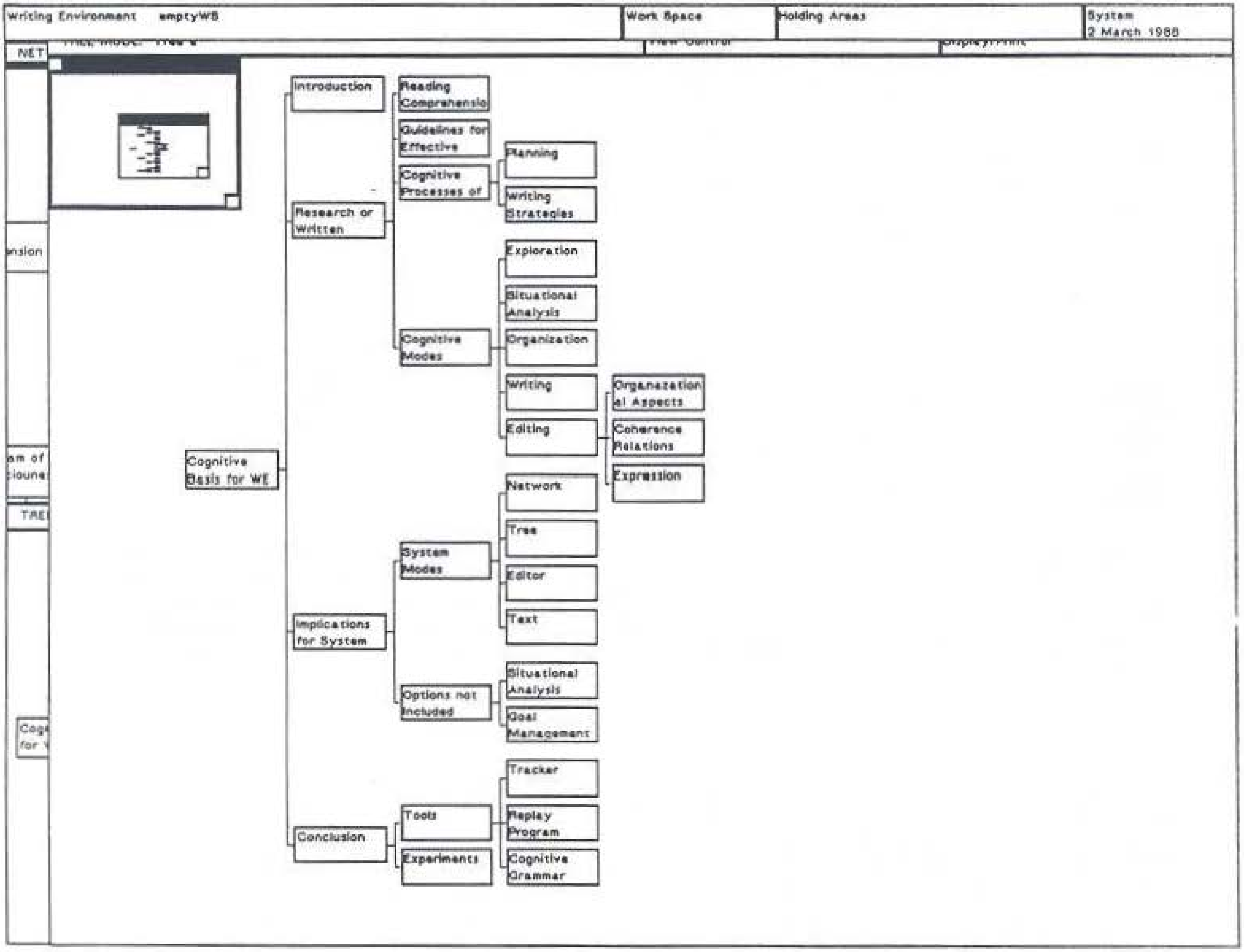
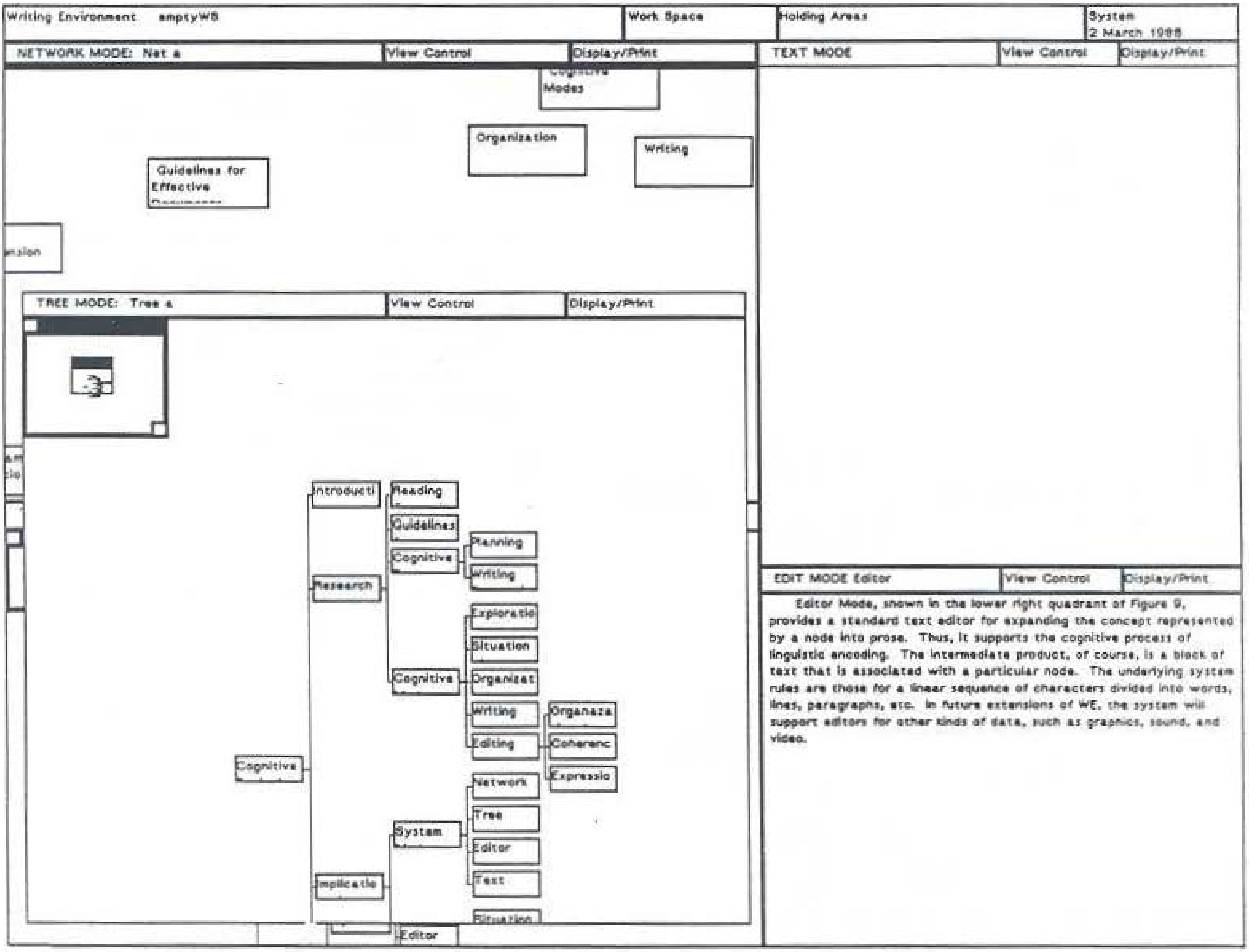
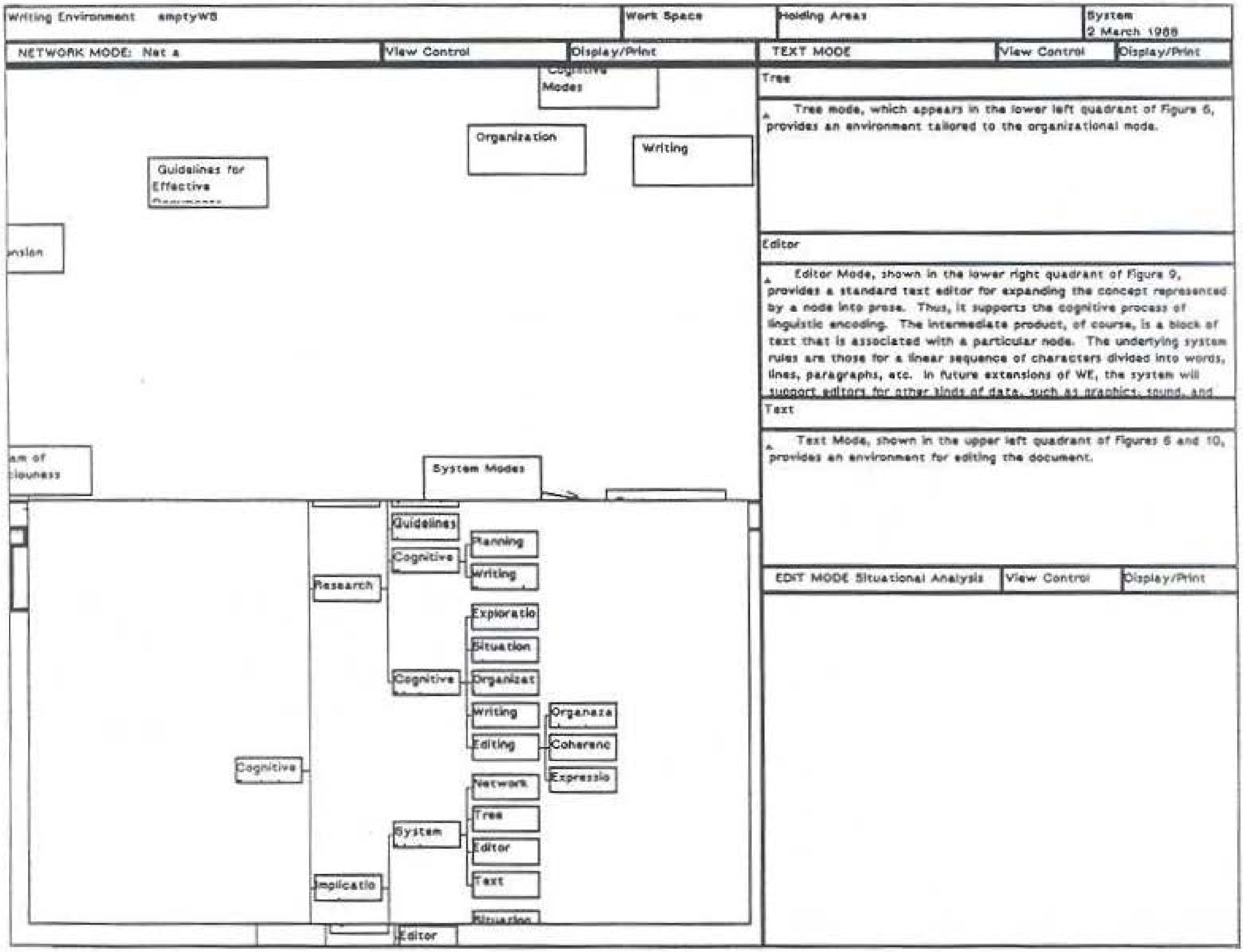


Figure 9:
WE Editor Mode



Editor Mode, shown in the lower right quadrant of Figure 9, provides a standard text editor for expanding the concept represented by a node into prose. Thus, it supports the cognitive process of linguistic encoding. The intermediate product, of course, is a block of text that is associated with a particular node. The underlying system rules are those for a linear sequence of characters divided into words, lines, paragraphs, etc. In future extensions of WE, the system will support editors for other kinds of data, such as graphics, sound, and video.

Figure 10:
WE Text Mode



Tree mode, which appears in the lower left quadrant of Figure 6, provides an environment tailored to the organizational mode.

Editor Mode, shown in the lower right quadrant of Figure 9, provides a standard text editor for expanding the concept represented by a node into prose. Thus, it supports the cognitive process of linguistic encoding. The intermediate product, of course, is a block of text that is associated with a particular node. The underlying system rules are those for a linear sequence of characters divided into words, lines, paragraphs, etc. In future extensions of WE, the system will support editors for other kinds of data, such as graphics, sound, and

Text Mode, shown in the upper left quadrant of Figures 6 and 10, provides an environment for editing the document.

Text mode constructs a representation of the continuous document by stepping through the tree – from top to bottom, left to right – interpreting each node label as a section heading for the block of text associated with that particular node. Writers traverse the tree, both forward and backwards, using a scroll bar attached to the side of the text mode window. As they move the scroll bar up and down, the labels and the blocks of text associated with the various nodes are moved into and out of the three areas of the text mode window. When they pause in their progression through the overall document, a second scroll bar attached to each of the three areas permits them to scroll through the text for that particular node. Thus, by scrolling to the bottom of one section and the top of the following section, writers can see how the text in two adjacent nodes fits together.

Within each area, they can edit the text for that node using the editor, just as in editor mode. They can also move text from one area/node to another, and they can edit section headings (node labels), as well. However, the node itself can't be deleted or moved from within text mode. This can be done only from tree mode.

While not its primary function, text mode also provides easy document browsing. Since it can be invoked not just from the root of the tree but from any node in the structure, the user can move around in the document quickly and easily using tree mode and then settle down to read a particular section using text mode. Thus, WE provides a form of hypertext.

Options Not Included in WE

Earlier, we discussed design decisions that led to incorporating various system functions in WE. Here, we describe several possible functions that we decided against. These include a possible mode for analyzing the rhetorical situation and, second, a mode for managing the various goals generated during writing.

Situational Analysis Mode

Writers must understand their readers and the rhetorical context for their document if they hope to communicate effectively. Consequently, we included in the cognitive modes shown in Figure 5a situational analysis mode that should be a part of any writer's planning. However, we did not include in WE a corresponding system mode. Instead, we drew the boundary of the system around the content of the document, *per se*. The system deals with ideas, relations, structures, text, and, soon, graphics. It does not help the writer analyze the rhetorical situation. For the present, we left this important concern to method and instruction.

One of us, in collaboration with Catherine F. Smith of Syracuse University, has developed a strategic method for writing [Smith & Smith, 1987] that includes three heuristic procedures to help writers turn implicit, dispersed knowledge of the rhetorical situation into explicit, usable insights. The first procedure helps writers identify the many different readers or kinds of readers that may read the document. The second helps them set priorities among readers and determine the limits of readers' expected prior knowledge of the document's subject matter. The third helps them evaluate change: how much change in knowledge and/or attitude should the document attempt to produce in order for the writer to attain his or her desired goals? These three heuristics are highly visual and could be incorporated into the system as an additional mode: situational analysis mode. At some future time, we may do so, but we want to gain more experience with the current system before extending its design to address extrinsic concerns.

Goal Management

Writing is a goal-directed activity. As noted above, Flower and Hayes suggest that writers generate a number of different goals as they relax and tighten constraints in order to produce different intermediate representations. We offered a somewhat different perspective. When writers adopt a particular mode of thinking, they do so in order to accomplish a specific task. That task is made concrete in the form of the intermediate products that can be developed in that mode. Thus, we view goals as an inherent part of the respective cognitive modes. Consequently, WE does not include separate functions for generating and managing goals, *per se*. Rather, it incorporates planning and goal-setting directly in the form of the specific, tangible products it supports in the respective system modes and the provisions it makes for their natural flow from one mode to another. Thus, the most important aspects of task management have been incorporated into system design rather than remaining a concern writers must consciously manage

Other Considerations

Space does not permit us to discuss a number of important, but less fundamental, design decisions. One of the most obvious is WE's spatial representation of structure and its direct manipulation controls. Thus, hierarchy is represented as a tree rather than as an outline. We regard the decision to use a spatial, versus linguistic, form as important, and we made it deliberately and with support from earlier cognitive studies. We are currently testing that assumption experimentally in a study of subjects' abilities to perceive, recall, and manipulate structures presented in different forms. We will review that literature as well as relevant decisions when we report those results.

Testing

In the first section of this paper, we reviewed the literature in cognitive psychology and composition theory in order to synthesize a cognitive basis for a computer writing environment. In the second, we showed how that basis influenced key design decisions for WE. While we believe our logic was sound, we also believe both the synthesis and the system should be tested. To help with this testing, we have developed three new tools.

First, we have included an automatic tracking function in WE. When turned on, it produces a detailed transcript for a session in which each action performed by the user is recorded along with the time and other relevant information, such as the location of a node for a create node operation. These data constitute a concurrent protocol that is gathered unobtrusively and in a machine-readable form, ready for analysis. Thus, these data avoid one of the most serious problems posed by think-aloud protocols – i.e., distortion of the user's cognitive processes [Nisbett & Wilson, 1977; Ericsson & Simon, 1980].

While these data can be analyzed directly, we use them with a second tool – a session replay program. Accepting the protocol data recorded by the tracker as input, the replay program reproduces the session so that the researcher and/or the user can observe it. Thus, we can watch a user's session unfold, in time that approximates the original session, "speeded up" or "slowed down," or we can manually step though the session, operation by operation. With this program, we can see factors such as the order in which the various system modes ere engaged; the operations that were used in combination; and the products that were constructed, their order of creation, and the particular transformations that turned one form into another. We can also observe patterns in the structure of ideas that led to recursive invocation of one mode or process from another. Thus, the replay program provides a valuable tool for analysis of protocol data by inspection.

It also provides a mechanism for gathering retrospective think-aloud protocols. This can be done by asking writers who produced the transcripts to observe their sessions and comment on their thinking and intentions for different operations or sequences. These protocols are, thus, gathered after-the-fact but in response to re-enactments of sessions completed just a short time before. While these protocols must be tested more thoroughly to establish their validity and reliability, we anticipate that the error introduced by re-enactment will be less than that produced by interference and delay for concurrent think-aloud protocols.

The third tool we are developing is a grammar to parse the protocols produced by the tracker. Since we consider this one of the most important tasks in our program of research, we will first describe the grammar itself and then its uses and implications.

In general, a grammar takes as its input a sequence of "terminal" symbols and produces as its output a parse tree that describes the structure of that sequence. The major constituents of the parse tree are "nonterminal" symbols that identify categories or patterns in the sequence of terminal symbols or in other lower level nonterminals. Thus, for a natural language such as English, the terminal symbols are the words and the nonterminals are categories, such as "noun" or "verb," or patterns, such as "noun phrase" or "verb phrase."

For our application, the terminals are the symbols produced by the protocol tracker that represent basic user actions, such as pointing to a particular node or selecting an option from a menu. The nonterminals identify patterns or categories, such as a "create node" operation comprised of the actions "point to the location for the node," "select" the create node option from the menu," and "type the name or label for the node." The resulting parse tree for some portion of the transcript identifies the kind of intermediate product being developed, the cognitive process being used, and the cognitive mode in which the writer is currently engaged.

To be more specific, our grammar is defined in terms of five levels of abstraction. The first level – the terminal symbols for the grammar – represents the user's actions. This is the protocol transcript produced by the tracker. The symbols representing those actions are mapped onto a second level of slightly more abstract symbols that identify operations, such as the create node operation described above. Operations are then mapped onto a third-level of symbols that represent intermediate products, such as isolated concepts, clusters, relations, structures, blocks of text, etc. At the fourth level, the grammar infers the cognitive processes being used by the writer to construct those products, such as recalling ideas from memory, associating them, or encoding them linguistically. Finally, the grammar infers the cognitive mode the writer is inhabiting at a particular time, such as exploring, organizing, or structural editing.

The grammar solves several problems posed by think-aloud protocols. First, its data reduction capabilities allow more efficient and extensive protocol analyses. A major problem posed by think-aloud protocols is the voluminous data they generate. The protocols generated by the WE tracker are also voluminous, but the grammar can reduce that information to manageable proportions. For example, a researcher interested in writers' global strategies might focus on their modal shifts. The grammar can produce a high-level representation of modal shifts for a session that would typically range from three or four to several dozen symbols – one for each shift. Since the data can be recorded and parsed automatically, the researcher can analyze a large number of protocols, for actual-use as well as experimental conditions. The grammar also makes practical longitudinal studies based on extensive protocol data.

Still another problem posed by think-aloud protocols is consistency of interpretation. Protocols are often incomplete, and subjects frequently describe their mental actions ambiguously. While techniques have been developed to increase the reliability of coders, the process is still frequently subjective. With our protocol grammar, the subjective element has been shifted from interpretation to rule definition. In order to write the rules that map symbols on one level onto symbols on another, we must interpret specific patterns. However, that interpretation is done once per pattern (within a given context) and it is explicit. Thus, the grammar rules can be debated, reconciled with subjects' verbal accounts, and modified; but once accepted, they become axiomatic. Thereafter, protocols will be interpreted by the grammar consistently and objectively, relative to those rules

Finally, the grammar constitutes a formal descriptive model of writers' cognitive interactions with the system. The grammar is a model since it characterizes writers' cognitive behavior with respect to WE. It is formal since it consists of a set of precise, logical rules for mapping from one set of well-defined symbols to another. It is descriptive since its symbols identify the cognitive modes engaged by the writer, the cognitive processes used, and the intermediate products defined or constructed.

In our discussion of Hayes and Flower, we suggested that to be considered valid a formal model should be tested and refined in response to actual protocols. The model we propose can be evaluated in several ways. First, since it is well-defined, it can be analyzed internally for consistency and ambiguity. That is, its rules can be analyzed to see if any contradict one another or if different rules interpret the same pattern differently. If so, rules can be modified or added to correct the grammar. Second, it can be calibrated with respect to think-aloud protocols. Since a session can be replayed and users asked to comment on their thinking, we can compare their verbal accounts with the characterizations produced by the grammar. If the two are inconsistent, we can probe writers further as to their intentions and, again, add or modify rules to make specific corrections. Third, we can test its adequacy. Since the grammar operates on concrete data – the protocols recorded by the tracker – any segments that cannot be interpreted by the grammar will reveal themselves in the form of symbol sequences that are not mapped to higher level symbols. Such instances will indicate that the model has not included some particular mental activity and will tell us where we need to add rules to do so.

A different kind of test involves utility. Does the grammar produce representations of writers' cognitive interactions with the system that are interesting and can be used to address significant questions? We believe so. We are just beginning to use these tools in a series of experiments and actual-use studies. Some of the questions that can be considered, and that we hope to answer, include the following:

- What cognitive processes are used in combination with one another?
- How are different processes distributed over the writing process as a whole?
- At what stage are various intermediate products created or transformed? Using which processes?
- What features of the conceptual structure trigger recursive invocation of one process from another? one mode from another?
- What are the specific differences in strategy between novice and expert writers?
- Which strategies produce more effective versus less effective documents?
- How do writers' strategies change over time?
    - What is the impact of instruction?
    - What is the impact of the writing system?
- Do the combinations of processes, products, goals, and constraints predicted by the concept of cognitive mode actually occur?

Thus, we believe our grammar/model can be refined in response to actual protocols and that it can address questions of sufficient interest for it to be considered useful. Like Hayes and Flower, we see it not as an end but as a starting point. Over the next several years we will test it and develop it.

Conclusion

In summary, we see our work as an integrated program of research that began with a description of the cognitive premises on which it is based. That cognitive basis was then used to guide the design of a computer writing environment that closely mirrors writers' mental function. Third, we developed new tools for studying writers working within a computer writing environment. Finally, we are designing experiments and actual-use studies to test the entire construct. The results will, no doubt, lead to refinements in the underlying cognitive basis, which, in turn, will lead to changes in the system, which will lead to . . . . The cycle of successive refinement we hope will lead to a better understanding of writing, thinking, and computing and their inherent interdependencies.

Acknowledgments

A number of organizations and people have contributed to the work reported here. We are grateful to the National Science Foundation (Grant # IRI-85-19517), and the Army Research Institute (Contract #MDA903-86-C-0345) for their support of various parts of this research. We wish to acknowledge our faculty colleagues, Stephen F. Weiss and Jay D. Bolter, for contributing ideas and perspective. Research Associate Gordon J. Ferguson was the principal technical designer of WE. We also wish to thank Prof. Catherine F. Smith of Syracuse University for her work on *The Strategic Method for Writing* and for sharing her views with us in numerous conversations. Finally, we wish to acknowledge the efforts and contributions of the students who have worked with us: Paulette Bush, Yen-Ping Shan, Irene Jenkins, Mark Rooks, Oliver Steele, Valerie Kierulf, and John Q. Walker. We are especially grateful to Greg Berg, who did much of the library work that underlies our review of research in cognitive psychology and composition theory.

References

Anderson, J. R. (1983). *The architecture of cognition.* Cambridge, MA: Harvard University Press.

Ausubel, D. P. (1963). *The psychology of meaningful verbal learning.* New York: Grune & Stratton.

Bereiter, C., & Scardamalia, M. (1987). *The psychology of written communication.* Hillsdale, NJ: Erlbaum.

Berkenkotter, C. (1983). Decisions and revisions: the planning strategies of a published writer. *College Composition and Communication*, 34, 156-169.

Bower, G. H., & Cirilo, R. K. (1985). Cognitive psychology and text processing. In T. A. van Dijk (Ed.), *Handbook of discourse analysis*, (vol. 1, pp. 71-106), London: Academic Press.

Bransford, J. D., & Franks, J. J. (1971). The abstraction of linguistic ideas. *Cognitive Psychology*, 2, 331-350.

Britton, B. K., Meyer, B. J. F., Hodge, M. H., & Glynn, S. M. (1980). Effects of the organization of text on memory: Tests of retrieval and response criterion hypotheses. *Journal of Experimental Psychology: Human Learning and Memory*, 6, 620-629.

Ericsson, K. A., & Simon, H. A. (1980). Verbal reports as data. *Psychological Review*, 87, 215-250.

Ericsson, K. A., & Simon, H. A. (1984). *Protocol Analysis: Verbal reports as data.* Cambridge, MA: MIT Press.

Flower, L. S., & Hayes, J. R. (1984). Images, plans, and prose: The representation of meaning in writing. *Written Communication*, 1, 120-160.

Hayes, J. R., & Flower, L. S. (1980). Identifying the organization of writing processes. In L. W. Gregg & E. R. Steinberg (Eds.), *Cognitive Processes in Writing* (pp. 3-30). Hillsdale, NJ: Erlbaum.

Hayes, J. R., & Flower, L. S. (1986). Writing research and the writer. *American Psychologist*, 41, 1106-1113.

Humes, A. (1983). Research on the composing process. *Review of Education Research*, 53, 201-216.

Kellog, T. T. (1984). Cognitive Strategies in writing. *Bulletin of the Psychometric Society*, 22, 287-292.

Kieras, D. E. (1980). Initial mention as a signal to thematic content in technical passages. *Memory and Cognition*, 8, 345-353.

Kintsch, W., & Greene, E. (1978). The role of culture-specific schemata in the comprehension and recall of stories. *Discourse Processes*, 1, 1-13.

Kintsch, W., & Keenan, J. M. (1973). Reading rate and retention as a function of the number of propositions in base structure of sentences. *Cognitive Psychology*, 5, 257-274.

Kintsch, W., & van Dijk, T. A. (1978). Toward a model of text comprehension and production. *Psychological Review*, 85, 363-394.

Matsuhashi, A. (1981). Pausing and planning: The tempo of written discourse production. *Research in the Teaching of English*, 15, 113-134.

Meyer, G. J. F. (1975). *The organization of prose and its effects on memory.* Amsterdam: North Holland Publishing Company.

Meyer, G. J. F., Brandt, D. M., & Bluth, G. J. (1980). Use of top-level structure in text: key for reading comprehension of ninth grade students. *Reading Research Quarterly*, 1, 72-103.

Nisbett, R. E., & Wilson, T. D. (1977). Telling more than we can know: verbal reports on mental processes. *Psychological Review*, 84, 231-259.

Schwartz, M. N. K., & Flammer, A. (1981). Text structure and title-effects on comprehension and recall. *Journal of Verbal Learning and Verbal Behavior*, 20, 61-66.

Smith, J. B., & Smith, C.F. (1987) *A strategic method for writing*. Chapel Hill, NC: UNC Department of Computer Science Technical Report TR87-024.

Thorndyke, P. W. (1977). Cognitive structures in comprehension and memory of narrative discourse. *Cognitive Psychology*, 9, 77-110.

Van Dijk, T. A. (1980). *Macrostructures*. Hillsdale, NJ: Erlbaum.

Williams, J. P., Taylor, M. B., & Ganger, S. (1981). Text variations at the level of the individual sentence and the comprehension of simple expository paragraphs. *Journal of Educational Psychology*, 73, 851-865.