Hypertext Writing Profiles and Visualisation

Margit Pohl Peter Purgathofer
Institute for Design and Technology Assessment
University of Technology Vienna
margit@igw.tuwien.ac.at purg@igw.tuwien.ac.at

Abstract

University students increasingly use hypertext to write their assignments. To employ hypertext effectively, more information about the hypertext authoring process is needed. There are features of hypertext which are not reflected in traditional theories of composition, especially the possibility to structure information visually. Our study indicates that graphical overview maps which can be edited are a rather attractive feature of hypertext authoring systems. Nevertheless, not all students profit from such features. Students employ different writing styles when they create hypertext documents. The majority of students experiments with hypertext's new features but others are still influenced by the model of the book.

Key Words: hypertext authoring, visualisation, written composition, organisation of knowledge

1. Introduction

Hypertext is increasingly used in university education. Many students write their assignments as hypertext documents. Some students find it difficult to do this (Pohl et al 1995). On the other hand, there are devices, for example tools for structuring knowledge visually, which might be helpful. A detailed analysis of the process of authoring hypertext should offer some information about the problems students encounter and possibilities to overcome these problems. Unluckily, no models about the hypertext authoring process exist so far. Models of the composition of linear text only have a limited value for the investigation of hypertext authoring.

One of the most influential models of written composition within the framework of cognitive psychology has been formulated by Hayes and Flower (see e.g. Hayes & Flower 1980, Flower & Hayes 1980). They describe writing as a process similar to problem solving. It is a goal-oriented activity during which several different constraints must be fulfilled more or less simultaneously. A similar model was developed by Bereiter and Scardamalia (1987). In a very general sense, the original Hayes-Flower model also encompasses the creation of hypertexts. Nevertheless, there are features of hypertext writing which cannot be captured by this model. It can be argued, for example, that the development of graphical overview maps (a feature

often found in hypertext documents) is an activity which cannot be integrated into the Hayes-Flower model without difficulties (Pohl & Purgathofer 2000). The same is probably true for the use of graphical elements in general. The hypertext "Writing Space" integrates text, graphical elements and other media in a way not possible in other media (Bolter 1991). Text becomes an element of visual space, and the visual organisation of knowledge conveys information. Manovich (2001) points out that the language of computer media is also heavily influenced by codes used in films. This tradition plays a dominant role in computer games but is also visible in other information systems. Movement in electronic space is, in some respects, a further development of the movement of a film camera. Increasingly, design for new media reflects a dynamic process, not a single fixed state (Maeda 2000). In this sense, it is more than the combination of graphics and text. Authors like Maeda present some first insights into the possibilities of this kind of graphical design.

There is some empirical evidence that structuring their ideas visually and interactively can help students to express their ideas (see section 3). Tools to support such activities - activities which are not reflected in the Hayes-Flower model - may form an important part of hypertext authoring systems. On the other hand, it is an open question whether all students profit from such tools. The following study tries to clarify these issues.

2. Writing

There is still fairly little research concerning writing within the framework of cognitive psychology (Eysenck & Keane 1990). Nevertheless, there has been some progress in the past two decades. It seems that the wide diffusion of computer technology plays an important role in this development. A considerable amount of research is motivated by the development of text editors or other software tools designed to support the writing process (see, for example, Hartley 1992, Sharples 1992). Models of the production of linear text, as they are developed in this kind of research, are certainly relevant for the understanding of writing hypertext. Nevertheless, there are modifications to be made to adapt these models to reflect the specific quality of creating hypertext documents.

One of the most well-known models of writing is the one developed by Hayes and Flower (see e.g. Hayes & Flower 1980, Flower & Hayes 1980). Hayes and Flower analysed a considerable number of protocols of expository writing activities and identified three subprocesses of composing texts: Planning, Translating and Reviewing. Planning consists of generating ideas, organising the material and setting goals, Reviewing consists of reading and editing. Translating refers to the activity of actually formulating sentences. Hayes and Flower assume that the writing process is similar to a problem-solving activity, therefore goal setting plays a central role in their model. They point out that their model is not a stage model in the sense that writers first engage in planning, then in translating and then in reviewing. The various subprocesses may occur fairly randomly during an individual writing session.

Hayes and Flower use protocol analysis as a method to investigate problems of the composing process. Protocol analysis is supposed to convey information about the sequence of activities an individual engages in during an experiment. It is obvious that these protocols

are incomplete because there are many phenomena about which a subject cannot or does not report. Nevertheless, Hayes and Flower (1980) assume that protocol analysis is a valuable instrument because it yields very rich data about the process of writing.

Hayes (1996) describes some important modifications to the original model (Hayes & Flower 1980). Still, this model cannot explain some specific traits of the hypertext authoring process, especially the integration of other media into hypertext documents. Nowadays, many hypertext documents contain, for example, graphical overview maps, although the original hypertext concept did not encompass such devices (Nelson 1992). Graphical user interfaces (GUIs) probably helped to make such devices popular. We assume that graphical overview maps are rather attractive because they represent the underlying structure of the hypertext document quite clearly. Activities related to the production of graphical overview maps are, for example, neither reflected in the model developed by Hayes and Flower (1980) nor in Hayes' revised model (1996). It must be mentioned that graphical forms of knowledge representation existed long before the introduction of computers. Some forms of diagrams were already used by Greek and Roman authors (Bonhoff 1993). During the last decades, concept mapping and other forms of visual structuring of knowledge were developed for pen and paper. The reason why such forms of knowledge representation were not considered in theories of composition is probably that the integration of text, pictures, and diagrams on paper is very difficult. Cognitive psychology has developed theories to explain the advantages of such forms of knowledge representation (Jonassen et al 1993), especially the theory of semantic networking. Such theories help to explain why devices such as clickable graphical overview maps might form a valuable part of a hypertext authoring system.

The use of graphical overview maps not only implies activities which are not considered in the Hayes-Flower model. It also has consequences for the definition of the subprocesses of the hypertext authoring process. In hypertext authoring it is not possible to distinguish between the three subprocesses in the Hayes-Flower (1980) model very neatly. Planning increasingly becomes part of the translation process. Concept maps or graphical overview maps are often used as planning tools even for linear texts on paper. Such methods help authors to create, clarify and order their ideas and concepts (Kellog 1992). The integration of such planning tools into hypertext authoring systems makes it difficult to distinguish between planning, translating and reviewing. The graphical overview map is an integral part of the final document whereas notes and concept maps produced during the planning stage of traditional writing are usually discarded after the document is finished. Many overview maps integrated into hypertext authoring systems can also be used for editing the document as such. Nodes and links can be created and deleted on the overview map. It is an open question whether these activities should be called planning, translating or reviewing because the overview map is at the same time a planning tool and a part of the final document. More obvious categories for analysing hypertext are 'creation of nodes', 'creation of links', 'writing text' or 'working on the overview map'. These categories do not refer to internal cognitive processes but reflect the technical properties of the system.

Hayes (1996) mentions that the original Hayes-Flower (1980) model owes very much to the work of Herbert Simon. Simon's assumptions about the process of human problem solving have been criticized by Winograd and Flores (1989). They point out that many everyday

decisions do not conform to Simon's assumptions about problem-solving processes. Especially his assumption that rational decision processes consist of searching for and evaluating various alternative solutions is not realistic. Winograd and Flores state that in many problem situations there is not enough time for a careful and scrupulous problem-solving behaviour. Decisions are rather made on the basis of background knowledge and previous experience. Sharples (1996) criticizes the Hayes-Flower (1980) model in a similar way. Following Simon, Hayes and Flower integrated his assumptions into their model. They, for example, posit that writing is a search for possible rhetoric solutions in long-term memory. Ideas and formulations are apparently already given. The problem is only to find the right idea and the right formulation.

The idea of writing as a problem-solving process was very influential. Still, there are contradictions this model cannot explain. The main contradiction seems to be that writing is on the one hand an analytic problem-solving process, yet also synthetic and productive (Sharples 1996). Related to this there are other contradictions as, e.g. that writers have to consider constraints – constraints which they have to transcend to produce creative and original texts. This does not conform to Simon's model of problem-solving because for him the space of possible solutions is always finite and given. Transcension of constraints would not be rational behaviour. Sharples suggests that the 'writing-as-problem-solving' model should be modified in such a way to explain these contradictions more easily. He sees writing as an open-ended design process, mediated by tools and resources. Writing and design have much in common insofar as both are creative processes.

There are important similarities between writing and design (Sharples 1996, p.136). Design problems are open-ended and cannot be fully specified. Sharples points out that writing, like design, does not lead to fixed sets of solutions of which one is optimal. There is, in principle, an infinite number of solutions, and no optimal solution can be determined with any certainty. The decision which solution to take is often a subjective one. There is no defined set of steps to reach a solution. In this sense the design process is endless. During the process of writing, problems are continually redefined, so that the final product may deviate considerably from the initial goal. Sharples posits that designers spend much time in identifying the problem, and problems and solutions emerge together. Design is always a practical activity changing concrete objects or existing social relationships.

Sharples (1996) adopts the Hayes and Flower model to a certain extent. His model also contains the processes of planning, reviewing and translating (called 'engagement' by Sharples). He modifies this model insofar as he adds another process which is called 'contemplation' ('Forming ideas, exploring and transforming conceptual spaces', Sharples 1996, p.143). Contemplation apparently is the activity which is responsible for the formulation of innovative and creative ideas. Sharples posits that this modification of the Hayes-Flower model makes it easier to resolve the above-mentioned contradictions. Writing is not only an analytic but also a synthetic, productive activity because contemplation can lead to the formulation of an indefinite set of solutions. These solutions often transcend the set of given conditions and constraints. In this sense, writing is more than problem-solving. It should be emphasized, however, that Sharples's model is not intended to replace the Hayes-Flower (1980) model. Writing as design still depends to a very large extent on problem solving

and the development of constraints.

Sharples himself points out (1996, p.147) that the concept of “writing as design” is related to the introduction of new writing technology. As mentioned above, electronic text and especially hypertext integrates text, pictures, audio and movies in ways not possible in other media. Text becomes part of graphic design or pictures, and the activities of writing and design become almost indistinguishable. The development of diagrams, pop-up fields, clickable maps or similar features of hypertext are based on a combination of writing and design. Text becomes part of the overall design and visual elements convey meaning (Schriver 1997). It might even be argued that these visual elements function like text. Typography on the one hand and techniques for the spatial representation of concepts and their interrelations play an important role in this process. It must be mentioned, however, that Sharple's argument is not restricted to electronic text. He assumes that the composition of any text (also traditional linear text on paper) is design because both are seen as creative and open-ended processes. The fact, that electronic text integrates text, graphics, audio and video very easily only made it more obvious that writing can be viewed as a design process.

Other authors also discuss the similarities between writing and graphical design/art. Kress and van Leeuwen (1996) try to develop a grammar of visual design. They identify major compositional structures and analyse how these are used to produce meaning. In this way, they want to establish a connection between the theory of language and the history of art. It is difficult to assess whether such approaches to integrate text and design/art may have a lasting effect on a definition of design, especially as there are many different definitions of design (see for example Buchanan & Margolin 1995).

We think that approaches which transcend the restrictions of the Hayes-Flower (1980) model are better able to reflect the specific character of the hypertext authoring process. Nevertheless, a detailed model of this process has yet to be formulated.

3. Visualisation

As mentioned above, the computer supports the convergence of writing and graphical design. The integration of various media is, for example, a feature of hypertext/hypermedia which makes it so attractive to very diverse groups of users. Graphical design is only one form of integration of media, although it is probably still the most important one. It offers many different possibilities to convey information to the users: pictures, diagrams, maps, typography, etc. Graphical browsers or (clickable) overview maps seem to be one of the most important graphical elements in hypertexts because they are a unique feature of this medium which does not exist in books or other older media. Today, many hypertext documents contain overview maps. This is no coincidence. The distributed, fragmented nature of hypertext emphasizes the importance of structure and relationships between the units of information contained in nodes. This structure is best represented in a topological manner. Chen (1999) and Chen and Czerwinski (1997) point out that spatial representation belongs to the most important ways of information visualisation in hypertext development. Spatial aspects of objects on the screen (especially distance and position) are meaningful and convey knowledge. Such forms of graphical design are only very seldom used in books.

Marshall, Shipman and Coombs (1994) argue that a visual/spatial metaphor for hypertext enables people to express nuances of structure more easily than in other hypertext systems. They posit that using spatial metaphors in hypertext systems takes advantage of the ability of humans to navigate in physical space. The system they developed, VIKI, is supposed to support implicit or emergent information structure. In this way, hypertext may help groups of cooperating individuals to develop a shared meaning of a domain. They state that "externalizing these relationships, assessments, and interpretive abstractions is a crucial part of reflecting on one's own understanding of a problem and communicating one's understanding to others during the course of collaborative work." (Marshall et al 1994, p.15). The practice of developing a shared understanding of multiple information sources becomes increasingly important (Marshall and Shipman 1997). Modern information systems provide an abundance of documents so that people lack the time to read all these documents thoroughly. Marshall and Shipman assume that hypertext systems using a spatial metaphor help people to organise and structure this information. Their empirical investigation supports this assumption.

Visualisation of document structure does not only help users to find their way in hypertext documents; it can also be used to support learning processes. Semantic networking or concept mapping helps students to analyse their own knowledge structures to integrate new information with what they already know. As mentioned above, in many hypertext authoring systems, especially in those used in an academic context, hypertext and semantic networking or concept mapping are integrated features. There is some evidence that the use of concept maps in learning environments can have beneficial effects. Kozma (1992), Kommers & de Vries (1992) and Fisher (1992) who used computer based concept mapping tools report positive results. Reader and Hammond (1994) argue that concept mapping is a more time consuming method than traditional forms of learning. Thus, students are motivated to spend more time going through the material and analysing it. This leads to a better performance of those students developing concept maps. McEneaney (1999, 2000) investigated navigational patterns in hypertext documents. He found out that there was a relationship between traversal patterns and scores reached on a search task. Those subjects who adopted a more linear approach had significantly lower scores than those subjects who adopted shallow hierarchical patterns of movement. He concluded that students who take advantage of the structure of the hypertext documents are more successful.

4. Description of the Project

Our research in the field of hypertext authoring is part of a long term study. The goal of this study is to gain insight into the application of hypertext-based information systems in teaching and learning. In the course of this project we ask our students to write their seminar papers in hypertextual form instead of the traditional, linear paper form. The students study computer science in their third or fourth year and are therefore acquainted with the use of computers. In addition, they are interested in the application of new methods and technologies.

The seminars cover a wide range of topics; technological and technology assessment issues as

well as social sciences, psychology, information theory, and design. The task of students in our seminars is, for example, to read a couple of scientific publications and present the underlying issues to the other participants. In addition, they have to prepare a written assignment, traditionally about 20 to 30 pages of text. The goal of such a seminar is to enable the students to discuss scientific texts from various points of view, and to teach students how to communicate their ideas in oral and written form.

One of the problems we had with this traditional approach was that it is hard or impossible to judge from the text whether the student has actually understood the issues, or was just copying quotes in a clever way. Another problem was that the papers written by students tend to be “archived”, while the information within might be a useful source of reference for further work.

Letting the students write (or design) hypertext instead of linear text has the potential to overcome both of these problems; when writing hypertext, students have to define a structure that reflects their own structural understanding of the issues involved, and how these issues are interconnected. Our assumption was that working interactively with the material would generate some higher level knowledge about the main ideas of the topics involved. Additionally, all text was searchable, and information was presented in a segmented, reusable way. That way, each document contributed to a growing body of hypertextual information.

The first steps were taken using HyperCard on the Macintosh. We found out, quite painfully, that students (who, as computer science students, love to program, but hate to write) get entangled in the programming necessary to generate the structure so much that it actually makes them change the content in favour of an easier way to implement structure. Our conclusion was to offer students an authoring tool specifically tailored to their needs, the needs of untrained hypertext authors. The system would implement the following concepts and features:

- A graphical overview editor that lets the author arrange the nodes and links with maximum freedom, thus building a semantic information network.
- A node editor that restricts the authors from making common mistakes (too much text, too many fonts and styles, too little structure).
- The possibility to change between node editing and structural editing at any given time.
- A meaningful direct manipulation interface to create nodes and define links between them without necessity to change from “writing” to “programming” in order to accomplish these basic tasks.

4.1 The Authoring Tools

Since no then existing hypertext environment satisfied our requirements, we decided to develop our own authoring tools. On top of Apple’s “HyperCard” for the Macintosh, we implemented a hypertext authoring tool, embedded into a simultaneously developed multi-user networking environment.

With these tools, hypertext authoring was elevated, from a complex task that is closely related to programming, to a high-level, interactive task embedded into a context of semantically appropriate tools and actions. Many constraints were introduced in order to liberate the authors from certain design decisions, and to constitute a common look and feel within the finished hypertext documents.

Basically, the tool consisted of two elements which are described in the following; a node editor, helping the authors to write text, create link anchors and embed media into nodes, and an overview editor where nodes and links could be created and arranged freely.

Figure 1 shows a typical node view of a node in work.

Fig. 1. Example of a node in the node editor view.

The node editor offered a fixed, simple layout consisting of a node header (top left), the document header (top right), a main text area (right) and a sidebar for subheads, comments and link placement. In the main text area, the authors had no choice of text fonts or styles other than plain or bold. While this decision might look like a harsh reduction of the authors' possibilities, we found that in fact it was expanding their possibilities: released from the need to choose a font and a style, authors were able to dedicate their time and attention to other decisions. Because the system was based on HyperCard, authors could place black-and-white graphics anywhere, regardless of the layout. We observed, however, that the visible borders separating the various parts of the layout were also used to crop inserted pictures, and that these images were also mostly aligned in respect to the text on a node (see Fig. 1).

The toolbar at the bottom of the screen offered drag&drop functionality to create and delete links via simple drag-and-drop operations. Fig. 2 explains the main set of tools.

Fig. 2: Node tools for authors explained.

As another restriction, link anchors had to consist of both an icon (showing its type) and a subtitle (explaining the target). Together, these two elements should give an impression of what a reader should expect when following a link. Authors could not, however, anchor links within the text (as it is most common within HTML).

On the very right, some basic navigation (title node, overview map, help and last visited node) was provided. Note that no “next node” and “previous node” type of navigation was provided. Earlier versions of the tool featured back and forward buttons for authors (not for readers), and we observed that authors arranged information according to the linear structure suggested by these buttons. Since this was an unwanted side effect, we prevented linear navigation for authors. Finally, there was a mode switch between author and reader view in the top right of the window. In reader view, all but the necessary tools are remove. Fig. 3 shows the same node as Fig. 1, but in reader mode and with an open text annotation.

Fig. 3: a node in reader mode with one text annotation open.

In addition to the above described toolbar functions, a number of administrative function were available via the usual set of menus. Among others, new nodes could be created via the menus.

The second major module of the authoring tools was the map editor, or overview map. The basic function of the overview map was to give authors an instrument to create links and nodes in a structural way, and let them arrange nodes freely within the resulting map. No automatic or semi-automatic arrangement was proposed; the layout of nodes and the resulting

network of links is completely under control of the author. When new nodes were created from the node editor via a menu command, their representation in the map was added at a random height near the left border. Because this positioning resulted in unreadable overview maps, authors were forced to find a proper position for every node. Fig. 4 shows a typical overview map of a finished hypertext, in author mode.

Fig 4: The overview map as authors see it.

Again, the visible tools provided for a simple and meaningful way to create and delete nodes and links. Nodes could be created via drag&drop from the node tool, the link tool let you define connections between nodes using a rubber-band metaphor. Nodes could be deleted by dragging them into the trash. Basic document information was available via the “i”-button. While there were no restrictions for the author in arranging of nodes on the map, the arrangement was also without any consequence for the individual nodes, or the content therein.

Links defined in the overview map had to be anchored on the node where the link originated, ie. assigned a link type icon and a link name, and placed within the node. For this, all unanchored links showed up in a separate list that is visible on every node where there are such links. The author could then select each link, define its respective type and place its individual anchors on the node.

To enable authors to test their work without constantly switching between author and reader modes (which took several seconds time), a modifier key was dedicated to the emulation of reader mode in author’s mode. Whenever an author held down the “option”-key, she was in a “reader-within-author”-mode that let her test her layout. For authors, that meant a strong motivation to constantly test the hypertext document for readability and usability.

4.2 Monitoring Tool

In order to enable us to evaluate the authoring process, all actions authors took, either on a node or in the overview map, were registered by a monitoring tool. The syntax used for event recording allowed for basic information to be provided with each event. An example for a resulting log is shown in Figure 5.

ET "structural vs. functional models",1289
NL 398,397,2887290994
NL 397,367,2887291001
NL 387,367,2887291006
ET "restaurant script",816
ET "knowledge organization",1145

Fig 5: A short snippet of an event recording result, showing a sequence of basic actions (edit node text, create three links, edit more node text)

In total, we were able to record around 150 protocols, 95 of which turned out to be correct and usable, containing more than 48.000 single events, or a statistical mean of 408 events per protocol. 55 protocols turned out to be either incomplete or erroneous due to version problems or software bugs. From these dynamic data we can retrieve some simple statistical information like the average node and link “age” (relative to the number of recorded events) in a document, the mode of node creation (from the node view vs. in the overview map), etc. In addition, we extracted some static information from the documents, like the number of nodes and links, and the compactness (a graph theory benchmark expressing the overall average path length when travelling within a graph).

The real challenge however is to find evaluation methods for the dynamic data that allows deeper insight into the authoring process. In the following, we will describe a method we used to evaluate these protocols.

4.3 Evaluation: Moving Window Analysis

The moving window analysis tries to form the basis for a quantitative analysis of the data. The problem we have with statistically analyzing the dynamic data is that they are impossible to normalize, reduce or categorize without losing some or most of the information within: the dynamic change. With the “moving window” method we try to generate a set of data from the protocols that makes it easier to apply statistical methods. The focus of this method is to understand the basic strategies of the author.

To get an analysis of granularity g of a protocol with N entries, we define a “window of observation” with a length of $l=2N/(g+1)$. We start by counting all different classes of actions within the first l actions of the protocol, resulting in an accumulation vector v_1 . Next we “move” the window by an offset of $\Delta=N/(g+1)$ - half the window size - and repeat the counting, getting v_2 . This procedure is reiterated until the g accumulation vectors are acquired. The result of this process, the accumulation vectors v_1-v_g or more conveniently the accumulation matrix V , can be visualized using a stacked bar graph. In this graph, each bar shows the relative importance of different classes of actions within an interval of $2(g+1)$ actions, overlapping each of the neighboring intervals by 50%. In this graph we can then see the relative importance each class of actions had for the author throughout the creation of the document.

The accumulation matrix V can be seen as some kind of redux of the authoring process. As we show later, different authoring strategies can be seen quite convincingly. Of course, V is a coarse reduction of the rich information within the original data, leaving behind much detail eventually worth taking into account. However, V can easily be normalized, enabling us to apply more sophisticated statistical methods like cluster analysis.

For the investigation described later in this text, we generated moving window analysis with granularity 19, counting five classes of actions: text editing (writing and deleting text), make (creating nodes and links), moving (positioning nodes in the overview map), delete (deleting nodes and links), and other actions (eg. renaming nodes).

5. Results

The aim of our investigation was to find out whether the hypertext authoring process has specific characteristics compared to composing linear text. In this context, we are especially interested in the use of graphical overview maps. If these overview maps are really utilised by the authors, the process of composing text might be changed quite radically. The development of overview maps might motivate authors to contemplate their documents on a more abstract level than when they write text on paper. Our practical experience with using hypertext in university education showed us that students who create hypertext documents often rely very heavily on the model of the book. Therefore, we wanted to analyse whether students could be motivated to use novel features of hypertext. It must be mentioned, however, that documents with a novel approach are not automatically “better” or more understandable than documents which resemble a book on a computer. Still, we think that students should explore the new possibilities offered by this medium and perhaps gain new insights in this process.

The data which we used were generated by the monitoring tool described in the previous section. The basis for the data were 95 documents. The categories of analysis were edit text, make, move, delete and other activities. “Edit text” refers to the traditional process of writing text in a text field on a node, “make” refers to making links or nodes, “move” describes those activities which are related to moving nodes around on the overview editor, “delete” refers to revision of the document (especially deleting nodes and links) and “other actions” are all actions not covered in the other categories (e.g. renaming nodes etc.). The categories include all the activities recorded by the monitoring tool. The categories distinguish between activities like writing text and other activities like making nodes and links or designing one’s overview map. We assume that students who write quite a lot of text and spend little time to design their overview map use books or assignments on paper as a model for their writing process. Students who concentrate on working on the overview map and make many links and nodes, on the other hand, seem to be more innovative and explore the possibilities offered by the new medium.

Figure 6 shows the average values of all analysed documents for 19 consecutive periods. The whole period of hypertext creation is segmented into these 19 phases. This figure is rather misleading as it conveys the impression that all activities are distributed evenly across all periods (from 1 to 19). The figures for the single documents are much more diverse. Students very often concentrate on one activity in one period or in a succession of several periods which makes the diagrams look more chaotic (see below). Fig. 6 also demonstrates that in our sample there were no overall temporal patterns in the process of composing hypertext documents.

Fig. 6: Average values of all analysed documents/relative proportion of activities

Figure 6 shows, on the other hand, that the visual appearance of their overview maps is very important for the students. A relevant part of their activities is made up of moving nodes around on the overview map (the links automatically move with the nodes). This is also supported by other data from our analysis. The monitoring tool we developed can distinguish between activities which took place ‘on the overview map’ and activities which took place in

the document itself. It was possible to make nodes ,on the overview map‘ by drag and drop or by clicking on a ,node symbol‘. Another method to make nodes would have been by using the menu which is the usual Hypercard method. In 87,85% of all cases students created nodes on the overview map. Again, this is an indication that the visualisation of the information is important for authors of hypertext documents, at least for the students of the sample we investigated.

We conducted a cluster analysis to find out whether it is possible to distinguish between different styles or writing profiles of hypertext writing. We used the five categories described above as variables for the cluster analysis. The cluster analysis used overall data for the whole document. The data for all the 19 periods were summed up and then analysed.

The cluster analysis yielded five clusters. We will describe these five clusters in the following section and discuss the most typical case from each cluster. We noticed that the smaller groups contain the more extreme cases.

Cluster 1

Cluster 1 forms an average sized group (n=16). Members from this cluster spent most of their time in making and deleting nodes and links. They wrote very little text and tended not to change the appearance of their overview maps. They apparently tried to explore the possibilities of the new medium by experimenting with nodes and links.

Fig. 7: Example for cluster 1 (relative proportion of activities)

Cluster 2

Cluster 2 is the biggest cluster (n=48). This cluster is not very homogeneous. The members of this cluster write fairly little text. The values for making and deleting nodes are average. It should be noticed, however, that they spend quite a lot of time moving nodes around on the overview map. This probably means that the appearance of the overview map is very important for them. These tendencies are not very pronounced, and most of the values are

very near the average of all 95 documents.

Fig. 8: Example for cluster 2 (relative proportion of activities)

Cluster 3

Cluster 3 is also a fairly large group (n=23). Members of this group write quite a lot of text and avoid changing the layout of their overview maps. The value for making and deleting nodes and links is fairly near the average of the whole group. It is not a group with a distinct profile although a tendency to write more text than the students who wrote the documents belonging to cluster 1 and 2 can be observed.

Fig. 9: Example for cluster 3 (relative proportion of activities)

Cluster 4

Cluster 4 is very small (n=3). The tendencies in this cluster are similar to cluster 3 but much more extreme. These documents contain a lot of text. Students avoid moving nodes and links

on the overview map as much as possible. They are probably not interested in the appearance of the overview maps of their documents. They seldom delete any nodes. The members of this group do not deviate very much from the model of linear text on paper. This is also supported by informal evidence from their documents. The documents have a clear hierarchical structure similar to the table of contents of a book. Non-hierarchical links are seldom used.

Fig. 10: Example for cluster 4 (relative proportion of activities)

Cluster 5

Cluster 5 is also a very small ($n=5$) and extreme group. The members of this group write fairly little text and seldom make or delete nodes. They spend most of their time moving nodes on the overview map. The value for „move“ deviates very much from the average. Informal evidence from the documents shows us that this group is very heterogeneous. Two of the documents are very confused. The overview maps contain quite a few nodes without links. One gets the impression that the students did not really know what to do. The other three documents have very clear and well-designed overview maps. The design of these overview maps conveys the impression that the students tried to express many of their ideas about the subject of the document with the help of the overview map.

Fig. 11: Example for cluster 5 (relative proportion of activities)

To a certain extent, our results can be compared to the work of van Waes (1992) and van Waes and Schellens (1998). They also conducted a cluster analysis to assess different writing styles. In contrast to our study, they compared writing with pen and paper and writing linear text using a computer. They found out that writing profiles are influenced by the media used. Authors who use pen and paper plan very extensively. When they start writing, only few and long breaks are observable. Authors who use a computer often adopt a fragmented style of writing. They usually do not concentrate on one activity (planning, transcribing, revision) but skip from one activity to another very quickly.

Our research interest in questions of hypertext style can also be compared to the work of Handler (1997). Handler studies the question of the style of scientific electronic documents from a linguistic point of view. He argues that there is no defined set of conventions for scientific electronic text. These conventions are emerging right now. He suggests a few variables which might be relevant for style conventions – e.g. size and role of textual nodes or function of links. Empirical studies might be a valuable input for theoretical discussions like this.

6. Conclusion

Several conclusions can be drawn from this study. First, our results seem to indicate that graphical overview maps as a means of structuring ideas in a visual way are, in general, a very attractive tool. Students belonging to our sample predominantly created links and nodes “on the overview map”. The reason for this might be that the overview map gives students a clear picture of the whole document and the relations of nodes to each other. Nevertheless, not all students profit from graphical methods of knowledge representation. Our research indicates that there are different styles of writing hypertext documents. A minority of students is influenced by the model of linear text. They concentrate on writing text and use the overview map like a table of contents. They edit their overview map only minimally. This implies that they do not use space as a means of conveying knowledge. The majority of students

experiments more with the possibilities hypertext authoring systems offer. They either concentrate on making nodes and links or on editing their overview maps. These results imply that the Hayes-Flower (1980) model cannot be applied easily to the hypertext authoring process. A third result of our investigation is that no overall temporal patterns could be observed. A plausible assumption would be that authors first create nodes and then write text. This is not the case. In general, our results seem to indicate that an easy-to-use tool to develop clickable overview maps can help students to express their ideas.

It is obvious that our results are influenced by the specific features of our authoring tool. The analysis of the creation of HTML-documents will probably lead to different results. A possible subject of further research could be the comparison of data based on different authoring tools. Another possible area of research is a comparison of the quality of the documents in the different clusters. It would be interesting to know whether different strategies of writing hypertext documents imply a different organisation and design of the documents. Consistent results in this area could form the basis for guidelines for the creation of hypertext documents.

Reference

- Bereiter, C., Scardamalia, M. (1987) *The Psychology of Written Composition*. Hillsdale, N.J.: Lawrence Erlbaum
- Bolter, J.D. (1991) *Writing Space. The Computer, Hypertext, and the History of Writing*. Hillsdale, N.J.: Lawrence Erlbaum
- Bonhoff, U.M. (1993) *Das Diagramm. Kunsthistorische Betrachtung über seine vielfältige Anwendung von der Antike bis zur Neuzeit*. Münster (Dissertation)
- Buchanan, R., Margolin, V. (eds.) (1995) *Discovering Design: Exploration in Design Studies*. Chicago: University of Chicago Press
- Chen, C. (1999) Visualising Semantic Spaces and Author Co-Citation Networks in Digital Libraries. *Information Processing and Management* **35**, 3, pp.401-420
- Chen, C. & Czerwinski, M. (1997) Spatial Ability and Visual Navigation: an Empirical Study. *The New Review of Hypermedia and Multimedia*, **3**, pp.67-89
- Eysenck, M.W., Keane, M.T. (1990) *Cognitive Psychology- A Student's Handbook*. Hove and London, UK: Lawrence Erlbaum
- Fisher, K. (1992) SemNet: A Tool for Personal Knowledge Construction. In P.A. Kommers, D.H. Jonassen, J.T. Mayes (Eds.) *Cognitive Tools for Learning*, pp.63-75, Berlin, Heidelberg, New York: Springer

- Flower, L., Hayes, J.R. (1980) The Dynamics of Composing: Making Plans and Juggling Constraints. In: L.W. Gregg, E.R. Steinberg (eds.). *Cognitive Processes in Writing*. Hillsdale, N.J.: Lawrence Erlbaum, p.31-50
- Handler, P. (1997) Stileigenschaften elektronisch vermittelter Wissenschaftstexte. In: D. Knorr, E.-M. Jakobs (eds) *Textproduktion in elektronischen Umgebungen*. Frankfurt/Main: Peter Lang
- Hartley, J. (1992) Technology and Writing. Readings in the Psychology of Written Communication. London and Philadelphia: Jessica Kingsley Publishers
- Hayes, J.R. (1996) A New Framework for Understanding Cognition and Affect in Writing. In: C.M. Levy, S. Ransdell (eds.) *The Science of Writing. Theories, Methods, Individual Differences and Applications*. Mahwah, N.J.: Lawrence Erlbaum, p.1-27
- Hayes, J.R., Flower, L. (1980) Identifying the Organization of Writing Processes. In: L.W. Gregg, E.R. Steinberg (eds.) *Cognitive Processes in Writing*. Hillsdale, N.J.: Lawrence Erlbaum, p.3-30
- Hayes, J.R., Nash, J.G. (1996) On the Nature of Planning in Writing. In: C.M. Levy, S.Ransdell (eds.) *The Science of Writing. Theories, Methods, Individual Differences and Applications*. Mahwah, N.J.: Lawrence Erlbaum, p.29-55
- Jonassen, D., Beissner, K., Yacci, M. (1993) Structural Knowledge. Techniques for Representing, Conveying, and Acquiring Structural Knowledge. Hillsdale, N.J., Hove and London: Lawrence Erlbaum
- Kaplan, N., Moulthrop, S. (1994) Where no mind has gone before. Ontological design for virtual spaces. In: *ECHT '94 Proceedings, ACM European Conference on Hypermedia Technology*, p.206-216
- Kellog, R.T. (1992) Designing Idea Processors for Document Composition. In J. Hartley, (ed.) *Technology and Writing. Readings in the Psychology of Written Communication*. London, Philadelphia: Jessica Kingsley Publishers, pp.181-199
- Kommers, P.A.M., de Vries, S.A. (1992) TextVision and the Visualisation of Knowledge: School-based Evaluation of its Acceptance at two Levels of Schooling. In P.A. Kommers, D.H. Jonassen, J.T. Mayes (eds.) *Cognitive Tools for Learning*. Berlin. Heidelberg, New York: Springer, pp.33-62
- Kozma, R.B. (1992) Constructing Knowledge with Learning Tool. In P.A. Kommers, D.H. Jonassen, J.T. Mayes (Eds.) *Cognitive Tools for Learning*. Berlin, Heidelberg, New York: Springer pp.23-32,
- Kress, G., van Leeuwen, T. (1996) Reading Images. The Grammar of Visual Design. London, New York: Routledge

- Maeda, J. (2001) *Maeda @ Media*. New York: Universe Books
- Manovich, L. (2001) *The language of new media*. Cambridge, Mass: MIT Press
- Marshall, C.C, Shipman, F.M., Coombs, J.H. (1994) VIKI: Spatial Hypertext Supporting Emergent Structure. In *Proceedings of the European Conference on Hypermedia Technologies* (Edinburgh, Scotland), pp. 13-23
- Marshall, C.C, Shipman, F.M (1997) Spatial Hypertext and the Practice of Information Triage. In *Proceedings of the Hypertext '97* (Southampton, UK), pp. 124-133
- McEneaney, J.E. (1999) Visualizing and Assessing Navigation in Hypertext. In: *Proceedings of the Hypertext '99 conference*, New York: ACM, p.61-70
- McEneaney, J.E. (2000) Navigational Correlates of Comprehension in Hypertext. In: *Proceedings of the Hypertext '00 conference*, New York: ACM, p.254-255
- Nelson, T. (1992) *Literary Machines 93.1*. Sausalito, CA: Mindful Press
- Pohl, M., Purgathofer, P., Prenner, P. (1995) Hypermedia in Education - Monitoring the Development of Hypermedia. In: *Psychology Teaching Review* (1995) **4**, 2, p.142-152
- Pohl, M., Purgathofer, P. (2000). Hypertext Authoring and Visualisation. In: *International Journal of Human-Computer Studies* (Special Issue on Information Visualisations), Vol. **53**, No.5, November 2000, p.809-825
- Reader, W. Hammond, N. (1994) Computer-Based Tools to Support Learning from Hypertext: Concept Mapping Tools and Beyond. *Computers Educ.*, **22**, pp.99-10
- Schrivver, K.A. (1997) *Dynamics of Document Design*. New York, Chichester, Brisbane: John Wiley
- Sharples, M. (1992) *Computers and Writing. Issues and Implementations*. Dordrecht, Boston, London: Kluwer
- Sharples, M. (1996). An Account of Writing as Creative Design. In: C.M. Levy, S. Ransdell (eds.) *The Science of Writing. Theories, Methods, Individual Differences, and Applications*. Mahwah, N.J.: Lawrence Erlbaum, p.127-148
- Waes, L.van (1992). The influence of the computer on writing profiles. In: H.P.Maat & M. Steehouder (eds.) *Studies of functional text quality*. Amsterdam, Rodopi
- Waes, L. van, Schellens, P.J. (1998) Writing Profiles. The effect of the word processor on pausing and revision patterns. Working Paper 98/261, UFSIA, University of Antwerp, Department of Business Economics

Winograd, T., Flores, F. (1989) Erkenntnis Maschinen Verstehen. Berlin: Rotbuch Verlag